A Nonlinear Model Predictive Control Algorithm for Obstacle Avoidance in Autonomous Ground Vehicles within Unknown Environments

Jiechao Liu, Paramsothy Jayakumar, Jeffrey L. Stein, and Tulga Ersal

Abstract—A nonlinear model predictive control algorithm is developed for obstacle avoidance in high-speed, large-size autonomous ground vehicles (AGVs) that perceive the environment only through information provided by on-board sensors. The mission of the AGV is to move from its initial configuration to the goal configuration safely. The resulting trajectory should be collision-free and the AGV should be dynamically safe. As a starting point, the scenario where the vehicle moves on a flat surface at a constant speed is considered. The nonlinear MPC algorithm generates steering commands for completing the mission while enforcing safety constraints. The first safety constraint is avoiding obstacles. This is fulfilled by constraining the position of the AGV inside a safe region established from sensor data. The second safety constraint is ensuring dynamical safety. This is translated into avoiding single tire lift-off, which is implemented by limiting the steering angle within a range obtained using a 14 DoF vehicle dynamics model. At each sampling time, at least one multi-phase optimal control problem (OCP) is formulated and solved on-line. The safe region is partitioned into multiple sub-regions, which can then be specified without using piecewise functions. The fact that the optimal trajectory traverses the sub-regions sequentially and hence the position constraints are different from phase to phase makes the OCP multi-phase. The multi-phase OCP is transcribed into a nonlinear programming problem using the hp-pseudospectral method, and solved using the interior-point method. Simulations of an AGV approaching multiple obstacles show the effectiveness of the proposed algorithm.

Index Terms—Collision avoidance, vehicle dynamics, model predictive control, autonomous ground vehicles.

I. INTRODUCTION

AUTONOMOUS ground vehicles (AGVs) are gaining importance and finding increased utility in both military and commercial applications. Although earlier AGV platforms were typically exclusively small ground robots, recent efforts have targeted passenger vehicles and larger size platforms, as well. For this size of vehicles, it becomes especially important to take the dynamical limitations of the vehicle into account to guarantee its dynamical safety during obstacle avoidance maneuvers [1]. Therefore, obstacle avoidance algorithms are needed that can ensure vehicle safety even if the vehicle is operating at its dynamic limits.

Many obstacle avoidance algorithms have been developed in the literature that allow for fast, continuous, and smooth motion of AGVs among unexpected obstacles. They can be classified into four categories: graph-search based methods [2], [3], virtual potential field and navigation function based methods [4], [5], meta-heuristic based methods [6], and mathematical optimization based methods [7], [8]. Among these categories, mathematical optimization based methods are particularly attractive, because they offer a rigorous and systematic approach to take vehicle dynamics and safety constraints into account.

A mathematical optimization approach can be used either in open-loop, if the environment is fully known a priori, or in closed-loop with a feedback controller for a more robust solution. Regarding the latter, the model predictive control (MPC) approach is one of the most widely adopted techniques [9]. Prior research has demonstrated successful applications of MPC to obstacle avoidance in AGVs [10], [11], [12], [13], [14], [15]. Some active safety methods leverage the MPC framework, as well, to ensure, for example, safe lane keeping or vehicle stability [16], [17], [18], [1].

The first applications of MPC to obstacle avoidance in AGVs assumed that the controller has full knowledge about the environment. They also were not concerned with the level of fidelity that the model used by the controller needs to possess for satisfactory performance, where the performance criteria in some cases also include the dynamical safety of the vehicle, such as no tire lift-off. Previous work by the authors aimed to address this gap by developing an MPC formulation that takes into account the information about the environment as provided by the on-board Light Detection and Ranging (LIDAR) sensor [19]. They also investigated the role of model fidelity and showed that the vehicle’s dynamical safety can be guaranteed by limiting the steering angle using a high fidelity model, which then allows the MPC to work with a low fidelity model for trajectory optimization [19]. However, this investigation was done using an exhaustive search approach on a coarse mesh of the control input.

In this paper, we extend the work presented in [19] and [20], and develop a novel nonlinear MPC algorithm for obstacle avoidance of AGVs that can achieve an optimal and smooth operation of the vehicle through the obstacle field.

Manuscript submitted February, 2015. The authors wish to acknowledge the financial support of the Automotive Research Center (ARC) in accordance with Cooperative Agreement W56HZV-14-2-0001 U.S. Army Tank Automotive Research, Development and Engineering Center (TARDEC) Warren, MI.

J. Liu, J. L. Stein, and T. Ersal are with the Department of Mechanical Engineering, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: ljch@umich.edu; stein@umich.edu; tersal@umich.edu).

P. Jayakumar is with the U.S. Army RDECOM-TARDEC, Warren, MI 48397 (email: paramsothy.jayakumar.cn@mail.mil).

Disclaimer: The views expressed in this article are those of the author and do not reflect the official policy or position of the Department of the Army, DOD, or the U.S. Government.
**Title:** A Nonlinear Model Predictive Control Algorithm for Obstacle Avoidance in Autonomous Ground Vehicles within Unknown Environments

**Performing Organization:** US Army RDECOM-TARDEC, 6501 E. 11 Mile Road, Warren, MI 48397-5000

**Report Date:** 24 APR 2015

**Distribution/Availability Statement:** Approved for public release; distribution unlimited

**Abstract:** See Report

**Security Classification:**
- a. Report: unclassified
- b. Abstract: unclassified
- c. This Page: unclassified

**Number of Pages:** 14

**Limitation of Abstract:** Same as Report (SAR)

**Form Approved OMB No. 0704-0188**
while ensuring vehicle safety. Optimality here refers to the optimality of the solution within the prediction horizon, taking into account all the information available at that moment in time, and not to the optimal solution that would have resulted if all environmental information was available for all times. Optimality is achieved by formulating the obstacle avoidance problem into an optimal control problem (OCP), which is then converted into a nonlinear programming (NLP) problem using the hp-pseudospectral method, and solved using the interior-point method. Two types of safety constraints are included in the formulation. The first type of safety constraint is avoiding obstacles. This is fulfilled by constraining the position of the AGV inside a safe region established from sensor data. The safe region is partitioned into multiple sub-regions, which can be specified without using piecewise functions. These specifications can then be used in the OCP formulation. The second type of safety constraint is ensuring dynamical safety. This is translated into avoiding single tire lift-off, which is implemented by limiting the steering angle within a range obtained using a 14 DoF vehicle dynamics model. Simulations of an AGV approaching multiple obstacles within a range obtained using a 14 DoF vehicle dynamics lift-off, which is implemented by limiting the steering angle dynamical safety. This is translated into avoiding single tire formulation. The second type of safety constraint is ensuring functions. These specifications can then be used in the OCP sub-regions, which can be specified without using piecewise sensor data. The safe region is partitioned into multiple position of the AGV inside a safe region established from is avoiding obstacles. This is fulfilled by constraining the included in the formulation. The first type of safety constraint the interior-point method. Two types of safety constraints are results. Conclusions are drawn in Section V.

Throughout the paper, the following assumptions are made:
(a) All obstacles of interest are at least the height of where the LIDAR is mounted on the vehicle, which is in front of the vehicle.
(b) The vehicle longitudinal speed is maintained to be constant.
(c) The vehicle travels on a constant-friction flat surface.

Because the formulated OCP is nonconvex, it is not guaranteed that the solution from the OCP solver is the unique global optimal solution over the prediction horizon. Thus, the terms “optimal trajectory”, “optimal states”, and “optimal control” in this paper refer to the local optimal solution generated by the OCP solver, which is the first minimum it finds.

The rest of the paper is organized as follows. Section II introduces the basic principle of MPC briefly. Section III presents the formulation of the multi-phase optimal control problem for obstacle avoidance, including safe region partition approach, details of the vehicle dynamics model used, maximum steering angle establishment, and solution techniques. Section IV presents and discusses the simulation results. Conclusions are drawn in Section V.

II. BASIC PRINCIPLE OF MPC

The idea of model predictive control is to utilize a model of the system to be controlled to predict and optimize future system behavior. MPC is an optimal control based state-feedback controller. The feedback law is obtained by an iterative on-line optimization over a moving finite prediction horizon. One advantage of using a moving time horizon is the ability to perform real-time optimization with hard constraints on plant variables [9].

The basic principle of MPC is illustrated in Fig. 1. At time \( t_0 \), starting from the state measurements, an optimal control sequence \( \zeta^*(t), t \in [t_0, t_0 + T_p] \) is computed by solving an open-loop, constrained, finite-time optimal control problem over the prediction horizon \( T_p \). The control calculations are based on both future predictions and current measurements. The optimal control sequence \( \xi^*(t) \) is bounded by the input saturations, and the resulting estimated optimal states \( \xi^*(t) \) satisfy the pre-defined constraints and minimize the cost function. Although the optimal control sequence is calculated over the horizon \( t \in [t_0, t_0 + T_p] \), only a portion of the computed control sequence \( \zeta^*(t), t \in [t_0, t_0 + T_e] \) is sent to the plant and executed, where \( T_e \) is called the execution horizon and represents the portion of the computed optimal sequence that is implemented. Due to model simplifications, model parameter uncertainties, and other types of noises and uncertainties, the actual states of the system \( \xi^*(t), t \in [t_0, t_0 + T_e] \) are highly likely to be different from the predicted value \( \xi^*(t), t \in [t_0, t_0 + T_e] \). In the next step, the optimal control problem is solved again over a shifted horizon based on the new state measurements. The feedback of measurement information to the optimization endows the whole procedure with a robustness typical of closed-loop systems. This process is thus repeated at each step until terminal requirements are satisfied.

III. NONLINEAR MPC ALGORITHM FOR OBSTACLE AVOIDANCE

In this section, details of the nonlinear MPC algorithm for obstacle avoidance are presented. The nonlinear MPC algorithm consists of two parts: the LIDAR data processor and the control commands generator. Fig. 2 shows the block diagram with the nonlinear MPC algorithm and the AGV in the loop.

The LIDAR data processor simplifies the obstacle shape, adds safety margin, and partitions the safe region. The outputs of the LIDAR data processor, the task information, and the estimated vehicle states are used in the formulation of the OCPs. The formulated OCPs are then solved and the control commands associated with the lowest cost solution is executed by the AGV. These steps are discussed in detail in the following sub-sections.

Three external inputs to the nonlinear MPC algorithm are required: task information, obstacle information, and estimated states. Within the scope of this paper, the task information is the specified target location and the desired vehicle speed,
maximum detection range obstacles within the detection range, the LIDAR returns the $\epsilon$ boundary in each radial direction at an angular resolution of used. The LIDAR returns the distance to the closest obstacle LIDAR sensor that is mounted in front of the vehicle is characteristics of the closest objects to the vehicle. A 2D algorithm. could also be generated by a high-level global path planning which is directly specified by the user. The task information obtained directly from the LIDAR can be noisy because the obstacle boundaries are not smooth and the detection results are not exact. An algorithm is required to identify the minimum number of lines for approximating the sequence of points considering these noises. The Ramer-Douglas-Peucker algorithm is an algorithm for reducing the number of points in a curve that is approximated by a series of points, which is widely used to perform simplification and denoising of range data acquired by a LIDAR [21]. As illustrated in Fig. 4, the numerous detected points, which are generated using a simulated LIDAR with added noise, can be simplified into two line segments represented by three points, which are a good approximation to the boundaries of the obstacle.

1) Line Simplification: The first step of the data processing is to reduce the number of points that defines the obstacle boundaries for further processing. However, the points obtained directly from the LIDAR can be noisy because the detection boundaries are not smooth and the detection results are not exact. An algorithm is required to identify the minimum number of lines for approximating the sequence of points considering these noises. The Ramer-Douglas-Peucker algorithm is an algorithm for reducing the number of points in a curve that is approximated by a series of points, which is widely used to perform simplification and denoising of range data acquired by a LIDAR [21]. As illustrated in Fig. 4, the numerous detected points, which are generated using a simulated LIDAR with added noise, can be simplified into two line segments represented by three points, which are a good approximation to the boundaries of the obstacle.

2) Safety Margin: A safety margin, $l_{SM}$, is added to the safe region to account for the size of the vehicle, detection noises, and differences between the predicted trajectory and the actual trajectory. Adding a safety margin allows for ignoring the vehicle size in the OCP formulation. The safe area is a polygon in general and it is a simple polygon when the safe region boundary is from a 2D LIDAR sensor. Thus, algorithms for performing polygon offsetting (inflating/deflating) in computer graphics can be adopted. Specifically, the Vatti’s clipping algorithm [22] implemented in the Clipper library developed by Angus Johnson [23] is used.

A safety margin, $l_{SM}$, is added to the safe region to account for the size of the vehicle, detection noises, and differences between the predicted trajectory and the actual trajectory. Adding a safety margin allows for ignoring the vehicle size in the OCP formulation. The safe area is a polygon in general and it is a simple polygon when the safe region boundary is from a 2D LIDAR sensor. Thus, algorithms for performing polygon offsetting (inflating/deflating) in computer graphics can be adopted. Specifically, the Vatti’s clipping algorithm [22] implemented in the Clipper library developed by Angus Johnson [23] is used.

As shown in Fig. 5a, the boundary of the safe region consists of three types of segments:

- LIDAR data segments, which specify the boundaries of the obstacles.
- Maximum LIDAR detection range segments, which are directions free from obstacles and are called “openings”.
- Obstacle’s laser shadow lines, which are along the rays from the LIDAR and are called “hypothetical openings”. The are called “hypothetical openings,” because in its current position and orientation the vehicle cannot know whether it is an actual opening or not due to the obstacle blocking its view.

To add the safety margin, the openings are expanded by the amount of specified safety margin first. Then, the entire region is shrunk by the amount of specified safety margin. This is to add safety margin only on the obstacle boundaries and
3) Region Partitioning: The safe region exemplified in Fig. 5b is very difficult, if not impossible, to be defined using a single function. Even if the function exists, it is not differentiable at some points, which would cause problems in the OCP solver that requires all functions to be twice continuously differentiable. To address this challenge, the safe region is partitioned into several sub-regions and a multi-phase optimal control problem formulation is used. After partitioning, each sub-region can be specified by a set of inequalities. The functions involved in these inequalities are not piecewise functions and are differentiable.

There are many approaches for partitioning the safe region to meet the specified requirements. Two of them are introduced in this paper. One approach is named the “polar partitioning”. The safe area is divided into sectors and triangles, where sectors are regions including an opening and triangles are regions including an obstacle boundary. Fig. 6a is the partitioning of the safe region shown in Fig. 5 using this approach. As an example, Region “OB4” is a triangle, which can be specified using three linear inequalities, whereas Region “OP3” is a sector, which is bounded by two lines. The third boundary of Region “OP3” is an arc; however, because of the limits on prediction horizon, this arc constraint will never be active and thus is ignored. This is an intuitive approach for partitioning the region from a 2D LIDAR sensor.

Another approach is named the “optimal convex partitioning” or simply “convex partitioning”. The interior of this safe area is decomposed into a minimum number of convex regions without introducing additional points inside the polygon. Several algorithms exist in the literature for performing this task. The dynamic programming algorithm by Keil and Soneyink is incorporated [24], which is efficient in decomposing simple polygons. Fig. 6b is an exemplified partitioning. Similarly, all the regions can be specified using a set of linear inequalities after partitioning. This approach is capable of partitioning a safe region in a more general form.

In either approach, a sub-region can be defined by

\[
\begin{bmatrix}
\vdots & \vdots & \vdots \\
\alpha_j & b_j & \vdots \\
\vdots & \vdots & \vdots \\
\end{bmatrix}
\begin{bmatrix}
x^{(i)}_1 \\
x^{(i)}_2 \\
\vdots \\
\end{bmatrix} 
\leq \begin{bmatrix}
c_j \\
\vdots \\
\end{bmatrix}, \quad j = 1, \ldots, L \tag{1}
\]

where \( i \) is the sub-region index and \( L \) is the total number of line segments bounding that sub-region; \( \alpha_j, b_j \), and \( c_j \) are coefficients calculated based on the two end points of the corresponding line segments; \((x, y)\) is a position in Cartesian coordinates.

Eq. (1) can be compacted in the following form:

\[
A^{(i)}_{L \times 1} x^{(i)}(t) + B^{(i)}_{L \times 1} y^{(i)}(t) = C^{(i)}_{L \times 1}, \quad t \in [T^{i-1}, T^i] \tag{2}
\]

where \( A^{(i)}_{L \times 1} \) is a vector with the \( j \)th entry being \( \alpha_j \). The definitions of \( B^{(i)}_{L \times 1} \) and \( C^{(i)}_{L \times 1} \) are similar.

After partitioning, the entire safe region can be specified by a structure variable SafeRegion. The definition of the structure variable SafeRegion is given by the following pseudo-code.

```plaintext
int N; // number of sub-regions
int L[N]; // vector of number of line segments

struct SafeRegion {
    double AM[N][N]; // adjacency matrix
    SubRegion SR[N]; // subregion specifications
};

struct SubRegion {
    char Type; // type of the subregion; "OP", "OB"
    int Index; // index of the subregion
    int HPNum; // number of hypothetical openings
    // end points of line segments (xs, ys, xe, ye)
    double EndPoints[L[Index]][4];
    // index of line segments representing hypothetical openings
    double HPIndex[HPNum];
};
```

The adjacency matrix for the partitioning shown in Fig. 6b is given below. When two sub-regions have a common edge,
the corresponding entry in the matrix is set to 1; otherwise, it is set to 0.

\[
\begin{pmatrix}
0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
\]

To avoid the obstacles and move towards the target, the trajectory should stay within the safe region and the last part of the predicted trajectory should lie within the sub-region of type “OP”. The following list presents the procedures of using the structure variable SafeRegion in the OCP formulation to meet the above requirement. The safe region partitioning shown in Fig. 6b is used as an example to elaborate.

(a) Identify the first region to traverse (“SR”). Region “OB1” is the “SR” in this example, because the vehicle heading at the current position is pointing upwards.

(b) Identify all regions with a feasible opening (“TRs”). A feasible opening is an arc segment, which can have an intersection with a feasible trajectory over a slightly longer period than the prediction horizon. They can be identified using the two extreme trajectories, which are shown in Fig. 7. They are obtained by simulating the 2 DoF vehicle model using steering controls that are at the limits of handling at each step at the measured initial states. In the example, the arc segments in “OP2” and “OP3” are feasible openings and hence “OP2” and “OP3” are “TRs”, whereas, “OP1” and “OP4” are not “TRs”, because the vehicle cannot make a sharp enough turn to move into those partitions.

(c) Find the sequence of regions from the “SR” to a “TR” for all “TRs”. This can be achieved by using Dijkstra’s algorithm and the adjacency matrix [25]. For example, with “OP3” being the “TR”, the region sequence is identified as “OB1 → OB2 → OB3 → OP3”. The illustration of the region sequence is shown in Fig. 7. With “OP2” being the “TR”, the region sequence is identified as “OB1 → OB2 → OB3 → OP2”.

For a region sequence from the “polar partitioning” as exemplified in Fig. 8a, a different region partition as shown in Fig. 8b can be obtained easily. This alternative partition approach is preferred when one of the boundaries separating two regions is almost along the vehicle heading direction.

The specifications of the regions in this partition are given by

\[
R_{min}^{i} \leq \sqrt{(x^{(i)}(t) - x(0))^2 + (y^{(i)}(t) - y(0))^2} \leq R_{max}^{i}
\]

\[
\Phi_{min}^{i} \leq \tan2(y^{(i)}(t) - y(0), x^{(i)}(t) - x(0)) \leq \Phi_{max}^{i}
\]

where \(R_{min}^{i}, R_{max}^{i}, \Phi_{min}^{i}, \) and \(\Phi_{max}^{i}\) are bounds calculated from the coordinates of end points specifying a region.

B. Control Command Generator

At each step of the MPC, a multi-phase optimal control problem is formulated for each of the “TRs”. The number of phases is the number of regions from “SR” to the “TR”. So, one or more OCPs is formulated and solved at each step. The formulation in general form is given by Eq. (4) - Eq. (10)

\[
\begin{align*}
\text{minimize} & \quad J = T \left[ \xi^{(N)}(T^N), \zeta^{(N)}(T^N), T^N \right] \\
\text{subject to} & \quad \xi^{(i)}(t) = \mathcal{V} \left( \xi^{(i)}(t), \zeta^{(i)}(t) \right) \\
& \quad \xi^{(i)}(T^1) = \xi^{(i-1)}(T^i-1) \\
& \quad \xi^{(0)}(T^0) = \xi_0 \\
& \quad \mathcal{S}^{(i)}(x^{(i)}(t), y^{(i)}(t)) \leq 0 \\
& \quad \delta_{f,\text{min}}(U_0) \leq \delta_{f,\text{max}}(U_0) \\
& \quad T^0 = 0, T^N = T_p, T^i, T^N < T_p \\
& \quad i = 1, \ldots, N
\end{align*}
\]

By minimizing the cost function specified in Eq. (4), subject to constraints defined in Eq. (5) - Eq. (10) for all phases, the optimal state trajectories \(\xi^{(i)}(t), t \in [T^{i-1}, T^i]\), the optimal control trajectories \(\zeta^{(i)}(t), t \in [T^{i-1}, T^i]\), and the times \(T^{i-1}, T^i, i = 1, \ldots, N\) are obtained, where \(N\) is the total number of phases.
Before we define the variables and explain the problem formulation in detail, let us go through the formulation at a high level. Eq. (5) is the dynamic model of the vehicle represented as a set of first order ordinary differential equations (ODEs). Eq. (6) sets the initial states of each phase as the final states of the previous phase. For the first phase, the initial states are the measured states. Eq. (7) defines the position constraints due to the obstacles perceived by the LIDAR sensor. A general form is given here. Eq. (8) and Eq. (9) represent the bounds on the steering angle and the steering rate, respectively. Eq. (10) specifies the time horizon for each phase. The overall prediction horizon is given by $t \in [0, T_p]$. There is a lower bound and an upper bound on the prediction time, which are described in Section IV.

1) Cost Function: The cost function defines the soft requirement, i.e., in what sense the trajectory is optimal. In this work, if the task is only to pass a target location without direction requirement, the trajectory is optimal when the end point of the predicted trajectory is close to the target, and the final heading angle is pointing to the target because a shorter distance-to-go is preferred. The cost function is defined as

$$ J = \frac{s_T}{s_0} + w_s \psi_{\text{diff}}^2 + w_d d $$

where

$$ s_0 = \sqrt{[x_g - x(0)]^2 + [y_g - y(0)]^2} $$

$$ s_T = \sqrt{[x_g - x(T_p)]^2 + [y_g - y(T_p)]^2} $$

$$ \psi_{\text{fg}} = \text{atan2}(y_g - y(T_p), x_g - x(T_p)) $$

$$ \psi_{\text{diff}} = \text{atan2}(\sin(\psi(T_p) - \psi_{\text{fg}}), \cos(\psi(T_p) - \psi_{\text{fg}})) $$

$$ d = \int_0^{T_p} [\dot{\psi}_f(t) + w_d \delta_f(t)] \, dt $$

Specifically, the cost function formulation includes three terms that are linearly combined using relative weights, $w_s$ and $w_d$. The first term is a ratio between distance $s_T$ and distance $s_0$, where $s_0$ is the distance between the initial position $[x(0), y(0)]$ and the goal $[x_g, y_g]$ as defined in Eq. (12), and $s_T$ is the distance between the end point of the predicted trajectory $[x(T_p), y(T_p)]$ and the goal as defined in Eq. (13). Visual representations of all variables are shown in Fig. 3b. The second term is the difference between the final heading angle $\psi(T_p)$ and the angle of the goal relative to the end point of the predicted trajectory $\psi_{\text{fg}}$ as defined in Eq. (15). The third term is a regulation term minimizing the control effort $d$ as defined in Eq. (16), where $\dot{\psi}_f$ is the steering rate, which is the control command to be optimized, $\delta_f$ is the front wheel steering angle, and $w_d$ is a weight.

If a particular direction of passing the target location in global coordinates is also required, the following cost function is used

$$ J = \frac{s_T}{s_0} + w_s \psi_{\text{diff}}^2 + w_s s + w_d d $$

where

$$ s = \int_0^{T_p} \left[ l_a x(t) + l_b y(t) + l_c \right]^2 \, dt $$(18)

The cost function specified by Eq. (17) has one more term than Eq. (11). This term is to minimize the integral of the distance to the line given by $l_a x + l_b y + l_c = 0$ over the prediction horizon. This line is passing through the goal $[x_g, y_g]$ along the desired direction $\psi_g$.

When the target position is within the sensor’s detection range, the term $\frac{s_T}{s_0}$ and $w_s \psi_{\text{diff}}^2$ are removed from the cost functions given in Eq. (11) and Eq. (17). Instead the following constraints are added to the OCP formulation.

$$ x_g - \sigma \leq x(T_p) \leq x_g + \sigma $$

$$ y_g - \sigma \leq y(T_p) \leq y_g + \sigma $$

where $\sigma$ is a small margin. If the vehicle is within this margin from the target position, then the target is considered to be reached.

2) Vehicle Models: Two different vehicle dynamics models are used in this work: a 14 DoF model to represent the plant and to generate offline the dynamic-safety-related look-up tables used in the MPC, and a 2 DoF model used in the MPC to predict trajectories. The schematics of the two representations are given in Fig. 9. The 14 DoF model consists of a single sprung mass connected to four unsprung masses. The suspensions between the sprung mass and unsprung masses are modeled as spring-damper systems. In the 2 DoF model, the left and right tires on each axle are lumped together. The equations for the 14 DoF model are omitted here for space limitations, but can be found in the literature [26].

$$ l_a = \sin(\psi_g) $$

$$ l_b = -\cos(\psi_g) $$

$$ l_c = -\sin(\psi_g)x_g + \cos(\psi_g)y_g $$

Fig. 9. Schematic of the vehicle models
The 2 DoF model is described by the following ODEs

\begin{align}
\dot{V} &= (F_{y,f} + F_{y,r})/M - U_0 r \\
\dot{r} &= (F_{y,f} L_f - F_{y,r} L_r)/I_{zz} \\
\dot{x} &= U_0 \cos \psi - (V + L_f r) \sin \psi \\
\dot{y} &= U_0 \sin \psi + (V + L_f r) \cos \psi
\end{align}

where $F_{y,f}$ and $F_{y,r}$ are tire lateral forces generated at the front axle and the rear axle, respectively. $U_0$ and $V$ are the longitudinal speed and lateral speed in the vehicle’s coordinate frame, respectively. $r$ is the yaw rate, $\psi$ is the yaw angle, $(x, y)$ is the vehicle’s front center location in global coordinates, $M$ is the vehicle mass, $I_{zz}$ is the moment of inertia, $L_f$ is the distance between the front axle and the vehicle’s CoG location, and $L_r$ is the distance between the rear axle and the vehicle’s CoG location.

By using the Pacejka Magic Formula ($MF$) tire model with pure slip, the lateral tire forces are presented as

\begin{align}
F_{y,f} &= MF_y(F_{z,f}, \alpha_f) \\
F_{y,r} &= MF_y(F_{z,r}, \alpha_r)
\end{align}

The exact form of the Pacejka tire model can be found in [27]. Fig. 10 shows the relationship between the tire lateral forces and the slip angle at different vertical loads described by the Pacejka tire model. In the 14 DoF vehicle model, Pacejka Magic Formula tire model with combined slip is used.

As concluded in [20], it is important to account for the longitudinal load transfer in the 2 DoF vehicle model when the vehicle travels at high speed. Thus, the following relationships are used with the 2DoF model to calculate the vertical loads on the front and rear axles taking into account the longitudinal load transfer effects:

\begin{align}
F_{z,f} &= \frac{M g L_f + M V r h_{CG}}{L_f + L_r} \\
F_{z,r} &= \frac{M g L_f - M V r h_{CG}}{L_f + L_r}
\end{align}

where $h_{CG}$ is the height of the vehicle CoG location above the ground.

In addition, the slip angles of front and rear tires are obtained from

\begin{align}
\alpha_f &= \tan^{-1} \left( \frac{V + L_f r - \delta_f}{U_0} \right) \\
\alpha_r &= \tan^{-1} \left( \frac{V - L_r r}{U_0} \right)
\end{align}

The steering rate $\zeta_f$ is used as the control command to be optimized and the steering angle $\delta_f$ is set as an additional state variable of the system. The reason is to make it possible to obtain a smooth steering angle sequence and impose a limit on the steering rate.

By setting the state vector as $\xi = [x \ y \ \psi \ V \ r \ \delta_f]$ and the control vector as $\zeta = \zeta_f$, the state-space equation for the 2 DoF nonlinear vehicle model can be written as

$$\dot{\xi} = f(\xi) + B \zeta$$

where

$$f(\xi) = \begin{bmatrix}
U_0 \cos \psi - (V + L_f r) \sin \psi \\
U_0 \sin \psi + (V + L_f r) \cos \psi \\
(F_{y,f} + F_{y,r})/M - U_0 r \\
(F_{y,f} L_f - F_{y,r} L_r)/I_{zz} \\
0
\end{bmatrix}$$

$$B^T = [0 \ 0 \ 0 \ 0 \ 1]$$

The steering profile in Fig. 11a is used as the input and a vehicle speed of 20 m/s is considered for a model comparison. As shown by the comparison of trajectories in Fig. 11b, yaw rates in Fig. 11c, and lateral accelerations in Fig. 11d, the 2 DoF vehicle model with longitudinal load transfer and nonlinear tire model is a very good approximation to the 14 DoF vehicle model.
in the structure variable SafeRegion. The specific form of Eq. (7) is either given by Eq. (2) or Eq. (3) depending on the scenario.

Regarding the safe region partitioning, there are still some questions, such as, what are the characteristics of a good partition, how to evaluate the goodness of the partition, and how to generate a good partition systematically and efficiently. Addressing these questions requires future work. In the simulations for this paper, the convex partition approach is used primarily, and the polar partition approach is used secondarily, if needed.

4) Dynamical Safety: In this study, ensuring the vehicle’s dynamical safety is translated to avoiding single tire lift-off. This is a conservative criterion used to prevent rollover [28]. This requirement could be taken into account directly to enforce a positive vertical load on all four tires at all times. However, vehicle models that could predict the vertical tire loads on all four tires would require a level of complexity whose computational load would be prohibitively high for the purpose of MPC. Therefore, another conservative approximation of the dynamical safety requirement is considered; namely, an upper bound on the steering angle magnitude as expressed by the following inequality constraint

$$|\delta_f(t)| \leq \delta_{f,\text{max}}(U_0) \tag{33}$$

where the maximum steering angle $\delta_{f,\text{max}}$ is a function of only the vehicle speed when the vehicle is assumed to move on a flat surface.

For all combinations of longitudinal speed ranging from 10 m/s to 30 m/s and maximum steering angle ranging from 0° to 14°, the corresponding minimum tire vertical loads are obtained using the 14 DoF vehicle model. The relationship is shown in Fig. 12a. If a minimum vertical load threshold is set, the relationship between the maximum steering angle and longitudinal speed can be extracted. Fig. 12b shows the relationship when $F_{z,\text{min}}$ is set as 500 N.

5) Solution Techniques: The formulated nonlinear multi-phase optimal control problems are solved using a two-step procedure. First, the continuous-time OCP is transcribed into a nonlinear programming (NLP) problem using a direct method called hp-pseudospectral method [29], [30], [31]. Second, the resulting NLP problem is solved using the interior point method [32].

The hp-pseudospectral method discretizes a continuous-time OCP into an NLP problem by approximating the state and variable-degree polynomial approximations of them within each interval. The differential-algebraic constraints of the OCP are enforced at a finite set of collocation points, where the collocation points are Legendre-Gauss-Radau (LGR) quadrature points. This method has been shown to be able to accurately approximate the solution to a general continuous-time OCP in a computationally efficient manner [31].

4) Dynamical Safety: In this study, ensuring the vehicle’s dynamical safety is translated to avoiding single tire lift-off. This is a conservative criterion used to prevent rollover [28]. This requirement could be taken into account directly to enforce a positive vertical load on all four tires at all times. However, vehicle models that could predict the vertical tire loads on all four tires would require a level of complexity whose computational load would be prohibitively high for the purpose of MPC. Therefore, another conservative approximation of the dynamical safety requirement is considered; namely, an upper bound on the steering angle magnitude as expressed by the following inequality constraint

$$|\delta_f(t)| \leq \delta_{f,\text{max}}(U_0) \tag{33}$$

where the maximum steering angle $\delta_{f,\text{max}}$ is a function of only the vehicle speed when the vehicle is assumed to move on a flat surface.

For all combinations of longitudinal speed ranging from 10 m/s to 30 m/s and maximum steering angle ranging from 0° to 14°, the corresponding minimum tire vertical loads are obtained using the 14 DoF vehicle model. The relationship is shown in Fig. 12a. If a minimum vertical load threshold is set, the relationship between the maximum steering angle and longitudinal speed can be extracted. Fig. 12b shows the relationship when $F_{z,\text{min}}$ is set as 500 N.

5) Solution Techniques: The formulated nonlinear multi-phase optimal control problems are solved using a two-step procedure. First, the continuous-time OCP is transcribed into a nonlinear programming (NLP) problem using a direct method called hp-pseudospectral method [29], [30], [31]. Second, the resulting NLP problem is solved using the interior point method [32].

The hp-pseudospectral method discretizes a continuous-time OCP into an NLP problem by approximating the state and variable-degree polynomial approximations of them within each interval. The differential-algebraic constraints of the OCP are enforced at a finite set of collocation points, where the collocation points are Legendre-Gauss-Radau (LGR) quadrature points. This method has been shown to be able to accurately approximate the solution to a general continuous-time OCP in a computationally efficient manner [31].

4) Dynamical Safety: In this study, ensuring the vehicle’s dynamical safety is translated to avoiding single tire lift-off. This is a conservative criterion used to prevent rollover [28]. This requirement could be taken into account directly to enforce a positive vertical load on all four tires at all times. However, vehicle models that could predict the vertical tire loads on all four tires would require a level of complexity whose computational load would be prohibitively high for the purpose of MPC. Therefore, another conservative approximation of the dynamical safety requirement is considered; namely, an upper bound on the steering angle magnitude as expressed by the following inequality constraint

$$|\delta_f(t)| \leq \delta_{f,\text{max}}(U_0) \tag{33}$$

where the maximum steering angle $\delta_{f,\text{max}}$ is a function of only the vehicle speed when the vehicle is assumed to move on a flat surface.

For all combinations of longitudinal speed ranging from 10 m/s to 30 m/s and maximum steering angle ranging from 0° to 14°, the corresponding minimum tire vertical loads are obtained using the 14 DoF vehicle model. The relationship is shown in Fig. 12a. If a minimum vertical load threshold is set, the relationship between the maximum steering angle and longitudinal speed can be extracted. Fig. 12b shows the relationship when $F_{z,\text{min}}$ is set as 500 N.

5) Solution Techniques: The formulated nonlinear multi-phase optimal control problems are solved using a two-step procedure. First, the continuous-time OCP is transcribed into a nonlinear programming (NLP) problem using a direct method called hp-pseudospectral method [29], [30], [31]. Second, the resulting NLP problem is solved using the interior point method [32].

The hp-pseudospectral method discretizes a continuous-time OCP into an NLP problem by approximating the state and variable-degree polynomial approximations of them within each interval. The differential-algebraic constraints of the OCP are enforced at a finite set of collocation points, where the collocation points are Legendre-Gauss-Radau (LGR) quadrature points. This method has been shown to be able to accurately approximate the solution to a general continuous-time OCP in a computationally efficient manner [31].
where $\mu$ is a small positive scalar called “barrier parameter”. As $\mu$ converges to zero, the solution to Eq. (37) should converge to a solution to Eq. (36). $B(\cdot)$ is a barrier function.

As an example, Fig. 13a shows the trajectory iterations in solving the problem given in Fig. 7. This is a four-phase problem. The initial guess is a straight line assuming equal length at each phase, which is not a feasible solution. Nevertheless, after 77 iterations, the solution converges to the optimal solution. Fig. 13b shows the corresponding objective value and maximum constraint violation at all steps.

IV. SIMULATION RESULTS AND DISCUSSION

In this section, numerical simulations of the developed nonlinear MPC obstacle avoidance algorithm with a 14 DoF vehicle model as the plant are presented. Table I gives part of the vehicle parameters used by the 14 DoF vehicle model and all the parameters used by the 2 DoF vehicle model.

The first set of simulations uses a LIDAR with detection range of 100 m. The results of the simulations are presented in Fig. 14. These results show that the developed algorithm can successfully navigate the vehicle through the specified obstacle field at 10 m/s, 15 m/s, 20 m/s, and 25 m/s. At these speeds, the vehicle avoids all obstacles, passes the target from the desired direction, and is dynamically safe as shown in Fig. 14c. However, the vehicle hits the second obstacle when the longitudinal speed is maintained at 30 m/s. This is because at this speed, the vehicle is not capable of making a turn at a smaller radius safely. A threshold of 500 N is set on the minimum tire vertical load and the corresponding maximum steering angle is set as a hard constraint in the OCP formulation. This constraint is active at most of the time during the maneuver as shown in Fig. 14b.

Three scenarios are considered in this section. In the first scenario, the vehicle is required to move from its initial location to a target location with the final heading angle required to be the same as its initial heading angle. Two obstacles are between the two locations. Vehicle speed ranging from 10 m/s to 30 m/s is considered.

In the second scenario, the vehicle has to traverse a dense obstacle field to reach the target location. There are 50 obstacles and each of them is 10 m × 10 m in size. The vehicle longitudinal speed is maintained at 20 m/s and there is no constraint on the final heading angle.

In the third scenario, the vehicle performs a double lane change maneuver at 15 m/s using the obstacle avoidance algorithm.

A. Scenario 1: Various Speeds

Table II summarizes the parameters of the nonlinear MPC algorithm, including weights in the cost function, safety margin, LIDAR detection range, length of prediction horizon, length of execution horizon, and maximum steering angle. The cost function given by Eq. 17 is used because the angle of passing the goal is specified. In the settings, the following relationship, which is used to ensure that all the predicted trajectories lie within the LIDAR detection range, is used

$$T_{p, \text{max}} = \frac{R_{\text{LIDAR}}}{U_0}$$

TABLE II

<table>
<thead>
<tr>
<th>$U_0$ (m/s)</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>25</th>
<th>30</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>$w_{\phi}$ (-)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$w_{d}$ (-)</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>$w_{s}$ (-)</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>$l_{SM}$ (m)</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>$R_{\text{LIDAR}}$ (m)</td>
<td>100</td>
<td>140</td>
<td>100</td>
<td>140</td>
<td>100</td>
<td>140</td>
</tr>
<tr>
<td>$T_{p, \text{max}}$ (s)</td>
<td>10.0</td>
<td>6.7</td>
<td>5.0</td>
<td>4.0</td>
<td>3.3</td>
<td>4.7</td>
</tr>
<tr>
<td>$T_{e}$ (s)</td>
<td>0.67</td>
<td>0.44</td>
<td>0.33</td>
<td>0.27</td>
<td>0.22</td>
<td>0.31</td>
</tr>
<tr>
<td>$\zeta_f, \text{max (}^\circ/s\text{)}$</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>$\delta_f, \text{max (}^\circ\text{)}$</td>
<td>10.5</td>
<td>5.14</td>
<td>3.18</td>
<td>2.24</td>
<td>1.72</td>
<td>1.72</td>
</tr>
</tbody>
</table>

The navigation at 30 m/s fails because the LIDAR detection range is not long enough and hence the prediction horizon
is too short to prepare the vehicle to avoid the obstacles sufficiently early. Fig. 15 shows the results of simulations with different LIDAR detection ranges. When a longer detection range of 140 m and a longer prediction horizon are used, the vehicle travels through the field safely.

Because the on-board sensors provide information about the environment only within the close proximity of the vehicle, the obstacle avoidance algorithm is not capable of determining the maximum speed that can be used to safely navigate through an obstacle field. The maximum speed should come from a high-level planner. Alternatively, a conservative lower bound on the prediction horizon or a conservative lower bound on the sensor detection range can be imposed to ensure that the obstacle avoidance maneuver is performed early enough.

These limits could be obtained from the trajectory for making a 90° turn when the initial steering angle is at the minimum bound, which is considered as the most extreme maneuver. The trajectories from speeds ranging from 10 m/s to 30 m/s are shown in Fig. 16. The time of completing
this maneuver is considered as the minimum prediction time, which is summarized in Fig. 17a. According to Eq. (38), the minimum detection range is given by Fig. 17b. As shown in the figure, the speed should be limited below 17 m/s when the LIDAR detection range is 100 m if this conservative bound is used.

B. Scenario 2: Dense Obstacle Field

This simulation is to test the capability of the algorithm within a dense obstacle field. In this simulation, the vehicle speed is maintained at 20 m/s and there is no constraint on the final heading angle. Hence, Eq. (11) is used as the cost function and the simulation parameters as the same as the one corresponding to 20 m/s in Table II except that $w_s$ is not used. Fig. 18 shows the simulation results. The vehicle clears the obstacle field and reaches the target successfully using the algorithm.

In this scenario, at most of the steps, there are multiple feasible openings as exemplified by Fig. 19. For each of the feasible openings, an OCP is formulated and solved. After all of them are solved, their objective function values are compared and the one with the smallest value is considered the best solution. In this example, the objective values of the calculated trajectories from right to left are 0.76, 0.74, 0.92, respectively. The smallest one is 0.74 and the control commands corresponding to the trajectory in the middle is sent to the plant.

C. Scenario 3: Double Lane Change

The last simulation is to test the capability of the algorithm in an on-road scenario. The vehicle performs a double lane change maneuver at 15 m/s using the nonlinear MPC
algorithm. Table III summarizes the parameters used. In this simulation, the weights $w_d$ and $w_s$ are increased to increase the vehicle’s tendency of following the original lane.

Fig. 20 shows the generated trajectory and the corresponding steering angle. Fig. 20a shows the trajectory of the CoG of the plant and the corresponding trajectories of the four corners of the vehicle. It can be seen that all the trajectories are within the white space, which means that the vehicle is free from collision. Fig. 20b is the corresponding steering sequence.

In this scenario, in most of the steps, there are no “openings” as defined in Fig. 5a. But there are “hypothetical openings”, which are lines connecting an obstacle and an opening that are long enough. The “TRs” are then defined as all regions with a feasible hypothetical opening. Fig. 21 shows the use of a hypothetical opening.

V. CONCLUSION

This paper presents the development of a novel nonlinear MPC algorithm for obstacle avoidance in autonomous ground vehicles using real-time sensor for environment detection. A multi-phase optimal control problem formulation is used to incorporate data from the on-board LIDAR sensor and the dynamic limitations of the vehicle to find an optimal solution. The resulting problem is solved using the pseudo-spectral method and the interior point method. Simulation results show that the developed method can yield a satisfactory performance under various scenarios.

The limitations of this work can be summarized as follows. The vehicle is assumed to travel on a flat terrain at a constant
speed. It is of interest to consider 3D terrains and include the vehicle speed as a second controlled variable. Furthermore, uncertainties in the model or sensor measurements are not yet considered, nor are moving obstacles. Addressing these questions is subject to future work. In addition, further investigations are required to study the impact of the objective function on the algorithm’s performance systematically and develop an approach for adapting the weights in the objective function based on sensor measurements.

REFERENCES


Jiechao Liu received the B.S.E. degree in electrical and computer engineering from Shanghai Jiao Tong University, China, in 2009, and the B.S.E. and M.S.E. degrees in mechanical engineering from the University of Michigan, Ann Arbor, MI, in 2009 and 2011, respectively. He is currently a Ph.D. candidate in the Department of Mechanical Engineering, University of Michigan, Ann Arbor. He is currently with the Automated Modeling Laboratory, University of Michigan, Ann Arbor, where he is researching methods in vehicle-dynamics-conscious real-time hazard avoidance in autonomous ground vehicles. His main research discipline is in system dynamics and control, and his research interests include vehicle dynamics and optimal control.

Paramsothy Jayakumar is a Senior Research Scientist, SAE Fellow, and a member of the Analytics Team at the U.S. Army Tank Automotive Research, Development, & Engineering Center (TARDEC) in Warren, Michigan. Prior to joining U.S. Army TARDEC, he worked for BAE Systems, Ford Motor Company, Altair Engineering, and Engineering Mechanics Research Corporation in the areas of multibody dynamics software development, vehicle dynamics modeling & simulation consulting, simulation technology development, durability load simulation, vehicle instrumentation & loads measurement, and road load engineering. Jayakumar has written more than 100 technical publications including journal articles and conference papers. His research in terramechanics and multibody dynamics won the best paper awards at the National Defense Industrial Association’s Ground Vehicle Systems Engineering and Technology Symposium in 2011 and 2012. He holds a U.S. patent for a system for virtual prediction of road loads and tire modeling. He was also instrumental in developing seven SAE standards for tire testing for the purpose of tire modeling for which he received the SAE 2014 James M. Crawford Technical Standards Board Outstanding Achievement Award. Jayakumar is a member of the U.S. Army Acquisition Corps, an Honorary Fellow of the Department of Mechanical Engineering at the University of Wisconsin Madison, and an Associate Editor for the ASME Journal of Computational and Nonlinear Dynamics. He received his M.S. and Ph.D. degrees in structural dynamics from Caltech, and B.Sc. Eng. (Hons, First Class) from the University of Peradeniya, Sri Lanka.

Jeffrey L. Stein received the B.S. degree in premedical studies from the University of Massachusetts, Amherst, MA, in 1973, and the S.B., S.M, and Ph.D. degrees in mechanical engineering from the Massachusetts Institute of Technology, Cambridge, MA, in 1976, 1976, and 1983, respectively. Since 1983 he has been with the University of Michigan, Ann Arbor, MI, where he is currently a Professor of Mechanical Engineering. His research interests include computer based modeling and simulation tools for system design and control, with applications to vehicle-to-grid integration, vehicle electrification, conventional vehicles, machine tools, and lower leg prosthetics. He has particular interest in algorithms for automating the development of proper dynamic mathematical models, i.e., minimum yet sufficient complexity models with physical parameters.

Tulga Ersal received the B.S.E. degree from the Istanbul Technical University, Istanbul, Turkey, in 2001, and the M.S. and Ph.D. degrees from the University of Michigan, Ann Arbor, MI, USA, in 2003 and 2007, respectively, all in mechanical engineering. He is currently an Assistant Research Scientist in the Department of Mechanical Engineering, University of Michigan, Ann Arbor. MI. His research interests include modeling, simulation, and control of dynamic systems, with applications to energy systems, multibody dynamics, vehicle systems, and biomechanics. Dr. Ersal is a member of the ASME. He is the recipient of Deans and Presidents Awards at the Istanbul Technical University, the Siemens Excellence Award, and three merit fellowships at the Istanbul Technical University and the University of Michigan. He also received the semi-plenary paper award at the 2012 Dynamic Systems and Control Conference.