Fuel Consumption Optimization of Heavy-Duty Vehicles With Grade, Wind, and Traffic Information

Chaozhe R. He[1]
Department of Mechanical Engineering,
University of Michigan,
Ann Arbor, MI 48109
e-mail: hchaozhe@umich.edu

Helmut Maurer
Institute of Computational and
Applied Mathematics,
University of Münster,
Münster D-48149, Germany
e-mail: maurer@math.uni-muenster.de

Gábor Orosz
Department of Mechanical Engineering,
University of Michigan,
Ann Arbor, MI 48109
e-mail: orosz@umich.edu

In this paper, we establish a mathematical framework that allows us to optimize the speed profile and select the optimal gears for heavy-duty vehicles (HDVs) traveling on highways while varying parameters. The key idea is to solve the analogous boundary value problem (BVP) analytically for a simple scenario (linear damped system with quadratic elevation profile) and use this result to initialize a numerical continuation algorithm. Then, the numerical algorithm is used to investigate how the optimal solution changes with parameters. In particular, we gradually introduce nonlinearities (air resistance and engine saturation), implement different elevation profiles, and incorporate external perturbations (headwind and traffic). This approach enables real-time optimization in dynamic traffic conditions, therefore may be implemented on-board.
[DOI: 10.1115/1.4033895]

1 Introduction

A large percentage of freight transport nowadays is carried out by HDVs. According to the U.S. Department of Energy [1] and Department of Transportation [2], more than 10 million registered HDVs deliver 70% of the domestic freight transport counted both in value and in weight, and they are responsible for 17% of the petroleum consumption in the transportation sector. Experiments have shown that given the same route, different driving profiles (speed and gear applied) are implemented by different truck drivers, resulting in a large variation in fuel consumption [3]. This implies that optimizing the driving profile has a large potential for saving fuel. However, when vehicles are driven in dynamic traffic environment, the optimization problem becomes very challenging due to lack of real-time information.

Recent developments in information and communication technologies present opportunities to solve this problem. In particular, vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication may be used to obtain real-time traffic information when optimizing the speed profiles and gear changes. Such systems have a clear advantage compared to sensory systems that only provide information about the immediate surroundings of the vehicle [4]. Moreover, V2V and V2I communication may provide information with relatively small latency compared to a service based on remote data aggregation.

Historically, the problem of obtaining the most fuel-efficient driving profile has been formulated as an optimal control problem. Most of the works in this field solved the problem numerically using dynamic programing (DP) [5–8]. For example, in Ref. [9] the longitudinal dynamics of the vehicle were considered, gear shifts were incorporated, and constraints were added to represent the speed limit. In Refs. [10,11], primitives, obtained from driving samples, were used in the optimization to develop a driver-assist system. Recently, researchers began to incorporate traffic information in the optimization. In Ref. [12], for example, traffic data were incorporated using rolling horizon optimal control (RHOC) (often called model predictive control or MPC). The RHOC problems formulated were typically solved by DP algorithms, which are not implementable on-board in real time due to the high computation demand. A direct solver based on a pseudospectral method was developed in Ref. [13] but its computation speed is still not feasible for on-board implementation.

Some researchers approached the problem from an analytical perspective. In Ref. [14], a necessary condition for optimality was derived using Pontryagin's maximum principle (PMP) while taking into account road elevation. In Refs. [15,16], the solutions were derived analytically for simplified models. In Ref. [17], different driving profiles resulted from PMP and DP were compared, and the results showed that PMP allows faster computation of the optimal solution. In Refs. [18,19], RHOC problem was formulated with geographical and traffic information, and the problem was solved using PMP. In Refs. [20,21], an RHOC problem was solved by using PMP while incorporating information about more than one vehicles ahead. These results showed that the analytical considerations in PMP can be used to speed up the optimization problem.

The arising optimal control problems contain many parameters that may be fixed to some reasonable values, and then, these problems can be solved using particular methods. However, a thorough and systematic study on how the optimal solution changes when the parameters are varied is not available in the literature. Since the performance of the optimal controller depends heavily on the parameter values, understanding how to tune the parameters is equally important as understanding the design framework. In this paper, we use analytical and numerical tools from bifurcation theory to investigate the effects of parameter variations on optimal control problems. This allows us to determine regions in parameter space with qualitatively different dynamics and control actions and quantify the trade-off between different control goals in the different regions.

We start with a modeling framework that allows one to optimize the driving profile to achieve better fuel economy given the elevation profile, headwind, and traffic information along the route. We also take into account the desired terminal time leading to multi-objective optimal control problem. Then, we present a systematic approach to analyze the optimal trajectory for fuel-efficient driving while varying parameters and quantify the trade-offs in different parameter regions. We convert the resulting optimal control problem to a BVP by using PMP. We first solve a simplified BVP analytically and then use numerical continuation to gradually change the parameters until the original nonlinear problem is reached. Also, in order to avoid sensitivity to initial conditions, we use collocation to solve the BVP that is embedded in our

[1]Corresponding author.

Contributed by the Design Engineering Division of ASME for publication in the Journal of Computational and Nonlinear Dynamics. Manuscript received October 23, 2015; final manuscript received May 17, 2016; published online July 27, 2016. Assoc. Editor: Paramsothy Jayakumar.

pseudo-arclength numerical continuation software. This allows us to increase the computational speed by distributing the sensitivity along the trajectory while reacting to changing parameters.

Mathematically, we need to deal with an optimal control problem that involves mixed state-control constraints and singular arcs. Optimal control problems with mixed constraints are challenging due to the nonsmooth and noncontinuous nature of the problem [22]. Necessary optimality conditions for this class of control problems can be found in Refs. [23–27]. In our case, the control variable appears linearly in the dynamics and cost functional. Then, in view of PMP, the optimal control is a concatenation of bang–bang and singular arcs which makes it difficult to generate numerical solutions. In Ref. [28], mixed control-state constraints were handled using saturation functions which can tolerate the discontinuity to some extent. In Ref. [29], numerical continuation was used to study the optimal solution while varying parameters in (smooth) optimal control problems occurring in biomedical imaging. Inspired by these works, in this paper we use the analytical solution of a simplified (linear damped) system to characterize the switching structure between different arcs while varying parameters. This allows us to identify the structure of the optimal controller that may switch between different "bang–bang" and "bang-singular-bang" scenarios as parameters vary. Moreover, using the analytical solution to initialize the numerical continuation allows us to speed up the solution process.

The rest of the paper is organized as follows: The optimal control problem is formulated in Sec. 2, and the necessary condition for optimal trajectories is presented in Sec. 3. We generate the analytical solution of a corresponding linear system in Sec. 4, which is used to initialize the numerical solver discussed in Sec. 5. We analyze the optimal solution with respect to the change of traffic condition in Sec. 6. Finally, we conclude the paper and propose some future research directions in Sec. 7.

2 Multi-Objective Driving Profile Optimization

In this section, we lay out a modeling framework that is used to optimize fuel economy of HDVs. This framework allows the use of different models to describe the vehicle dynamics, a wide variety of fuel consumption maps, and real-time traffic information.

2.1 Optimization Problem. Let us denote the distance (arc-length) traveled by the vehicle as s and the speed of the vehicle by v. The goal is to find the scalar input a_d that minimizes the objective function

$$J_0 = \int_0^{t_f} q(v, a_d)\, dt + \sigma_0\, t_f + \int_0^{t_f} r(v, s) dt \tag{1}$$

subject to the constraints

$$\begin{bmatrix} \dot{s} \\ \dot{v} \end{bmatrix} = \begin{bmatrix} v \\ f(s, v, a_d) \end{bmatrix} \tag{2}$$

$$\begin{bmatrix} s(0) \\ v(0) \end{bmatrix} = \begin{bmatrix} 0 \\ v_0 \end{bmatrix}, \quad \begin{bmatrix} s(t_f) \\ v(t_f) \end{bmatrix} = \begin{bmatrix} s_f \\ v_f \end{bmatrix} \tag{3}$$

$$0 \le a_d \le a_U(v) \tag{4}$$

where the dot represents the derivative with respect to time t. The initial time is considered to be 0, while the terminal time is denoted by t_f and it is considered to be unknown.

The objective functional J_0 in Eq. (1) consists of three parts. The first term represents the total fuel consumption, where the fuel consumption rate $q(v, a_d)$ is a function of the vehicle speed v and control input a_d. The second-term $\sigma_0 t_f$ represents the total cost corresponding to the terminal time t_f with weight σ_0. The third term is a penalty term related to traffic. It includes penalties related to the speed limit and traffic speed.

The dynamic system (2) describes a longitudinal vehicle dynamics which will be specified below. The boundary conditions (3) fix the total arc-length of the route s_f, the initial speed v_0, and the final speed v_f, while Eq. (4) gives a speed-dependent upper bound for the control input a_d. To avoid braking (since it dissipates energy), we require the control input to be non-negative.

2.2 Vehicle Dynamics. The longitudinal dynamics of the HDV is derived using classical mechanics. We assume that no slip occurs on the wheels and that the flexibility of the tires and the suspension can be neglected. Then using the power law, we obtain

$$m_{\text{eff}} \dot{v} = -m\, g \sin\phi - \gamma\, m\, g \cos\phi - k(v + v_w)^2 + \frac{\eta}{R} T_e \tag{5}$$

see Refs. [30,31], where the effective mass $m_{\text{eff}} = m + I/R^2$ contains the mass of the vehicle m, the moment of inertia I of the rotating elements, and the wheel radius R. Furthermore, g is the gravitational constant, ϕ is the inclination angle, γ is the rolling resistance coefficient, k is the air drag constant, v_w is the speed of the headwind, η is the gear ratio (that includes the final drive ratio and the transmission efficiency), and T_e is the engine torque. See Appendix A (Table 2) for parameter values used in this paper, which are for a ProStar truck, a class 8 HDV manufactured by Navistar, Lisle, IL [32]. When units are not spelled out, quantities should be understood in SI units.

Based on Eq. (5), we have

$$f(s, v, a_d) = -\alpha \sin\phi - \beta \cos\phi - \kappa\,(v + v_w)^2 + a_d \tag{6}$$

in Eq. (3) where

$$\alpha = \frac{mg}{m_{\text{eff}}}, \quad \beta = \frac{\gamma\, mg}{m_{\text{eff}}}, \quad \kappa = \frac{k}{m_{\text{eff}}}, \quad a_d = \frac{\eta T_e}{m_{\text{eff}} R} \tag{7}$$

Note that the control input a_d is a rescaled torque, with unit of acceleration (m/s^2), and by choosing the appropriate gear ratio η, one can calculate the corresponding engine torque T_e.

The inclination angle ϕ can be calculated from the elevation profile $h(s)$ that gives the elevation as a function of the distance traveled s. Often elevation is given as a function of the direct distance d. The relationship between s, d, and h is illustrated in Fig. 1. It can be seen that $h'(s) = \sin\phi$ and $h'(d) = \tan\phi$. One needs the arc-length parameterization $d(s)$ to obtain $h(s)$, but since $\phi < 0.05$ rad here we use the approximation $\cos\phi \approx 1$. Moreover, for simplicity we consider no headwind $v_w = 0$. Thus, Eq. (6) can be simplified to

$$f(s, v, a_d) = -\alpha\, h'(s) - \beta - \kappa\, v^2 + a_d \tag{8}$$

Throughout this paper, we consider a straight road with the simple elevation profile

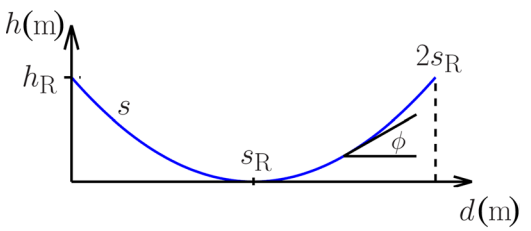

Fig. 1 Elevation h as a function of distance d and arc-lengths s

$$h(s) = h_R \left(\frac{s - s_R}{s_R} \right)^2 \tag{9}$$

shown in Fig. 1. We also consider $h_R \ll s_R$, in particular, we use $h_R = 30\,\text{m}$ and $s_f = 2s_R = 4000\,\text{m}$. Nonetheless, the methods presented can be applied when using other profiles as well.

We emphasize that using a_d as the control input enables us to decouple the optimization of the speed profile and the gear selection: the rescaled torque a_d is derived first, and then, the optimal gear is selected to minimize fuel consumption.

2.3 Fuel Consumption Map. In order to keep the problem analytically tractable, we use a static fuel consumption map, $q(v, a_d)$, that specifies the fuel consumption rate (with unit (g/s)) for a given input a_d and a given speed v. Fuel consumption maps are typically given as a function of the engine speed ω_e and engine torque T_e, that is, $q(\omega_e, T_e)$. Dividing this with the engine power $P_d = T_e\omega_e = m_{\text{eff}} a_d v$, we obtain the brake-specific fuel consumption (BSFC)

$$\text{BSFC} = \frac{q(\omega_e, T_e)}{T_e\omega_e} = \frac{q(v, a_d)}{m_{\text{eff}} a_d v} \tag{10}$$

where we used $a_d = \eta T_e / m_{\text{eff}} R$ and $v = R\omega_e/\eta$; cf. Eq. (7). Small BSFC values typically imply good fuel economy [33]. Previous efforts on fuel economy optimization usually assumed fixed gear ratio, which resulted in a one-to-one relationship between $q(\omega_e, T_e)$ and $q(v, a_d)$ [16,34]. In this paper, we generate a map with the gear changes involved.

Given a control input a_d at a certain speed v, different gears set the engine to different working points, and therefore, yield different BSFC values. We choose the gear that gives the least BSFC among all the available gears and generate the working zone for each gear in the (v, a_d) plane as shown in Fig. 2(*a*) for a MaxxForce 13 diesel engine with a ten-speed transmission used in a ProStar truck manufactured by Navistar, Lisle, IL [32]. We found that for any given (v, a_d) point, there is a single optimal gear ratio, so one can map the fuel consumption from the (ω_e, T_e)-plane to the (v, a_d)-plane using the associated gear, which is shown by the contours in Fig. 2(*b*). It can be observed that the contours of the fuel consumption map are similar to the isopower curves ($P_d = m_{\text{eff}} a_d v$). To obtain an analytical model, we fit the data using the Willans approximation

$$q(v, a_d) = p_2 v\, a_d + p_1 v + p_0 \tag{11}$$

see Ref. [35]. By applying least-square fitting, we obtain $p_2 = 1.8284 \pm 0.0019\,\text{gs}^2/\text{m}^2$, $p_1 = 0.0209 \pm 0.0006\,\text{g/m}$, $p_0 = -0.1868 \pm 0.0068\,\text{g/s}$. The corresponding contours are shown in Fig. 2(*c*). With the gear applied, the BSFC in (v, a_d)-plane is shown in Fig. 2(*d*), where crosses indicate the minimal points for different gears.

We assume that gear changes occur instantaneously, and the engine's state jumps along isopower curves during gear change. The blank regions in Fig. 2 correspond to (v, a_d) combinations that are not accessible by the engine. The corresponding black-dashed boundary at the top consists of two sections: a constant section at $a_{\max}$ for low speed and an isopower curve at $P_{\max}$ for higher speed. Since $P = T_e\omega_e = m_{\text{eff}} a_d v$, we have

$$a_U(v) = \min\{a_{\max},\, U/v\}, \qquad U = \frac{P_{\max}}{m_{\text{eff}}} \tag{12}$$

in Eq. (4). In this paper, we use $a_{\max} = 2\,\text{m/s}$, $P_{\max} = 300.65\,\text{kW} \Rightarrow U = 10.14\,\text{m}^2/\text{s}^3$ that are acquired through data fitting. We rewrite the constraints (4) and (12) into the form

$$\begin{aligned} 0 \le \frac{a_d}{\min\{a_{\max},\, U/v\}} \le 1 \\ \iff \begin{cases} C_1(v, a_d) := \dfrac{a_d}{\min\{a_{\max},\, U/v\}} - 1 \le 0 \\ C_2(v, a_d) := -\dfrac{a_d}{\min\{a_{\max},\, U/v\}} \le 0 \end{cases} \end{aligned} \tag{13}$$

We note that the mixed constraint trivially satisfies the regularity condition $\partial_{a_d} C_i(v, a_d) \neq 0$, $i = 1, 2$, which allows to obtain the corresponding Lagrange multipliers. Moreover, having the two constraints in a similar form results in multipliers of same scale, which is convenient for numerical computation.

2.4 Penalty on Traffic. We consider two types of penalty on traffic: one related to the speed limit, and the other related to the traffic flow speed, that is, in Eq. (1), we consider

$$r(s, v) = r_1(v_{\text{lim}}(s), v) + r_2(v_{\text{traf}}(s), v) \tag{14}$$

Here, we assume the speed limit penalty

$$r_1(v_{\text{lim}}(s), v) = \rho_1 \left\{ \sec\left[\frac{\pi}{2} \left(\frac{2v}{v_{\text{lim}}(s)} - 1 \right)^n \right] - 1 \right\} \tag{15}$$

where the speed limit $v_{\text{lim}}(s)$ is given along the route and as a function of the arc-length s, and ρ_1 is a constant weight. The function (15) is illustrated in Fig. 3(*a*), with $v_{\text{lim}} \equiv 30\,\text{m/s}$ and $\rho_1 = 0.1\,\text{g/s}$, for different values of n. It can be seen that as parameter n increases, the function becomes more "square shaped." In this paper, we use the power $n = 10$.

For traffic penalty, we use a quadratic function

$$r_2(v_{\text{traf}}(s), v) = \rho_2 (v_{\text{traf}}(s) - v)^2 \tag{16}$$

which is illustrated in Fig. 3(*b*) for the weight $\rho_2 = 1\,\text{gs/m}^2$. Indeed, the traffic penalty increases when the vehicle speed moves away from the traffic flow speed $v_{\text{traf}}(s)$ that is given as a function of the arc-length s. We remark that we use an Eulerian description to describe the traffic flow, that is, v_{traf} shows the "averaged" flow speed at location s and does not necessarily correspond to the

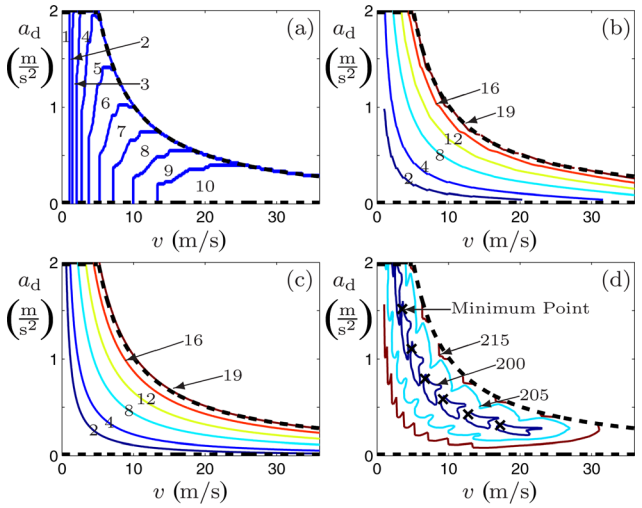

Fig. 2 Contours in the plane of speed *v* and control input a_d. (*a*) Optimal gear ratios; (*b*) experimental fuel consumption contours, with unit (g/s); (*c*) fitted fuel consumption contours, with unit (g/s); and (*d*) BSFC with optimal gear ratios applied, with unit (g/(kW hr)). The black crosses represent points with minimal BSFC value. In all the four panels, black-dashed curves indicate the boundaries of the domain accessible by the engine for all the gears.

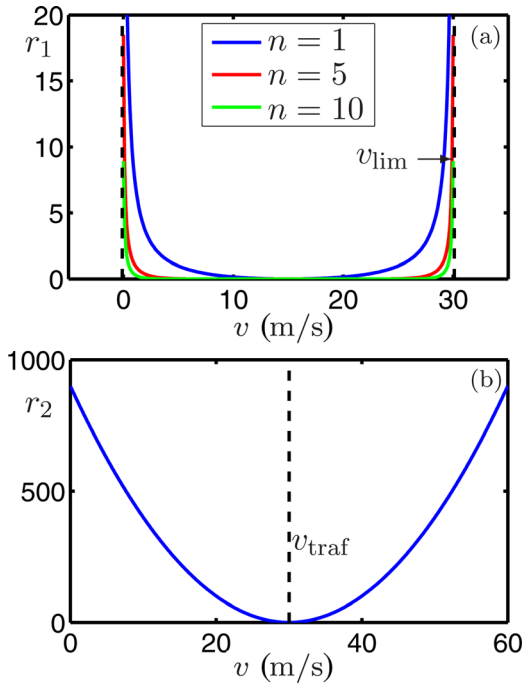

Fig. 3 Traffic penalty given by Eqs. (14)–(16)

speed of an individual vehicle (that would require a Lagrangian description [36]). When increasing the weight ρ_1, one intends to adapt more to the traffic flow speed.

3 Optimal Control Problem and Necessary Conditions for Optimality

Substituting Eq. (11) into Eq. (1), we obtain

$$J_0 = \int_0^{t_f} (p_2 v a_d + p_1 v + p_0)\,\mathrm{d}t + \sigma_0\, t_f + \int_0^{t_f} r(s,v)\,\mathrm{d}t \tag{17}$$

Since $p_1 v$ is independent of the control input a_d and $\int_0^{t_f} p_1 v\,\mathrm{d}t = p_1 s_f$ is constant, these two terms may be dropped. We define $\sigma = p_0 + \sigma_0$, drop the subscript of p_2, and redefine the objective function as

$$J = \int_0^{t_f} p\, a_d v\,\mathrm{d}t + \sigma\, t_f + \int_0^{t_f} r(s,v)\,\mathrm{d}t \tag{18}$$

To minimize this, we need to determine a piecewise continuous control function $a_d : [0, t_f] \to \mathbb{R}$ that minimizes the functional (18) subject to the dynamic system (2) (or specifically Eq. (8)), the boundary conditions (3), and the mixed state-control constraints (4) (or specifically Eq. (13)). For systems with mixed constraints, necessary conditions of optimality were derived in Refs. [24,27,37] in the form of PMP, and the following analytical framework is based on these works.

Let $x = [s, v]^{\mathrm{T}}$ be the state vector of the system and $\lambda = [\lambda_s, \lambda_v]^{\mathrm{T}}$ be the associated costate vectors. The standard Hamiltonian function is defined as

$$\begin{aligned} H(x,\lambda,a_d) = {} & p a_d v + \sigma + r(v,s) \\ & + \lambda_s v + \lambda_v(-\alpha h'(s) - \beta - \kappa v^2 + a_d) \end{aligned} \tag{19}$$

Then, the *augmented* Hamiltonian that takes into account the mixed constraints is given by

$$\mathcal{H}(x,\lambda,\mu,a_d) = H(x,\lambda,a_d) + \mu^{\mathrm{T}} C(x,a_d) \tag{20}$$

where $C = [C_1, C_2]^{\mathrm{T}}$ is given by Eq. (13), and $\mu = [\mu_1, \mu_2]^{\mathrm{T}}$ are the Lagrange multipliers.

Let x^*, a_d^* denote a local minimum pair of the optimal control problem. Then, according to Ref. [24] there exists a piecewise continuous costate $\lambda^* : [0, t_f] \to \mathbb{R}^2$ and a piecewise continuous multiplier $\mu^* : [0, t_f] \to \mathbb{R}^2$ such that the following conditions hold:

(1) adjoint equations

$$\dot{\lambda}^* = -\partial_x \mathcal{H}(x^*, \lambda^*, \mu^*, a_d^*) \tag{21}$$

(2) minimum condition for the standard Hamiltonian

$$H(x^*(t), \lambda^*(t), a_d^*(t)) = \min_{u \in \Omega(t)} H(x^*(t), \lambda^*(t), u) \tag{22}$$

with (momentarily) admissible control set

$$\Omega(t) = \{u \,|\, 0 \le u \le \min\{a_{\max},\, U/v^*(t)\}\}$$

(3) local minimum condition for the augmented Hamiltonian

$$\partial_{a_d} \mathcal{H}(x^*, \lambda^*, \mu^*, a_d^*) = 0 \tag{23}$$

(4) complementarity condition

$$\mu^* \ge 0, \quad \mu^{*\mathrm{T}} C(x^*, a_d^*) = 0 \tag{24}$$

(5) transversality condition (for free terminal time t_f)

$$H(x^*(t_f), \lambda^*(t_f), a_d^*(t_f)) = 0 \tag{25}$$

Note that because C_1 and C_2 cannot be active at the same time, the complementarity condition (24) actually leads to $\mu_i^* C_i = 0$, $i = 1, 2$ when either of the constraints is active and, consequently, the constraint qualification condition is trivially satisfied [22]. Also note that we do not get a terminal condition for $\lambda(t_f)$, since the terminal state $x(t_f)$ is fixed.

From now on, we abuse the notation and drop the * for the optimal solution. Using Eqs. (19) and (20) the adjoint equation (21) becomes

$$\begin{bmatrix} \dot{\lambda}_s \\ \dot{\lambda}_v \end{bmatrix} = \begin{bmatrix} \lambda_v \alpha h''(s) - \partial_s r(v,s) \\ -\lambda_s + 2\lambda_v \kappa v - p a_d - \partial_v r(v,s) - \mu_1 \partial_v C_1 - \mu_2 \partial_v C_2 \end{bmatrix} \tag{26}$$

where ∂_s and ∂_v denote the partial derivatives with respect to s and v. Since the control variable a_d appears linearly in the Hamiltonian (19), the minimum condition (22) gives the controller

$$a_d(t) = \begin{cases} \min\{a_{\max}, U/v(t)\}, & \text{if} \quad \zeta(t) < 0 \\ a_d^{\text{sing}}(t), & \text{if} \quad \zeta(t) = 0,\ t \in I_s \subset [0, t_f] \\ 0, & \text{if} \quad \zeta(t) > 0 \end{cases} \tag{27}$$

where $\zeta(t) = \zeta(v(t), \lambda_v(t))$ is the *switching function* given by

$$\zeta(v, \lambda_v) = \partial_{a_d} H = pv + \lambda_v \tag{28}$$

while the control input a_d^{sing} along the singular arc is described further below. To determine the multipliers μ_1 and μ_2, we use the local minimum condition (23), which gives

$$\mu_1 = -\frac{\zeta}{\partial_{a_d} C_1}, \quad \mu_2 = -\frac{\zeta}{\partial_{a_d} C_2} \tag{29}$$

Note that the sign condition $\mu_1 \geq 0$ and $\mu_2 \geq 0$ in Eq. (24) are in accordance with the sign of the switching function ζ in Eq. (28), and we exploited that C_1 and C_2 cannot be active at the same time.

Now, we derive a formula of the singular control a_d^{sing} in Eq. (27). We achieve this by differentiating Eq. (28) with respect to time until a_d appears explicitly. The derivation is carried out for the case with no traffic penalty ($r(s,v) \equiv 0$ in (14), but it can easily be generalized when adding traffic penalty. Assume that control takes values in the interior of the control set, that is, $C_i(v, a_d) < 0$ for $i = 1, 2$, in a certain time interval I_s. Then, the singular arc is given by

$$\zeta = pv + \lambda_v = 0 \tag{30}$$

cf. Eqs. (27) and (28). Differentiating this with respect to time and using Eqs. (2), (8), and (26), we obtain

$$\dot{\zeta} = p(-\alpha h'(s) - \beta - \kappa v^2) - \lambda_s + \lambda_v 2\kappa v = 0 \tag{31}$$

Note that the control variable a_d drops out in accordance with the theory of singular control, see Ref. [38]. Substituting $\lambda_v = -pv$, we attain

$$\dot{\zeta} = p(-\alpha h'(s) - \beta - 3\kappa v^2) - \lambda_s = 0 \tag{32}$$

Using again Eqs. (2), (8), and (26), the second derivative of ζ is computed as

$$\ddot{\zeta} = p(-\alpha h''(s)v - 6\kappa v(-\alpha h'(s) - \beta - \kappa v^2 + a_d)) - \lambda_v \alpha h''(s) = 0 \tag{33}$$

Substituting $\lambda_v = -pv$ again, we obtain

$$\ddot{\zeta} = -p6\kappa v(-\alpha h'(s) - \beta - \kappa v^2 + a_d) = -p6\kappa v \cdot \dot{v} = 0 \tag{34}$$

The control a_d appears explicitly in the second derivative of ζ and, hence, a singular arc is of *first-order*, see Ref. [38]. According to Eq. (34), a singular arc is characterized by the condition $\dot{v} \equiv 0$ of constant speed which yields the singular control

$$a_d^{sing}(s, v) = \alpha h'(s) + \beta + \kappa v^2 \tag{35}$$

Moreover, note that the strict generalized Legendre–Clebsch condition, which is a higher order necessary condition for the singular control to be optimal, holds since

$$-\frac{\partial \ddot{\zeta}}{\partial a_d} = 6p\kappa v > 0 \tag{36}$$

see Ref. [38]. To interpret the physical meaning of singular control (35), we give the following two lemmas.

LEMMA 1. *Assume that $\kappa \neq 0$ in Eq. (8), no constraints are applied to control input a_d, and no traffic penalties are considered ($r(s, v) \equiv 0$ in Eq. (14)). Then, the necessary and sufficient condition to maintain $\zeta = 0$ has a constant speed $v \equiv \sqrt[3]{\sigma/(2p\kappa)}$, which can be calculated by setting $\dot{v} \equiv 0$ and using Eq. (25).*

LEMMA 2. *If the gradient $h'(s)$ is such that $0 \leq \alpha h'(s) + \beta + \kappa v^2 < \min\{a_{max}, U/v\}$ for $s \in [0, s_f]$ and we assume $v(0) = v(t_f) = v_0$, then the singular control input (35) keeps the speed constant along the whole route.*

According to these lemmas, if the combined effect of the grade, rolling resistance, and wind is small, then limiting the acceleration (by traveling with constant speed) is the best strategy to minimize the fuel consumption. On the other hand, it will be demonstrated below that when the external effects are significant, the optimal controller switches between the minimum and maximum of a_d.

In the rest of this paper, we will use the vector notation

$$X = [s, v, \lambda_s, \lambda_v]^T \tag{37}$$

Then, the BVP (2), (3), (8), (25), and (26) can be summarized as

$$\begin{bmatrix} \dot{s} \\ \dot{v} \\ \dot{\lambda}_s \\ \dot{\lambda}_v \end{bmatrix} = \begin{bmatrix} v \\ -\alpha h'(s) - \beta - \kappa v^2 + a_d \\ \lambda_v \alpha h''(s) - \partial_s r(v, s) \\ -\lambda_s + 2\lambda_v \kappa v - p a_d - \partial_v r(v, s) - \mu_1 \partial_v C_1 - \mu_2 \partial_v C_2 \end{bmatrix} \tag{38}$$

$$\underbrace{\begin{bmatrix} s(0) \\ s(t_f) - s_f \\ v(0) - v_0 \\ v(t_f) - v_f \\ H(s(t_f), v(t_f), \lambda_s(t_f), \lambda_v(t_f), a_d(t_f)) \end{bmatrix}}_{\mathbf{B}(X(0), X(t_f))} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{39}$$

There is no general existence and uniqueness conditions for BVPs, even for smooth dynamic systems [39]. Therefore, whether a solution exists is unknown when setting the parameter values. Even if (27), (38), and (39) consist of smooth subsystems, switches make the system nonsmooth. Moreover, varying the parameter σ changes the boundary condition (39), and consequently, changes the optimal solution. We also remark that Eqs. (27), (38), and (39) essentially give a multipoint boundary value problem (MBVP) since the switching times $t_k \in (0, t_f)$, where the switching function $\zeta(t)$ vanishes, are not known a priori. Standard numerical methods like the shooting method cannot be applied directly to solve this problem as the system may become very sensitive to initial conditions due to the switches at the interim points. In this paper, we bypass this issue by using smoothing techniques and apply collocation to obtain the solutions. These combined with pseudo-arclength continuation allow us to trace the optimal solution while varying parameters. To obtain an initial guess of the solution, we analyze the switching structure of the optimal controller for a simplified linearized model.

Alternatively, one can discretize the problem by using a large number of grid points and solve the resulting nonlinear programing problem, e.g., by the interior-point optimization code IPOPT [40] that is implemented using the applied modeling programing language [41]. We use this direct "discretize and then optimize" approach to check the solutions provided by the proposed "collocation and continuation" method explained above. However, we remark though that the direct method requires much finer time mesh to obtain the optimal solution. In Sec. 4, we present the analytical solution for a simplified problem where nonlinearities are neglected. This analytical solution will be used to initialize our numerical solver proposed in Sec. 5.

4 Analytical Solution of the Linear Damped System

In this section, we simplify the BVP (27), (38), and (39) to a linear system with simple constraints and derive the analytical solution. This allows us to characterize how the optimal solution changes with the parameter σ. Moreover, the analytical solution will be utilized in Sec. 5 to initialize the numerical continuation when solving the original nonlinear BVP. Here, the air drag is substituted by linear damping, i.e., κv^2 is replaced by $\kappa v_0 v$,

where $v_0 = v(0)$. Furthermore, traffic penalty is omitted by setting $r(s,v) \equiv 0$ in Eq. (14). Finally, the constraints (13) are substituted by $0 \le a_d \le a_{max}$. Thus, Eq. (38) is simplified to the affine equations:

$$\begin{bmatrix} \dot{s} \\ \dot{v} \\ \dot{\lambda}_s \\ \dot{\lambda}_v \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -2\alpha \dfrac{h_R}{s_R^2} & -\kappa v_0 & 0 & 0 \\ 0 & 0 & 0 & 2\alpha \dfrac{h_R}{s_R^2} \\ 0 & 0 & -1 & \kappa v_0 \end{bmatrix} \begin{bmatrix} s \\ v \\ \lambda_s \\ \lambda_v \end{bmatrix} + \begin{bmatrix} 0 \\ 2\alpha \dfrac{h_R}{s_R} - \beta + a_d \\ 0 \\ -p a_d \end{bmatrix} \tag{40}$$

which can be solved analytically, given the control input a_d. According to Eq. (27), the input a_d stays at its maximum a_{max} or minimum 0 when $\zeta \neq 0$. Therefore, we obtain the solution in the form when a_d is constant

$$\begin{bmatrix} s(t) \\ v(t) \\ \lambda_s(t) \\ \lambda_v(t) \end{bmatrix} = \underbrace{\begin{bmatrix} F_1(s_0, v_0, t, a_d) \\ F_2(s_0, v_0, t, a_d) \\ F_3(\lambda_{s0}, \lambda_{v0}, t, a_d) \\ F_4(\lambda_{s0}, \lambda_{v0}, t, a_d) \end{bmatrix}}_{\mathbf{F}(X_0, t, a_d)} \tag{41}$$

where we used the abbreviated notation: $s(0) = s_0$, $v(0) = v_0$, $\lambda_s(0) = \lambda_{s0}$, $\lambda_v(0) = \lambda_{v0}$, $X_0 = [s_0, v_0, \lambda_{s0}, \lambda_{v0}]^T$. Note that, when a_d changes in time, we may still get an expression similar to Eq. (41), but the right-hand side becomes a functional. In this case, the dynamics of s, v and the dynamics of λ_s, λ_v are not decoupled anymore, but the control law (27) makes them coupled through switching. Even though we cannot guarantee that a unique solution exists, we assume that the boundary condition (39) allows at least one solution.

If the system does not satisfy the conditions in Lemma 2, the control input switches between the cases in Eq. (27). Switches may occur in six different ways as illustrated in Fig. 4 where the switching surface $\zeta = 0$ is also depicted. The control input a_d either switches from the maximum a_{max} to minimum 0 (trajectories 1 and 2) or vice versa (trajectories 4 and 5). Trajectories 3 and 6 do not cross the switching surface but attach to it and leave to the same side, so the control input stays at the minimum (trajectory 3) or at the maximum (trajectory 6). We call them *grazing scenarios*. Trajectories 1 and 5 are named *traverse scenarios* since these trajectories go through the switching surface. On the other hand, trajectories 2 and 4 attach to the plane, travel along it, and leave it on the other side. We name these *tangential scenarios*. We remark that the switching structure is determined by the boundary conditions, and in general, multiple switches may occur.

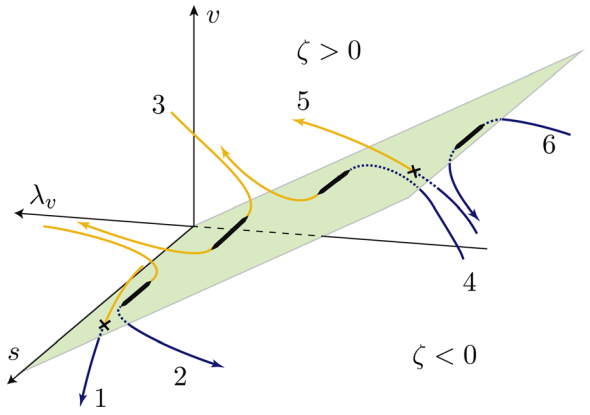

Fig. 4 Six possible scenarios involving one switch. For trajectories 1 and 2, a_d switches from maximum to minimum (i.e., $\zeta = pv + \lambda_v$ switches from negative to positive). For trajectories 4 and 5, a_d switches from minimum to maximum (i.e., $\zeta = pv + \lambda_v$ switches from positive to negative). Trajectories 1 and 5 represent transverse scenarios, while trajectories 2 and 4 show tangential scenarios. Trajectories 3 and 6 are for the grazing scenarios, where the minimum and the maximum control input is maintained, respectively. Bold segments and crosses indicate $\zeta = 0$.

By solving the affine equation (40) analytically, the linear systems (27), (39), and (40) can be transformed to a system of nonlinear algebraic equations. By using the notation defined in Eqs. (39) and (41), for the traverse scenario we obtain

$$\begin{aligned} &\mathbf{B}(X_0, X_f) = 0 \\ &X_1 = \mathbf{F}(X_0, t_1, a_{d1}) \\ &X_f = \mathbf{F}(X_1, t_f - t_1, a_{df}) \end{aligned} \tag{42}$$

where $a_{d1}, a_{df} \in \{0, a_{max}\}$ such that $a_{d1} \neq a_{df}$. Solving these equations, we obtain $t_1, t_f, \lambda_{s0}, \lambda_{v0}$. On the other hand, for the tangential and grazing scenarios, we have

$$\begin{aligned} &\mathbf{B}(X_0, X_f) = 0 \\ &p\dot{v} + \dot{\lambda}_v = 0, \quad \text{for} \quad t \in [t_1, t_2] \\ &X_1 = \mathbf{F}(X_0, t_1, a_{d1}) \\ &X_f = \mathbf{F}(X_2, t_f - t_2, a_{df}) \end{aligned} \tag{43}$$

where $a_{d1}, a_{df} \in \{0, a_{max}\}$ such that $a_{d1} \neq a_{df}$ for tangential scenarios and $a_{d1} = a_{df}$ for grazing scenarios. Solving these equations, we obtain $t_1, t_2, t_f, \lambda_{s0}, \lambda_{v0}$. The overall solution of the MBVP will be a series of segments that are given analytically.

As an example, we consider the case $v(t_f) = v(0) = 25\,\mathrm{m/s}$ and $a_{max} = 0.6\,\mathrm{m/s^2}$. We set a_{max} at this value since it is close to the minimum value of U/v in the velocity range we consider, cf. Fig. 2. With these parameters, all the types of switches shown in Fig. 4 can be obtained except trajectory 6. Meanwhile, solutions with multiple switches also appear.

The range of parameter σ is divided into six domains, and the appearing six different types of solutions are shown in Fig. 5. Dashed curves represent analytical solutions, while solid curves represent numerical simulations for the same initial condition, and they match very well. We compared these trajectories with those using a direct method "discrete and optimize" method mentioned in Sec. 3, and they essentially give the same results.

When $\sigma \in [0, 1.05]\mathrm{g/s}$, the optimal solution is of transverse scenario 1, and the control input switches from minimum to maximum, see case A in Fig. 5. When $\sigma \in [1.05, 3.71]\,\mathrm{g/s}$, the optimal solution is of attached scenario 2, and the control input changes from minimum to maximum but the trajectory attaches to the surface $\zeta = 0$ in the middle, see case B in Fig. 5. When $\sigma \in [3.71, 4.33]$ (g/s), the optimal solution is of grazing scenario 3, and the control input starts from minimum and ends at minimum while attaching to the plane $\zeta = 0$ in the middle, see case C in Fig. 5. When $\sigma \in [4.33, 9.43]$ (g/s), trajectories with two switches exist so that the transverse scenario 5 is followed by the attached scenario 3, see case D in Fig. 5. When $\sigma \in [9.43, 12.73]$ (g/s), the optimal solution is of attached scenario 4, see case E in Fig. 5. Finally, when $\sigma \in [12.73, +\infty)\,\mathrm{g/s}$, the optimal solution is of transverse scenario 5, see case F in Fig. 5.

The terminal time t_f and total fuel consumption Q are plotted in Fig. 6 as a function of σ, where the total fuel consumption is defined as

$$Q = \int_0^{t_f} q(v, a_d)\,\mathrm{d}t = \int_0^{t_f} (p_2 v a_d + p_1 v + p_0)\,\mathrm{d}t \tag{44}$$

cf. Eqs. (1) and (11). The six domains explained above are separated by the black vertical lines, with the numbers 1–5 indicating

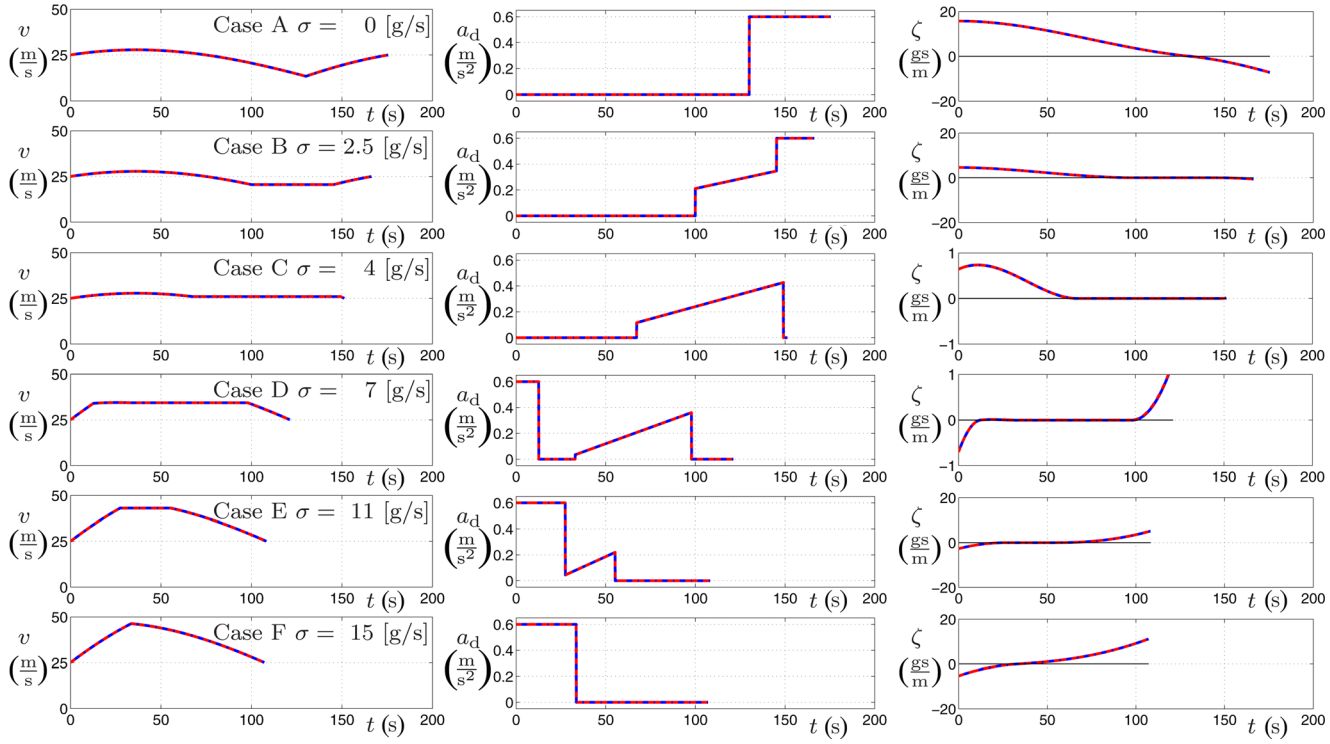

Fig. 5 Time evolution of the speed v (left column), the corresponding control input a_d (middle column), and the switching variable ζ (right column) for different values of the parameters σ as indicated. The rows correspond to the points A–F marked in Fig. 6. Red-dashed curves represent analytical solutions, while solid curves represent numerical solutions and they match very well.

the solution type corresponding to those in Fig. 4, while letters A–F correspond to the cases in Fig. 5. Note that the solution $\sigma > 12.73$ g/s, the trajectories for s, v, and a_d are essentially the same as those for $\sigma = 15$ g/s.

The analysis of this simplified problem shows that by varying the system parameters, the optimal solution may change qualitatively. To experience the trade-off between travel time and fuel consumption, one shall set $\sigma \in [3, 15]$, where, indeed, better fuel economy leads to a longer traveling time. Meanwhile, it is possible to use the analytical solution to start the numerical continuation and gradually steer the system to the original nonlinear system. We will introduce the numerical technique in Sec. 5.

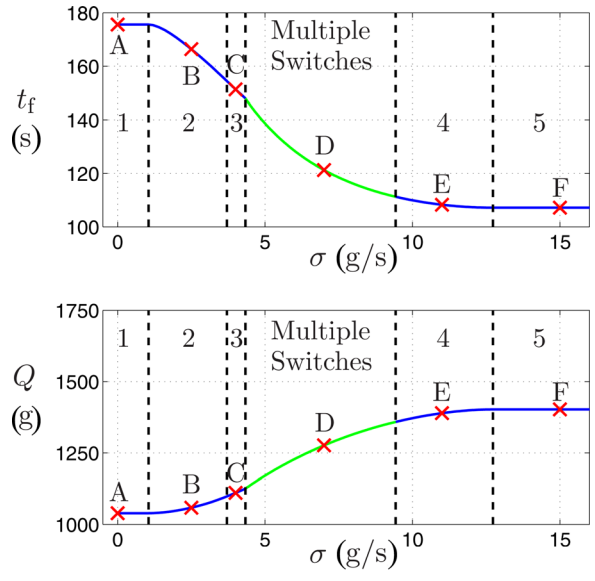

Fig. 6 The upper panel gives the terminal time t_f as a function of σ, while the lower panel shows the fuel consumption Q (44) as a function of σ, for $v(t_f) = v(0) = 25\,\mathrm{m/s}$, $a_{max} = 0.6\,\mathrm{m/s^2}$ for the systems (27) and (40). The dashed vertical lines separate six regions of qualitatively different solutions, with the numbers corresponding to those in Fig. 4. Points A–F correspond to the cases in Fig. 5.

5 Numerical Solution of the Full Nonlinear System

In this section, we investigate the original nonlinear BVP (27), (38), and (39) using numerical continuation. This technique was originally developed to compute solutions of systems of parameterized nonlinear equations [42]. The idea is to start from a solution with certain set of system parameters and gradually change the parameters until the target parameters are reached, see Appendix B for a brief overview of the method. We start from the solution of the simplified systems (27), (39), and (40) and add nonlinearities gradually until we reach (27), (38), and (39) by varying the parameters. Besides the nonlinear bang–bang controller (27), the original problem also has other types of nonlinearity: nonlinear input constraint $a_U(v)$ in Eq. (13), air drag in Eq. (8), and traffic penalty $r(v, s)$ in Eq. (1).

Instead of starting from $U \to \infty$, we set $U = 100\ \mathrm{m^2/s^3}$, so that $U/v \gg a_{max} = 0.6\,\mathrm{m/s^2}$, cf. Fig. 2. Then, we decrease U until we reach $U = P_{max}/m_{eff} = 10.14\mathrm{m^2/s^3}$. After that we change the maximum acceleration gradually from $a_{max} = 0.6\,\mathrm{m/s^2}$ to $a_{max} = 2\,\mathrm{m/s^2}$, though this last step does not change the optimal trajectories in the cases considered here. The nonlinear air drag term is added by varying $\tilde{\kappa}$ from 0 to κ in $\tilde{\kappa}v^2 + (\kappa - \tilde{\kappa})v_0 v$. The traffic penalty (14) is introduced in similar manners. For speed limit r_1, initially we set v_{lim} in Eq. (15) to be large and gradually decrease it to the target value while ρ_1 is kept fixed. For penalty r_2 corresponding to the deviation from the traffic speed, we introduce it gradually by increasing ρ_2.

Since continuation requires a certain level of smoothness, we derive an approximate system by smoothing Eqs. (27), (38), and (39). Specifically, the smoothed version of constraint (12) is written as

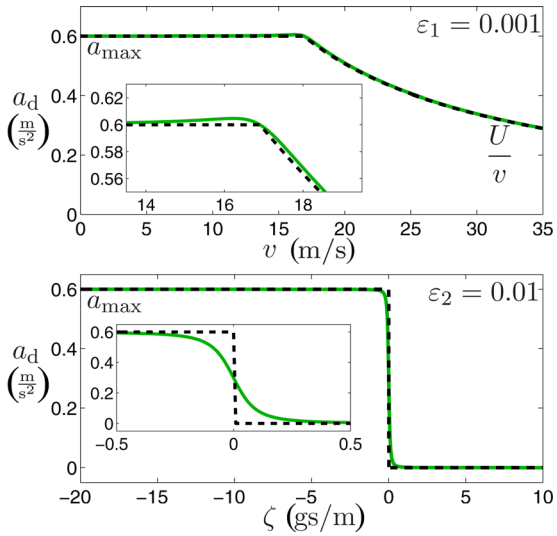

Fig. 7 Visualization of the nonsmooth functions (12) and (27) (black dashed) and the corresponding smooth functions (45) and (46) (solid)

$$\tilde{a}_{\mathrm{U}}(v) = \frac{a_{\max} + U/v}{2} - \frac{(a_{\max} - U/v)^2}{2\sqrt{\varepsilon_1 + (a_{\max} - U/v)^2}} \tag{45}$$

while the smoothed version of the switching rule (27) is given as

$$a_{\mathrm{d}} = \frac{1}{2} a_{\mathrm{U}}(v) \left(1 - \frac{\zeta}{\sqrt{\varepsilon_2 + \zeta^2}}\right) \tag{46}$$

where ε_1 and ε_2 are the small parameters with units ($\mathrm{m}^2/\mathrm{s}^4$) and ($\mathrm{g}^2\ \mathrm{s}^2/\mathrm{m}^2$), respectively. In Fig. 7, the nonsmooth functions (12) and (27) (black-dashed curves) are compared to Eqs. (45) and (46) (solid curves), for $\varepsilon_1 = 0.001\ \mathrm{m}^2/\mathrm{s}^4$, $\varepsilon_2 = 0.01\ \mathrm{g}^2\ \mathrm{s}^2/\mathrm{m}^2$, and $a_{\max} = 0.6\mathrm{m}/\mathrm{s}^2$. Indeed, the smoothed curves approximate the nonsmooth ones well.

Our BVP solver is based on pseudo-arclength continuation algorithm. We discretize the BVP using collocation method [43] which results in a large system of nonlinear equations that also depends on parameters. Then, a continuation algorithm is used to solve these nonlinear equations while varying parameters. To speed up the computation, we use adaptive steps. When solving BVPs with singular arcs with other methods like the shooting method, see Refs. [22,39,44], difficulties often arise due to the sensitivity with respect to initial values. The collocation method can bypass this problem because it tunes the whole solution, that is, distributes the sensitivity along the whole trajectory. In this paper, we use the collocation method with 400 points $\varepsilon_1 = 10^{-6}\mathrm{m}^2/\mathrm{s}^4$ and $\varepsilon_2 = 0.01\mathrm{g}^2\mathrm{s}^2/\mathrm{m}^2$.

The time evolution of the system (27), (38), and (39) is shown in Fig. 8 for different values of σ as indicated (different from those used in Figs. 5 and 6 for better illustration of the switching structure). The solid curves are the trajectories acquired by the our BVP solver based on pseudo-arclength continuation method using the smoothed controller (45) and (46), while the dashed curves are generated by a direct method using the nonsmooth controller (27). The two trajectories in each panel are close to each other, implying that the smoothed controller represents the original nonsmooth one with a good accuracy. Note that when applying our method we used 400 points to represent the trajectory, whereas for the direct method we used 10,000 points which increases the computational demand significantly. The numerical solutions of the nonlinear system maintain the same trend as the analytical solutions of the linear system. That is, as σ increases, the controller starts with minimum to maximum bang–bang type controller. Then, a singular arc appears. Eventually, it becomes a bang–bang type again but with maximum to minimum switch. We also plot the terminal time t_{f} and the total fuel consumption Q (cf. Eq. (44)) as a function of σ in Fig. 9. Compared with Fig. 6, besides the change in values of terminal time and total fuel consumption, the transition region expands in the σ direction due to the nonlinearities.

We summarize the above results in Table 1, where we show the terminal time t_{f} and the total fuel consumption Q for different σ

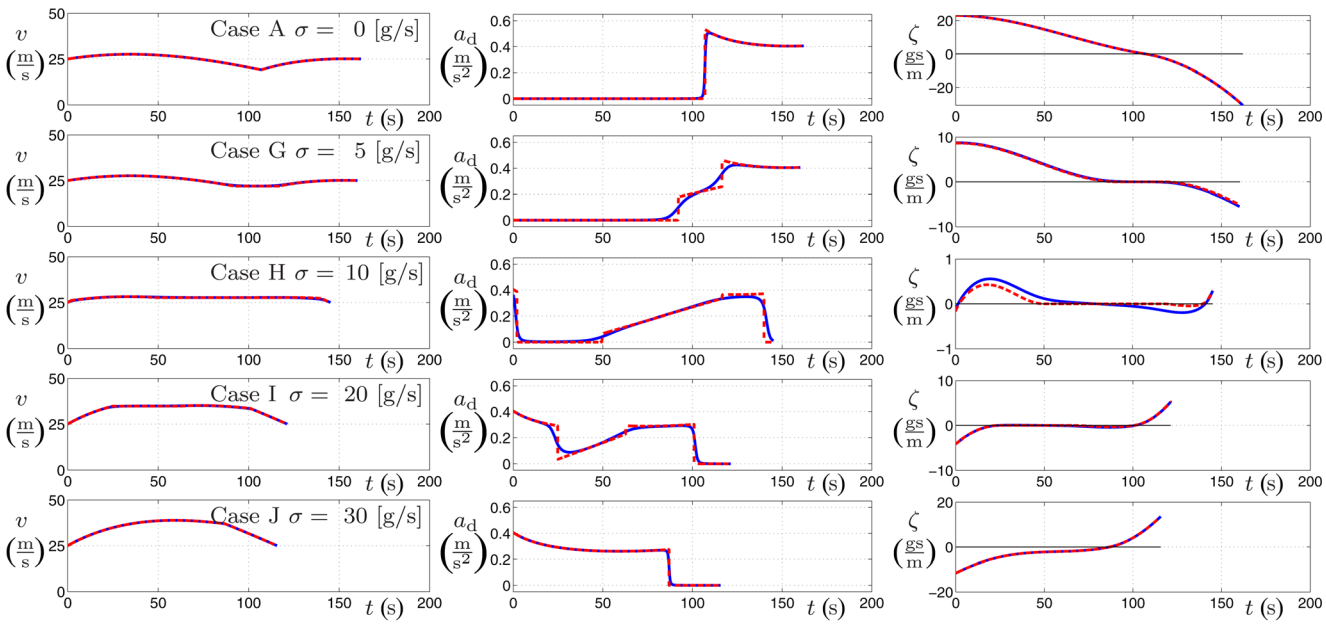

Fig. 8 Time evolution of the speed v (left column), the corresponding control input a_{d} (center column), and the switching variable ζ for different values of the parameter σ (right column) as indicated. The rows correspond to the points A, G, H, I, and J marked in Fig. 9. The solid curves are associated with the trajectories acquired by our BVP solver based on pseudo-arclength collocation method using the smoothed controllers (45) and (46), and the dashed curves are those generated by direct method.

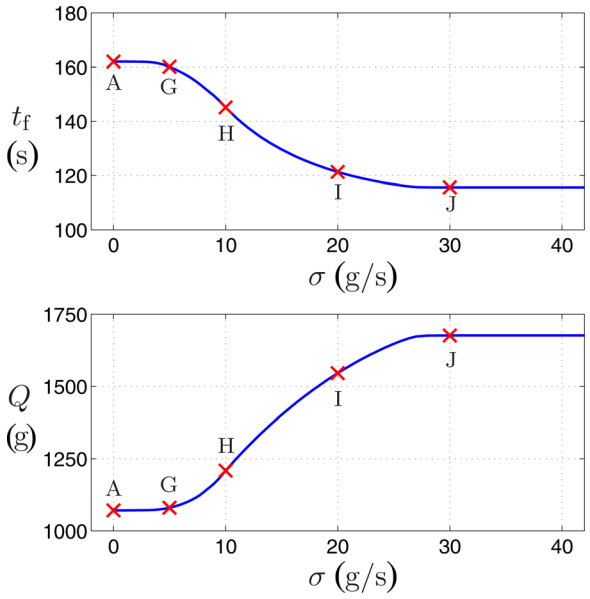

Fig. 9 The terminal time t_f (upper panel) and the fuel consumption Q (44) (lower panel) as a function of σ for $v(t_f) = v(0) = 25\,\mathrm{m/s}$. The corresponding trajectories are shown in Fig. 8 for the points marked A, G, H, I, and J.

Table 1 Terminal time t_f and fuel consumption Q for multiple σ values (cf. Fig. 8), compared with the fuel consumption and time of cruise control, with conditions $v(0) = v(t_f) = 25\,\mathrm{m/s}$

	t_f (s)	Q (g)
$\sigma = 0$ (case A)	162.1	1071.1
$\sigma = 5$ (case G)	160.1	1080.2
$\sigma = 10$ (case H)	145.2	1208.9
$\sigma = 20$ (case I)	121.3	1545.7
$\sigma = 30$ (case J)	115.6	1676.2
Cruise control	160.0	1222.3

values for $v(0) = v(t_f) = 25\,\mathrm{m/s}$. For comparison, we also show the results for constant speed (that can be maintained using standard cruise control). When σ is small, the optimal solution consumes approximately 11.9% less fuel compared to the constant speed scenario. On the other hand, when the weight on terminal time is large, the truck reaches its destination earlier but consumes more fuel. Again, we can identify the region $\sigma \in [3, 30]$, where one has a trade-off between travel time and fuel consumption. We remark that in order to maintain the constant speed, braking and large engine torque may be needed (i.e., the constraints (13) may be violated). Therefore, the constant speed driving profile is not necessarily in the function space for the optimal control problem. Finally, we remark that increasing the speed of headwind will increase the length of the time domain where maximum available control input is applied.

6 Traffic Information

In this section, we investigate the effect of the penalty term on traffic. In particular, we study the effects of the speed limit and the traffic flow speed separately. We first study the penalty on speed limit by setting $\rho_1 = 0.1$ g/s and $\rho_2 = 0$ gs/m^2 and using constant v_{lim} along the route, cf. Eq (15). We start from the solution with $\sigma = 30$ g/m (cf. Fig. 8, case J, where the maximum speed reaches ≈39 m/s) and continue the solution while changing v_{lim} from 40 m/s to 30 m/s. The results are shown in Fig. 10, where the same notation is used as in Fig. 8. As the speed limit decreases, the maximum speed of the optimal speed profile decreases in response. Notice that to achieve the optimal profile, complicated switching structure may be required for the control input. For example, when $v_{\mathrm{lim}} = 30$ m/s, minimum control, singular control, and maximum control are all needed.

In order to investigate the effect of traffic flow penalty, we set $\sigma = 30$ g/s and $\rho_1 = 0$ g/s, and consider v_{traf} to be constant along the route, cf. Eq. (16). We start from the solution with $\sigma = 30$ g/m (cf. Fig. 8, case J) and change ρ_2 from 0 to 1. The results with different v_{traf} values are shown in Fig. 11, where again the same notation is used as in Fig. 8. Four different values of the constant traffic speed v_{traf} are considered along the road, and the results imply that the optimal trajectories settle down to different speed profiles. When v_{traf} is small, the speed profile is similar to the solution without traffic penalty, cf. Fig. 8, case A. This is because, according to the results in Sec. 5, the solution is the one with the lowest average speed, so the controller cannot bring the vehicle to lower speed. Note that lower desired traffic speed v_{traf} can be achieved by adjusting problem parameters (e.g., the boundary condition (3) and the constraints (4)) but such analysis is beyond the scope of this paper. As v_{traf} increases, the speed profile comes closer to v_{traf} and the penalty cost is reduced. However, as a result of the multi-objective optimization considering fuel consumption, travel time, and traffic speed, the optimal profile requires nontrivial control action including multiple switches between maximum, minimum, and singular control.

Finally, in Fig. 12 we show the ratio between traffic cost and the total cost (1) as a function of the weight ρ_2. When it is achievable by control, the speed profile gets closer to the traffic flow speed (as in the second two rows of Fig. 11) and therefore $(v - v_{\mathrm{traf}})^2$ goes to zero as ρ_2 increases. As a result, the ratio between traffic cost and the total cost will converge to a constant value, as shown by the black dotted and magenta dashed–dotted curves in Fig. 12. On the other hand, if the traffic flow speed cannot be followed by the vehicle using any feasible control (as in the first two rows of Fig. 11), then the speed profile remains

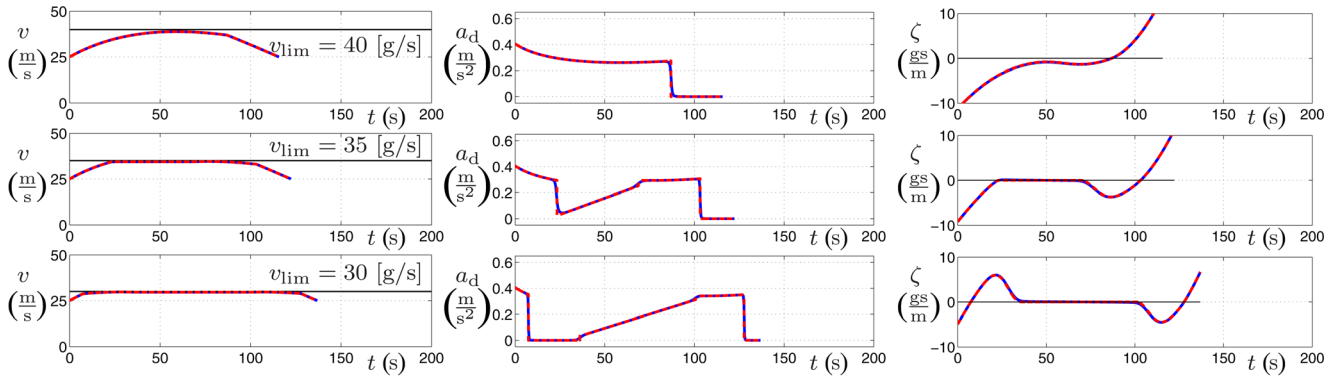

Fig. 10 Effect of the speed limit penalty with different values of v_{lim} as indicated. The same notation is used as that in Fig. 8.

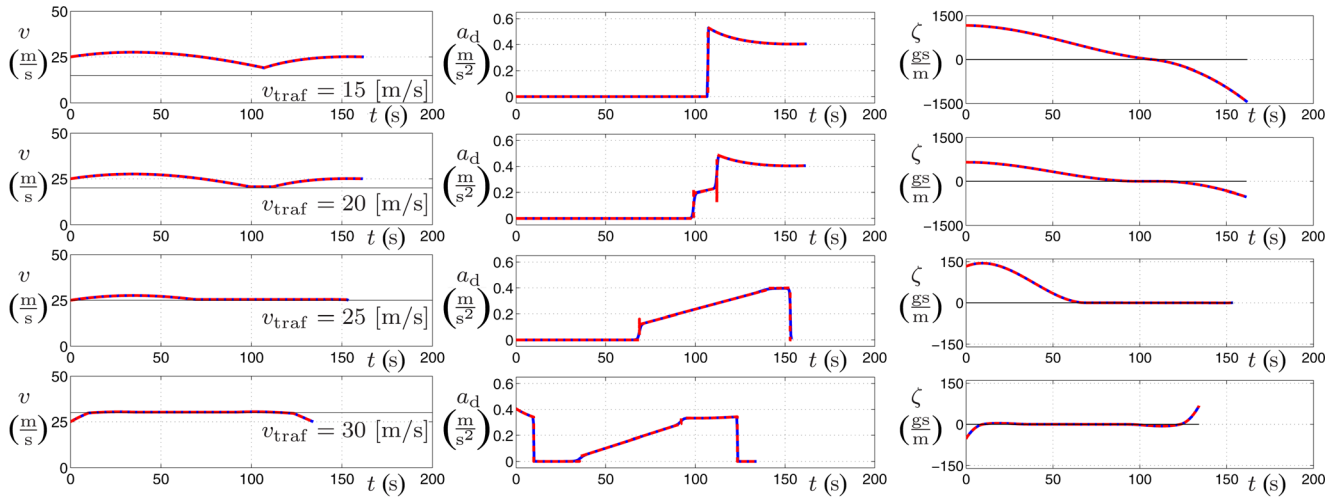

Fig. 11 Traffic flow penalty with different values of v_{traf} as indicated. The same notation is used as that in Fig. 8.

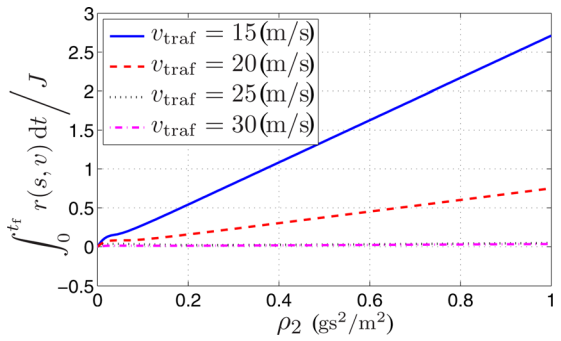

Fig. 12 The ratio between the traffic cost and the total cost (1) as a function of the weight ρ_2

significantly different from the traffic flow as ρ_2 increases. As a result, the ratio between traffic cost and the total cost will keep growing as ρ_2 increases, since $(v - v_{\text{traf}})^2$ converges to a constant value, as demonstrated by the solid and dashed curves in Fig. 12. These results show that increasing the weight on traffic may not force the vehicle to approach the desired speed given by the traffic conditions. Also, we emphasize that this desired speed does not correspond to the speed of a particular vehicle. To take into account such effects (e.g., to address safety), a different description is needed and it is left for future research.

7 Conclusion and Future Work

In this paper, we proposed a framework for fuel economy optimization of HDVs that can incorporate road elevation, headwind, desired terminal time, and traffic information. We established a systematic approach in order to solve the arising multi-objective control problem while varying the system parameters. First, we solved a simplified problem analytically that allowed us to characterize the switching structure of the resulting bang–bang controller or bang-singular-bang controller. Then, we used this knowledge to initialize our numerical continuation software.

We demonstrated that varying the weight on the desired terminal time causes qualitative changes in the switching structure of the controller. We also identified a parameter region where one can balance the fuel economy and the traveling time. Moreover, we investigated the effects of traffic dynamics and identified the conditions (in terms of the traffic speed and the weigh on the traffic cost) that allow the truck to balance between fuel economy and adaptation to traffic conditions. These were achieved by nontrivial control actions that cannot be obtained intuitively.

For fuel-efficient driving in heavy traffic conditions, reacting to the motion of individual vehicles in the neighborhood may be important. Merging such a Lagrangian description with the Eulerian description used in the paper is a challenging task and is left for future research. Future goals also include implementing our method using a rolling horizon setup and incorporating V2V information from multiple vehicles ahead.

Acknowledgment

The authors would like to thank the Navistar company for providing the source data used in this paper.

Appendix A: Table of Parameter of Navistar Truck

Table 2 Data of a 2012 Navistar ProStar truck [32]

Parameter	Value
Mass (m)	29,484 (kg)
Air drag coefficient (k)	3.84 (kg/m)
Tire rolling radius (R)	0.504 (m)
Tire rolling resistance coefficient (γ)	0.006
Maximum acceleration ($a_{\max}$)	2 (m/s^2)
Engine rotational inertia (I)	39.9 (kg m^2)
Gravitational constant (g)	9.81 (m/s^2)
Number of forward gears	10
First gear ratio/efficiency	12.94/0.97
Second gear ratio/efficiency	9.29/0.97
Third gear ratio/efficiency	6.75/0.97
Fourth gear ratio/efficiency	4.9/0.97
Fifth gear ratio/efficiency	3.62/0.97
Sixth gear ratio/efficiency	2.64/0.97
Seventh gear ratio/efficiency	1.90/0.97
Eighth gear ratio/efficiency	1.38/0.98
Ninth gear ratio/efficiency	1/0.99
Tenth gear ratio/efficiency	0.74/0.98
Final drive ratio/efficiency	4.17/0.98

Appendix B: Pseudo-Arclength Continuation Method

Our BVP solver developed is based on pseudo-arclength continuation. Given a BVP problem, the solver first uses a collocation method to transform the ODE to a set of algebraic equations. Consider the BVP in the form

$$\begin{aligned} \dot{x} &= f(x,t;p), \quad t \in [0,t_f] \\ 0 &= g(x(0),x(t_f);p) \end{aligned} \tag{B1}$$

where $x \in \mathbb{R}^n$ is the state, and $p \in \mathbb{R}$ is the system parameter. The boundary condition g contains n equations, assuming that the terminal time t_f is given.

We define N collocation points $x(t_i) = x_i$ at time $t_i = i \cdot (t_f/N-1)$, $i = 0,1,\ldots,N-1$. Using the *trapezoidal rule*, we arrive at the algebraic equations

$$\begin{aligned} x_{i+1} - x_i &= \frac{t_f}{2(N-1)}\left[f(x_{i+1},t_{i+1};p) + f(x_i,t_i;p)\right] \\ 0 &= g(x_0,x_N;p) \end{aligned} \tag{B2}$$

g. Therefore, the collocation method of order N transforms the original BVP to an algebraic equation of dimension Nn that can be written to the form

$$F(\mathbf{x},p) = 0 \tag{B3}$$

where $\mathbf{x} \in \mathbb{R}^{Nn}$ contains all the components $x_i \in \mathbb{R}^n$, $i = 0,\ldots,N-1$. Note that the terminal time t_f may be a variable as well, and in this case, g shall contain $n+1$ equations, and the resulting nonlinear equations will be of dimension $Nn+1$.

The key idea behind continuation is that one may solve the BVP for a particular value of p and then use the corresponding solution as an initial guess when solving the BVP for the nearby parameter $p + \delta p$. This way the solution can be continued while the parameter is varied [45]. Continuation is based on the implicit function theorem, which guarantees that Eq. (B3) can be solved for $\mathbf{x}(p)$. However, there may be points where implicit function theorem is violated. For example, in Fig. 13, the curve (B3) is shown in the $(p, \|\mathbf{x}\|)$-plane. The implicit function theorem is violated at the fold points where the partial derivative of F with respect to $\mathbf{x}$ is 0. In order to be able to continue the curve through these points, we apply the so-called pseudo-arclength continuation method [46]. Throughout the description of the algorithm, we use the subscript k to denote the iteration number. In the kth iteration, the point $(\mathbf{x}^{(k)}, p^{(k)})$ lies on the curve with tangent vector $\mathbf{z}^{(k)}$. Let us consider the arc-length step δs.

Then, we follow the steps:

(1) Compute the predicted value for the $(k+1)$th step

$$\begin{bmatrix} \mathbf{x}_{\text{pred}}^{(k+1)} \\ p_{\text{pred}}^{(k+1)} \end{bmatrix} = \begin{bmatrix} \mathbf{x}^{(k)} \\ p^{(k)} \end{bmatrix} + \mathbf{z}^{(k)} \delta s \tag{B4}$$

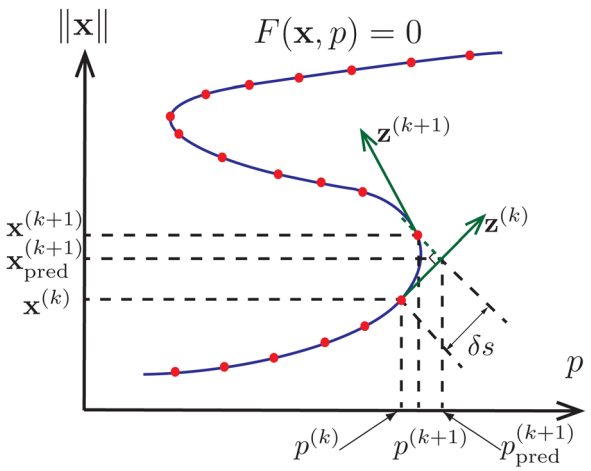

Fig. 13 Concept of the pseudo-arclength method

(2) Solve the following system of augmented nonlinear equations:

$$\begin{aligned} 0 &= F(\mathbf{x}^{(k+1)}, p^{(k+1)}) \\ 0 &= \mathbf{z}^{(k)\mathrm{T}} \cdot \begin{bmatrix} \mathbf{x}^{(k+1)} - \mathbf{x}_{\text{pred}}^{(k+1)} \\ p^{(k+1)} - p_{\text{pred}}^{(k+1)} \end{bmatrix} \end{aligned} \tag{B5}$$

for $(\mathbf{x}^{(k+1)}, p^{(k+1)})$ given $(\mathbf{x}_{\text{pred}}^{(k+1)}, p_{\text{pred}}^{(k+1)})$.

(3) Find the tangent $\mathbf{z}^{(k+1)}$ to the curve for the $(k+1)$th point by solving the linear equations

$$\begin{bmatrix} \dfrac{\partial F}{\partial \mathbf{x}}\left(\mathbf{x}^{(k+1)}, p^{(k+1)}\right), \dfrac{\partial F}{\partial p}\left(\mathbf{x}^{(k+1)}, p^{(k+1)}\right) \\ \mathbf{z}^{(k)\mathrm{T}} \end{bmatrix} \mathbf{z}^{(k+1)} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \tag{B6}$$

with $\|\mathbf{z}^{(k+1)}\| = 1$.

In this paper, we use the classical Newton method [47] to solve the nonlinear algebraic equation (B5). These steps are demonstrated in Fig. 13. It can be seen that the pseudo-arclength continuation method may allow one to continue the curve through the fold points. Finally, we remark that the step size δs can be adapted during the process, e.g., can be chosen larger where Eq. (B3) is "flat."

References

[1] Davis, S. C., Diegel, S. W., and Boundy, R. G., 2013, *Transportation Energy Data Book*, 32 ed., U.S. Department of Energy, Washington, DC.

[2] U.S. Department of Transportation, 2014, "2012 Commodity Flow Survey United States," Technical Report No. EC12TCF-US.

[3] Saltsman, B., 2014, "Impacts of Connectivity and Automation on Vehicle Operations," Global Symposium on Connected Vehicle and Infrastructure, University of Michigan Transportation Research Institute.

[4] Caveney, D., 2010, "Cooperative Vehicular Safety Applications," IEEE Control Syst. Mag., **30**(4), pp. 38–53.

[5] Hooker, J. N., 1988, "Optimal Driving for Single-Vehicle Fuel Economy," Transp. Res. Part A, **22**(3), pp. 183–201.

[6] Monastyrsky, V. V., and Golownykh, I. M., 1993, "Rapid Computation of Optimal Control for Vehicles," Transp. Res. Part B, **27**(3), pp. 219–227.

[7] Lattemann, F., Neiss, K., Terwen, S., and Connolly, T., 2004, "The Predictive Cruise Control: A System to Reduce Fuel Consumption of Heavy Duty Trucks," SAE Paper No. 2004-01-2616.

[8] Hellström, E., 2010, "Look-Ahead Control of Heavy Vehicles," Ph.D. thesis, Linköping University, Linköping, Sweden.

[9] Mensing, F., Bideaux, E., Trigui, R., and Tattegrain, H., 2013, "Trajectory Optimization for Eco-Driving Taking Into Account Traffic Constraints," Transp. Res. Part D, **18**, pp. 55–61.

[10] Jiménez, F., López-Covarrubias, J. L., Cabrera, W., and Aparicio, F., 2013, "Real-Time Speed Profile Calculation for Fuel Saving Considering Unforeseen Situations and Travel Time," IET Intell. Transp. Syst., **7**(1), pp. 10–19.

[11] Jiménez, F., and Cabrera-Montiel, W., 2014, "System for Road Vehicle Energy Optimization Using Real Time Road and Traffic Information," Energies, **7**(6), pp. 3576–3598.

[12] Kohut, N., Hedrick, K., and Borrelli, F., 2009, "Integrating Traffic Data and Model Predictive Control to Improve Fuel Economy," 12th IFAC Symposium on Control in Transportation Systems, pp. 155–160.

[13] Xu, S., Li, S. E., Deng, K., Li, S., and Cheng, B., 2014, "A Unified Pseudospectral Computational Framework for Optimal Control of Road Vehicles," IEEE/ASME Trans. Mechatronics, **20**(4), pp. 1–12.

[14] Schwarzkopf, A. B., and Leipnik, R. B., 1977, "Control of Highway Vehicles for Minimum Fuel Consumption Over Varying Terrain," Transp. Res., **11**(4), pp. 279–286.

[15] Chang, D. J., and Morlok, E. K., 2005, "Vehicle Speed Profiles to Minimize Work and Fuel Consumption," J. Transp. Eng., **131**(3), pp. 173–182.

[16] Fröberg, A., Hellström, E., and Nielsen, L., 2006, "Explicit Fuel Optimal Speed Profiles for Heavy Trucks on a Set of Topographic Road Profiles," SAE Technical Paper No. 2006-01-1071.

[17] Saerens, B., Rakha, H., Diehl, M., and Van den Bulck, E., 2013, "A Methodology for Assessing Eco-Cruise Control for Passenger Vehicles," Transp. Res. Part D, **19**, pp. 20–27.

[18] Kamal, M. A. S., Mukai, M., Murata, J., and Kawabe, T., 2011, "Ecological Vehicle Control on Roads With Up-Down Slopes," IEEE Trans. Intell. Transp. Syst., **12**(3), pp. 783–794.

[19] Kamal, M. S., Mukai, M., Murata, J., and Kawabe, T., 2013, "Model Predictive Control of Vehicles on Urban Roads for Improved Fuel Economy," IEEE Trans. Control Syst. Technol., **21**(3), pp. 831–841.

[20] Wang, M., Daamen, W., Hoogendoorn, S. P., and van Arem, B., 2014, "Rolling Horizon Control Framework for Driver Assistance Systems—Part I:

Mathematical Formulation and Non-Cooperative Systems," Transp. Res. Part C, **40**, pp. 271–289.
[21] Wang, M., Daamen, W., Hoogendoorn, S. P., and van Arem, B., 2014, "Rolling Horizon Control Framework for Driver Assistance Systems—Part II: Cooperative Sensing and Cooperative Control," Transp. Res. Part C, **40**, pp. 290–311.
[22] Hartl, R. F., Sethi, S. P., and Vickson, R. G., 1995, "A Survey of the Maximum Principles for Optimal Control Problems With State Constraints," SIAM Rev., **37**(2), pp. 181–218.
[23] Vinter, R., 2000, *Optimal Control* (Systems & Control: Foundations & Applications), Springer Science & Business Media, Heidelberg, Germany.
[24] Clarke, F., and De Pinho, M., 2010, "Optimal Control Problems With Mixed Constraints," SIAM J. Control Optim., **48**(7), pp. 4500–4524.
[25] Osmolovskii, N., and Maurer, H., 2012, *Applications to Regular and Bang-Bang Control: Second-Order Necessary and Sufficient Optimality Conditions in Calculus of Variations and Optimal Control*, Vol. DC 24, SIAM Publications, Philadelphia, PA.
[26] Maurer, H., Büskens, C., Kim, J.-H., and Kaya, Y., 2005, "Optimization Methods for the Verification of Second Order Sufficient Conditions for Bang-Bang Controls," Optim. Control Appl. Methods, **26**(3), pp. 129–156.
[27] Maurer, H., and Pickenhain, S., 1995, "Second-Order Sufficient Conditions for Control Problems With Mixed Control-State Constraints," J. Optim. Theory Appl., **86**(3), pp. 649–667.
[28] Graichen, K., and Petit, N., 2008, "Constructive Methods for Initialization and Handling Mixed State-Input Constraints in Optimal Control," J. Guid. Control Dyn. **31**(5), pp. 1334–1343.
[29] Wyczalkowski, M., and Szeri, A. J., 2003, "Optimization of Acoustic Scattering From Dual-Frequency Driven Microbubbles at the Difference Frequency," J. Acoust. Soc. Am., **113**(6), pp. 3073–3079.
[30] Alam, A., 2014, "Fuel-Efficient Heavy Duty Vehicle Platooning," Ph.D. thesis, Kungliga Tekniska Högskolan, Stockholm, Sweden.
[31] Orosz, G., and Shah, S. P., 2012, "A Nonlinear Modeling Framework for Autonomous Cruise Control," ASME Paper No. DSCC2012-MOVIC2012-8871.
[32] Navistar, 2011, "Maxxforce 11 and 13 Liter Engines," Navistar, Lisle, IL.
[33] Heywood, J. B., 2002, *Internal Combustion Engine Fundamentals*, McGraw-Hill, New York.
[34] Ozatay, E., Özgüner, Ü., Onori, S., and Rizzoni, G., 2012, "Analytical Solution to the Minimum Fuel Consumption Optimization Problem With the Existence of a Traffic Light," 5th Annual Dynamic Systems and Control Conference and 11th Motion and Vibration Conference, pp. 837–846.
[35] Guzzella, L., and Onder, C. H., 2004, *Introduction to Modelling and Control of Internal Combustion Engine Systems*, Springer, Heidelberg, Germany.
[36] Orosz, G., Wilson, R. E., and Stépán, G., 2010, "Traffic Jams: Dynamics and Control," Philos. Trans. R. Soc., A, **368**(1928), pp. 4455–4479.
[37] Hestenes, M. R., 1966, *Calculus of Variations and Optimal Control Theory*, Wiley, Hoboken, NJ.
[38] Krener, A. J., 1977, "The High Order Maximal Principle and Its Application to Singular Extremals," SIAM J. Control Optim., **15**(2), pp. 256–293.
[39] Stoer, J., Bulirsch, R., Bartels, R., Gautschi, W., and Witzgall, C., 1993, *Introduction to Numerical Analysis*, Vol. 2, Springer, Heidelberg, Germany.
[40] Wächter, A., and Biegler, L. T., 2006, "On the Implementation of an Interior-Point Filter Line-Search Algorithm for Large-Scale Nonlinear Programming," Math. Program., **106**(1), pp. 25–57.
[41] Gay, D. M., and Kernighan, B., 2002, *AMPL: A Modeling Language for Mathematical Programming*, second edition, Duxbury Press/Brooks/Cole, Boston, MA.
[42] Allgower, E. E. L., and Georg, K., 2003, *Introduction to Numerical Continuation Methods* (Classics in Applied Mathematics), Vol. 45, Society for Industrial and Applied Mathematics, Philadelphia, PA.
[43] Rao, A. V., 2009, "Survey of Numerical Methods for Optimal Control," Adv. Astronaut. Sci., **135**(1), pp. 497–528.
[44] He, C. R., and Orosz, G., 2014, "Fuel Consumption Optimization of Heavy-Duty Vehicles: An Analytical Approach," ASME Paper No. DSCC2014-6362.
[45] Roose, D., and Szalai, R., 2007, "Continuation and Bifurcation Analysis of Delay Differential Equations," *Numerical Continuation Methods for Dynamical Systems*, B. Krauskopf, H. Osinga, and J. Galán-Vioque, eds., Springer, Heidelberg, Germany, pp. 359–399.
[46] Mittelmann, H. D., 1986, "A Pseudo-Arclength Continuation Method for Nonlinear Eigenvalue Problems," SIAM J. Numer. Anal., **23**(5), pp. 1007–1016.
[47] Jorge, N., and Wright, S. J., 1999, *Numerical Optimization* (Operations Research and Financial Engineering), Vol. 2, Springer, Heidelberg, Germany.