Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine CONFERENCE PAPER ARCHIVE

IFAC PapersOnLine 58-18 (2024) 392–399

Real-time MPC with Control Barrier Functions for Autonomous Driving using Safety Enhanced Collocation ⋆

Jean Pierre Allamaa [*,**] **Panagiotis Patrinos** [**]
Toshiyuki Ohtsuka [***] **Tong Duy Son** [*]

[*] *Siemens Digital Industries Software, 3001, Leuven, Belgium (e-mail: {jean.pierre.allamaa, son.tong}@siemens.com)*
[**] *Dept. Electr. Eng. (ESAT) - STADIUS research group, KU Leuven, 3001 Leuven, Belgium (e-mail: panos.patrinos@esat.kuleuven.be)*
[***] *Department of Informatics, Kyoto University, Kyoto, Japan (e-mail: ohtsuka@i.kyoto-u.ac.jp)*

Abstract: The autonomous driving industry is continuously dealing with safety-critical scenarios, and nonlinear model predictive control (NMPC) is a powerful control strategy for handling such situations. However, standard safety constraints are not scalable and require a long NMPC horizon. Moreover, the adoption of NMPC in the automotive industry is limited by the heavy computation of numerical optimization routines. To address those issues, this paper presents a real-time capable NMPC for automated driving in urban environments, using control barrier functions (CBFs). Furthermore, the designed NMPC is based on a novel collocation transcription approach, named RESAFE/COL, that allows to reduce the number of optimization variables while still guaranteeing the continuous time (nonlinear) inequality constraints satisfaction, through regional convex hull approximation. RESAFE/COL is proven to be 5 times faster than multiple shooting and more tractable for embedded hardware without a decrease in the performance, nor accuracy and safety of the numerical solution. We validate our NMPC-CBF with RESAFE/COL on digital twins of the vehicle and the urban environment and show the safe controller's ability to improve crash avoidance by 91%. Supplementary visual material can be found at `https://youtu.be/_EnbfYwljp4`.

Copyright © 2024 The Authors. This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/)

Keywords: Real-time NMPC, Control Barrier Function, Autonomous Driving

1. INTRODUCTION

Advancements in autonomous driving (AD) and automated vehicles (AV) have been weakened by the acceptability of such intelligent components by human beings. This is mainly due to the safety-linked nature of AVs that are required to navigate with optimal performance, even in safety-critical situations such as urban driving in the presence of obstacles. Therefore, the AD motion planning module often needs to control the system at the limit of states and inputs. Nonlinear model predictive control (NMPC) is one solid control strategy allowing optimal and deterministic motion planning that optimizes over possibly conflicting stability, safety, and control objectives. There exists abundant work in the literature to tackle AD collision avoidance and NMPC has proven to be one promising approach, however, most of the work relies on simplified models, limited scenarios, and consider distance constraints with Euclidean norms. Although successful, the scalability of those approaches towards different applications, models (for e.g. shorter horizon), and embedded hardware deployment, is tight. In particular, the usage of a fast NMPC in the control loop is hindered by two main challenges: 1) lack of general methods to tackle urban driving with long planning horizons, and 2) the computation time of the underlying numerical optimization.

One possible contained approach to combine a mid-level planner for collision-free trajectories, and a low-level control for trajectory tracking is known as control barrier functions (CBF). CBF approaches provide safety guarantees for the trajectories of the continuous-time nonlinear system beyond the predicted horizon in the NMPC through set invariance that exploits a Lyapunov-like condition for safety. CBFs have shown to be advantageous over other methods, when used within the MPC formulation (Zeng et al. (2021); Son and Nguyen (2019)), or as a safety filter for end-to-end control approaches (Cosner et al. (2022)), yet not real-time capable. There are proven computationally efficient CBF formulations for collision avoidance as in He et al. (2021), but they do not scale well for urban driving and long planning horizons. Therefore, in this paper we present a real-time capable NMPC-CBF, for simultaneous path tracking and collision avoidance allowing to safely navigate through urban environments, while satisfying spatio-temporal input/state constraints.

⋆ This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement ELO-X No 953348 and the Flemish Agency for Innovation and Entrepreneurship (VLAIO) under research project No. HBC.2021.0939 (BECAREFUL).

2405-8963 Copyright © 2024 The Authors. This is an open access article under the CC BY-NC-ND license.
Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2024.09.058

In most MPC applications, the continuous-time model evolution equations are discretized using an integrator such as the Runge-Kutta 4^{th} order method (RK4). The original optimal control problem (OCP) can be transcribed into a nonlinear programming problem (NLP) using a popular approach known as Direct Multiple Shooting (DMS). For long predictions and dense time sampling, the discrete approaches result in a high number of optimization variables, making MPC challenging for real-time and embedded control, and tradeoffs are often made between the computational load and the closed-loop safety and performance. To circumvent the weakness of direct and discrete approaches, the collocation transcription method is used, allowing to reduce the number of optimization variables by posing the states and input trajectories as polynomials most famously with Pseudospectral collocation (PSC) that benefits from a fast convergence rate (Huntington (2007)). However, similar to DMS, PSC enforces safety constraints only at finite nodes. To tackle this issue, collocation based on B-splines with convex hulls have been used for collision avoidance as in Cichella et al. (2021). However, those approaches might lead to conservatism, and are 1) slow to converge in comparison with spectral methods and 2) not tractable for real-time constraint satisfaction due to the multidimensional Bernstein polynomials.

Near the limits of performance where optimality is crucial, the safety of the predicted trajectory is not guaranteed with DMS and PSC. This can lead to uncomfortable and dangerous situations such as emergency braking or potentially, crash with other road users specifically in safety-critical conditions. Failure to generate safe control inputs limits the user's acceptability of such controllers, therefore this work presents a safe yet non-conservative MPC approach for autonomous driving. In this paper, we extend the collocation framework with safety envelope developed in Allamaa et al. (2023), to include general nonlinear constraints, and present a transcription framework for smooth MPC problems, called RESAFE/COL: Regional Envelope for SAFety Enhanced COLlocation. It allows solving for the original continuous-time OCP by generating safe trajectories of states and inputs as splines, satisfying the constraints over the complete prediction horizon through computationally efficient manipulation of the splines' coefficients. In addition, we introduce the concept of regional convex hulls to reduce the conservatism on the splines' extrema approximation, allowing dynamic solutions near the limits of handling towards a natural longitudinal and lateral driving. The contributions of this paper are three-fold:

(1) Collocation framework with safety envelope over nonlinear constraints,
(2) Control barrier function development for safe autonomous driving in urban environment,
(3) Validation of RESAFE/COL with CBF in safety-critical situations, integrated in a real-time NMPC.

This paper is organized as follows: in Sec. 2 we provide a background on a collocation method for NMPC and exponential control barrier functions. In Sec. 3 we introduce the transcription method for real-time NMPC based on regional safety envelopes. We follow with Sec. 4 in which we introduce the NMPC-CBF formulation for AD, and present the results in Sec. 5, before concluding in Sec. 6.

2. BACKGROUND

The navigation of the vehicle relies on a special set of splines that can guarantee a safe tracked trajectory in the look-ahead of the NMPC. The transcription into splines using collocation enhances the tractability of the NMPC problem for real-time applications, without decreasing the performance. In this section we first start by giving a background on the spectral orthogonal collocation with safety envelope, and briefly introduce control barrier functions.

2.1 Safety Envelope for Spectral Orthogonal Collocation

We consider the nonlinear continuous Bolza problem, that optimizes a cost function J over the states x, u with a stage cost l and terminal cost $\phi(x(t_f))$, to satisfy a set of (nonlinear) constraints g, g_f, by propagating the dynamics $f(x,u)$ over a time horizon $[0, t_f]$:

$$\begin{cases} \min\limits_{x(.),u(.)} J(x,u) = \phi(x(t_f)) + \int_{t_0}^{t_f} l(x(t),u(t))dt \\ \text{subject to } \dot{x}(t) = f(x(t),u(t)), \\ \qquad\qquad g(x(t),u(t)) \leq 0, \\ \qquad\qquad g_f(x(t_f)) \leq 0. \end{cases} \tag{1}$$

We focus on the collocation approach to transcribe the OCP into an NLP as it does not require an embedded ODE solver. Instead, the problem is a dynamic optimization problem with polynomial representation of the state and control in each finite element. In particular, we enforce the control and states to be continuous on $[t_0, t_f]$, and transcribe the OCP into an NLP using spectral orthogonal collocation SOCSE as in Allamaa et al. (2023). For smooth problems, an accurate solution is obtained using a small number N of collocation nodes, and the approach benefits from a spectral accuracy of convergence, faster than any power of $1/N$ (Huntington (2007)).

We define the orthogonal collocation scheme based on the truncated Legendre-series (TLS) of degree M as in (2) to approximate the solutions $x(\tau)$ and $u(\tau)$ of the continuous time OCP in a compact representation through the coefficients α rather than a discrete set of points:

$$x(\tau) = \sum_{k=0}^{M} \alpha_{k,x} \mathcal{L}_k(\tau) = \alpha_x^\top \mathbf{L}_M v(\tau), (\tau \in [-1,1]), \tag{2}$$

and similarly for control inputs $u(\tau)$. The matrix $\mathbf{L}_M \in \mathbb{R}^{(M+1)\times(M+1)}$ is upper triangular, formed by the coefficients of $\mathcal{L}_k$ with respect to the normalized time horizon $\tau \in [-1,1]$ of $t \in [t_0, t_f]$ in the OCP. In particular, the spanning basis are Legendre polynomials $\mathcal{L}_k$, which are orthogonal, satisfy $\mathcal{L}_k(1) = 1$ for any degree k, and in Huntington (2007), they are given by:

$$\mathcal{L}_k(\tau) = \frac{1}{2^k k!} \frac{d^k}{d\tau^k}[(\tau^2 - 1)^k]. \tag{3}$$

An optimization over the polynomial with an orthogonal basis 1) enhances numerical stability and, 2) shrinks the higher order terms to zero if they are not significant on the output. Moreover, the TLS is parametrized by the coefficients $\alpha_x = [\alpha_0 \cdots \alpha_M]^\top \in \mathbb{R}^{(M+1)\times N_x}$, $\alpha_u \in \mathbb{R}^{(M+1)\times N_u}$, and $v(\tau) = \begin{bmatrix} 1 & \tau & \tau^2 & \cdots & \tau^M \end{bmatrix}^\top$ is a vector with a geometric progression of the normalized time instance $\tau = (2t/t_f - 1)$ with $t_0 = 0$. In addition, N_x, N_u are

the number of state and control variables respectively and M is the degree of the fitting polynomial. The dynamics equations of the OCP are satisfied by differentiating (2) at the N collocation points τ_i to satisfy the dynamics:

$$\dot{x}(\tau_i) = \alpha_x^\top \mathbf{L}_M \dot{v}(\tau_i) = \frac{t_f}{2} f(x(\tau_i), u(\tau_i)). \tag{4}$$

Note that $dt/d\tau = t_f/2$ accounts for the timescale transformation. The N collocation points τ_i, are neither random, nor uniform, but are specifically set as the Legendre-Gauss-Lobatto (LGL) nodes, defining the spectral grid of a Lobatto Pseudospectral method (LPM). The collocation points $\tau_i, i = 2, \ldots, N-1$ are uniquely chosen to be the roots of $\dot{\mathcal{L}}_{N-1}$, the derivative of the Legendre polynomial $\mathcal{L}_{N-1}$ of degree $N-1$, c.f (3). The two remaining collocation points are the boundaries $[-1, 1]$ such that T =: $\begin{bmatrix}-1 & \text{Roots of } \dot{\mathcal{L}}_{N-1} & 1\end{bmatrix}$ are the collocation nodes (Allamaa et al. (2023)). LPM can achieve exact estimation up to machine precision, as long as the true polynomial is up to degree $2N-3$, with a relatively small number of nodes (Huntington (2007)) and a truncation error $O(h^{2N-2})$. Importantly, from Allamaa et al. (2023) and with the particular choice of parametrization as in (2), the resulting TLS is contained in its convex envelope, having as its boundaries the maximum and minimum elements of its convex hull $\mathcal{P}_M \in \mathbb{R}^{(M+1)}$:

$$\min\{\mathcal{P}_M\} \leq x(\tau) \leq \max\{\mathcal{P}_M\}, \quad \forall \tau \in [-1, 1], \tag{5}$$

$$\text{where } \mathcal{P}_M = \mathcal{B}\mathcal{E}^\top \mathbf{L}_M^\top \alpha_x = \mathcal{C}_M \alpha_x. \tag{6}$$

As proven in Cargo and Shisha (1966), the matrix $\mathcal{B} \in \mathbb{R}^{(M+1)\times(M+1)}$ is a mapping from the coefficient of a polynomial to the Bernstein extrema estimations b_j, such that for a polynomial $P(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_M t^M$ of degree $M \geq 0$ with real coefficients $a_j, j = 0, \ldots, M$, and defined over the time interval $[0, 1]$:

$$b_j = \sum_{k=0}^{j} a_k \binom{j}{k} \Big/ \binom{M}{k}, \; j = 0, 1, \ldots, M, \tag{7a}$$

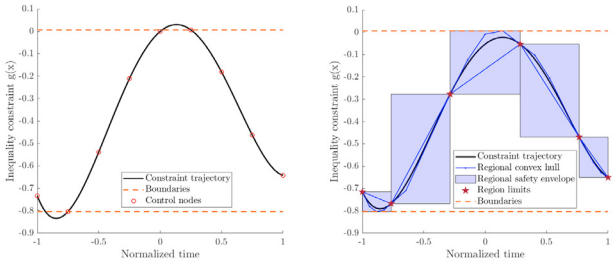
$$b = \mathcal{B}a, \; \mathcal{B} \in \mathbb{R}^{(M+1)\times(M+1)}. \tag{7b}$$

The lower triangular matrix $\mathcal{E} = \mathcal{E}^{-1:1} \in \mathbb{R}^{(M+1)\times(M+1)}$ is a time transformation for the vector $v(\tau)$ in (2) allowing a conversion from any interval of normalized time instances $[\tau_k, \tau_{k+1}] \subset [-1, 1]$ with $\tau = (2\tau_* - 1)$ to $\tau_* \in [0, 1]$, and particularly for the case of the boundaries such that $\tau_k = -1, \tau_{k+1} = 1$. Using the Binomial theorem, the entry on the i^{th} row and j^{th} column of the matrix $\mathcal{E}^{\tau_k:\tau_{k+1}}$ is:

$$\mathcal{E}_{ij}^{\tau_k:\tau_{k+1}} = \sum_{i=0}^{M} \sum_{j=0}^{i} \binom{i}{j} (\tau_{k+1} - \tau_k)^j (\tau_k)^{i-j}, \tag{8a}$$

$$\text{such that: } v(\tau) = \begin{bmatrix}1 & \tau & \cdots & \tau^M\end{bmatrix}^\top = \mathcal{E}^{\tau_k:\tau_{k+1}} v(\tau_*). \tag{8b}$$

The extrema approximation is tight and exact, if the true maximum and minimum are at the boundary elements of $\mathcal{P}_M$, and that is when the TLS is monotonic within the time interval. Similarly, the bounds on the states' and inputs' derivatives can be calculated by rewriting $\mathbf{L}_M$ in (2) and (5) in terms of the coefficients of $\mathcal{L}_k^{(i)}(\tau)$ instead of $\mathcal{L}_k(\tau)$. Similar matrices $\mathcal{C}_{M,i}$ can be calculated offline, allowing the efficient evaluation of the convex hull of the i^{th} derivative within the NLP. The previous result is exploited to transcribe the linear inequality constraints of the continuous time OCP into convex constraints on the spline coefficients α. However, there remains a question on how to efficiently exploit $\alpha, \mathcal{P}_M$, as to include general nonlinear constraints within the optimization routine, which will be elaborated in this paper.



(a) DMS and PSC (b) RESAFE/COL (ours)

Fig. 1. Continuous constraint satisfaction with RESAFE/COL with one finite element in comparison with DMS and PSC: DMS and PSC optimize over finite control nodes, and RESAFE/COL over the Legendre-series' coefficients. In Fig. 1a, DMS and PSC fail to capture the internode dynamics and violate the continuous-time constraint. In Fig. 1b, RESAFE/COL, through regional safety envelopes accurately estimates the (nonlinear) constraint's extrema, and returns a continuous-time feasible solution

2.2 Exponential Control Barrier Function

Motivated by Control Lyapunov Functions (CLF) for guaranteeing the stability of a system, CBFs are employed to guarantee the safety of a control action through forward invariance conditions. A CBF extension for non-affine nonlinear systems with relative degrees higher than 1 has been derived in the work of Son and Nguyen (2019).

Consider a nonlinear dynamic system $\dot{x} = f(x, u)$, a set $\mathcal{C} \in \mathbb{R}^n$ is forward invariant, if for every $x(0) = x_0 \in \mathcal{C}$, the state trajectory $x(t) \in \mathcal{C}, \forall t$. Moreover, the continuously differentiable function $h(x) : \mathbb{R}^n \to \mathbb{R}$ is a CBF iff:

$$\mathcal{C} = \{x \in \mathbb{R}^n | h(x) \geq 0\}, \tag{9a}$$

$$\partial\mathcal{C} = \{x \in \mathbb{R}^n | h(x) = 0\}, \tag{9b}$$

$$\text{Int}(\mathcal{C}) = \{x \in \mathbb{R}^n | h(x) > 0\}. \tag{9c}$$

To guarantee forward invariance of a safety constraint $h(x)$ within a finite-horizon MPC, Son and Nguyen (2019) derive the condition for an exponential CBF:

$$L_f^r h(x, u) + k_1 h(x) + k_2 L_f h(x) + \cdots + k_{r-1} L_f^{r-1} h(x) \geq 0. \tag{10}$$

The Lie derivative L_f, and high-order Lie derivatives L_f^k of the barrier function $h(x)$ with respect to the vector function $f(x, u)$ are:

$$L_f h(x, u) = \frac{\partial h(x)}{\partial x} f(x, u), \tag{11}$$

$$L_f^k h(x, u) = \frac{\partial L_f^{k-1} h(x, u)}{\partial x} f(x, u). \tag{12}$$

with $L_f^0 h(x, u) = h(x)$. Furthermore, r is such that the constraint $h(x)$ has a relative degree $1 \leq r \leq n$. In other words, the input does not explicitly appear in the first r Lie derivatives and that is $L_f^i h(x, u) = L_f^i h(x)$ for $0 \leq i \leq r-1$. Moreover, there exists a vector $[k_1, k_2, \ldots, k_{r-1}]$ such that the safety condition $h(x) \geq 0$ is satisfied.

3. REGIONAL ENVELOPE FOR NONLINEAR CONSTRAINTS

In the standard direct approaches such as DMS or PSC, path inequality constraints are only enforced at the shooting or collocation points (Huntington (2007)). In some work such as Ruof et al. (2023), temporary numerical constraints violations are tolerated. However, for automotive industry standards or for other safety-critical robotic applications, this might not always be permissible. In this section, we present a transcription approach for NMPC, allowing to capture the extrema of the states, inputs, and nonlinear constraints in their convex hulls. Those extrema would serve a basis for a sped-up numerical optimization.

For non-smooth systems, and solutions that cannot be fit with a single TLS, the normalized NMPC horizon $[-1, 1]$ can be divided into smaller finite elements, to create a piecewise polynomial, namely a Legendre-spline. On each element, every state and input is parametrized by a TLS, the coefficients of which serve as optimization variables α. Without loss of generality, we assume one finite element over the time horizon, and seek to minimize the approximation error on the spline's extrema. Similar to Theorem 4 in Rivlin (1970):

Proposition 1. The error on the regional extrema approximation for a truncated Legendre-series, decreases proportionally to $1/k^2$ for an increasing number of regions k within one finite element on $[0, t_f]$.

Instead of estimating the spline's extrema on the full time horizon as in SOCSE, we propose a regional approximation defined by the regional convex hulls $\mathcal{P}_M^k$. The time interval $[-1, 1]$ is divided into K regions as: $[t_0, t_1, \ldots, t_{k+1}]$, that are not necessarily equidistant, and define a region k to be a sub-interval $[t_k, t_{k+1}] \subset [-1, 1]$, with $t_{k+1} \geq t_k$. At the expense of computing additional mapping matrices C_M^k, but offline, we can compute the regional convex hulls as:

$$\mathcal{P}_M^k = \mathcal{B}(\mathcal{E}^{t_k:t_{k+1}})^\top \mathbf{L}_M^\top \alpha = \mathcal{C}_M^k \alpha. \tag{13}$$

From practice, a reasonable choice is to set the region bounds t_k as the LGL nodes of the Legendre polynomial of order K, where K is a hyperparameter for RESAFE/COL. The regional extrema approximation using a linear mapping over α are shown in Figure 1, and can be applied to state and input box constraints, as well as path and general nonlinear constraints. The transfer to a regional safety envelope formulation has a three-fold benefit:

- the conservatism on the spline extrema approximation is reduced,
- the open-loop trajectory, can be more dynamic with change of derivative, as large high-order coefficients lead to a conservative extrema approximation if only one region is used,
- regional convex hulls are computed using the original optimization variables α, allowing for time varying constraints without the burden of additional optimization variables.

Therefore, RESAFE/COL is a generalization of SOCSE with more than one region, and allows the inclusion of nonlinear constraints as will follow.

Theorem 2. Using the regional convex hull $\mathcal{P}^k$ of the state (or input) spline, the nonlinear constraint $g(x)$ is included in its continuous safety envelope defined by:

$$\begin{aligned} \min\{g(\mathcal{P}^k)\} - \frac{d_k^2}{2} \max_{x \in \mathcal{X}^k} \left| \frac{d^2 g}{dx^2} \right| \\ \leq g(x) \leq \\ \max\{g(\mathcal{P}^k)\} + \frac{d_k^2}{2} \max_{x \in \mathcal{X}^k} \left| \frac{d^2 g}{dx^2} \right|, \end{aligned} \tag{14}$$

where $\{\mathcal{X}^k : x(t) | x(t_k) \leq x(t) \leq x(t_{k+1})\}$ and $d_k = \max\{\bar{\mathcal{P}}_{i+1}^k - \bar{\mathcal{P}}_i^k\}$, where $\bar{\mathcal{P}}_{i+1}^k$ is a sorted $\mathcal{P}^k$ with $i = [1, \ldots, M]$.

Proof. Let $\xi = \text{argmax}(g(x))$ for $x \in \mathcal{X}^k$. Given the definition of d_k, it stands that $d_k \geq |\mathcal{P}_i^k - \xi|$.

Through the generalized 2^{nd} order Taylor series expansion:

$$g(\mathcal{P}_i^k) = g(\xi) + g'(\xi)(\mathcal{P}_i^k - \xi) + \frac{g''(\beta)}{2}(\mathcal{P}_i^k - \xi)^2,\ \beta \in \mathcal{X}^k. \tag{15}$$

- If $g'(x) \neq 0$ for $x \in \mathcal{X}^k$, then the extrema are at the boundaries and the extrema's approximation is tight and that is $\max g(x) = \max\{g(\mathcal{P}^k)\}$, which verifies the upper bound in (14) and the extrema is an element of the convex hull.
- If there is a change of curvature, $g'(x) = 0$ for $x \in \mathcal{X}^k$, with $g'(\xi) = 0$, the extrema is contained in an enlarged convex hull as:

$$\begin{aligned} g(\xi) = g(\mathcal{P}_i^k) - \frac{g''(\beta)}{2}(\mathcal{P}_i^k - \xi)^2 \\ \leq \max\{g(\mathcal{P}^k)\} + \frac{d_k^2}{2} \cdot \max_{x \in \mathcal{X}^k} |g''(x)|. \end{aligned} \tag{16}$$

The lower bound of (14) is proven similarly. □

Note that for clarity, Theorem 2 was proven for the scalar case with $g(x) : \mathbb{R} \to \mathbb{R}$, however, it extends to multidimensional systems by replacing $\max_{x \in \mathcal{X}^k} |\frac{d^2 g}{dx^2}|$ with the maximum eigenvalue of the Hessian of the nonlinear constraint $g(x, u)$. Moreover, as the number of regions increases, d_k decreases and converges to zero, thus the approximation with the regional convex hull gets tighter.

4. APPLICATION OF RESAFE/COL TO AUTONOMOUS DRIVING

In this section, we apply the derived transcription approach to the case of automated driving in an urban environment in the presence of road users. We first start by introducing the vehicle model in the NMPC prediction, then follow with the collision avoidance protocol and present the resulting NMPC formulation and the NLP.

4.1 Vehicle model

We represent the car dynamics in the NMPC by a 3 DoF dynamic single-track model, in the curvilinear error frame:

$$\begin{aligned} M\dot{v}_x &= (F_{xf}\cos\delta + F_{xr} - F_{yf}\sin\delta - F_{res} + Mrv_y), \\ M\dot{v}_y &= (F_{xf}\sin\delta + F_{yr} + F_{yf}\cos\delta - Mrv_x), \\ I_z\dot{r} &= (L_f(F_{yf}\cos\delta + F_{xf}\sin\delta) - L_r F_{yr}), \\ \dot{s} &= (v_x\cos\theta - v_y\sin\theta)/(1 - \kappa_c w), \\ \dot{w} &= v_x\sin\theta + v_y\cos\theta, \\ \dot{\theta} &= r - \dot{\psi}_c = r - \kappa_c\dot{s}. \end{aligned} \tag{17}$$

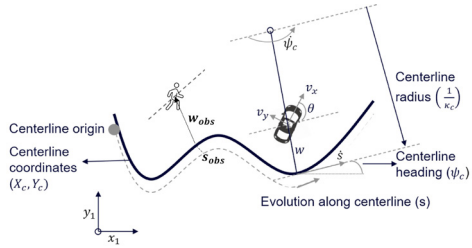


Fig. 2. Transformation to Curvilinear frame

The first three equations of (17) dictate the dynamics (velocities v_x, v_y and yaw rate r) in the car body frame, and the last three represent the vehicle's kinematics in the curvilinear frame as presented in Figure 2. Moreover, M is the vehicle's mass, I_z is the inertia about the z-axis, and L_r, L_f are the distances from the center of gravity to the rear and front axles. For control purposes, we formulate the kinematics in the curvilinear or Frenet frame, that is equivalent to an error frame with respect to the desired path. That is, s tracks the evolution along the path, w and θ are the distance and heading deviation from the path respectively. This transformation allows an easier navigation on curved roads as it is only parametrized by the road curvature $\kappa_c(s)$ and exposes convex constraints over the path deviation as a tube $w_l \leq w \leq w_r$ where w_l and w_r are the left and right track limits. Additionally, it allows for spatial prediction in the NMPC horizon. For smoother driving style, the NMPC dynamics model is augmented with the input rates $u = [\dot{t}_r, \dot{\delta}]$, the derivatives of the steering angle control δ and the normalized acceleration t_r over the longitudinal forces $F_{x,\{r,f\}}$. The lateral forces F_y are assumed linear with respect to the wheel slip angle. As explained in Allamaa et al. (2023), we fuse the dynamic model in (17) with the kinematic one to allow the model to be numerically stable near zero speeds. Hence, the single-track curvilinear dynamics between the state vector $x = [v_x, v_y, r, s, w, \theta, \delta, t_r]^\top$ and input u are:

$$\dot{x} = f(x,u) = \lambda f_{dyn}(x,u) + (1-\lambda) f_{kin}(x,u). \tag{18}$$

The control objective of the NMPC is a path following navigation with zero reference deviation from the path with velocity tracking as a non-zero term in the reference x^r, considering energy efficiency in terms of inputs. Therefore, we set the quadratic stage cost with weighting matrices $Q \in \mathbb{R}^{8\times 8} \succeq 0, R \in \mathbb{R}^{2\times 2} \succ 0$ as:

$$l(x(\tau), u(\tau)) = \|(x(\tau) - x^r(\tau))\|_Q^2 + \|u(\tau)\|_R^2. \tag{19}$$

4.2 CBF for collision avoidance

In order to combine the safety constraint with the NMPC's state prediction, and input and state constraints, we opt to incorporate the CBF within the NMPC framework. We consider collision avoidance for AD, in an environment with multiple static and dynamic agents. In the work of Reiter et al. (2023), ellipsoidal safety constraints have shown be superior to other geometric shapes. For this reason, we augment obstacles through ellipsoidal boundaries, centered at (s_{obs}, w_{obs}), with axes lengths a and b. We set the safety constraint in the Curvilinear frame, by projecting the obstacles' Cartesian position in the global frame on the path frame driven by the autonomous vehicle. That is, s_{obs} and w_{obs} are the obstacle's evolution along the path and its deviation from the closest point on the path respectively. The safety constraint is formulated as:

$$h(x) = \frac{(s - s_{obs})^2}{a^2} + \frac{(w - w_{obs})^2}{b^2} - 1 \geq 0. \tag{20}$$

Finally, the exponential CBF constraint is:

$$h_{CBF}(x,u) = L_f^2 h(x,u) + k_1 L_f h(x) + k_2 h(x) \geq 0, \tag{21}$$

$$L_f h(x) = \frac{2(s - s_{obs})\dot{s}}{a^2} + \frac{2(w - w_{obs})\dot{w}}{b^2}, \tag{22}$$

$$L_f^2 h(x,u) = \frac{2(s - s_{obs})\ddot{s} + 2\dot{s}}{a^2} + \frac{2(w - w_{obs})\ddot{w} + 2\dot{w}}{b^2}. \tag{23}$$

In Son and Nguyen (2019), a change of variable was required to get rid of the second order derivatives. In comparison, and given that each state trajectory is a TLS defined by (2), we can efficiently calculate $\dot{s}, \dot{w}, \ddot{s}, \ddot{w}$ by differentiating the Vandermonde vector $v(\tau)$ without relying on the dynamics. That is, the k^{th} derivative of the spline, can be written in a compact form as $d^k x/dt^k = \frac{2}{t_f} \cdot \alpha \mathcal{L}(d^k v(\tau)/d\tau^k)$ where $d^k v(\tau)/d\tau^k$ is trivially calculated given the geometric form of the vector $v(\tau)$. In relation with (2), the barrier function in (20) is multidimensional but quadratic, hence with a diagonal Hessian. Therefore, the maximum eigenvalue of the constraint's Hessian can be calculated as the maximum of $\{2/a^2, 2/b^2\}$.

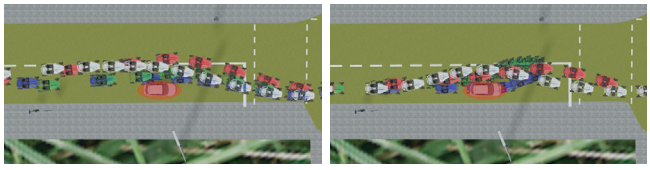
4.3 NMPC formulation

We cast the continuous time OCP into an NLP using RESAFE/COL. That is, the integral of the Lagrange term $l(x(\tau), u(\tau)) =: l(\tau)$ in the OCP cost function is given by the exactness of the Gauss quadrature rule as:

$$\int_{-1}^{1} l(\tau) d\tau = w_1 l(-1) + w_N l(1) + \sum_{i=2}^{N-1} w_i l(\tau_i). \tag{24}$$

The dynamics are satisfied according to (4). Moreover, the inequality constraints on the optimization variables α_x, α_u, the coefficients of the TLS are set through the regional envelopes in (13) and (14). This approach benefits from the offline computation of the mapping matrices $\mathcal{C}_M^k \in \mathbb{R}^{(M+1)\times(M+1)}, \forall k \in [1, \ldots, K]$, according to (13), given the spline degree M and the number of regions K. The box constraints on the states and inputs are replaced by linear constraints over α_x, α_u, and the nonlinear constraints maintain their structure, but are applied over the regional convex hull of the trajectories $\mathcal{P}_{M,x}^k, \mathcal{P}_{M,u}^k$ and their derivatives $\mathcal{P}_{M,\dot{x}}^k, \mathcal{P}_{M,\ddot{x}}^k$. In other words, each continuous time constraint is transcribed by replacing the variable with its convex hull. The resulting NLP is:

$$\begin{cases} \min\limits_{\alpha_x, \alpha_u} \dfrac{t_f}{2}\Big[\sum\limits_{i=2}^{N-1} w_i l(\tau_i) + \phi(\alpha_x^\top \mathbf{L}_M v(1)) \\ \qquad + w_1 l(x_0, \alpha_u^\top \mathbf{L}_M v(-1)) + w_N l(1)\Big] \\ \text{s.t. } \alpha_x^\top \mathbf{L}_M v(-1) = \bar{x}_0, \\ \qquad \alpha_x^\top \mathbf{L}_M \dot{v}(\tau_i) = \dfrac{t_f}{2} f(\alpha_x^\top \mathbf{L}_M v(\tau_i), \alpha_u^\top \mathbf{L}_M v(\tau_i)) \\ \qquad \underline{x} \leq (\mathcal{C}_M^k \alpha_x)^\top \leq \overline{x}, \ \forall k \in [1, \ldots, K], \\ \qquad \underline{u} \leq (\mathcal{C}_M^k \alpha_u)^\top \leq \overline{u}, \ \forall k \in [1, \ldots, K], \\ \qquad h(\mathcal{C}_M^k \alpha_x) \geq 0 \quad \text{from (13), (20)}, \\ \qquad h_{CBF}(\mathcal{C}_M^k \alpha_x) \geq 0 \quad \text{from (14), (21)}, \end{cases} \tag{25}$$



(a) Long look-ahead (b) Short look-ahead

Fig. 3. Closed-loop control with collision avoidance using CBF with RESAFE/COL (white) in comparison with standard position avoidance with RESAFE/COL (red), DMS (green) and PSC (blue): MPC-CBF ensures a natural double lane change. For a horizon length of 1.75 s (Fig. 3b), only RESAEF/COL formulations avoid the obstacle safely

where $l(\tau_i) = l(\alpha_x^\top \mathbf{L}_M v(\tau_i), \alpha_u^\top \mathbf{L}_M v(\tau_i)), \forall i \in [1, N]$, the evaluation of the cost function at a particular collocation point. It is important to note that the barrier constraint and the CBF constraints are applied to the convex hulls of the states and their derivatives, but the second order terms from (14) are left out as slack variables. The NLP is solved with a Sequential Quadratic Programming (SQP) method using the OSQP QP solver (Stellato et al. (2020)) as SQP is suited for embedded applications such a real-time control for AD.

We set the major-axes of the ellipses enlarging the static and dynamic obstacles to $a = 3\,\text{m}$, $b = 2\,\text{m}$, the coefficients of the CBF constraint $k_1 = 1.6, k_2 = 1.1$ and the weighting matrices of the cost function $Q = \text{diag}(3.1, 10, 10, 0, 5.2, 48, 0.9, 1.5), R = \text{diag}(1, 1)$. Moreover, the control action is executed every 50 ms. Finally, we distinguish between two situations based on the prediction horizon t_f: long look-ahead with $t_f = 3\,\text{s}$ and short look-ahead for safety-critical situations with $t_f = 1.75\,\text{s}$.

5. RESULTS AND COMPARISON WITH DMS AND PSC

We define four vehicles, each represented as a Digital Twin of an electrical SimRod vehicle. That is, each vehicle is a high-fidelity 15DoF simulator in Simcenter Amesim, and navigates in the traffic scenario simulator Simcenter Prescan. The four vehicles are defined as:

- Vehicle 1, labeled RESAEF/COL+CBF: white vehicle (black plots), NMPC with the exponential CBF and position constraints (20),(21), transcribed using RESAFE/COL, with TLS of degree $M = 5$.
- Vehicle 2, labeled PSC: blue vehicle (blue plots), NMPC with position constraint (20), transcribed using Pseudospectral collocation, and polynomials of order 5.
- Vehicle 3, labeled RESAFE/COL: red vehicle (red plots), NMPC with position constraint (20), transcribed using RESAFE/COL, and TLS of degree 5.
- Vehicle 4, labeled DMS: green vehicle (green plots), NMPC with position constraint (20), transcribed using Direct Multiple Shooting.

The closed-loop performance of the different controllers is shown in Figure 3 in which the formulations with RESAFE/COL outperform DMS and PSC, specifically in the event of a safety-critical situation with a short look-ahead

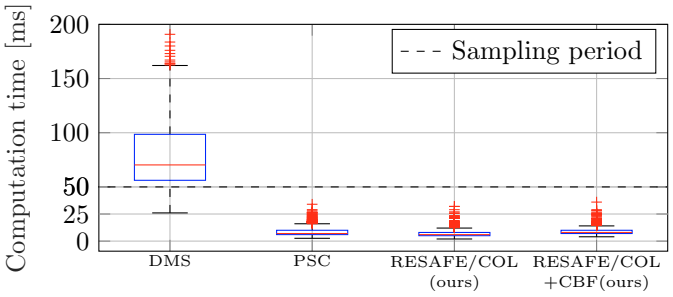


Fig. 4. Comparison of the different numerical methods and collision avoidance approaches with a long horizon length, over all instances of the closed-loop scenarios

of the MPC. Moreover, the developed RESAFE/COL approach is intended for real-time platforms. Therefore, we first assess the required computational effort in comparison with DMS and PSC. For this, we run a multitude of scenarios, to mimic an urban driving style: path following at 20 m/s, intersections, roundabouts, and a multitude of road users such as dynamic vehicles, pedestrians, and dangerous objects to test the safety-critical aspect. For a considerably large horizon length $t_f = 3\,\text{s}$, the NMPC is tasked to solve in a receding horizon at the sampling period of 50 ms. In the boxplot of the computation time in Figure 4, the collocation methods have a clear computational advantage over DMS. Albeit solving to convergence, the PSC method does not satisfy the continuous path constraints fully, and therefore does not reflect the original continuous time OCP, ending up crashing with the obstacle at high-speed. Using RESAFE/COL on the other hand, both open-loop trajectories computed at every instance, and the closed-loop performance, satisfy the collision constraints described in the continuous time OCP. Using only the barrier function (20) as a constraint, RESAFE/COL with a mean of computation time at 7 ms is 7 times faster than DMS, and all the computation instances are well below the sampling period of 50 ms. Similarly, by including the CBF constraint (21), despite an increase in the computation time to 11 ms. It is worth nothing that the computation time using RESAFE/COL is not highly sensitive to the change of active set during the appearance of obstacles. This is reflected by the small spread in computation time in comparison to the large spread using DMS.

In addition, we perform a sensitivity analysis over the number of regions utilized for extrema estimation in Theorem 2 and (25), with the results shown in Figure 5. The percentage of crash avoidance is calculated as the ratio between the total time the vehicle is outside the enlarged obstacles' ellipse over the total duration in which the vehicle detects an obstacle closer than 30 meters away. For the percentage of crash avoidance, we consider the closed-loop performance of the vehicle. The NMPC for the vehicles with DMS (green) and PSC (blue) remain unchanged with 60 shooting nodes, and polynomial of order 5 respectively. As the OCP with the CBF constraint is more complex than the position avoidance counterpart, the computation time is more sensitive to the number of regions. That is due to tightening on the extrema of both the barrier constraint and CBF constraints, which activates the NLP constraints.

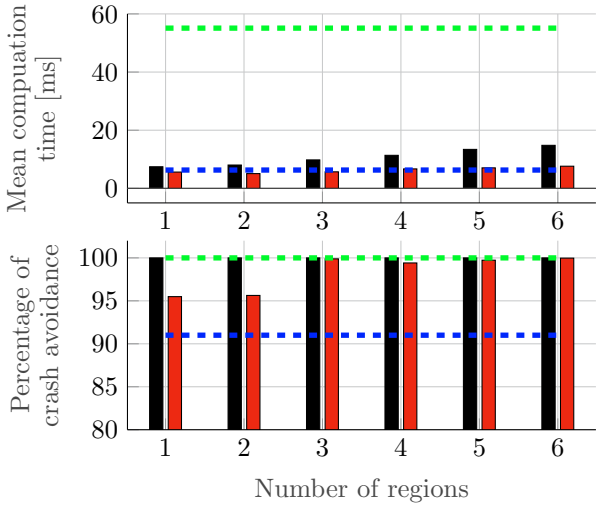


Fig. 5. Sensitivity of computation time and closed-loop performance in terms of collision avoidance to the number of regions in RESAFE/COL, with a large horizon length: RESAFE/COL with CBF (black); RESAFE/COL with position constraint only (red), PSC order 5 (blue) and DMS with 60 nodes (green)

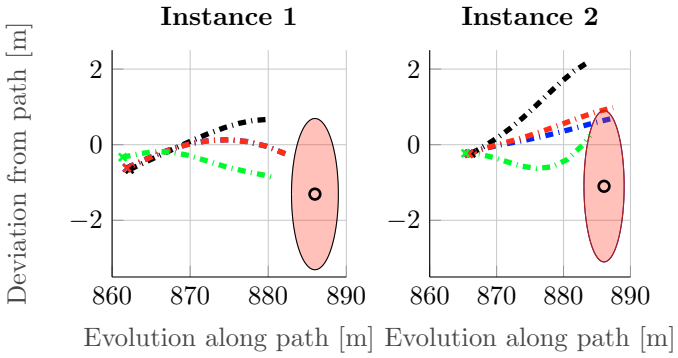


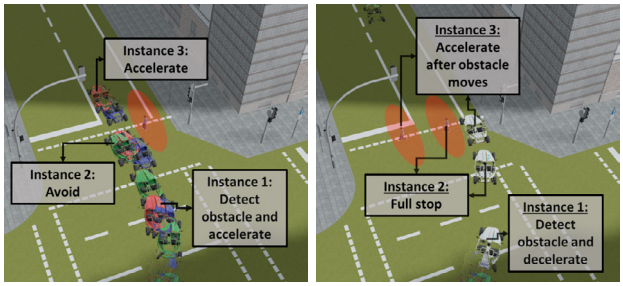
Fig. 6. Open-loop trajectory during collision avoidance using forward invariant CBF with RESAFE/COL (black) in comparison with position constraint with RESAE/COL (red), PSC(blue) and DMS (green)

However, computation time remains well below the sampling period of 50 ms confirming the real-time feasibility of the complex OCP, and a performance improvement between 5 and 7 times in terms of computation time in comparison with DMS. As the formulation through a TLS and the exploitation of the convex-hull are intended to not deteriorate the control performance, we also visualize the percentage of crash avoidance in Figure 5. With a dense sampling of 60 nodes, DMS is able to maintain a 100% crash avoidance in closed-loop, still at the risk of constraint violation between the shooting nodes in the open-loop prediction. As from Proposition 1, the extrema approximation with RESAFE/COL decreases as the number of regions increases, the percentage of crash avoidance with RESAFE/COL increases, and reaches 99.5% starting 3 regions. That is, the NMPC solves for more dynamic trajectories near the limit of handling, with less conservatism as the number of regions increases. Finally, with PSC, the percentage of crash avoidance is at 91% with noticeable continuous time OCP constraints' violations in the open-loop trajectories as the constraints are only satisfied at 6 collocation points.

5.1 Comparison between position and CBF constraints

For the vehicles with the position (barrier) constraint (20) only, Vehicle 3 (red) initiates the avoidance maneuver earlier than Vehicles 2 and 4 as seen in Figure 3a, and has a safer escape than Vehicles 2 and 4 that pass tightly along the obstacles' ellipse. The inclusion of the CBF constraint does not considerably change the closed-loop performance of the system in terms of collision avoidance. This is due the long horizon length, allowing the vehicle to react early to upcoming obstacles. Among the approaches with RESAFE/COL, Vehicle 3 exhibits a larger deviation from the path and a longer recovery time after crossing the parked vehicle. The applied control effort are therefore slightly larger and more aggressive than with the CBF counterpart. In order to verify the benefit of the CBFs constraints in terms of satisfying the collision avoidance in the infinite horizon ahead, we reduce the horizon length from $t_f = 3\,\text{s}$ to $t_f = 1.75\,\text{s}$. That is, at full speed of $20\,\text{m/s}$, the NMPC can only predict for a distance of $35\,\text{m}$ ahead. In Figure 3b, the closed-loop performance of all 4 vehicles is visualized, in a safety-critical situation around a parked vehicle highlighted in red. The first two direct realizations are that PSC and DMS fail completely to avoid the obstacle. That is, even when the open-loop trajectories are safe, the model mismatch between the NMPC prediction model and the Digital Twin, causes both vehicles to overestimate their dynamic abilities. Once close to the obstacle, both vehicles engage in an emergency brake maneuver. In the case of PSC, and given that the collision avoidance constraint is satisfied at finite points, the NMPC barely commands the vehicle to steer away from the obstacle, causing a rear collision. Unlike the NMPC formulated with RESAFE/COL (red, white) that steer the vehicle away on time. For the two NMPCs formulated with RESAFE/COL, the controller with the CBF constraint (white), and for the same enlarged size of the vehicle, is more cautious and engages in a slow-down, avoid, and recover maneuver that resembles the ISO 3888-1 standard double lane change. The controller with only the position constraint avoids the obstacle, but just tightly at the limits of the ellipse, without decreasing the speed of the vehicle. The open-loop prediction at two instances near the obstacle are shown in Figure 6: the CBF constraint ensures a forward invariance of the predicted trajectory beyond the NMPC horizon, allowing a smoother and earlier deviation from the obstacle.

Furthermore, we test the capabilities of the controller at an intersection. Following a highly curved turn, an obstacle suddenly appears in front of all vehicles as depicted in Figure 7. The ability of the CBF formulation to incorporate mid-level motion planning by altering the velocity profile, and low-level control tracking is put to highlight. With the standard collision avoidance constraint, all three vehicles avoid the obstacle through high steering and continuous acceleration. Albeit being a successful collision avoidance, the evasive maneuver is uncomfortable and unnatural. One could follow a hierarchical framework with a planner for velocity decision-making, however, this generally comes in



(a) Position constraint (b) CBF constraint

Fig. 7. Collision avoidance at an intersection with and without CBF. The CBF formulation embeds both motion planning and control as it performs a stop-and-go

the form of a non-scalable rule-based or heuristic module. Alternatively, with the CBF constraint, the vehicle performs a stop-and-go maneuver, by slowing down to the pedestrian, coming to a natural and smooth full stop without an emergency brake, and then accelerating to recover the velocity and path tracking as the obstacle moves. This distinction between velocity and path tracking allowing such a decision-making is possible through the path following formulation that plans over space rather than time as in Figure 2.

Table 1: NMPC's performance during safety-critical maneuvers with a short prediction horizon of 1.75 s, at 20 m/s

Method	Mean computation time [ms]	Percentage of crash
DMS	37.6	14.29
PSC	7.8	4.09
RESAFE/COL	7.7	1.28
RESAFE/COL+CBF	**12.3**	**0**

Finally, we showcase the computational and control performance advantages offered by RESAFE/COL in a safety-critical situation. From Table 1, RESAFE/COL is approximately 5 times faster than DMS for the same formulation, and 3 time faster for the more complex formulation with CBF. The computational benefit is also complemented by the success of collision avoidance: while it is at 100% with RESAFE/COL+CBF similar to the case of long horizon, RESAFE/COL with position constraint is able to decrease the total percentage of in-crash instances by 91% from 14.29% to 1.28%.

6. CONCLUSION

A safety-critical model predictive controller based on control barrier functions has been proposed in this paper to tackle urban autonomous vehicle driving in the presence of road users. We presented RESAFE/COL a novel collocation based transcription approach allowing to satisfy the nonlinear constraints of the continuous time OCP over the full prediction horizon. We validate the approach in high-fidelity vehicle and environment simulators. The navigation of the vehicle relies on a special set of splines that can guarantee a safe tracked trajectory in the look-ahead of the NMPC, with spatio-temporal predictions in the error frame. The proposed transcription approach makes the problem more tractable for real-time applications without decreasing the performance, as it is shown to outperform the direct multiple shooting method in terms of computation time and delivers a safer and more stable navigation than standard collision avoidance formulations that suffer from the shortness of the NMPC horizon.

REFERENCES

Allamaa, J.P., Patrinos, P., Van Der Auweraer, H., and Son, T.D. (2023). Safety envelope for orthogonal collocation methods in embedded optimal control. In *2023 European Control Conference (ECC)*, 1–7. doi: 10.23919/ECC57647.2023.10178116.

Cargo, G. and Shisha, O. (1966). Bernstein form of a polynomial. *Journal of Research of the National Bureau of Standards Section B Mathematics and Mathematical Physics*, 70B(1), 79. doi:10.6028/jres.070B.005.

Cichella, V., Kaminer, I., Walton, C., Hovakimyan, N., and Pascoal, A.M. (2021). Optimal Multivehicle Motion Planning Using Bernstein Approximants. *IEEE Transactions on Automatic Control*, 66(4), 1453–1467. doi: 10.1109/TAC.2020.2999329.

Cosner, R.K., Yue, Y., and Ames, A.D. (2022). End-to-End Imitation Learning with Safety Guarantees using Control Barrier Functions. In *2022 IEEE 61st Conference on Decision and Control (CDC)*, 5316–5322. doi: 10.1109/CDC51059.2022.9993193. ISSN: 2576-2370.

He, S., Zeng, J., Zhang, B., and Sreenath, K. (2021). Rule-Based Safety-Critical Control Design using Control Barrier Functions with Application to Autonomous Lane Change. In *2021 American Control Conference (ACC)*, 178–185. IEEE. doi:10.23919/ACC50511.2021.9482848.

Huntington, G.T. (2007). *Advancement and analysis of Gauss pseudospectral transcription for optimal control problems*. Ph.D. thesis, Massachusetts Institute of Technology.

Reiter, R., Nurkanović, A., Frey, J., and Diehl, M. (2023). Frenet-Cartesian model representations for automotive obstacle avoidance within nonlinear MPC. *European Journal of Control*, 74, 100847. doi:https://doi.org/10.1016/j.ejcon.2023.100847.

Rivlin, T.J. (1970). Bounds on a polynomial. *Journal of Research of the National Bureau of Standards, Section B: Mathematical Sciences*, 47.

Ruof, J., Mertens, M.B., Buchholz, M., and Dietmayer, K. (2023). Real-time spatial trajectory planning for urban environments using dynamic optimization. In *2023 IEEE Intelligent Vehicles Symposium (IV)*, 1–7. doi:10.1109/IV55152.2023.10186535.

Son, T.D. and Nguyen, Q. (2019). Safety-Critical Control for Non-affine Nonlinear Systems with Application on Autonomous Vehicle. In *2019 Conference on Decision and Control (CDC)*, 7623–7628. IEEE. doi:10.1109/CDC40024.2019.9029446.

Stellato, B., Banjac, G., Goulart, P., Bemporad, A., and Boyd, S. (2020). OSQP: an operator splitting solver for quadratic programs. *Mathematical Programming Computation*, 12(4), 637–672. doi:10.1007/s12532-020-00179-2.

Zeng, J., Zhang, B., and Sreenath, K. (2021). Safety-Critical Model Predictive Control with Discrete-Time Control Barrier Function. In *2021 American Control Conference (ACC)*, 3882–3889. IEEE. doi:10.23919/ACC50511.2021.9483029.