Optimal train control via switched system dynamic optimization

Weifeng Zhong[a], Qun Lin[b], Ryan Loxton[b] and Kok Lay Teo[b]

[a]School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing, China; [b]School of Electrical Engineering, Computing, and Mathematical Sciences, Curtin University, Perth, Australia

ARTICLE HISTORY
Compiled April 16, 2019

ABSTRACT
This paper considers an optimal train control problem with two challenging, non-standard constraints: a speed constraint that is piecewise-constant with respect to the train's position, and control constraints that are non-smooth functions of the train's speed. We formulate this problem as an optimal switching control problem in which the mode switching times are decision variables to be optimized, and the track gradient and speed limit in each mode are constant. Then, using control parameterization and time-scaling techniques, we approximate the switching control problem by a finite-dimensional optimization problem, which is still subject to the challenging speed limit constraint (imposed continuously during each mode) and the non-smooth control constraints. We show that the speed constraint can be transformed into a finite number of point constraints. We also show that the non-smooth control constraints can be approximated by a sequence of conventional (smooth) inequality constraints. The resulting approximate problem can be viewed as a nonlinear programming problem and solved using gradient-based optimization algorithms, where the gradients of the cost and constraint functions are computed via the sensitivity method. A case study using data for a real subway line shows that the proposed method yields a realistic optimal control profile without the undesirable control fluctuations that can occur with the pseudospectral method.

KEYWORDS
Optimal train control; switched system; control parameterization; time-scaling transformation; state-dependent control constraint

1. Introduction

Rail transit systems consume vast amounts of energy, of which 70-90% is due to train traction [4]. Accordingly, there is now a large body of work on optimal train driving strategies for the purpose of reducing tractive energy consumption. The aim is to find the optimal control law—that is, the optimal sequence of tractive and braking forces applied to the train—such that the tractive energy consumption is minimized while the train moves from one station to the next within a given time frame. In general, the railway line consists of various segments with different gradients and thus the line resistance varies along the track. Moreover, when the train is moving, the running resistance—normally a quadratic function of speed—also affects the train, resulting in motion that is governed by complex nonlinear dynamics. There are also typically two

CONTACT Weifeng Zhong. Email: wfzhong@bjtu.edu.cn

core constraints that must be satisfied: the train's speed is prohibited from exceeding the speed limit in each track segment because of operational safety, and the tractive and braking forces must be restricted within maximum physical limits, which may depend nonlinearly on the train's speed. Therefore, obtaining the optimal train driving strategy requires solving a nonlinear optimal control problem with complex state and control constraints.

There are two types of dynamic models for describing the train's motion in optimal train control: time-based models [2, 5, 6, 18], where time is the independent variable of the system, and position-based models [1, 7, 10, 16, 17, 20], where position is the independent variable. Since the line resistance and speed limit constraints are both functions of position, the latter models are more common in the literature. They lead to an optimal control problem in standard form that can be readily solved, either analytically with the Pontryagin maximum principle [1, 7, 10] or numerically by implementing various approximation schemes [16, 17, 20]. However, position-based models usually contain the reciprocal of the train's speed or kinetic energy, which means the differential equations are undefined when the speed or kinetic energy equals zero. This can lead to numerical difficulties when solving the train differential equations near the initial point, where the speed and kinetic energy are indeed zero because the train starts from rest [7, 16, 17, 20].

Time-based models do not include a reciprocal term and hence they are unaffected by these numerical difficulties. However, they lead to a more complex optimal control problem with three non-standard features, as we now describe. First, the line resistance, which governs the train dynamic equations, is piecewise constant (and thus non-smooth) with respect to the train's position, which is a state variable in time-based models. Second, the speed constraint is also piecewise constant with respect to the train's position because the speed limit can change along the track. Third, the upper bounds for the tractive and braking controls both depend on the train's speed, with different profiles for low and high speeds and a non-smooth transition point [17, 18].

In several previous studies using time-based train control models, the three non-standard features described above were circumvented by simplifying the line conditions, running resistance or state constraints [2, 5, 6]. In [18, 19], the full model without simplifications—including varying line gradients, piecewise-constant speed limits, and state-dependent control constraints—was tackled using the Gauss pseudospectral method (GPM). However, the control profiles obtained by the GPM and reported in [18, 19] fluctuate rapidly in some sections along the track, making them unrealistic to implement in practice. Similar fluctuations appear in [16], where the GPM was applied to solve the optimal train control problem for a position-based model. As explained in [18, 19], these fluctuations are likely caused by the presence of singular arcs in the respective optimal control problems [11, 12].

Given the undesirable control fluctuations experienced with the pseudospectral method, in this paper we propose an alternative approach based on the control parameterization and time-scaling methods [8, 14, 15]. We first formulate the optimal train control problem (with time-based model) as an optimal switching control problem, where the line gradient and speed limit are constant within each subsystem but can change from subsystem to subsystem. Then, by applying control parameterization and the time-scaling transformation to each subsystem, the tractive and braking control variables are approximated by piecewise constant functions whose heights and switching time points are regarded as decision variables. In this way, the optimal train control problem is approximated by a constrained finite-dimensional optimization problem, albeit one with two complex sets of constraints: the speed limit constraints and the non-smooth bound constraints governing the tractive and braking controls. To handle the speed constraints, we use the analytical solution of the train differential equations to equivalently convert these constraints (essentially an infinite number of point constraints) into a finite number of point constraints. To handle the non-smooth control constraints, we first smooth the sharp corner between the upper bound profiles for low and high speeds to yield a set of approximate constraints, which are then transformed into a sequence of conventional inequality constraints. The resulting optimization problem can be viewed as a nonlinear programming (NLP) problem and solved by gradient-based optimization algorithms, such as the sequential quadratic programming (SQP) algorithm. Gradient formulae for the cost and constraint functions are derived using the sensitivity method. The method has been tested using data for the Yizhuang subway line in Beijing and the results show that the proposed approach can efficiently solve the complex optimal train control problem, with all constraints satisfied and without any control fluctuations.

The rest of this paper is organized as follows. Section 2 introduces the time-based train dynamics and the corresponding optimal control problem. Then, Section 3 introduces an equivalent switched system formulation for the train control problem. Based on the switched system formulation, Section 4 presents the key computational procedures for generating the optimal control profiles for the tractive and braking forces. Section 5 demonstrates the performance of our method using data for the Yizhuang line and Section 6 concludes the paper.

2. The optimal train control problem

The motion of a point-mass train with time as the independent variable can be described as follows [10, 19]:

$$\dot{x}_1(t) = x_2(t), \tag{1}$$

$$\dot{x}_2(t) = \frac{1}{m\rho}\big[u_1(t) + u_2(t) - r_b(x_2(t)) - r_l(x_1(t))\big], \tag{2}$$

and

$$x_1(0) = 0,\ x_2(0) = 0, \tag{3}$$

where $x_1(t)$ is the train's position along the track at time t, $x_2(t)$ is the train's speed at time t, m is the train's mass, ρ is a factor that depends on the train's rotary mass, $u_1(t)$ is the train's tractive force, $u_2(t)$ is the train's braking force, $r_b(x_2(t))$ is the basic resistance caused by mechanical friction and air, and $r_l(x_1(t))$ is the line resistance caused by gravity.

In general, the control forces $u_1(t)$ and $u_2(t)$ are continuous subject to the following constraints:

$$0 \le u_1(t) \le u_1^{\max}(x_2(t)), \tag{4}$$

$$-u_2^{\max}(x_2(t)) \le u_2(t) \le 0, \tag{5}$$

where both $u_1^{\max}(x_2(t))$ and $u_2^{\max}(x_2(t))$ are non-smooth functions of the speed. The precise formulas for $u_1^{\max}(x_2)$ and $u_2^{\max}(x_2)$ depend on the specific train under consideration, but they typically have the shape shown in Fig. 1, where there are two distinct

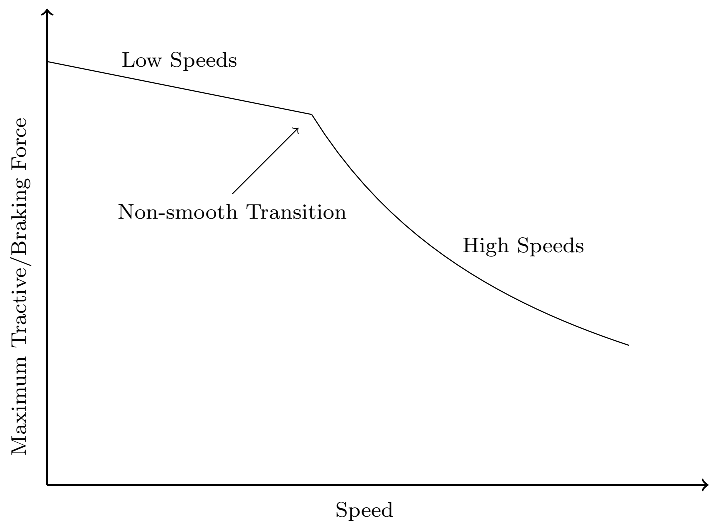

Figure 1. Shape of the upper bounds for the tractive and braking forces.

profiles for low and high speeds with a non-smooth transition point. The behavior at low speeds is less variable and normally linear. The basic resistance $r_b(x_2)$ is described by the Davis formula [3] as given below:

$$r_b(x_2) = a + bx_2 + cx_2^2,$$

where $a \geq 0$, $b \geq 0$ and $c > 0$ are coefficients determined by the train's characteristics. Furthermore, the line resistance can be expressed as:

$$r_l(x_1) = mg \sin \beta(x_1) \approx mg \tan \beta(x_1), \tag{6}$$

where g is the acceleration due to gravity, $\beta(x_1)$ is the slope angle at x_1 (measured anti-clockwise from the horizontal), and $\tan \beta(x_1)$ is the line gradient at x_1. Note that $\beta(x_1)$ is positive along uphill sections of the track and negative along downhill sections. See Fig. 2 for a diagram showing the forces acting on the train. The approximation in (6), which holds on non-steep tracks where $\beta(x_1)$ is not too far from zero, is often used when the railway line data is expressed in terms of track gradients.

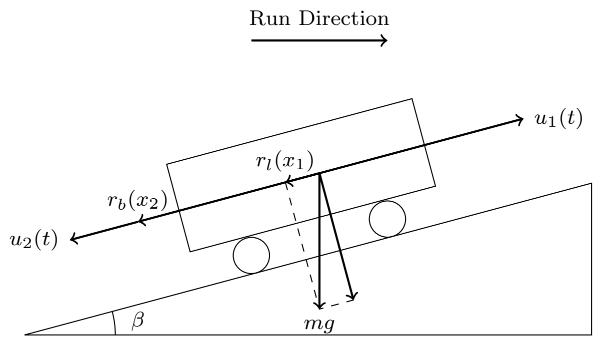

Figure 2. Forces acting on the train.

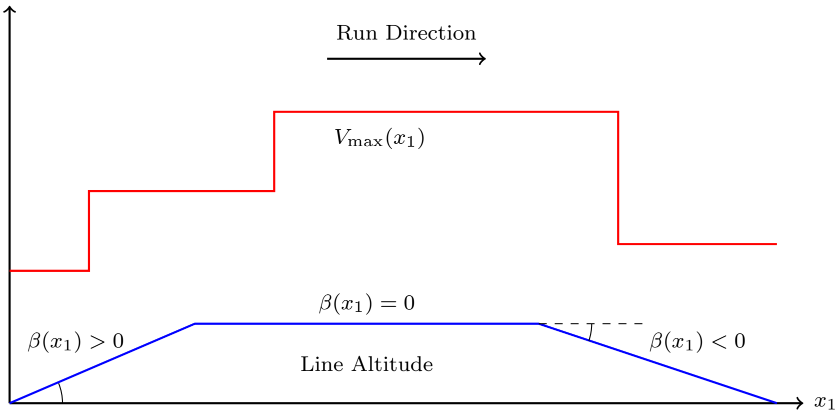

Figure 3. An example of a track with three line gradients and four speed limits.

The train's speed is subject to the following bound constraint:

$$0 \leq x_2(t) \leq V_{\max}(x_1(t)), \tag{7}$$

where $V_{\max}(x_1(t))$ is a piecewise-constant function with respect to $x_1(t)$, defining separate speed limits along different sections of the track. A simple scenario involving three line gradients and four speed limits is shown in Fig. 3.

Let T be the trip time determined by the timetable. The position and speed at the terminal point of the route must satisfy

$$x_1(T) = L, \;\; x_2(T) = 0, \tag{8}$$

where L is the length of the route. The tractive energy consumed by the train during the trip is given by

$$J = \int_0^T u_1(t)x_2(t)dt. \tag{9}$$

Then, the optimal train control problem can be formally stated as follows.

Problem P. Given the train dynamics (1)-(2) with initial conditions (3) and terminal conditions (8), find a control law $\boldsymbol{u} = [u_1, u_2]^\top$, such that the objective function (9) is minimized subject to the control constraints (4)-(5) and the state constraint (7).

3. Switched system model

Since the track consists of a finite set of straight-line gradients, the line resistance $r_l(x_1)$ is piecewise-constant with respect to x_1. The upper speed limit $V_{\max}(x_1)$ is also piecewise-constant because different segments of track may have different speed limits. Furthermore, the control boundary functions $u_1^{\max}(x_2)$ and $u_2^{\max}(x_2)$ are speed-dependent and non-smooth. These characteristics make the optimal train control problem difficult to solve using traditional optimal control methods. In this section, we will re-formulate Problem P as a switched system optimization problem.

To begin, let the track $[0, L]$ be divided into N subsections in such a way that the line gradient and speed limit in each subsection are constant. The dividing positions along the line satisfy

$$0 = x_1^0 < x_1^1 < \cdots < x_1^{N-1} < x_1^N = L,$$

where x_1^i, $i = 1, \ldots, N-1$, are fixed switching points, and x_1^0 and x_1^N represent the start and end points of the line, respectively. Let t_i denote the corresponding switching time at x_1^i, $i = 1, \ldots, N-1$. Then t_i, $i = 1, \ldots, N-1$, satisfy

$$x_1(t_i) = x_1^i, \ i = 1, \ldots, N, \tag{10}$$

and

$$0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = T.$$

Equation (10) can be viewed as a set of interior point constraints for the new decision variables t_i, $i = 1, \ldots, N-1$.

On each interval $[t_{i-1}, t_i]$, the line resistance $r_l(x_1(t))$ is constant because the track gradient is constant. Thus, the original system (1)-(2) can be viewed as a switched system:

$$\dot{x}_1(t) = x_2(t), \ t \in [t_{i-1}, t_i), \tag{11}$$

$$\dot{x}_2(t) = \frac{1}{m\rho}\left[u_1(t) + u_2(t) - r_b(x_2(t)) - r_l^i\right], \ t \in [t_{i-1}, t_i), \tag{12}$$

where $r_l^i = r_l(x_1(t))$, $t \in [t_{i-1}, t_i)$, is the constant line resistance in the ith subsystem. The state constraint (7) becomes

$$0 \leq x_2(t) \leq V_{\max}^i, \ t \in [t_{i-1}, t_i), \tag{13}$$

where $V_{\max}^i = V_{\max}(x_1(t))$, $t \in [t_{i-1}, t_i)$, is the constant speed limit in the ith subsystem. The cost function (9) can then be expressed as

$$J_N = \sum_{i=1}^{N} \int_{t_{i-1}}^{t_i} u_1(t) x_2(t) dt. \tag{14}$$

Thus, Problem P can be restated as the following switching control problem, in which the line resistance and speed limits are constant in each subsystem.

Problem $\boldsymbol{P_N}$. Given the switched system (11)-(12) with the initial conditions (3) and terminal conditions (8), find a control law $\boldsymbol{u} = [u_1, u_2]^\top$ and switching times t_i, $i = 1, \ldots, N-1$, such that the objective function (14) is minimized subject to the control constraints (4)-(5), interior point constraints (10), and state constraints (13).

4. Solution procedure

4.1. Control parameterization

To solve Problem P_N, we partition each subsystem $[x_1^{i-1}, x_1^i]$ into a set of smaller subsections $[x_1^{i,j-1}, x_1^{i,j}]$, $j = 1, \dots, N_i$, of equal length satisfying

$$x_1^{i-1} = x_1^{i,0} < x_1^{i,1} < \cdots < x_1^{i,N_i-1} < x_1^{i,N_i} = x_1^i,$$

where N_i is the number of subsections and $x_1^{i,j}$ represents the jth dividing point in $[x_1^{i-1}, x_1^i]$. The value of N_i can be chosen as

$$N_i = \left\lceil (x_1^i - x_1^{i-1})/L_b \right\rceil,$$

where $\lceil \cdot \rceil$ is the ceiling function and L_b is a given base length. In this way, the length of each subsection never exceeds L_b, since $(x_1^i - x_1^{i-1})/N_i \le L_b$.

Denote the switching time point at $x_1^{i,j}$ as t_i^j. Then

$$t_{i-1} = t_i^0 < t_i^1 < \cdots < t_i^{N_i-1} < t_i^{N_i} = t_i.$$

Thus, the time interval $[t_{i-1}, t_i]$ is also partitioned into N_i subintervals $[t_i^{j-1}, t_i^j]$, $j = 1, \dots, N_i$.

Let $Q = \sum_{i=1}^N N_i$ and $\tau_{N_1+\cdots+N_{i-1}+j} = t_i^j$, $j = 0, \dots, N_i$, $i = 1, \dots, N$. Clearly, $\tau_0 = 0$ and $\tau_Q = T$. The control function $\boldsymbol{u}$ can be approximated by a piecewise constant function in the form given below:

$$\boldsymbol{u}(t) \approx \boldsymbol{u}_Q(t) = \sum_{k=1}^{Q} \boldsymbol{\delta}^k \chi_{[\tau_{k-1}, \tau_k)}(t), \tag{15}$$

where $\boldsymbol{\delta}^k = [\delta_1^k, \delta_2^k]^\top$ is the constant control value on the kth subinterval $[\tau_{k-1}, \tau_k)$ and $\chi_{[\tau_{k-1}, \tau_k)}(t)$ is the characteristic function defined by

$$\chi_{[\tau_{k-1}, \tau_k)}(t) = \begin{cases} 1, & \text{if } t \in [\tau_{k-1}, \tau_k), \\ 0, & \text{if } t \notin [\tau_{k-1}, \tau_k). \end{cases}$$

This is the control parameterization approach for approximating an optimal control problem by a finite-dimensional optimization problem [8, 14]. Obviously, smaller values of L_b lead to larger values of Q, and the approximation accuracy of $\boldsymbol{u}_Q(t)$ improves as $Q \to \infty$.

With $\boldsymbol{u}(t)$ taking the form of (15), the switched system (11)-(12) becomes

$$\dot{x}_1(t) = x_2(t), \; t \in [\tau_{k-1}, \tau_k), \tag{16}$$

$$\dot{x}_2(t) = \frac{1}{m\rho}\left[\delta_1^k + \delta_2^k - r_b(x_2(t)) - \tilde{r}_l^k\right], \; t \in [\tau_{k-1}, \tau_k), \tag{17}$$

where $\tilde{r}_l^k = r_l(x_1(t))$, $t \in [\tau_{k-1}, \tau_k)$, is the constant line resistance during the kth subinterval.

The control bound constraints (4)-(5) become

$$0 \leq \delta_1^k \leq u_1^{\max}(x_2(t)),\ t \in [\tau_{k-1}, \tau_k),\ k = 1, \ldots, Q, \tag{18}$$

$$-u_2^{\max}(x_2(t)) \leq \delta_2^k \leq 0,\ t \in [\tau_{k-1}, \tau_k),\ k = 1, \ldots, Q, \tag{19}$$

and the state constraint (13) becomes

$$0 \leq x_2(t) \leq \tilde{V}_{\max}^k,\ t \in [\tau_{k-1}, \tau_k), \tag{20}$$

where $\tilde{V}_{\max}^k = V_{\max}(x_1(t))$, $t \in [\tau_{k-1}, \tau_k)$, is the constant speed limit in the kth subinterval.

The cost function (14) now takes the form given below:

$$J_{N,Q} = \sum s_{k=1}^{Q} \int_{\tau_{k-1}}^{\tau_k} \delta_1^k x_2(t) dt = \sum_{k=1}^{Q} \delta_1^k \big[x_1(\tau_k) - x_1(\tau_{k-1})\big]. \tag{21}$$

Let $\tilde{x}_1^{N_1+\cdots+N_{i-1}+j} = x_1^{i,j}$ for each $i = 1, \ldots, N$, $j = 0, \ldots, N_i$. Then $\tilde{x}_1^k$ is the position at switching time point τ_k and thus we have the interior-point constraints

$$x_1(\tau_k) = \tilde{x}_1^k,\ k = 1, \ldots, Q, \tag{22}$$

where each τ_k is a decision variable to be determined.

Clearly, $\tilde{x}_1^k > \tilde{x}_1^{k-1}$ and thus $\tau_k - \tau_{k-1} > 0$, since the train cannot instantaneously move between two distinct points. In fact, given that the maximum speed during the kth subsection is $\tilde{V}_{\max}^k$,

$$\tau_k - \tau_{k-1} \geq \frac{\tilde{x}_1^k - \tilde{x}_1^{k-1}}{\tilde{V}_{\max}^k} > 0,\ k = 1, \ldots, Q. \tag{23}$$

Problem P_N can then be rewritten as follows.

Problem $\mathbf{P}_{N,Q}$. Given the dynamics (16)-(17) with the initial conditions (3) and terminal conditions (8), choose $\boldsymbol{\delta} = [(\boldsymbol{\delta}^1)^\top, \ldots, (\boldsymbol{\delta}^Q)^\top]^\top$ and $\boldsymbol{\tau} = [\tau_1, \ldots, \tau_Q]^\top$, such that the cost function (21) is minimized subject to the control bound constraints (18)-(19), state constraints (20), interior point constraints (22) and switching time constraints (23).

In Problem $\mathrm{P}_{N,Q}$, the control values $\boldsymbol{\delta}$ and switching times $\boldsymbol{\tau}$ are regarded as decision parameters to be determined optimally. To solve this problem using gradient-based optimization techniques, the derivatives of the cost and constraint functions with respect to $\boldsymbol{\delta}$ and $\boldsymbol{\tau}$ are needed. However, as discussed in [8], it is difficult to obtain and implement the derivatives with respect to the switching times τ_k, $k = 1, \ldots, Q$. Thus, in the next subsection, we will employ the time-scaling transformation [15] to map the variable switching time points in $[0, T]$ into fixed time points in the new time horizon $[0, Q]$.

4.2. Time-scaling transformation

Define $\theta_k = \tau_k - \tau_{k-1}$, $k = 1, \dots, Q$, and consider the following time-scaling transformation [15]:

$$t(s) = \begin{cases} \sum\limits_{l=1}^{\lfloor s \rfloor} \theta_l + \theta_{\lfloor s \rfloor + 1}(s - \lfloor s \rfloor), & \text{if } s \in [0, Q), \\ T, & \text{if } s = Q, \end{cases} \tag{24}$$

where $s \in [0, Q]$ is a new time variable and $\lfloor \cdot \rfloor$ is the floor function. Evaluating (24) at $s = k$ gives

$$t(k) = \sum_{l=1}^{k} \theta_l = \tau_k, \ k = 1, \dots, Q,$$

and thus the time-scaling transformation maps $s = k$ to the kth switching time $t = \tau_k$. Let $\boldsymbol{\theta} = [\theta_1, \dots, \theta_Q]^\top$ be a new decision vector replacing $\boldsymbol{\tau} = [\tau_1, \dots, \tau_Q]^\top$. Then the following constraints are required:

$$\theta_k = \tau_k - \tau_{k-1} \geq \frac{\tilde{x}_1^k - \tilde{x}_1^{k-1}}{\tilde{V}_{\max}^k} > 0, \ k = 1, \dots, Q, \tag{25}$$

$$\sum_{k=1}^{Q} \theta_k = T. \tag{26}$$

Under transformation (24), the state variables $x_1(t)$ and $x_2(t)$ become

$$y_1(s) = x_1(t(s)), \ y_2(s) = x_2(t(s)), \ s \in [0, Q].$$

Thus, system (16)-(17) can be recast as

$$\dot{y}_1(s) = \theta_k y_2(s), \ s \in [k-1, k), \tag{27}$$

$$\dot{y}_2(s) = \frac{\theta_k}{m\rho}\left[\delta_1^k + \delta_2^k - r_b(y_2(s)) - \tilde{r}_l^k\right], \ s \in [k-1, k), \tag{28}$$

for $k = 1, \dots, Q$, subject to the initial conditions

$$y_1(0) = 0, \ y_2(0) = 0, \tag{29}$$

and terminal conditions

$$y_1(Q) = L, \ y_2(Q) = 0. \tag{30}$$

The state constraints (20) become

$$0 \leq y_2(s) \leq \tilde{V}_{\max}^k, \ s \in [k-1, k), \ k = 1, \dots, Q, \tag{31}$$

and the control bound constraints (18)-(19) become

$$0 \leq \delta_1^k \leq u_1^{\max}(y_2(s)),\ s \in [k-1,k),\ k = 1,\dots,Q, \tag{32}$$

$$-u_2^{\max}(y_2(s)) \leq \delta_2^k \leq 0,\ s \in [k-1,k),\ k = 1,\dots,Q. \tag{33}$$

Furthermore, the interior point constraints (22) become

$$y_1(k) = \tilde{x}_1^k,\ k = 1,\dots,Q. \tag{34}$$

Finally, the cost function (21) is transformed into

$$G_{N,Q} = \sum_{k=1}^{Q} \int_{k-1}^{k} \theta_k \delta_1^k y_2(s) ds = \sum_{k=1}^{Q} \delta_1^k \big[y_1(k) - y_1(k-1)\big]. \tag{35}$$

Now, Problem $\mathrm{P}_{N,Q}$ can be reformulated equivalently as follows.

Problem $\mathbf{S}_{\boldsymbol{N,Q}}$. Given the system (27)-(28) with initial conditions (29), find $\boldsymbol{\delta}$ and $\boldsymbol{\theta}$ such that (35) is minimized subject to the constraints (25)-(26) and (30)-(34).

Problem $\mathrm{S}_{N,Q}$ is a finite-dimensional optimization problem with $\boldsymbol{\delta}$ and $\boldsymbol{\theta}$ as decision variables. Although simpler than Problem $\mathrm{P}_{N,Q}$ (which has variable switching times), Problem $\mathrm{S}_{N,Q}$ still has two features that prevent it from being solved directly using standard gradient-based optimization techniques:

(a) The constraints (31) restrict the state variable $y_2(s)$ at an infinite number of time points in the horizon $[0, Q]$; and

(b) The control bound constraints (32) and (33) are state-dependent and non-smooth.

Regarding issue (a), state constraints like (31) are typically handled using constraint transcription [8, 14] or exact penalty methods [9]. However, these methods introduce approximations and require manually adjusting at least one approximation parameter to ensure convergence. In the next section, we show that such approximation techniques are unnecessary because $y_2(s)$ is monotonic on each subinterval and thus the infinite number of point constraints in (31) can be expressed equivalently as a finite set of point constraints.

Regarding issue (b), constraints (32) and (33) are more complex than (31) because they include both the control and the state and are defined by non-smooth functions (recall Fig. 1). Nevertheless, in Section 4.4, we show how to handle these non-smooth constraints by introducing a smooth approximation scheme.

4.3. State constraints

Since the basic resistance has the form $r_b(y_2(s)) = a + by_2(s) + cy_2^2(s)$, equation (28) can be rewritten as follows:

$$\dot{y}_2(s) = \frac{\theta_k}{m\rho}(\delta_1^k + \delta_2^k - a - by_2(s) - cy_2^2(s) - \tilde{r}_l^k),\ s \in [k-1,k].$$

By completing the square on the right-hand side, this equation becomes

$$\dot{y}_2(s) = -\frac{\theta_k c}{m\rho}\left(\left(y_2(s) + \frac{b}{2c}\right)^2 + \frac{\omega_k}{4c^2}\right), \ s \in [k-1, k], \tag{36}$$

where

$$\omega_k = 4c\big(a + \tilde{r}_l^k - \delta_1^k - \delta_2^k\big) - b^2.$$

Thus, if $\omega_k \geq 0$, then

$$\dot{y}_2(s) = -\frac{\theta_k c}{m\rho}\left(\left(y_2(s) + \frac{b}{2c}\right)^2 + \frac{\omega_k}{4c^2}\right) \leq 0, \ s \in [k-1, k],$$

which implies that $y_2(s)$ is non-increasing on $[k-1, k]$. This leads to the following result when $\omega_k \geq 0$:

$$0 \leq y_2(s) \leq \tilde{V}_{\max}^k, \ s \in [k-1, k] \quad \Longleftrightarrow \quad y_2(k) \geq 0, \ y_2(k-1) \leq \tilde{V}_{\max}^k,$$

which shows that the state constraint (31) for the kth subsystem is equivalent to just two point constraints, one at either end of the subsystem.

To deduce a similar result when $\omega_k < 0$, we need the analytical solution for $y_2(s)$, which was derived in [19]. The analytical solution depends critically on the sign of ω_k and in Appendix A we improve the results in [19] and give a thorough analysis of each case $\omega_k = 0$, $\omega_k > 0$, and $\omega_k < 0$. In particular, when $\omega_k < 0$, the solution of (36) is

$$y_2(s) = \frac{\sqrt{|\omega_k|}\phi_k^+(y_2(k-1))}{c\phi_k^+(y_2(k-1)) - c\phi_k^-(y_2(k-1))\exp\left(-\theta_k\sqrt{|\omega_k|}(s-k+1)/m\rho\right)} - \frac{\sqrt{|\omega_k|}}{2c} - \frac{b}{2c},$$

where

$$\phi_k^{\pm}(y_2(k-1)) = 2cy_2(k-1) + b \pm \sqrt{|\omega_k|}.$$

Clearly, $y_2(s)$ is non-increasing on $[k-1, k]$ if $\phi_k^-(y_2(k-1)) \geq 0$ and non-decreasing if $\phi_k^-(y_2(k-1)) < 0$. Hence,

$$0 \leq y_2(s) \leq \tilde{V}_{\max}^k, \ s \in [k-1, k] \iff \begin{cases} y_2(k) \geq 0, \ y_2(k-1) \leq \tilde{V}_{\max}^k, \ \text{if } \phi_k^-(y_2(k-1)) \geq 0, \\ y_2(k) \leq \tilde{V}_{\max}^k, \ y_2(k-1) \geq 0, \ \text{if } \phi_k^-(y_2(k-1)) < 0. \end{cases}$$

The above arguments lead to the following proposition.

Proposition 4.1. *State constraint (31) is equivalent to the following interior point constraints:*

$$0 \leq y_2(k) \leq \min\{\tilde{V}_{\max}^k, \tilde{V}_{\max}^{k+1}\}, \ k = 1, \ldots, Q-1. \tag{37}$$

Proposition 4.1 shows that the infinite-index state constraints (31) for Problem $\mathrm{S}_{N,Q}$ can be converted into $2 \times (Q-1)$ interior point constraints. With this transformation, we can avoid the well-known constraint transcription method, which is the standard approach for handling state constraints and requires introducing an approximation governed by two adjustable parameters.

4.4. Non-smooth control bound constraints

Recall from Fig. 1 that the bounds for the tractive and braking controls each consist of two regimes, one for low speeds and one for high speeds, with the regimes joining at a non-smooth transition point. For example, the bound constraints in [18] are

$$u_1^{\max}(y_2) = \begin{cases} 310, & \text{if } 0 \le y_2 \le 36, \\ 310 - 5(y_2 - 36), & \text{if } 36 < y_2 \le 80, \end{cases} \tag{38}$$

$$u_2^{\max}(y_2) = \begin{cases} 260, & \text{if } 0 \le y_2 \le 60, \\ 260 - 5(y_2 - 60), & \text{if } 60 < y_2 \le 80, \end{cases} \tag{39}$$

where $u_1^{\max}(y_2)$ and $u_2^{\max}(y_2)$ are measured in kN and y_2 is measured in km/h.

In general, the control bounds $u_1^{\max}(y_2)$ and $u_2^{\max}(y_2)$ are non-increasing functions of y_2 that are smooth everywhere except at the respective transition points p_1^* and p_2^*, which mark the transition from low to high speeds. In the example above, $p_1^* = 36$ and $p_2^* = 60$.

The sharp corners at p_1^* and p_2^* will pose challenges for gradient-based optimization methods. Thus, we approximate $u_1^{\max}(y_2)$ and $u_2^{\max}(y_2)$ as follows:

$$u_1^{\max}(y_2) \approx u_{1,\alpha}^{\max}(y_2) = \begin{cases} u_1^{\max}(y_2), & \text{if } y_2 < p_1^* - \alpha,\ y_2 > p_1^* + \alpha, \\ \psi_{1,\alpha}(y_2), & \text{if } p_1^* - \alpha \le y_2 \le p_1^* + \alpha, \end{cases}$$

$$u_2^{\max}(y_2) \approx u_{2,\alpha}^{\max}(y_2) = \begin{cases} u_2^{\max}(y_2), & \text{if } y_2 < p_2^* - \alpha,\ y_2 > p_2^* + \alpha, \\ \psi_{2,\alpha}(y_2), & \text{if } p_2^* - \alpha \le y_2 \le p_2^* + \alpha, \end{cases}$$

where α is a small positive number, and $\psi_{i,\alpha}(y_2)$, $i = 1, 2$, are cubic arcs satisfying

$$\psi_{i,\alpha}(p_i^* - \alpha) = u_i^{\max}(p_i^* - \alpha), \quad \psi_{i,\alpha}(p_i^* + \alpha) = u_i^{\max}(p_i^* + \alpha),$$

and

$$\left.\frac{d\psi_{i,\alpha}}{dy_2}\right|_{p_i^*-\alpha} = \left.\frac{du_i^{\max}}{dy_2}\right|_{p_i^*-\alpha}, \quad \left.\frac{d\psi_{i,\alpha}}{dy_2}\right|_{p_i^*+\alpha} = \left.\frac{du_i^{\max}}{dy_2}\right|_{p_i^*+\alpha}.$$

These conditions ensure that $u_{1,\alpha}^{\max}(y_2)$ and $u_{2,\alpha}^{\max}(y_2)$ are continuously differentiable on $[0, +\infty)$. The approximations $u_{1,\alpha}^{\max}(y_2)$ and $u_{2,\alpha}^{\max}(y_2)$ for (38) and (39) are shown in Fig. 4. With the approximations above, the upper constraints (32) become

$$\delta_1^k \le u_{1,\alpha}^{\max}(y_2(s)), \ s \in [k-1, k), \ k = 1, \ldots, Q,$$

and the lower constraints (33) can be approximated as

$$-u_{2,\alpha}^{\max}(y_2(s)) \le \delta_2^k, \ s \in [k-1, k), \ k = 1, \ldots, Q.$$

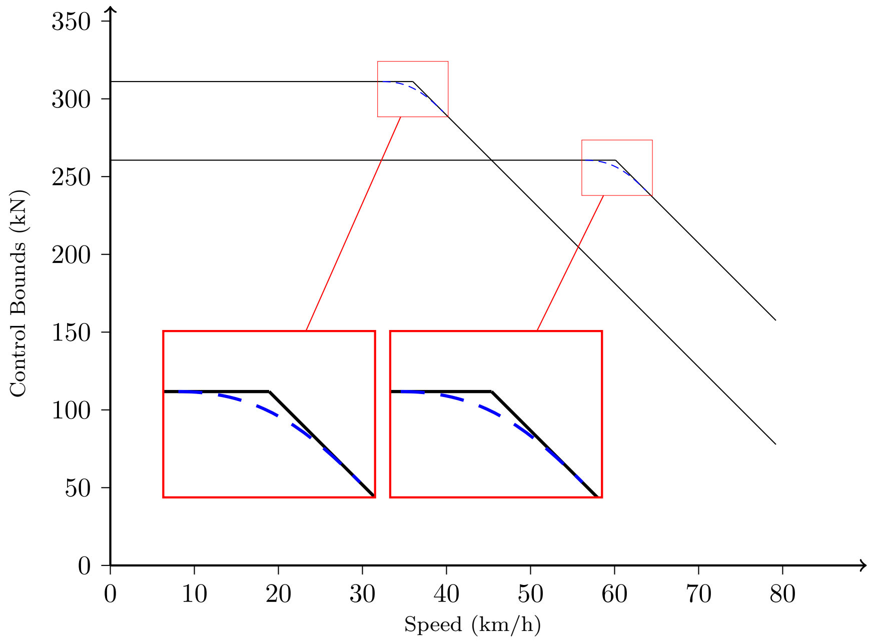

Figure 4. Shapes of the original (non-smooth) and approximate (smooth) control bounds for the locomotive in [18].

Rearranging the above inequalities yields

$$h_1^k(s) = \delta_1^k - u_{1,\alpha}^{\max}(y_2(s)) \leq 0, \ \ s \in [k-1, k), \tag{40}$$

$$h_2^k(s) = -\delta_2^k - u_{2,\alpha}^{\max}(y_2(s)) \leq 0, \ \ s \in [k-1, k), \tag{41}$$

where $k = 1, \dots, Q$. These are joint state-control constraints imposed at an infinite number of time points. By using the constraint transcription method [8, 14], constraints (40) and (41) can be approximated by the following integral constraints:

$$\int_{k-1}^{k} \theta_k \varphi_\varepsilon(h_1^k(s)) ds \leq \gamma, \ \ k = 1, \dots, Q, \tag{42}$$

$$\int_{k-1}^{k} \theta_k \varphi_\varepsilon(h_2^k(s)) ds \leq \gamma, \ \ k = 1, \dots, Q, \tag{43}$$

where $\varphi_\varepsilon(\cdot)$ is defined by

$$\varphi_\varepsilon(\eta) = \begin{cases} \eta, & \text{if } \eta > \varepsilon, \\ (\eta + \varepsilon)^2 / 4\varepsilon, & \text{if } \eta \in [-\varepsilon, \varepsilon], \\ 0, & \text{if } \eta < -\varepsilon, \end{cases}$$

and $\gamma > 0$ and $\varepsilon > 0$ are adjustable parameters.

Thus, the control bound constraints (32)-(33) have been approximated by the conventional constraints (42)-(43). This leads to the following approximation for Problem $\mathrm{S}_{N,Q}$.

Problem $\mathbf{S}_{\boldsymbol{N,Q}}^{\boldsymbol{\alpha,\varepsilon,\gamma}}$. Given the system (27)-(28) with initial conditions (29), find $\boldsymbol{\delta}$ and $\boldsymbol{\theta}$ such that (35) is minimized subject to the constraints (25)-(26), (30), (34), (37), and (42)-(43).

In principal, this problem can be viewed as a nonlinear programming problem. To solve such problems efficiently using existing nonlinear programming methods (e.g., sequential quadratic programming), the gradients of the cost and constraint functions with respect to $\boldsymbol{\delta}$ and $\boldsymbol{\theta}$ are required. We derive these gradients in the next subsection using the sensitivity method [8, 13].

4.5. Gradient formulae

To calculate the gradients of the cost and constraint functions with respect to $\boldsymbol{\delta}$ and $\boldsymbol{\theta}$, we first differentiate (28) to yield the the following linear differential equations for the sensitivity functions $\partial y_2(s)/\partial\delta_i^q$ and $\partial y_2(s)/\partial\theta_q$ on each subinterval $[k-1,k]$:

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\partial y_2(s)}{\partial\delta_i^q}\right) = \frac{\theta_k}{m\rho}\left(\sigma_{kq} - b\frac{\partial y_2(s)}{\partial\delta_i^q} - 2cy_2(s)\frac{\partial y_2(s)}{\partial\delta_i^q}\right),\ i=1,2, \tag{44}$$

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\partial y_2(s)}{\partial\theta_q}\right) &= -\frac{\theta_k}{m\rho}\left(b\frac{\partial y_2(s)}{\partial\theta_q} + 2cy_2(s)\frac{\partial y_2(s)}{\partial\theta_q}\right)\\ &\quad + \frac{\sigma_{kq}}{m\rho}\big(\delta_1^k+\delta_2^k - a - by_2(s) - cy_2^2(s) - \tilde{r}_l^k\big),\end{aligned} \tag{45}$$

where

$$\sigma_{kq} = \begin{cases}1, & \text{if } k=q,\\ 0, & \text{if } k\neq q.\end{cases}$$

The initial conditions for (44)-(45) are

$$\frac{\partial y_2(s)}{\partial\delta_i^q} = 0,\ \ \frac{\partial y_2(s)}{\partial\theta_q} = 0,\ \ s\in[0,q-1].$$

Using the integrating factor method, the solutions of (44)-(45) satisfy

$$I_k(s)\frac{\partial y_2(s)}{\partial\delta_i^q} = \frac{\partial y_2(k-1)}{\partial\delta_i^q} + \frac{\theta_k\sigma_{kq}}{m\rho}\int_{k-1}^{s} I_k(\eta)d\eta,\ i=1,2, \tag{46}$$

$$I_k(s)\frac{\partial y_2(s)}{\partial\theta_q} = \frac{\partial y_2(k-1)}{\partial\theta_q} + \frac{\sigma_{kq}}{m\rho}\int_{k-1}^{s} I_k(\eta)\big(\delta_1^k+\delta_2^k - a - by_2(\eta) - cy_2^2(\eta) - \tilde{r}_l^k\big)d\eta, \tag{47}$$

where

$$\begin{aligned}I_k(s) &= \exp\left(\frac{\theta_k}{m\rho}\int_{k-1}^{s}(2cy_2(\eta)+b)d\eta\right)\\ &= \exp\left(\frac{2c}{m\rho}(y_1(s)-y_1(k-1)) + \frac{\theta_k b}{m\rho}(s-k+1)\right).\end{aligned}$$

The sensitivity functions can then be calculated numerically by applying a cumulative integration scheme, such as Simpson's rule, to the integral terms in (46) and (47). Alternatively, the sensitivity functions can also be obtained by differentiating the analytical solutions in Appendix A, but the algebra becomes very messy. The tedious algebraic manipulations can be avoided by using numerical approximation.

Now, differentiating (27) with respect to $\boldsymbol{\delta}$ and $\boldsymbol{\theta}$ yields the the following differential equations for the sensitivity functions $\partial y_1(s)/\partial \delta_i^q$ and $\partial y_1(s)/\partial \theta_q$ on $[k-1, k]$:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\partial y_1(s)}{\partial \delta_i^q}\right) &= \theta_k \frac{\partial y_2(s)}{\partial \delta_i^q}, \ i = 1, 2, \\
\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\partial y_1(s)}{\partial \theta_q}\right) &= \theta_k \frac{\partial y_2(s)}{\partial \theta_q} + \sigma_{kq} y_2(s),
\end{aligned}$$

with initial conditions

$$\frac{\partial y_1(s)}{\partial \delta_i^q} = 0, \ \frac{\partial y_1(s)}{\partial \theta_q} = 0, \ s \in [0, q-1].$$

Hence,

$$\frac{\partial y_1(s)}{\partial \delta_i^q} = \theta_k \int_{k-1}^{s} \frac{\partial y_2(\eta)}{\partial \delta_i^q} d\eta, \ i = 1, 2, \tag{48}$$

$$\frac{\partial y_1(s)}{\partial \theta_q} = \theta_k \int_{k-1}^{s} \frac{\partial y_2(\eta)}{\partial \theta_q} d\eta + \frac{\sigma_{kq}}{\theta_k}(y_1(s) - y_1(k-1)). \tag{49}$$

As with $\partial y_2(s)/\partial \delta_i^q$ and $\partial y_2(s)/\partial \theta_q$, these integrals can be evaluated numerically using standard numerical integration methods. Once the sensitivity functions have been determined using the equations above, the cost function (35) can be differentiated to yield

$$\frac{\partial G_{N,Q}}{\partial \delta_i^q} = \sum_{k=1}^{Q} \delta_1^k \left(\frac{\partial y_1(k)}{\partial \delta_i^q} - \frac{\partial y_1(k-1)}{\partial \delta_i^q}\right) + \begin{cases} y_1(q) - y_1(q-1), & \text{if } i = 1, \\ 0, & \text{if } i = 2, \end{cases}$$

and

$$\frac{\partial G_{N,Q}}{\partial \theta_q} = \sum_{k=1}^{Q} \delta_1^k \left(\frac{\partial y_1(k)}{\partial \theta_q} - \frac{\partial y_1(k-1)}{\partial \theta_q}\right).$$

Moreover, the terminal and interior point constraints (30), (34) and (37) are all in the form $y_1(k)$ or $y_2(k)$ minus a constant, and thus their gradients can be immediately obtained by evaluating the sensitivity functions at each time point $k = 1, \ldots, Q$.

Finally, for the approximate control bound constraints (42) and (43), denote

$$G_{u_i}^k = \int_{k-1}^{k} \theta_k \varphi_\varepsilon(h_i^k(s)) ds - \gamma \leq 0, \ k = 1, \ldots, Q, \ i = 1, 2.$$

Then, we have

$$\begin{aligned}
\frac{\partial G_{u_i}^k}{\partial \delta_1^q} &= \theta_k \int_{k-1}^{k} \frac{\partial \varphi_\varepsilon(h_i^k(s))}{\partial h_i^k} \left((2-i)\sigma_{kq} - \frac{\partial u_{i,\alpha}^{\max}(y_2(s))}{\partial y_2} \frac{\partial y_2(s)}{\partial \delta_1^q} \right) ds, \\
\frac{\partial G_{u_i}^k}{\partial \delta_2^q} &= \theta_k \int_{k-1}^{k} \frac{\partial \varphi_\varepsilon(h_i^k(s))}{\partial h_i^k} \left((1-i)\sigma_{kq} - \frac{\partial u_{i,\alpha}^{\max}(y_2(s))}{\partial y_2} \frac{\partial y_2(s)}{\partial \delta_2^q} \right) ds, \\
\frac{\partial G_{u_i}^k}{\partial \theta_q} &= -\theta_k \int_{k-1}^{k} \frac{\partial \varphi_\varepsilon(h_i^k(s))}{\partial h_i^k} \frac{\partial u_{i,\alpha}^{\max}(y_2(s))}{\partial y_2} \frac{\partial y_2(s)}{\partial \theta_q} ds + \sigma_{kq} \int_{q-1}^{q} \varphi_\varepsilon(h_i^q(s)) ds,
\end{aligned}$$

where $i = 1, 2$, and

$$\frac{\partial \varphi_\varepsilon(\eta)}{\partial \eta} = \begin{cases} 1, & \text{if } \eta > \varepsilon, \\ (\eta + \varepsilon)/2\varepsilon, & \text{if } \eta \in [-\varepsilon, \varepsilon], \\ 0, & \text{if } \eta < -\varepsilon. \end{cases}$$

4.6. Computational procedure

In summary, the computational procedure for solving Problem $\mathrm{S}_{N,Q}^{\alpha,\varepsilon,\gamma}$ is given below.

Algorithm I. Solving Problem $\mathrm{S}_{N,Q}^{\alpha,\varepsilon,\gamma}$.

1. *Initialization: Set*

$$0 \to \delta_1^k,\ 0 \to \delta_2^k,\ (\tilde{x}_1^k - \tilde{x}_1^{k-1})/\tilde{V}_{\max}^k \to \theta_k,\ k = 1, \ldots, Q.$$

2. *State trajectories: For each subinterval $[k-1, k]$, use the analytical solutions in Appendix A to evaluate $y_1(s)$ and $y_2(s)$ at discrete time points $s_j = k - 1 + j/M_k$, $j = 1, \ldots, M_k$, where M_k is the number of points and $1/M_k$ is the discretization steplength.*
3. *Cost and constraints: Use $y_1(s)$ and $y_2(s)$ from Step 2 to compute the cost and constraint function values.*
4. *Sensitivity functions: For each subinterval $[k-1, k]$, use $y_1(s)$ and $y_2(s)$ from Step 2 to compute the sensitivity functions $\partial y_1(s)/\partial \delta_i^q$, $\partial y_1(s)/\partial \theta_q$, $\partial y_2(s)/\partial \delta_i^q$, and $\partial y_2(s)/\partial \theta_q$, $i = 1, 2$, $q = 1, \ldots, Q$, at discrete time points $s_j = k - 1 + j/M_k$, $j = 1, \ldots, M_k$.*
5. *Gradients: Use $\partial y_1(s)/\partial \delta_i^q$, $\partial y_1(s)/\partial \theta_q$, $\partial y_2(s)/\partial \delta_i^q$, and $\partial y_2(s)/\partial \theta_q$ from Step 4 to determine the gradients for the cost and constraint functions.*
6. *Optimization: Use a gradient-based optimization solver (e.g., fmincon in Matlab) together with the information in Steps 2-5 to calculate a search direction and update $\boldsymbol{\delta}$ and $\boldsymbol{\theta}$ accordingly.*
7. *Return to Step 2.*

5. Case study

To test the computational approach described in Section 4, we consider the Yizhuang subway line in Beijing. There are 14 stations along this line and we choose the segment between Songjiazhuang and Xiaocun stations to define our test problem.

Table 1. Line gradients between Songjiazhuang and Xiaocun stations.

Start Point (m)	End Point (m)	Gradient ($\tan\beta(x_1)$)
0	160	−0.002
160	470	−0.003
470	970	0.0104
970	1370	0.003
1370	1880	−0.008
1880	2500	0.003
2500	2631	−0.002

Table 2. Speed limits between Songjiazhuang and Xiaocun stations.

Start Point (m)	End Point (m)	Speed Limit (km/h)
0	150	50
150	480	85
480	1161	65
1161	2501	85
2501	2631	60

The total length of the segment is $L = 2631$ metres and the operational timetable stipulates a terminal time of $T = 190$ seconds. The line gradients and speed limits are listed in Tables 1 and 2, respectively [18]. Note that the switching points for the line gradients are different to the switching points for the speed limits. The complete set of switching points is:

$$x_1^i \in \{0, 150, 160, 470, 480, 970, 1161, 1370, 1880, 2500, 2501, 2631\}.$$

Thus, the number of track sections is $N = 11$. According to the procedure in Section 4.1, each section $[x_1^{i-1}, x_1^i]$ is decomposed into N_i subsections using a base length of $L_b = 60$ metres, giving

$$N_1 = 3,\ N_2 = 1,\ N_3 = 6,\ N_4 = 1,\ N_5 = 9,\ N_6 = 4,$$
$$N_7 = 4,\ N_8 = 9,\ N_9 = 11,\ N_{10} = 1,\ N_{11} = 3,$$

and $Q = 52$ subsections in total.

For this example, the train mass is $m = 2.78 \times 10^5$ kg, the rotatory mass factor is $\rho = 1.0$, the basic resistance is $r_b(y_2) = 3.9476 + 0.0022294y_2^2$ kN, and the maximum tractive and braking control bounds are given by (38) and (39), respectively. Using *fmincon* in Matlab as the nonlinear optimization solver, we ran Algorithm I with $\alpha = 1$, $\varepsilon = 0.1$, $\gamma = 0.01$, and $M_k = 10$ on a laptop computer with 8G RAM and Intel Core i5-7200U@2.5GHz processor. The optimal control and speed trajectories are shown in Fig. 5, where the solid blue line is the sum of the tractive and braking forces, the dot-dashed lines represent the speed limits and control bounds, and the solid black line is the track altitude. The figure clearly shows that the optimal control signal satisfies the speed limit and control force constraints. Moreover, the optimal control is similar to the optimal four-stage strategy obtained in [1, 10] via Pontryagin's maximum principle:

$$\text{Maximum Traction} \rightarrow \text{Hold Speed} \rightarrow \text{Coast} \rightarrow \text{Maximum Brake}.$$

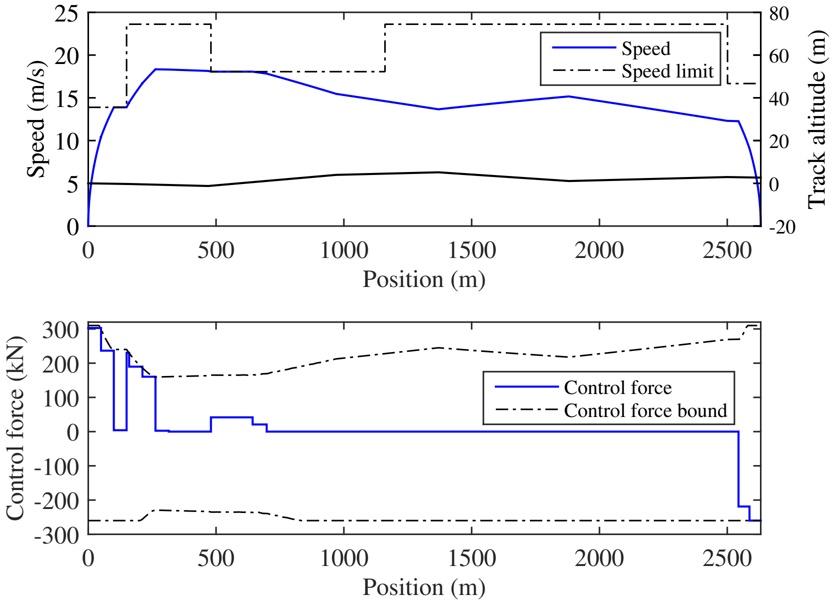

Figure 5. Optimal speed and control trajectories from Algorithm I ($T = 190$ s).

Here "Coast" means that neither tractive nor braking forces are applied to the train.

For comparision, we solved the same problem using three other algorithms:

- Algorithm II – same as Algorithm I except that the gradients are computed using *fmincon*'s finite difference approximation scheme instead of Steps 4 and 5;
- Algorithm III – same as Algorithm I except that the state and sensitivity functions are evaluated numerically by applying Runge-Kutta methods (*ode45* in Matlab) to the respective differential equations; and
- Algorithm IV – Gauss pseudospectral method in GPOPS, a Matlab-based optimal control package [12, 18].

Algorithms II-IV were run on the same computer as Algorithm I. Algorithms I-III are all based on the control parameterization method, whereby the control signal is discretized and the state is considered a function of the control rather than an independent decision variable. Algorithm IV, in contrast, involves discretizing the state in addition to the control and treating both of them as decision variables in the optimization problem, subject to equality constraints defined by the discretized differential equations. The approximation process for Algorithm IV involves transforming the independent variable t in each subsystem into a new variable $\tau \in [-1, 1]$ and then discretizing the state and control variables, along with the cost and constraint functions, at the Legendre-Gauss collocation points in $[-1, 1]$. This yields a nonlinear programming problem that, like the approximate problem obtained via control parameterization, can be solved using standard nonlinear programming algorithms; the GPOPS implementation of Algorithm IV uses $SNOPT$ as the optimization solver. More details on the Gauss pseudospectral method and its applications to solving optimal train control problems can be found in [11, 12, 16–18]. In our simulations, we used 40 collocation points for each subsystem.

The differences between the four algorithms are summarized in Table 3. Note that

Table 3. Summary of the four algorithms used in the case study.

	Algorithm			
	I	**II**	**III**	**IV**
Discretization Scheme	Control Parameterization	Control Parameterization	Control Parameterization	Gauss Pseudospectral
Discretized Variables	Control	Control	Control	Control, State
State Equations	Analytic Solution	Analytic Solution	Approximation (Runge-Kutta)	Approximation (Algebraic Constraints)
Gradient Computation	Sensitivity Functions (Equations (46)-(49))	Finite Differences	Sensitivity Functions (Runge-Kutta)	Automatic Differentiation
Optimization Solver	*fmincon*	*fmincon*	*fmincon*	*SNOPT*
Optimization Tolerance	10^{-6}	10^{-6}	10^{-6}	10^{-8}

Table 4. Performance of the four algorithms used in the case study.

		Algorithm			
		I	**II**	**III**	**IV**
$T = 190$ s	Energy Consumption (kJ)	5.560333×10^4	5.558236×10^4	5.559763×10^4	5.543154×10^4
	Computation Time (s)	20.4	109	242	15.2
$T = 170$ s	Energy Consumption (kJ)	7.055643×10^4	7.049544×10^4	7.054315×10^4	7.020780×10^4
	Computation Time (s)	13.1	96	194	11.2

"Optimization Tolerance" is the tolerance used by the optimization solver (either $fmincon$ or $SNOPT$) for the decision variables, cost, and constraint functions.

We compared Algorithms I-IV for two versions of the train control problem: one using the scheduled time of $T = 190$ seconds, and the other using a shorter final time of $T = 170$ seconds. The minimum energy consumptions and computation times are reported in Table 4. The computation times for Algorithm I are close to those of Algorithm IV, and these algorithms are much quicker than Algorithms II and III. All four algorithms yield almost the same tractive energy consumption and the actual trip times are identical to the target terminal times of 190 and 170 seconds. The control and speed trajectories from Algorithms II-III are similar to Algorithm I (see Fig. 5 for $T = 190$ and Fig. 7 for $T = 170$), but the control trajectories from Algorithm IV are very different: there is severe fluctuation during the Speed-Hold stage, as shown in Fig. 6 and Fig. 8. This is obviously unrealistic to implement in practice, irrespective of whether the train is controlled by a human driver or an automatic train control system. Our new method does not yield any control fluctuation.

6. Conclusion

This paper has discussed a time-based switched system formulation for the optimal train control problem with variable line gradients and speed limit constraints. To solve this problem, we proposed a numerical approach consisting of the following key elements: control parameterization for discretizing the control signals, a time-scaling

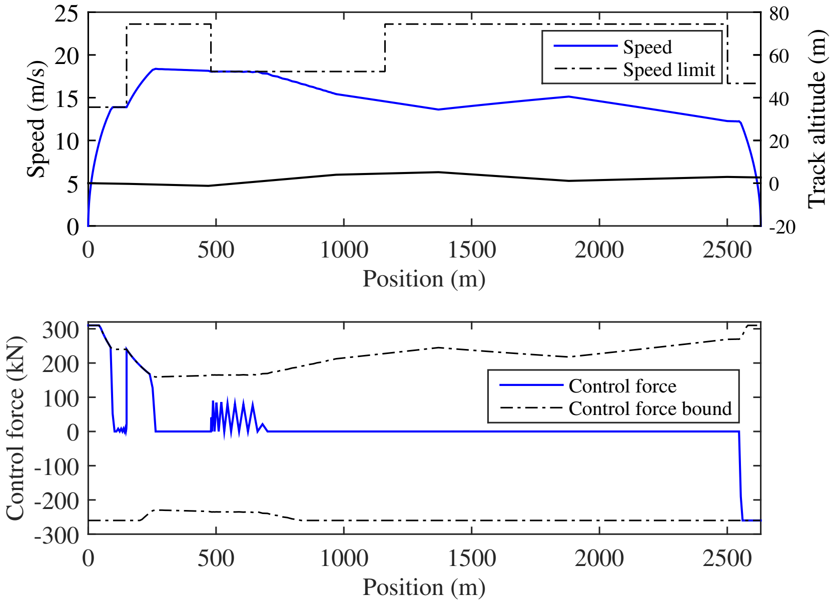

Figure 6. Optimal speed and control trajectories from Algorithm IV ($T = 190$ s).

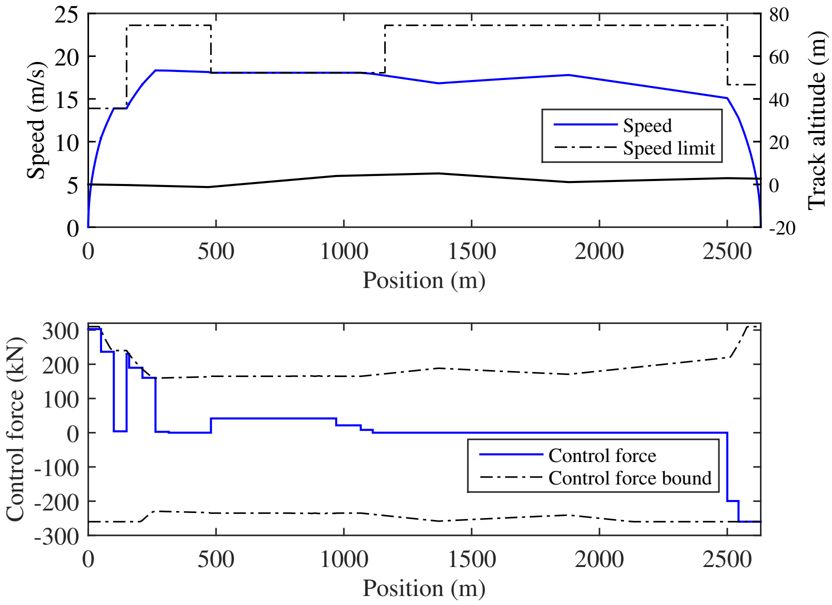

Figure 7. Optimal speed and control trajectories from Algorithm I ($T = 170$ s).

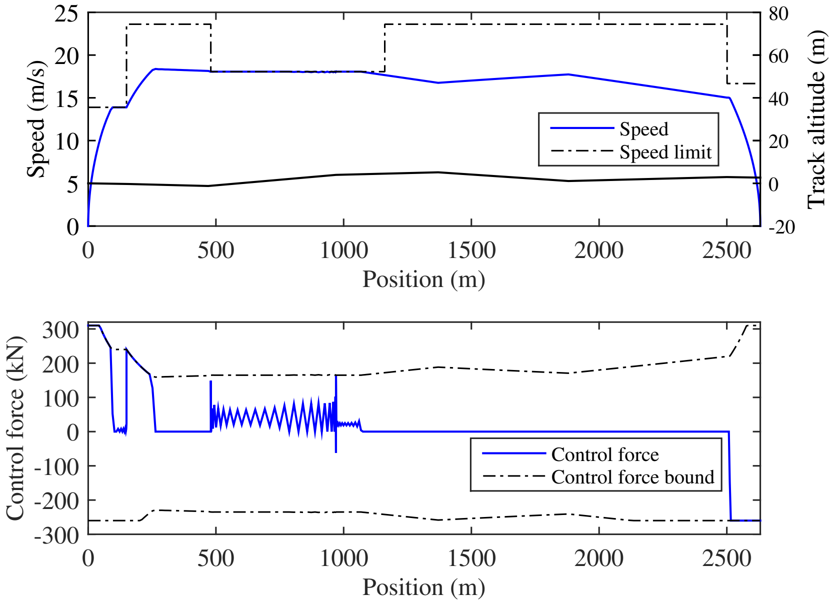

Figure 8. Optimal speed and control trajectories from Algorithm IV ($T = 170$ s).

transformation for converting the variable subsystem switching times into fixed integer points, and a smooth approximation scheme for the non-smooth control bounds. The "infinite-index" speed limit constraints are normally very challenging, but we showed that by exploiting the structure of the analytical solution to the train differential equations, the speed limits can be reformulated as a finite number of standard point constraints. The end result is a nonlinear programming problem that can be solved by gradient-based optimization algorithms such as sequential quadratic programming, for which many efficient practical implementations are available. The case study results for the Yizhuang subway line show that the proposed approach is effective at handling the complex speed limit and control bound constraints in the train control problem. Moreover, compared with the pseudospectral method, our new method can avoid control fluctuations during singular arcs, without any sacrifice to the tractive energy consumption or computational time. This is a key advantage because, as recognized in [18, 19] and observed in our case study, control fluctuation can be an issue with pseudospectral methods.

Funding

This work was supported by the National Key Research and Development Program of China under Grant 2016YFB1200601, the National Natural Science Foundation of China under Grant 61120106009 and the China Scholarship Council under Grant 201707095078.

Appendix A. Analytical solutions for the position and speed

From (36), the speed satisfies the following differential equation:

$$\dot{y}_2(s) = -\frac{\theta_k c}{m\rho}\Big(\Big(y_2(s) + \frac{b}{2c}\Big)^2 + \frac{\omega_k}{4c^2}\Big),\ s \in [k-1, k],$$

where

$$\omega_k = 4c\big(a + \tilde{r}_l^k - \delta_1^k - \delta_2^k\big) - b^2.$$

We consider three cases: $\omega_k > 0$, $\omega_k = 0$, and $\omega_k < 0$. Note that five cases were considered in [19]; our new solution expressions derived below show that the two cases for $\omega_k = 0$ in [19] can be combined, and likewise the two cases for $\omega_k < 0$ can also be combined.

When $\omega_k > 0$, the differential equation can be written as

$$\Big(\big(y_2(s) + b/2c\big)^2 + \frac{\omega_k}{4c^2}\Big)^{-1}\dot{y}_2(s) = -\frac{\theta_k c}{m\rho},\ s \in [k-1, k].$$

Integrating both sides and using the substitution $v = y_2 + b/2c$ gives

$$\int_{k-1}^{s} \Big(\big(y_2(\eta)+b/2c\big)^2+\frac{\omega_k}{4c^2}\Big)^{-1}\dot{y}_2(\eta)d\eta = \int_{y_2(k-1)+b/2c}^{y_2(s)+b/2c} \frac{1}{v^2 + \omega_k/4c^2}dv = -\frac{\theta_k c(s-k+1)}{m\rho},$$

and thus the speed $y_2(s)$ must satisfy

$$\arctan\Big(\frac{2cy_2(s) + b}{\sqrt{\omega_k}}\Big) = -\frac{\theta_k\sqrt{\omega_k}(s-k+1)}{2m\rho} + \Delta_k(y_2(k-1)), \tag{A1}$$

where

$$\Delta_k(y_2(k-1)) = \arctan\Big(\frac{2cy_2(k-1) + b}{\sqrt{\omega_k}}\Big).$$

If the range of the right-hand side of (A1) is not within the interval $(-\pi/2, \pi/2)$, then the speed differential equation does not have a solution over the entire subinterval $[k-1, k]$. However, since $\Delta_k(y_2(k-1)) \in (-\pi/2, \pi/2)$ and θ_k, $\sqrt{\omega_k}$, m, and ρ are all positive, the right-hand side of (A1) satisfies

$$\begin{aligned}-\frac{\theta_k\sqrt{\omega_k}}{2m\rho} + \Delta_k(y_2(k-1)) &\le -\frac{\theta_k\sqrt{\omega_k}(s-k+1)}{2m\rho} + \Delta_k(y_2(k-1)) \\ &\le \Delta_k(y_2(k-1)) < \frac{\pi}{2},\ s \in [k-1, k],\end{aligned}$$

and thus a solution only exists when

$$-\frac{\theta_k\sqrt{\omega_k}}{2m\rho} + \Delta_k(y_2(k-1)) > -\frac{\pi}{2}. \tag{A2}$$

Thankfully, this condition is almost always satisfied in practice because the mass m is large relative to the other parameters and therefore,

$$\Delta_k(y_2(k-1)) - \frac{\theta_k\sqrt{\omega_k}}{2m\rho} \approx \Delta_k(y_2(k-1)) > -\frac{\pi}{2},$$

as required. Moreover, if the right-hand side of (A1) approaches $-\pi/2$, then $y_2(s)$ will become negative on the kth subinterval, violating the state constraints. Hence, feasible trajectories will satisfy (A2). Assuming (A2) holds, the speed $y_2(s)$ is obtained by solving equation (A1):

$$y_2(s) = \frac{\sqrt{\omega_k}}{2c}\tan\Big(-\frac{\theta_k\sqrt{\omega_k}}{2m\rho}(s-k+1) + \Delta_k(y_2(k-1))\Big) - \frac{b}{2c}.$$

Then, the position $y_1(s)$ is obtained by integrating both sides of (27), yielding

$$\begin{aligned} y_1(s) &= y_1(k-1) + \theta_k \int_{k-1}^{s} y_2(\eta)d\eta \\ &= y_1(k-1) + \frac{m\rho}{c}\ln\Big(\cos\Big(-\frac{\theta_k\sqrt{\omega_k}}{2m\rho}(s-k+1) + \Delta_k(y_2(k-1))\Big)\Big) \\ &\quad - \frac{m\rho}{c}\ln\big(\cos(\Delta_k(y_2(k-1)))\big) - \frac{\theta_k b}{2c}(s-k+1),\ \ s \in [k-1,k]. \end{aligned}$$

When $\omega_k = 0$, the speed differential equation is

$$\dot{y}_2(s) = -\frac{\theta_k c}{m\rho}\Big(y_2(s) + \frac{b}{2c}\Big)^2,\ \ s \in [k-1,k].$$

If $y_2(k-1) = -b/2c$, then clearly $y_2(s) = -b/2c$ is the solution of this equation (although this solution clearly violates the state constraints, which prohibit negative speeds). Thus, we assume that $y_2(k-1) \neq -b/2c$ and let $s^* > k-1$ denote the first time at which $y_2(s) = -b/2c$, where $s^* = +\infty$ if $y_2(s)$ never reaches $-b/2c$ in the kth subinterval. For $s < s^*$, the differential equation can be written as

$$\frac{1}{\big(y_2(s) + b/2c\big)^2}\dot{y}_2(s) = -\frac{\theta_k c}{m\rho},\ \ s \in [k-1,k],$$

and thus integrating both sides yields

$$\int_{k-1}^{s} \frac{1}{\big(y_2(\eta) + b/2c\big)^2}\dot{y}_2(\eta)d\eta = -\frac{\theta_k c}{m\rho}(s-k+1).$$

Hence, by evaluating the integral on the left-hand side, we obtain

$$\frac{1}{y_2(s) + b/2c} = \frac{\theta_k c}{m\rho}(s-k+1) + \frac{1}{y_2(k-1) + b/2c}.$$

Simplifying gives the following expression for $y_2(s)$, which holds for all $s < s^*$:

$$y_2(s) = \frac{m\rho(y_2(k-1)+b/2c)}{m\rho + \theta_k c(y_2(k-1)+b/2c)(s-k+1)} - \frac{b}{2c}.$$

This expression incorporates the case when $y_2(k-1) = -b/2c$ and is always well-defined for non-negative speeds because in this case

$$m\rho + \theta_k c(y_2(k-1)+b/2c)(s-k+1) > 0.$$

Moreover, when $y_2(k-1) > -b/2c$, the speed $y_2(s) \neq -b/2c$ for all $s \in [k-1,k]$ and thus the derivation above is valid over the entire subinterval $[k-1,k]$, since

$$y_2(s) = \frac{m\rho(y_2(k-1)+b/2c)}{m\rho + \theta_k c(y_2(k-1)+b/2c)(s-k+1)} - \frac{b}{2c} > -\frac{b}{2c}.$$

The corresponding solution for $y_1(s)$ is:

$$\begin{aligned} y_1(s) &= y_1(k-1) + \theta_k \int_{k-1}^{s} y_2(\eta)d\eta \\ &= y_1(k-1) + \frac{m\rho}{c}\ln\Big(1 + \frac{\theta_k c}{m\rho}(y_2(k-1)+b/2c)(s-k+1)\Big) - \frac{\theta_k b}{2c}(s-k+1). \end{aligned}$$

Finally, for $\omega_k < 0$, if $2cy_2(k-1)+b = \sqrt{|\omega_k|}$, then $y_2(s) = \sqrt{|\omega_k|}/2c - b/2c$ is the solution of the differential equation. Hence, assume $2cy_2(k-1)+b \neq \sqrt{|\omega_k|}$ and let $s^* > k-1$ denote the first time at which $y_2(s) = \sqrt{|\omega_k|}/2c - b/2c$. Then for $s < s^*$, the differential equation can be rewritten as

$$\bigg(\Big(y_2(s) + \frac{b}{2c}\Big)^2 - \Big(\frac{\sqrt{|\omega_k|}}{2c}\Big)^2\bigg)^{-1}\dot{y}_2(s) = -\frac{\theta_k c}{m\rho},$$

or equivalently,

$$\Big(\frac{2c}{2cy_2(s)+b-\sqrt{|\omega_k|}} - \frac{2c}{2cy_2(s)+b+\sqrt{|\omega_k|}}\Big)\dot{y}_2(s) = -\frac{\theta_k\sqrt{|\omega_k|}}{m\rho}.$$

Hence, integrating both sides gives

$$\ln\left|\frac{2cy_2(s)+b-\sqrt{|\omega_k|}}{2cy_2(s)+b+\sqrt{|\omega_k|}}\right| - \ln\left|\frac{2cy_2(k-1)+b-\sqrt{|\omega_k|}}{2cy_2(k-1)+b+\sqrt{|\omega_k|}}\right| = -\frac{\theta_k\sqrt{|\omega_k|}(s-k+1)}{m\rho},$$

and

$$\ln\left|\frac{2cy_2(s)+b-\sqrt{|\omega_k|}}{2cy_2(s)+b+\sqrt{|\omega_k|}} \cdot \frac{2cy_2(k-1)+b+\sqrt{|\omega_k|}}{2cy_2(k-1)+b-\sqrt{|\omega_k|}}\right| = -\frac{\theta_k\sqrt{|\omega_k|}(s-k+1)}{m\rho}.$$

Since $2cy_2(s)+b-\sqrt{|\omega_k|}$ and $2cy_2(k-1)+b-\sqrt{|\omega_k|}$ have the same sign for $s < s^*$, when the speed is non-negative, this can be simplified to

$$\frac{2cy_2(s)+b-\sqrt{|\omega_k|}}{2cy_2(s)+b+\sqrt{|\omega_k|}}=\frac{2cy_2(k-1)+b-\sqrt{|\omega_k|}}{2cy_2(k-1)+b+\sqrt{|\omega_k|}}\exp\Big(-\frac{\theta_k\sqrt{|\omega_k|}(s-k+1)}{m\rho}\Big).$$

Solving for $y_2(s)$ yields

$$y_2(s)=\frac{\sqrt{|\omega_k|}\phi_k^+(y_2(k-1))}{c\phi_k^+(y_2(k-1))-c\phi_k^-(y_2(k-1))\exp\left(-\theta_k\sqrt{|\omega_k|}(s-k+1)/m\rho\right)}-\frac{\sqrt{|\omega_k|}}{2c}-\frac{b}{2c},$$

where $\phi_k^{\pm}(y_2(k-1))=2cy_2(k-1)+b\pm\sqrt{|\omega_k|}$. This solution incorporates the case when $2cy_2(k-1)+b=\sqrt{|\omega_k|}$ and is well-defined for non-negative speeds because the denominator satisfies

$$\begin{aligned}c\big(2cy_2(k-1)+b+\sqrt{|\omega_k|}\big)-c\big(2cy_2(k-1)+b-\sqrt{|\omega_k|}\big)\exp\left(-\theta_k\sqrt{|\omega_k|}(s-k+1)/m\rho\right)\\ =c\big(2cy_2(k-1)+b\big)\big(1-\exp\left(-\theta_k\sqrt{|\omega_k|}(s-k+1)/m\rho\right)\big)\\ +\ c\sqrt{|\omega_k|}\big(1+\exp\left(-\theta_k\sqrt{|\omega_k|}(s-k+1)/m\rho\right)\big)>0.\end{aligned}$$

It is also clear that if $2cy_2(k-1)+b\neq\sqrt{|\omega_k|}$, then s^* must be infinite and the solution exists over the entire subinterval. For the position $y_1(s)$ when $\omega_k<0$, the analytical formula is

$$\begin{aligned}y_1(s)&=y_1(k-1)+\theta_k\int_{k-1}^{s}y_2(\eta)d\eta\\ &=y_1(k-1)+\frac{m\rho}{c}\ln\left(\frac{\phi_k^+(y_2(k-1))\exp\left(\theta_k\sqrt{|\omega_k|}(s-k+1)/m\rho\right)-\phi_k^-(y_2(k-1))}{2\sqrt{|\omega_k|}}\right)\\ &\qquad-\frac{\theta_k(\sqrt{|\omega_k|}+b)}{2c}(s-k+1).\end{aligned}$$

References

[1] A. Albrecht, P. Howlett, P. Pudney, X. Vu, and P. Zhou, *The key principles of optimal train control — part 1: Formulation of the model, strategies of optimal type, evolutionary lines, location of optimal switching points*, Transportation Research Part B: Methodological 94 (2016), pp. 482 – 508. Available at https://www.sciencedirect.com/science/article/pii/S0191261515002064.

[2] I.A. Asnis, A.V. Dmitruk, and N.P. Osmolovskii, *Solution of the problem of the energetically optimal control of the motion of a train by the maximum principle*, USSR Computational Mathematics and Mathematical Physics 25 (1985), pp. 37–44.

[3] J.W. Davis Jr., *The tractive resistance of electric locomotives and cars*, General Electric Review 29 (1926), pp. 2–24.

[4] A. González-Gil, R. Palacin, P. Batty, and J. Powell, *A systems approach to reduce urban rail energy consumption*, Energy Conversion and Management 80 (2014), pp. 509 – 524. Available at http://www.sciencedirect.com/science/article/pii/S0196890414001058.

[5] P. Howlett, *The optimal control of a train*, Annals of Operations Research 98 (2000), pp. 65–87. Available at https://doi.org/10.1023/A:1019235819716.

[6] K. ICHIKAWA, *Application of optimization theory for bounded state variable problems to the operation of train*, Bulletin of JSME 11 (1968), pp. 857–865.
[7] E. Khmelnitsky, *On an optimal control problem of train operation*, IEEE Transactions on Automatic Control 45 (2000), pp. 1257–1266.
[8] Q. Lin, R. Loxton, and K.L. Teo, *The control parameterization method for nonlinear optimal control: A survey*, Journal of Industrial and Management Optimization 10 (2014), pp. 275–309.
[9] Q. Lin, R. Loxton, K.L. Teo, Y.H. Wu, and C. Yu, *A new exact penalty method for semi-infinite programming problems*, Journal of Computational and Applied Mathematics 261 (2014), pp. 271–286. Available at http://www.sciencedirect.com/science/article/pii/S0377042713006298.
[10] R.R. Liu and I.M. Golovitcher, *Energy-efficient operation of rail vehicles*, Transportation Research Part A: Policy and Practice 37 (2003), pp. 917–932. Available at http://www.sciencedirect.com/science/article/pii/S0965856403000752.
[11] M.A. Patterson and A.V. Rao, *Gpops-ii: A matlab software for solving multiple-phase optimal control problems using hp-adaptive gaussian quadrature collocation methods and sparse nonlinear programming*, ACM Trans. Math. Softw. 41 (2014), pp. 1:1–1:37. Available at http://doi.acm.org/10.1145/2558904.
[12] A.V. Rao, D.A. Benson, C. Darby, M.A. Patterson, and C. Francolin, *Algorithm 902: Gpops, a matlab software for solving multiple-phase optimal control problems using the gauss pseudospectral method*, ACM Transactions on Mathematical Software 37 (2010), pp. 22:1–22:39.
[13] O. Rosen and R. Luus, *Evaluation of gradients for piecewise constant optimal control*, Computers and Chemical Engineering 15 (1991), pp. 273–281.
[14] K.L. Teo and C.J. Goh, *A unified computational approach to optimal control problems.*, Longman Scientific and Technical, Essex,UK, 1991.
[15] K.L. Teo, L.S. Jennings, H.W.J. Lee, and V. Rehbock, *The control parameterization enhancing transform for constrained optimal control problems*, The Journal of the Australian Mathematical Society. Series B. Applied Mathematics 40 (1999), p. 314–335.
[16] P. Wang and R.M. Goverde, *Multiple-phase train trajectory optimization with signalling and operational constraints*, Transportation Research Part C: Emerging Technologies 69 (2016), pp. 255 – 275. Available at http://www.sciencedirect.com/science/article/pii/S0968090X16300766.
[17] Y. Wang, B.D. Schutter, T.J. van den Boom, and B. Ning, *Optimal trajectory planning for trains – a pseudospectral method and a mixed integer linear programming approach*, Transportation Research Part C: Emerging Technologies 29 (2013), pp. 97–114. Available at http://www.sciencedirect.com/science/article/pii/S0968090X13000193.
[18] H. Ye and R. Liu, *A multiphase optimal control method for multi-train control and scheduling on railway lines*, Transportation Research Part B: Methodological 93 (2016), pp. 377 – 393. Available at http://www.sciencedirect.com/science/article/pii/S0191261516301084.
[19] H. Ye and R. Liu, *Nonlinear programming methods based on closed-form expressions for optimal train control*, Transportation Research Part C: Emerging Technologies 82 (2017), pp. 102 – 123. Available at http://www.sciencedirect.com/science/article/pii/S0968090X1730164X.
[20] W. Zhong and H. Xu, *Energy-efficient train operation optimization using control parameterization method*, Journal of the China Railway Society 39 (2017), pp. 72–79.