Multi-train trajectory optimization for energy-efficient timetabling

Wang, Pengling; Goverde, Rob M.P.

DOI
10.1016/j.ejor.2018.06.034

Publication date
2019

Document Version
Accepted author manuscript

Published in
European Journal of Operational Research

Multi-Train Trajectory Optimization for Energy-Efficient Timetabling

Pengling Wang*, Rob M.P. Goverde

Department of Transport and Planning, Delft University Technology, Delft, the Netherlands

Abstract

This paper proposes a novel approach for energy-efficient timetabling by adjusting the running time allocation of given timetables using train trajectory optimization. The approach first converts the arrival and departure times to time window constraints in order to relax the given timetable. Then a train trajectory optimization method is developed to find optimal arrival/departure times and optimal energy-efficient speed profiles within the relaxed time windows. The proposed train trajectory optimization method includes two types, a single-train trajectory optimization (STTO), which focuses on optimizing individual train movements within the relaxed arrival and departure time windows, and a multi-train trajectory optimization (MTTO), which computes multi-train trajectories simultaneously with a shared objective of minimizing multi-train energy consumption and an additional target of eliminating conflicts between trains. The STTO and MTTO are re-formulated as a multiple-phase optimal control problem, which has the advantage of accurately incorporating varying gradients, curves and speed limits and different train routes. The multiple-phase optimal control problem is then solved by a pseudospectral method. The proposed approach is applied in case studies to fine-tune two timetables, for a single-track railway corridor and a double-track corridor of the Dutch railway. The results suggest that the proposed approach is able to improve the energy efficiency of a timetable.

Keywords: Transportation, Energy efficient timetabling, Train trajectory optimization, Pseudospectral method

1. Introduction

Improving energy efficiency is an important issue for railways, even though rail is already more energy efficient than most other transport modes. One promising means is to optimize train operations by using energy-efficient driving strategies, which does not need extra investment for more infrastructure. Even with a small amount of energy saved in each train run, the total energy costs saved by the whole railway network are huge. Energy-efficient driving is highly relevant to the running times given in the timetable. A running time between two stations contains two parts, the technically minimum running time and the running time supplement. The running time supplement is the extra running time on top of the technically minimum running time between two stations which is included in the timetable primarily to manage disturbances in operations and to recover from small delays (Lusby et al., 2017). However, if a train is punctual then these supplements can be used for energy-efficient driving (Scheepmaker et al., 2017). Therefore, the timetable has a direct influence on energy-efficiency of train operations.

The research on energy-efficient timetabling has just drawn attentions recently, whilst the research on energy-efficient train operations has been studied for decades. There are comprehensive surveys on relevant areas, which include Wang et al. (2011), Albrecht et al. (2016), Scheepmaker et al. (2017), and Yang et al. (2016). Wang et al. (2011) reviewed the numerical approaches for solving the train trajectory optimization problem. Albrecht et al. (2016) focused on the state-of-the-art of using the Pontryagin's Maximum Principle (PMP) to find the key principles of optimal train control. Scheepmaker et al. (2017) and Yang et al. (2016) provided surveys on the energy-efficient train control and energy-efficient train timetable problems, where Scheepmaker et al. (2017) focused on general railway systems, and Yang et al. (2016) focused on urban rail.

*Corresponding author
Email addresses: P.L.Wang@tudelft.nl (Pengling Wang), R.M.P.Goverde@tudelft.nl (Rob M.P. Goverde)

The literature review presented below focuses on energy-efficient timetabling and energy-efficient operation of multiple trains, with a focus on recent publications that extend the recent review paper of Scheepmaker et al. (2017).

1.1. Review of energy-efficient timetabling

The quality of a railway timetable can be measured by several key performance indicators (KPIs): journey time efficiency, timetable feasibility, robustness, and energy efficiency (Goverde et al., 2016). Among those KPIs, energy efficiency is a secondary objective, particularly in dense railway networks, and can, therefore, be considered as a fine-tuning step after the time allowances have been set based on feasibility and robustness. The scientific literature on railway timetabling mainly considers macroscopic optimization models (Cacchiani and Toth, 2012; Lusby et al., 2017). The periodic event-scheduling problem model introduced by Serafini and Ukovich (1989) has been widely applied in cyclic timetabling (Peeters, 2003; Hansen and Pachl, 2014). Graph-based models (Cacchiani et al., 2008; Yang et al., 2009) and mixed integer linear programming (MILP) models (Brännlund et al., 1998; Cacchiani and Toth, 2012) were developed for generating non-cyclic timetables. Those macroscopic optimization models focus on finding optimal train orders, but are not concerned about how to the get accurate input parameters, such as running times, to set up the macroscopic model, while energy-efficient timetabling requires a microscopic level of details (Caimi et al., 2011).

One stream of energy-efficient timetabling is to find running time allocations that are beneficial for energy-efficient driving. Scheepmaker and Goverde (2015) developed an EZR model (energy-efficient operation or in Dutch 'EnergieZuinig Rijden') to find the energy-efficient trajectory of an individual train trip. The EZR model was adopted to analyze energy efficiency of practical train movements and a real-world timetable. The results show that using a more uniform allocation of the running time supplements leads to extra energy savings and an improvement on punctuality compared to the method of tightening the timetable. Su et al. (2013) developed an optimization model that determines both an energy-efficient driving strategy and an optimal distribution of the running time supplements in the timetable. The authors first explicitly calculated the energy efficient train control strategy per trip. Then the model distributes the running time supplements among consecutive trips in order to minimize the total energy consumption. Goverde et al. (2016) proposed a three-level timetable design method, which constructs a stable robust conflict-free timetable with optimal train orders first and then adjusts time supplements between the stops of regional trains in a corridor between two main stations for energy-efficiency. The time supplements are re-allocated to optimize energy-saving train operation, taking into account the stochastic dwell times. The problem is formulated as a multi-stage, multi-criteria decision problem that is solved by dynamic programming. Canca and Zarzo (2017) researched the energy-efficient timetabling problem in rapid railway transit networks. The timetable design problem was formulated as a mixed integer non-linear optimization, while the train energy consumption was described as a function of running times, and solved by a sequential mixed integer linear solving process.

Another stream is to find an optimal timetable which maximizes the utilization of regenerative energy. Gupta et al. (2016) proposed a two-step optimization model to optimize timetables in metro railway networks. The first model minimizes the train energy consumption and finds associated optimal trip times. The second model maximizes the transfer of regenerative braking energy. Huang et al. (2016, 2017) proposed a timetabling method by optimizing passenger trip time and operational energy consumption (including regenerative energy). The strategy is to adjust the headway on different operating lines to acquire a service-level-based and energy-efficient timetable.

1.2. Review of multi-train operation optimization

Energy-efficient train operation research aims at searching for the optimal trajectory (speed-distance profile) and driving strategies of a train that minimize energy consumption caused by the train movements (Howlett and Pudney, 1995).

Multi-train trajectory optimization (MTTO) optimizes multiple train movements together with a shared objective to find optimal control strategies for every involved train. Saving energy consumption is a common-used optimization objective.

Yang et al. (2012) provided a mathematical model for multiple trains on a railway network. The model aims at minimizing total energy consumption and running times of all trains, while satisfying the constraints to ensure the feasibility of multi-train operations, which include headway constraints, vehicle speed limit constraints, passenger riding comfort constraints, and dwell time constraints. The control strategies of every

involved train are the decision variables of the multi-train model. A genetic algorithm (GA) integrated with simulation was designed to find the optimal control strategies. Li and Lo (2014a,b) proposed a multi-train model for metro lines. The model uses the switching time and speeds of control regimes (acceleration, cruising, coasting, and braking) as the decision variables, aims at minimizing the net energy consumption of multiple trains, and takes into account constraints of the number of trains, cyclist times, switching times, turnaround times, vehicle speed limits and dwell times. In Li and Lo (2014a), a GA method was designed to find optimal switching time and speeds and jointly optimize the timetables and speed profiles. In Li and Lo (2014b), the model is transformed to a convex optimization problem by using a linear approximation method, and solved using the Kuhn-Tucker conditions for dynamic train scheduling. Su et al. (2014) formulated an integrated energy-efficient optimization model to realize the optimal control of multi-trains. The model is for a railway corridor with one type of cyclic train operation. The model aims at minimizing the energy consumption caused by traction, and takes into account the cycle time, train dynamic movements, traction and braking forces, and speed limits constraints. The headway between adjacent cyclic operations is first determined according to the passenger demands. Then an iteration algorithm is used to find the optimal cycle time and the optimal number of trains, which minimize the timetable's energy consumption. The optimal control strategy for each train is computed by an independent optimal train-control algorithm. Zhou et al. (2017) incorporated train speed control into the timetable design process. A space-time-speed grid network for joint train routing, timetabling and trajectory optimization is constructed as a path finding problem, which is solved by a dynamic programming algorithm.

As shown in the literature, researchers have proven that the MTTO has benefits in solving timetabling or rescheduling problems because the MTTO approaches optimize multi-train movements simultaneously, which achieves a global optimization of energy and capacity usage while taking into account the operational interactions between adjacent trains. The train trajectory and timetable are closely related and both of them have a direct influence on the energy efficiency of train operations.

1.3. Contributions of this paper

This paper contributes to a novel approach of energy-efficient timetabling that is developed from a train trajectory optimization view. The presented work belongs to the first stream of energy-efficient timetabling, by adjusting running time allocations of an existing timetable in order to maximize the possibility of energy-efficient driving. The method (a) focuses on a timetable for a railway corridor, (b) calculates energy-efficient speed profiles for every train in the timetable, taking into account the practical situation that trains use different routes and platforms at stations, constraints of train dynamic movements, vehicle characteristics, varying speed limits and gradients, and interactions between trains; and (c) is capable of adjusting running time allocation and producing a new energy-efficient timetable.

The method is based on an existing timetable. It firstly converts the given timetable into flexible arrival and departure time window constraints, which are described as a timetable constraint set (TCS). Then an energy-efficient train trajectory optimization (ETTO) method is developed to find the optimal energy-efficient time-distance paths within those TCSs. The ETTO method include two parts: one is a single-train trajectory optimization (STTO), which focuses on optimizing an individual train movement within the relaxed arrival and departure time windows, and the other is a multi-train trajectory optimization (MTTO), which computes multi-train trajectories simultaneously with a shared objective of minimizing multi-train energy consumption and an additional target of eliminating conflicts between trains. Compared to the STTO method, the MTTO method can optimize the energy consumption of multiple trains together with constraints to avoid conflicts between trains. The STTO and MTTO are re-formulated as a multiple-phase optimal control problem (Rao, 2003; Wang and Goverde, 2016b,c), which has the advantage of accurately incorporating varying gradients, curves and speed limits and different train routes. It is then solved by a pseudospectral method (Rao et al., 2010).

The multiple-phase optimal control problem formulation and the pseudospectral method have been used in Wang and Goverde (2016a,c, 2017) for single-/multiple- train trajectory optimization. The advantage of this model is an accurate and flexible representation of changing speed limits, gradients and time/speed constraints at stations (Wang and Goverde, 2016a). This paper extends the methods to optimize trajectories for multiple trains on a railway corridor composed of single and/or double tracks, and implements the trajectory optimization method in energy-efficient timetable adjustment.

The proposed approach integrates the train trajectory optimization and energy-efficient timetabling, which does not need pre-calculations of the relationships between running times and energy consumptions like the methods in (Su et al., 2013; Goverde et al., 2016). Compared to the integrated methods on the timetable design level (Kraay et al., 1991; Yang et al., 2012; Li and Lo, 2014a,b; Yin et al., 2016), the proposed approach provides a more microscopic description of train movements and more accurate results of energy-efficient speed profiles. The regenerative braking is not taken into account in this paper, although it is a significant research direction. Regenerative braking is most efficient for high frequent (metro-like) systems or when using high-voltage electricity systems. The aim of the studies on regenerative braking is to synchronize braking and accelerating trains close to mutual stops (Gupta et al., 2016; Huang et al., 2016, 2017). For mainline railway networks as considered in our paper, the frequency of trains is relatively low and the trains are already constrained by many network constraints, so adding this kind of synchronization to trains on opposite directions of a double-track line makes the (NP-complete) timetabling (and traffic management) even more complicated. Moreover, periodic timetables such as the Dutch railway timetable are usually symmetric, which means that trains from the same line in opposite directions meet in a station at the same time (i.e. brake simultaneously and accelerate again simultaneously) rather than having an offset such that one train accelerates while the other brakes. This holds in particular for single-track lines, since the departing (accelerating) train has to wait until the opposite arriving (braking) train has already stopped. Something similar applies to trains from different directions that meet at a station, in which case passengers might want to transfer between the two trains, and therefore the objective in timetabling for mainline railway networks is usually the opposite: synchronizing the trains so that they arrive at the same time, rather than having one train depart when the other arrives. These are the reasons that we did not consider regenerative braking. There might be situations where regenerative braking is also worthwhile on mainline railway networks, but this would typically be another fine-tuning step for specific train pairs.

The proposed approach is demonstrated by application to two case studies of opposite-trains running on a Dutch partially single-track railway corridor and following-trains running on a Dutch double-track railway corridor. The results suggest that our method is able to produce optimal energy-efficient speed profiles and improve the energy efficiency of a timetable.

The remainder of the paper is organized as follows: Section 2 provides a description of the energy-efficient timetabling problem and an introduction to the energy-efficient timetabling strategy. Section 3 presents the train trajectory optimization method. Section 4 presents the modelling and solution methods for the train trajectory optimization problem. Section 5 illustrates the approach in case studies. Finally Section 6 ends the paper with conclusions.

2. Problem description

In this paper, energy-efficient timetabling is taken as a fine-tuning step after the time allowances have been set based on feasibility and robustness. As the three-level railway timetabling framework presented by Goverde et al. (2016), the timetable design can be done in three successive levels: microscopic, macroscopic, and a corridor fine-tuning level. Feasibility and stability are optimized at the microscopic level; efficiency and robustness are optimized at the macroscopic network level; and energy efficiency and robustness are optimized at the corridor fine-tuning level. The corridor fine-tuning level optimizes the speed profiles of all trains on each corridor while maintaining the scheduled event times at the corridor ends provided by the micro-macro model. The arrival and departure times at the intermediate stops are optimized for energy saving. The optimization must not influence the feasibility, stability, efficiency and robustness of the network which was optimized in the previous level.

A timetable example for a railway corridor is shown in Fig. 1 (a). The timetable lists the times when a service is scheduled to arrive at and depart from specified locations. The railway corridor may consist of single-track and/or double-track lines. Trains involved in this timetable are in the same and/or different directions. The stations in the downstream direction are named with successive numbers $1, 2, \ldots, Z$. Denote by $\mathbf{Z} = \{1, 2, ..., Z\}$ the set of stations where Z is the number of stations on this corridor. Denote by $\mathbf{I_{all}}$ the set of trains operating on this corridor, $\mathbf{I_d}$ the set of trains in the downstream direction from station 1 to Z and $\mathbf{I_u}$ the set of trains in the upstream direction from station Z to 1, and $\mathbf{Z}_i$ the stations on train $i \in \mathbf{I_{all}}$'s journey. Define the train sets $\mathbf{I_o} = \{(i,j)|\ i$ and j are from opposite directions, $i, j \in \mathbf{I_{all}}\}$ and $\mathbf{I_s} = \{(i,j)|\ i$ and j are in the same direction, $i, j \in \mathbf{I_{all}}\}$.

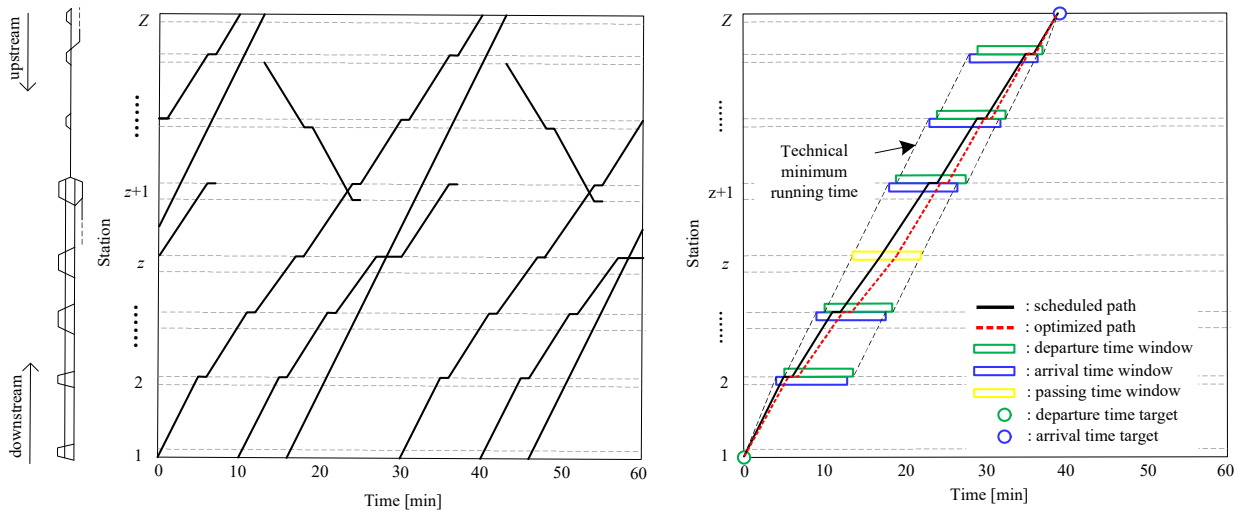

Figure 1: Examples of timetable, time targets and time windows.

To optimize the arrival and departure times at the intermediate stops for energy saving, an energy-efficient timetabling strategy is proposed. The strategy includes two steps:

Step 1: Relax the timetable by converting the time targets to time window constraints;

Step 2: Find optimal arrival/departure/passing-through (A/D/P) times and optimal energy-efficient speed profiles within those time windows.

The first step is to relax the given timetable. The time targets of all the trains in intermediate stations are converted into time window constraints, so that Step 2 can adjust the allocation of running time supplements. Trains' arrival and departure times at corridor ends maintain their original values. The lower bounds of the time windows represent the earliest possible A/D/P times assuming the train leaves the first station on the scheduled time. The upper bounds represent the latest possible A/D/P times in order to arrive at the last station on time. The lower bounds are computed with technical minimum running (dwell) times in the train's travel direction, while the upper bounds are computed with technical minimum running (dwell) times backwards. An example of time windows of a single train is shown in Fig. 1 (b).

The time windows and corresponding speed windows are formulated with a timetable constraint set (TCS), which is a sequence of time and speed constraints at timetable points (stations). We classify a train's event at a station into three types: arrival, departure and pass-through. Denote by e an event, $e \in \{a, d, p\}$ where a refers to arrival, d refers to departure, and p refers to pass-through. For train $i \in \mathbf{I_{all}}$, the TCS is written as

$$\mathbf{TCS}_i = \left\{ \left(k_{i,z,e}, [t^{\min}_{i,z,e}, t^{\max}_{i,z,e}], [v^{\min}_{i,z,e}, v^{\max}_{i,z,e}] \right) \right\}_{z \in \mathbf{Z}_i}, \tag{1}$$

where $k_{i,z,e}$ is the location for train i at station z of event e, and $t^{\min}_{i,z,e}$, $t^{\max}_{i,z,e}$, $v^{\min}_{i,z,e}$ and $v^{\max}_{i,z,e}$ are respectively the lower and upper bounds of time and speed for e at $k_{i,z,e}$. Define $\mathbf{K_{TCS}}_i$ as the set of timetable points of train i. If $k_{i,z,e}$ is a stop station, the speed window $[v^{\min}_{i,z,e}, v^{\max}_{i,z,e}] = [0, 0]$, and $k_{i,z,a} = k_{i,z,d}$. If $k_{i,z,e}$ is a pass-through station, $v^{\min}_{i,z,e}$ and $v^{\max}_{i,z,e}$ are respectively the lower and upper passing-through speeds. Albrecht et al. (2013) proposed a train path envelope (TPE) to describe the time and speed constraints at timetable points, signals, and conflict points. The TPE is allocated to one train only and the TPEs do not overlap each other, so that as long as a train stays in its TPE there will be no conflicts. In this paper we allow overlapping TCSs while the conflicts are resolved with the multi-train trajectory optimization algorithm (Section 3).

Step 2 is to find optimal energy-efficient A/D/P times and speed profiles within those time windows (An example of the optimal A/D/P times is presented with the red dashed lines in Fig. 1 (b)). The reason of keeping A/D/P times within the time windows produced by Step 1 is to avoid changing train sequences.

Consequently, the journey times and capacity usage are not influenced. A train trajectory optimization method (Section 3) is developed to find optimal energy-efficient A/D/P times and speed profiles.

3. Train trajectory optimization method

A single-train trajectory optimization (STTO) is developed firstly to find every single train's optimal energy-efficient A/D/P times and speed profiles within the train's TCS. In order to re-allocate the running times, the STTO no longer focuses on the train movements between two adjacent stops like the classical ETTO problem does (Khmelnitsky, 2000; Howlett, 2000; Liu and Golovitcher, 2003), but it concerns the whole corridor between the first and last station, which may contain a few intermediate stops and passing-through stations. Note that the TCSs, as well as the trajectories computed by the STTO, can not promise conflict-free train paths. Therefore it is necessary to detect conflicts.

The conflict detection is based on the blocking time theory (Hansen and Pachl, 2014). The time duration over which a block section is allocated exclusively to a specific train is called blocking time. The blocking time of a track section begins with issuing a train its movement authority for this section and ends after the train has completely left the section and all signalling appliances have been reset so that a movement authority can be issued to another train to enter the same section. The blocking time of a track is thus longer than the train occupies the section. During this blocking time the track is thus blocked for other trains. The conflict detection process first computes the successive blocking times for each train over a railway line. The visualization of these blocking times in a time-distance graph is called a blocking time diagram. Then conflicts are represented by overlapping blocking times in the time-distance domain (Goverde et al., 2013, 2016). If there are any overlaps in the blocking diagram, a multi-train trajectory optimization (MTTO) is applied to jointly compute the conflicting trains' trajectories. The MTTO eliminates conflicts between trains.

In summary, the train trajectory optimization method follows the three steps to find optimal A/D/P times and speed profiles:

Step 1: Compute the speed profiles and time-distance paths for every single train separately with the STTO;

Step 2: Conflict detection. If there is no conflict between the time-distance paths computed in Step 1, then take the time-distance paths from Step 1 as the output, otherwise go to Step 3;

Step 3: Compute the speed profiles and time-distance paths for the conflicting trains simultaneously with the MTTO, and output the computed time-distance paths.

The train trajectory optimization model is extended based on the the optimal control model presented in (Wang and Goverde, 2016a). Denote by $\mathbf{I}$ the set of trains which require speed profiles. If $\mathbf{I}$ only contains one train, the following model is the STTO model. If $\mathbf{I}$ contains more than one train, it becomes a MTTO model.

The cost function is to minimize the total energy costs of the trains in $\mathbf{I}$, that is

$$\text{minimize} \qquad J = \sum_{i \in \mathbf{I}} J_i, \quad J_i = \int_{K_i^0}^{K_i^f} f_i(s_i)\ ds_i, \tag{2}$$

subject to the operational constraints to every single train $i \in \mathbf{I}$

$$\begin{cases} \dfrac{dv_i(s_i)}{ds_i} = \dfrac{\theta_1(s_i) f_i(s_i) - (1-\theta_1(s_i)) b_i(s_i) - R_{train,i}(v_i) - R_{line,i}(s_i)}{\rho_i \cdot m_i \cdot v_i(s_i)}, \\ \dfrac{dt_i(s_i)}{ds_i} = \dfrac{1}{v_i(s_i)}, \end{cases} \tag{3}$$

$$\begin{cases} 0 \le f_i(s_i) \le F_i^{\max}, \\ 0 \le b_i(s_i) \le B_i^{\max}, \\ 0 \le f_i(s_i) \cdot v_i(s_i) \le P_i^{\max}, \\ 0 \le v_i(s_i) \le V_i^{\max}(s_i), \\ A_i^{\min} \le \dfrac{dv_i(s_i)}{dt_i(s_i)} \le A_i^{\max}, \end{cases} \tag{4}$$

$$\begin{cases} v_{i,z,e}^{\min} \le v_i(k_{i,z,e}) \le v_{i,z,e}^{\max}, \\ t_{i,z,e}^{\min} \le t_i(k_{i,z,e}) \le t_{i,z,e}^{\max}, \quad \forall\ z \in \mathbf{Z}_i,\ k_{i,z,e} \in \mathbf{K_{TCS}}_i, \end{cases} \tag{5}$$

$$D_{i,z}^{\min} \le t_i(k_{i,z,d}) - t_i(k_{i,z,a}) \le D_{i,z}^{\max}, \quad \forall\ z \in \mathbf{Z}_i,\ k_{i,z,a}, k_{i,z,d} \in \mathbf{K_{TCS}}_i. \tag{6}$$

$$\frac{t_i(k_{i,z,d})}{6} - \left\lfloor \frac{t_i(k_{i,z,d})}{6} \right\rfloor = 0, \quad \forall\ z \in \mathbf{Z}_i. \tag{7}$$

$$\begin{cases} \left(t_i(q) - t_j(q)\right)^2 \ge h_{i,j,q}^2, & \forall (i,j) \in \mathbf{I_s}, q \in \mathbf{Signal}_{i,j}, \\ \left(t_i(q) - t_j(q)\right)^2 \ge h_o^2, & \forall (i,j) \in \mathbf{I_o}, q \in \mathbf{StationSignal}_{i,j}, \\ \left(t_i(l_0) - t_j(l_0)\right)\left(t_i(l_f) - t_j(l_f)\right) \ge \sigma, & \forall (i,j) \in \mathbf{I_o}, [l_0, l_f] \in \mathbf{SingleTrack}_{i,j}. \end{cases} \tag{8}$$

The STTO model minimizes the cost function (2), subject to constraints (3)-(7), when $\mathbf{I}$ includes only one train. The MTTO model minimizes the cost function (2), subject to constraints (3)-(8) (or (9)), when $\mathbf{I}$ includes more than one train.

The cost function (2) reflects the objective of minimizing energy consumption, where K_i^0 and K_i^f are respectively the start and end positions of train i's journey, s_i is the traversed distance [m], and $f_i(s_i)$ is the traction force [kN]. In this work only the traction energy consumption is considered. For trains that use regenerative braking an additional term can be added to the integral, $J_i = \int_{K_i^0}^{K_i^f} \left(f_i(s_i) + \eta b_i(s_i)\right)\ ds_i$, where $\eta \in [0,1]$ is the recuperation coefficient which determines the efficiency of the regenerative braking system.

Equations (3) are the differential equations of train movements. Distance is adopted as the independent variable because gradients and speed limits occur as functions of distance rather than of time. In equations (3), $v_i(s_i)$ is the velocity of train i [m/s], ρ_i is the rotating mass factor, m_i is the mass of train i [t], $b_i(s_i)$ is the braking force of train i [kN], $R_{train,i}(v_i) = \alpha_i + \beta_i \cdot v_i + \gamma_i \cdot v_i^2$ is the train resistance force [kN] with coefficients α_i, β_i and γ_i, $R_{line,i}(s_i)$ is the line resistance force [kN], which is a function of position and consists of grade resistance and curve resistance, $t_i(s_i)$ is the traversed time of train i [s], and $\theta_1(s_i) \in \{0,1\}$ is a binary parameter. $\theta_1(s_i)$ is adopted to prevent traction and braking controls to be applied simultaneously.

Equations (4) are the path constraints of train i's vehicle performance parameters, speed limit and riding-comfort. $F_i^{\max}$, $B_i^{\max}$ and $P_i^{\max}$ are the upper bound of traction force, maximum braking force and maximum traction power of train i. The maximum traction force equals $\min\{F_i^{\max}, P_i^{\max}/v_i(s_i)\}$. $V_i^{\max}(s_i)$ is the speed limit at position s_i, including static and temporary speed restrictions. $A_i^{\min}$ and $A_i^{\max}$ are the lower and upper bound of acceptable acceleration.

Equations (5) are the event constraints which represent that the A/D/P times and speeds should respect the time and speed constraints of every timetable point in $\mathbf{K_{TCS}}_i$. Equation (6) represents the dwell time constraints. The dwell times are restricted by the minimal and maximal dwell times, where the minimal dwell times should provide sufficient time for boarding and alighting. In (6), $D_{i,z}^{\min}$ and $D_{i,z}^{\max}$ are the minimal and maximal dwell times of train i at station z. The minimal dwell time constraints are adopted to ensure enough alighting and boarding times for passengers. The maximal dwell time constraints are adopted to ensure efficient utilization of station platforms and avoid long waiting time. Equation (7) defines the departure times as integral multiples of 6 seconds (in some countries 30 seconds or 60 seconds) for more precise information to drivers and dispatchers. In Equation (7), the unit of $t_i(k_{i,z,d})$ is in seconds. $\lfloor \cdot \rfloor$ refers to rounding down to the nearest whole unit.

If $\mathbf{I}$ contains more than one train, Equations (8) are used by the MTTO model to avoid conflicts between trains. Equations (8) are based on practical operational constraints:

1. a time interval between every two adjacent trains in the same direction is required for safe train separation (Fig. 2 ①),

2. for opposite trains, there are two important types of headway requirements: (i) "depart-arrive" headway, the headway between the departure of a train and the arrival of an opposing train from the same line (Fig. 2 ②); and (ii) "arrive-depart" headway, the headway between the arrival of a train and the departure of an opposing train towards the same line (Fig. 2 ③), and
3. opposite trains must not travel on a single-track section between two adjacent stations at the same time, otherwise a head-on conflict occurs (Fig 2. ④).

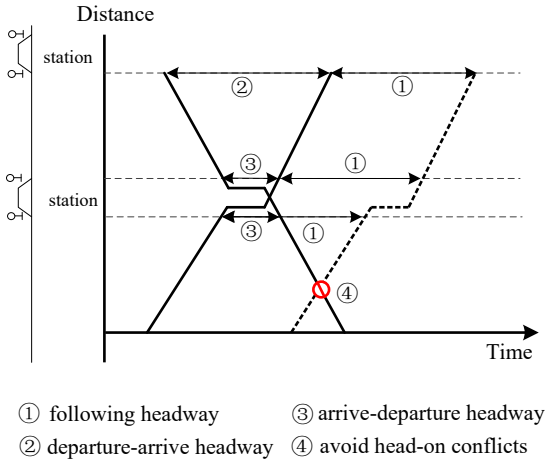

Figure 2: Necessary headways between trains.

In Equations (8), the first inequality represents the safe following headway constraint for trains in the same directions. Denote by $\mathbf{Signal}_{i,j}$ the set of signal locations on both train i and j's journeys for $i, j \in \mathbf{I_s}$, $t_i(q)$ and $t_j(q)$ refer to the time of train i and j passing through $q \in \mathbf{Signal}_{i,j}$, and $h_{i,j,q}$ is the minimum headway time between train i and j at location q. The left hand sides of the first two equations are squared to include the influence of train sequences. The second inequality is for safe "depart-arrive" or "arrive-depart" headway constraints. $\mathbf{StationSignal}_{i,j}$ is the set of signals at station boundaries on both train i and j's journeys for $i, j \in \mathbf{I_o}$. h_o is the minimum "depart-arrive" or "arrive-depart" headway time. Here "station signal" refers to a signal at station boundaries, either the inbound side or the outbound side. The point where the headway time applies is usually the station but we put the headway times at station signals. The reason is that trains may use different platforms which correspond to different phases in the multiple phase optimal control problem formulation, that is introduced later. We want to keep the headway constraints at one location where both i and j are passing through to simplify the formulation.

If the timetable is a periodic timetable, conflict-avoiding constraints shall take into account that the trains run every T_{period} (cycle time in seconds) (Hansen and Pachl, 2014). So that the first two equations in (8) become

$$\begin{cases} h_{i,j,q}^2 \leq \left(t_i(q) - t_j(q)\right)^2 \leq (T_{\text{period}} - h_{i,j,q})^2, \forall (i,j) \in \mathbf{I_s}, q \in \mathbf{Signal}_{i,j} \\ h_o^2 \leq \left(t_i(q) - t_j(q)\right)^2 \leq (T_{\text{period}} - h_o)^2, \forall (i,j) \in \mathbf{I_o}, q \in \mathbf{StationSignal}_{i,j}. \end{cases} \tag{9}$$

The third inequality in (8) is used to avoid any crossover of two trains' time-distance paths on single-track segments in order to avoid head-on conflicts between two opposite trains. Define any section between l_0 and l_f, which opposite trains are allowed to use, as a single-track. $\mathbf{SingleTrack}_{i,j}$ is the set of single-tracks on both train i and j's journeys for $i, j \in \mathbf{I_o}$. $[l_0, l_f]$ refers to a piece of single track (l_0 and l_f are the distances of the start and end locations along the route). $t_i(l_0)$, $t_i(l_f)$, $t_j(l_0)$ and $t_j(l_f)$ refer to the time of train i and j passing through l_0 and l_f, and σ is a positive value close to 0 representing a safe margin for train separation.

4. Modelling and solution methods

This section presents the modelling and solution methods for the MTTO.

4.1. Modelling

For the STTO modelling and solution methods see Wang and Goverde (2016a). The advantage of the MTTO presented in Section 3 is that it enables optimizing the energy consumption of multi-trains together with constraints to avoid conflicts between trains. However, it gives rise to three key questions.

First, Equations (3) are the differential equations of the train movements. Every train's traversed distance is taken as the independent variable of its own differential equations. However, the trains' traversed distances are not always the same since they are in different directions or follow different routes, as the example of red and blue routes show in Fig. 3. It is necessary to unify the independent variable in order to formulate train movements together.

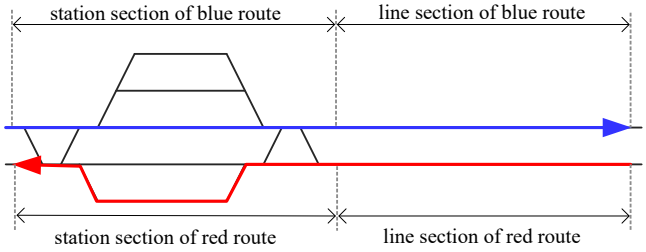

Figure 3: Example of two trains following different routes.

Second, trains are assigned different routes. Let "station sections" be sequences of connected blocks within a single station starting and ending at so-called station boundaries, and "line sections" be sequences of connected blocks within a single-track or double-track line (examples of station sections and line sections are shown in Fig. 3). Each train route is composed of a sequence of station sections and line sections, guiding the train through each station and each line. An accurate route model is the fundamental base of computing the train trajectories. However, trains are assigned different routes, while the lengths of different station sections within a station might be different, and the lengths of different line sections within a double-track line might also be different. The length differences make the multi-train trajectory optimization more complicated.

Last but not least, different trains have different speed limits and gradients since they travel on different routes, and the speed limits and gradients change along train routes. Besides, the MTTO model contains a lot of time and speed constraints at timetable points (in Equations (5)-(7)) and signals (in Equations (8)-(9)). How to deal with the varying speed limits and gradients and time and speed constraints at specific points is a difficult question.

This section provides solutions to these three questions. The independent variable is firstly unified, the section lengths are normalized, and the MTTO model is re-formulated as a multiple-phase optimal control problem to capture the varying speed limits and gradients and time and speed constraints.

Unification of the independent variable:

Take distance s as the independent variable, which increases in the downstream direction. Let

$$ds_i = ds, \quad ds_j = -ds, \quad \forall\, i \in \mathbf{I_d}, j \in \mathbf{I_u}.$$

$ds_j = -ds$ because train j is in the upstream direction.

Normalization of section lengths:

Our model takes into account the actual situation that trains use different routes and different tracks at stations, whilst the route lengths might be different. A normalization vector λ_i ($i \in \mathbf{I}$) is adopted for every station section or line section of train $i \in \mathbf{I}$ to normalize section lengths. With the introduction of λ_i, we scale all routes with respect to a fixed reference 'main route', so that all routes have unified and continuous kilometer points. λ_i is used in the formulation of cost function and dynamic constraints to scale the section lengths. The section lengths will be re-scaled again in the final solution for each train, so that the computed trajectories respect the actual route lengths.

The normalization follows 3 steps:

Step 1: Define station boundaries to divide the railway corridor into station regions and line regions.

Step 2: Define a main route in the downstream direction. The main route covers the whole corridor, and the kilometer points of the main route are continuous.

Step 3: Get the information of the lengths of every station section and line section of all the train routes and the main route, and compute λ_i for each section.

For a station section in train $i \in \mathbf{I}$'s route at station $z \in \mathbf{Z}_i$,

$$\lambda_{\text{Stationsec}_i^z} = \frac{L_{\text{Stationsec}_i^z}}{L_{\text{Stationsec}_{\text{main}}^z}},\ \forall\ i \in \mathbf{I}, z \in \mathbf{Z}_i, \tag{10}$$

where Stationsec_i^z refers to the station section, which train i uses at station z, $\text{Stationsec}_{\text{main}}^z$ refers to the station section on the main route at station j, $L_{\text{Stationsec}_i^z}$ and $L_{\text{Stationsec}_{\text{main}}^z}$ are respectively the lengths of Stationsec_i^z and $\text{Stationsec}_{\text{main}}^z$.

For a line section on train $i \in \mathbf{I}$'s route between adjacent stations z and $z+1$ ($z, z+1 \in \mathbf{Z}_i$),

$$\lambda_{\text{Linesec}_i^{z,z+1}} = \frac{L_{\text{Linesec}_i^{z,z+1}}}{L_{\text{Linesec}_{\text{main}}^{z,z+1}}},\ \forall\ i \in \mathbf{I}, z, z+1 \in \mathbf{Z}_i, \tag{11}$$

where $\text{Linesec}_i^{z,z+1}$ refers to the line section on train i's route between station z and $z+1$, $\text{Linesec}_{\text{main}}^{z,z+1}$ refers to the line section within the main route between station z and $z+1$, $L_{\text{Linesec}_i^{z,z+1}}$ and $L_{\text{Linesec}_{\text{main}}^{z,z+1}}$ are respectively the lengths of $\text{Linesec}_i^{z,z+1}$ and $\text{Linesec}_{\text{main}}^{z,z+1}$.

Multiple-phase optimal control problem formulation:

The multiple-phase optimal control problem divides the railway corridor into multiple segments. Each segment is a phase, where any particular phase has its own cost function (minimizing energy costs), dynamic model (train dynamic movement model), path constraints (vehicle characteristics, speed limits, and riding comfort), boundary conditions (time and speed constraints at timetable points), and event constraints (to avoid conflicts). The complete trajectory is then obtained by properly linking adjacent phases via linkage conditions (continuous speeds and times). The total cost function is the sum of the cost functions within each phase. The optimal trajectory is then found by minimizing the total cost functional subject to the constraints within each phase and the linkage constraints connecting adjacent phases. The advantages of this modelling method are multiple: it gives an accurate description of varying speed limits and gradients and the time and speed restrictions at timetable points.

Assume that the infrastructure data combine track information such as speed, gradient, the signalling system (location of signals) and operational information like routes and timetable points. The first step is to divide the railway corridor into several segments. The division-points are one of three types:

1. *Critical points of speed limits and gradients.* The points of changing speed limits or gradients are used to partition the whole corridor, so that each interval has a unique speed limit and gradient value.
2. *Timetable points* that have time or speed limitations indicated by every train's TCS.
3. *Signal positions in both directions.* The signaling system consists of a series of railway signals that divide a railway line into a series of sections, or "blocks", which are important elements in managing train movements. For example, each block can only be occupied by one train at a time.

Fig. 4 gives an example, where the corridor is divided into 16 segments. Each segment between two adjacent division-points is a phase for the multiple-phase optimal control problem. Within each phase, the gradient and speed limit are constant, but their values may be different from the ones in other phases. The boundary points of a phase also might be a timetable point or a signal, where time and/or speed restrictions apply on the train operations.

Denote by $r \in \mathbf{R} = \{1, \ldots, R\}$ a phase, R is the number of phases of this railway corridor, $s_0^{(r)}$ the initial location of phase r, and $s_f^{(r)}$ the terminal location of phase r, $s_0^{(r)} < s_f^{(r)}$. Define $\mathbf{A} = \left\{(m,n) | s_f^{(m)} = s_0^{(n)}, m, n \in \mathbf{R}\right\}$

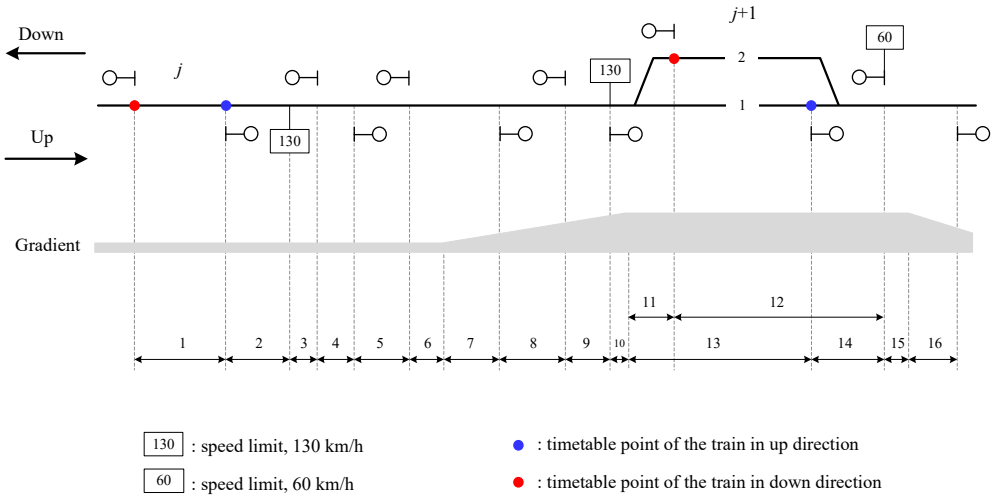

Figure 4: Example of partitioning in phases.

as the set of adjacent phases, and $\mathbf{I}^{(r)}$ as the set of trains passing through phase r, $\mathbf{I}^{(r)} \subseteq \mathbf{I}$. Let $v_i^{(r)}(s)$, $t_i^{(r)}(s)$ be train i's speed and time in phase r, $f_i^{(r)}(s)$, $b_i^{(r)}(s)$ be train i's traction and braking forces in phase r, $V_i^{\max,(r)}(s)$ be the speed limit of train i in phase r, and $R_{line,i}^{(r)}(s)$ be the line resistance caused by the constant gradient within phase r. ξ_i and $\lambda_i^{(r)}$ are introduced for later mathematical formulations. $\xi_i = 0$ if $i \in \mathbf{I_d}$, otherwise $\xi_i = 1$, and $\lambda_i^{(r)}$ equals the value of λ_i in phase r.

The MTTO model is re-formulated as a multiple-phase optimal control problem:

$$\begin{aligned} \text{Minimize} \quad J = \sum_{r=1}^{R} J^{(r)}, J^{(r)} = \sum_{i \in \mathbf{I}^{(r)}} J_i^{(r)}, \\ J_i^{(r)} = \lambda_i^{(r)} \int_{s_0^{(r)}}^{s_f^{(r)}} f_i^{(r)}(s)\, ds, \end{aligned} \tag{12}$$

subject to the dynamic constraints:

$$\begin{cases} \dfrac{dv_i^{(r)}(s)}{ds} = (-1)^{\xi_i} \lambda_i^{(r)} \dfrac{\theta_1(s_i) f_i^{(r)}(s) - (1-\theta_1(s_i)) b_i^{(r)}(s) - R_{train,i}(v_i^{(r)}) - R_{line,i}^{(r)}(s)}{\rho_i \cdot m_i \cdot v_i^{(r)}(s)}, \\ \dfrac{dt_i^{(r)}(s)}{ds} = (-1)^{\xi_i} \lambda_i^{(r)} \dfrac{1}{v_i^{(r)}(s)}, \end{cases} \quad \forall \ r \in \mathbf{R}, i \in \mathbf{I}^{(r)}, \tag{13}$$

the path constraints:

$$\begin{cases} 0 \le f_i^{(r)}(s) \le F_i^{\max}, \\ 0 \le b_i^{(r)}(s) \le B_i^{\max}, \\ 0 \le f_i^{(r)}(s) \cdot v_i^{(r)}(s) \le P_i^{\max}, \\ 0 \le v_i^{(r)}(s) \le V_i^{\max,(r)}, \\ A_i^{\min} \le \dfrac{dv_i^{(r)}(s)}{dt_i^{(r)}(s)} \le A_i^{\max}, \end{cases} \quad \forall\ r \in \mathbf{R}, i \in \mathbf{I}^{(r)}, \tag{14}$$

the boundary conditions (if any):

$$\begin{cases} v_{i,z,e}^{\min} \le v_i^{(r)}(k_{i,z,e}) \le v_{i,z,e}^{\max}, \\ t_{i,z,e}^{\min} \le t_i^{(r)}(k_{i,z,e}) \le t_{i,z,e}^{\max}, \\ \frac{t_i^{(r)}(k_{i,z,d})}{6} - \left\lfloor \frac{t_i^{(r)}(k_{i,z,d})}{6} \right\rfloor = 0, \end{cases} \quad \forall\, r \in \mathbf{R}, i \in \mathbf{I}^{(r)}, k_{i,z,e} \in \mathbf{K}_{\mathbf{TCS}_i} \cap \{s_0^{(r)}, s_f^{(r)}\}, \tag{15}$$

the event constraints (if any):

$$\begin{cases} \left(t_i^{(r)}(q) - t_j^{(r)}(q)\right)^2 \ge h_{i,j,q}^2, \\ \qquad\qquad \forall\, r \in \mathbf{R}, (i,j) \in \mathbf{I_s} : i,j \in \mathbf{I}^{(r)}, q \in \mathbf{Signal}_{i,j} \cap \{s_0^{(r)}, s_f^{(r)}\}, \\ \left(t_i^{(r)}(q) - t_j^{(r)}(q)\right)^2 \ge h_o^2, \\ \qquad\qquad \forall\, r \in \mathbf{R}, (i,j) \in \mathbf{I_o} : i,j \in \mathbf{I}^{(r)}, q \in \mathbf{StationSignal}_{i,j} \cap \{s_0^{(r)}, s_f^{(r)}\}, \\ \left(t_i^{(r)}(s_0^{(r)}) - t_j^{(r)}(s_0^{(r)})\right)\left(t_i^{(r)}(s_f^{(r)}) - t_j^{(r)}(s_f^{(r)})\right) \ge \sigma, \\ \qquad\qquad \forall\, r \in \mathbf{R}, (i,j) \in \mathbf{I_o} : i,j \in \mathbf{I}^{(r)}, [s_0^{(r)}, s_f^{(r)}] \in \mathbf{SingleTrack}_{i,j}, \end{cases} \tag{16}$$

and the linkage conditions of all adjacent phases:

$$\begin{cases} v_i^{(m)}(s_f^{(m)}) - v_i^{(n)}(s_0^{(n)}) = 0, \\ D_{i,s_f^{(m)}}^{\min} \le t_i^{(n)}(s_0^{(n)}) - t_i^{(m)}(s_f^{(m)}) \le D_{i,s_f^{(m)}}^{\max}, \end{cases} \quad \forall (m,n) \in \mathbf{A}, i \in \mathbf{I}^{(m)} \cap \mathbf{I}^{(n)}. \tag{17}$$

The cost function (12) aims at minimizing the cost functions over all phases, $J^{(r)}$ is the sum of the cost function(s) of the train(s) going through phase r. A single train's cost function $J_i^{(r)}$ aims at minimizing energy consumption within phase r. Each phase $r \in \mathbf{R}$ adopts the dynamic constraints (13) to represent the dynamic movements of train(s) going through segment $[s_0^{(r)}, s_f^{(r)}]$. The dynamic constraints are stated with s as the independent variable. $(-1)^{\xi_i}$ is adopted to eliminate the influence of travelling directions. $\lambda_i^{(r)}$ is adopted for the normalization of section lengths. Path constraints (14) are used to represent the operational constraints of vehicle characteristics, speed limits and riding comfort. Inequalities (15) work as the boundary conditions of phase r if $k_{i,z,e} \in \mathbf{K}_{\mathbf{TCS}_i}$ is the initial or terminal point of that phase. Timetable points are adopted as division-points, so phase r's initial or terminal point might be a timetable point, where time and speed restrictions on train operations apply. For a phase that has more than one train going through, event constraints (16) are required to avoid conflicts between trains. They are developed based on (8) or (9).

The linkage conditions (17) are to make sure that the train's speed-distance and time-distance trajectories are continuous. If the linkage point of two successive phases is a stop point of a train i, $t_i^{(m)}(s_f^{(m)})$ and $t_i^{(n)}(s_0^{(n)})$ represent the arrival and departure times of train i at $s_f^{(m)}$, and $D_{i,s_f^{(m)}}^{\min}, D_{i,s_f^{(m)}}^{\max}$ are the lower bound and upper bound of dwell times of train i at $s_f^{(m)}$ and $s_0^{(n)}$. Otherwise, $D_{i,s_f^{(m)}}^{\min} = D_{i,s_f^{(m)}}^{\max} = 0$.

4.2. Pseudospectral method

GPOPS (Rao et al., 2010) is adopted to solve the multiple-phase optimal control problem. GPOPS is a MATLAB software program for solving multiple-phase optimal control problems using the pseudospectral method. In general, pseudospectral methods transcribe the continuous-time optimal control problem into a nonlinear programming problem, after which a nonlinear programming solver is adopted to directly solve the problem. GPOPS uses the Radau pseudospectral method (Garg, 2011), which takes the Legendre-Gauss-Radau (LGR) points for collocation of the dynamic constraints, and for quadrature approximation of the integrated Lagrange cost term. The Lagrange polynomial approximation of the state, however, uses the LGR points plus the final point. In addition, GPOPS offers a function that implements an hp-adaptive mesh refinement algorithm that iteratively determines a mesh that accurately distributes the collocation points. GPOPS transcribes the continuous-time multiple-phase optimal control problem into a discrete NLP problem. The resulting NLP is then solved by SNOPT (Gill et al., 2005). For detailed mathematical descriptions, we recommend Wang and Goverde (2016a); Rao et al. (2010); Ye and Liu (2016).

5. Case studies

5.1. Case descriptions

The proposed approach is applied in case studies on a Dutch single-track corridor and a double-track corridor. The infrastructure data is provided by the Dutch infrastructure manager Prorail. The infrastructure characteristics consist of a description of all track sections, points, speed signs and signals over the entire track layout. Since we dont have gradient data, we assume flat tracks. The single-track rail corridor is about 20 km long, between Den Helder (Hdr) and Schagen (Sgn) in the north of the Netherlands, see Fig. 5 (a). There are 4 stations along this corridor: Den Helder (Hdr), Den Helder Zuid (Hdrz), Anna Paulowna (Ana), and Schagen (Sgn). The double-track corridor is between Utrecht-'s-Hertogenbosch, which is one of the busies rail lines in the Netherlands. This railway corridor is about 50 km long, including two main tracks, divided into one long corridor for each traffic direction, 8 passenger stations: Utrecht (Ut), Utrecht Lunetten (Utl), Houten (Htn), Houten Castellum (Htnc), Culemborg (Cl), Geldermalsen (Gdm), Zaltbommel (Zbm) and 's-Hertogenbosch (Ht) (Fig. 5 (b)).

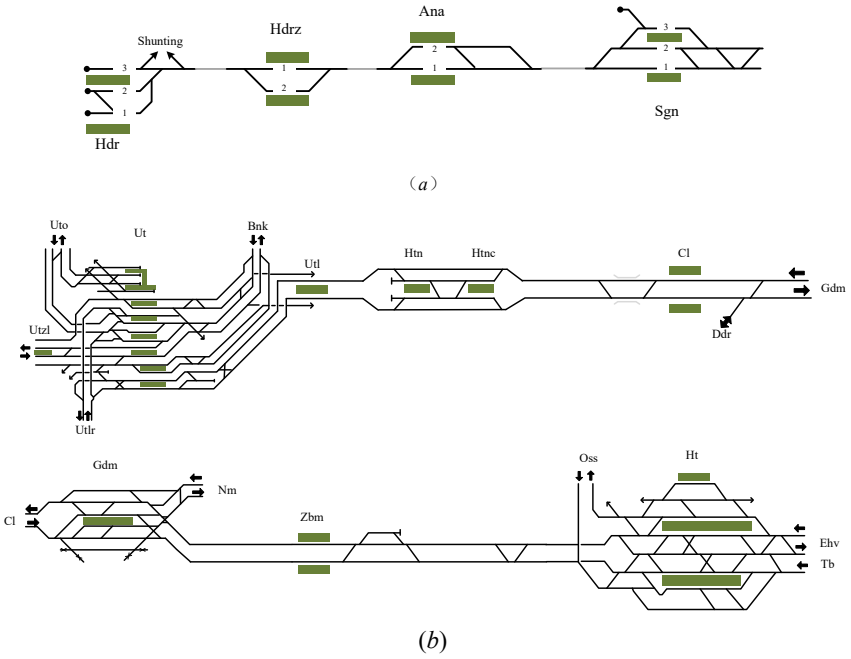

Figure 5: The partial single-track and double-track corridors used in the case studies.

Fig. 6 shows the pre-defined timetables for the single-track and the double-track corridor (one direction from Ht to Ut). The time-distance diagrams are based on the timetables in use in 2016, which are periodic timetables with train services repeating every half hour. The trains in black solid lines in Fig. 6 are chosen for the case studies as the traffic within half hour timetables. On the double-track corridor, the trains in different directions are independent from each other, so only the trains from Ht to Ut are taken into account. By optimizing the six trains' time paths, the timetables for the whole day are optimized. To simplify the illustration, we name the trains as T_1 (regional train from Hdr to Sgn), T_2 (regional train from Sgn to Hdr), T_3 (regional train from Gdm to Ut), T_4 (regional train from Ht to Ut), T_5 (intercity train from Ht to Ut) and T_6 (intercity train from Ht to Ut), as shown in Fig. 6. The light blue rectangles in Fig. 6 refer to multi-track lines, where the trains use different tracks. Trains use the same tracks on the other regions. On the single-track corridor between Hdr and Sgn, T_1 and T_2 operate in different directions and the two trains are scheduled to meet at Ana station. The double track corridor has two intercity trains (T_5 and T_6) and two regional trains (T_3 and T_4) running in the direction from Ht to Ut, and T_5 overtakes T_4 at Gdm station.

The static parameters of the regional and intercity trains are listed in Table 1. The traction force and train resistance curves are shown in Fig. 7. Since the braking rates are the only accessible data characterizing the braking behavior, we let the braking force be equal to the braking rate times train mass.

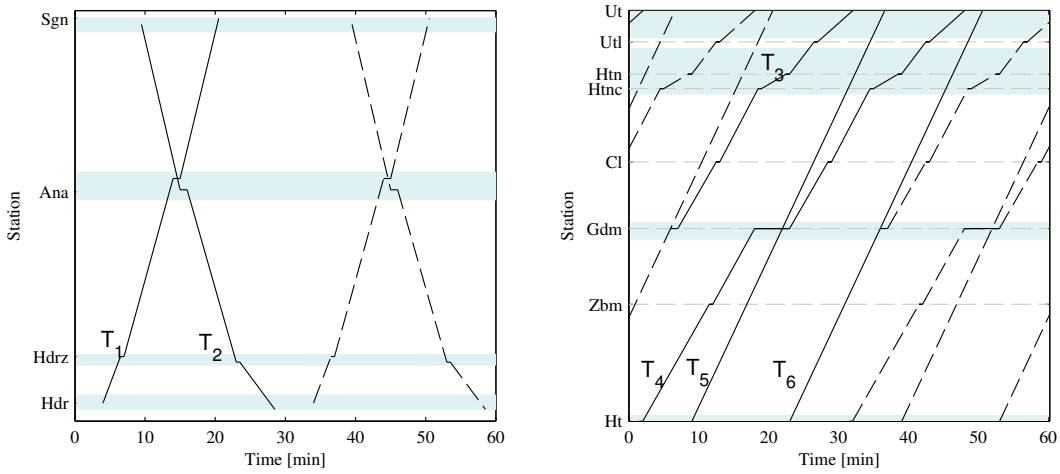

Figure 6: Original timetables of single-track and double track corridors.

Table 1: Basic parameters of Regional and Intercity train.

Property	Value	
	Regional	Intercity
Train mass [t]	220	391
Rotating mass factor [-]	1.06	1.06
Maximum traction power [kW]	1918	2157
Maximum traction force [kN]	170	214
Maximum braking deceleration [m/s^2]	-0.8	-0.66

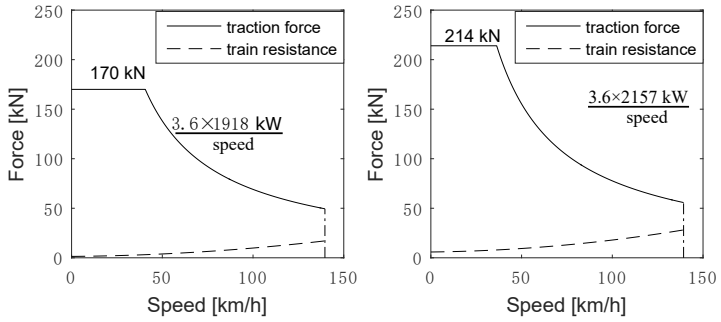

Figure 7: Traction force and line resistance of Regional and Intercity train.

The proposed methods were tested in two scenarios. The first is the scheduled scenario. The STTO is adopted to compute the speed profiles of the six trains separately, which assumes the trains operate according to the given timetables. Arrival and departure target points constraints are used to model the trains' punctuality. The second scenario is the energy-efficient timetable adjustment case. The A/D/P time target constraints are converted to time window constraints first, then the STTO is adopted to compute the speed profiles and time paths of the six train separately. After that, the conflicts are detected by checking the overlaps on blocking diagrams. If there are conflicts between the train paths, the MTTO method is adopted to re-compute the speed profiles and time-distance paths of the conflicting trains.

The timetables we study are periodic timetables, so Equations (9) are used by the MTTO model to ensure enough headways between trains. The parameters in Equations (8) and (9) were set as $T_{\text{period}} = 1800$ s, $h_{i,j,q} = 120$ s, $h_o = 15$ s, and $\sigma = 5$ s. The value of $h_{i,j,q}$ is set according to the network statement 2017 by Prorail (ProRail, 2017). We use default headway norms that should be sufficient for the final timetable to be conflict-free and contain some buffer. This can be checked by computing the blocking time diagram for the final solution to prove that the timetable is conflict-free, or if not to increase some headway times in the same manner as presented in Bešinović et al. (2016). The value of h_o ensures enough time interval for signal clearing and setting. σ is a positive value to avoid any crossover of opposite trains' path. The dwell time windows at the meeting station, Ana, and the overtaking station, Gdm, are respectively [1 min, 2 min], and [3 min, 6 min]. The minimal and maximal dwell times at other stations are 0.5 min and 1 min.

The experiments were carried out with GPOPS 4.1 on a laptop equipped with a 3.2 GHz Pentium R processor. As for the parameter settings of GPOPS, we used 'complex' as the string to indicate the differentiation method, a tolerance of 1e-3, an iteration of 2, and didn't use autoscaling. The explanation of the parameters can be found in (Rao et al., 2011).

5.2. Results and analyses

The TCS time windows of train T_1, T_2, T_3 and T_4 are presented in Fig. 8. The TCSs of T_5 and T_6 are composed by scheduled departure and arrival times. The running times of T_5 and T_6 between Ht and Ut are the same as the ones provided by the original timetable. Fig. 9, 11, 13 and Table 2 present the optimized results for the two trains, T_1 and T_2, on the single-track corridor. Fig. 10, 12, 14, and Table 3 present the optimized results for the four trains, T_3, T_4, T_5 and T_6, on the double-track corridor.

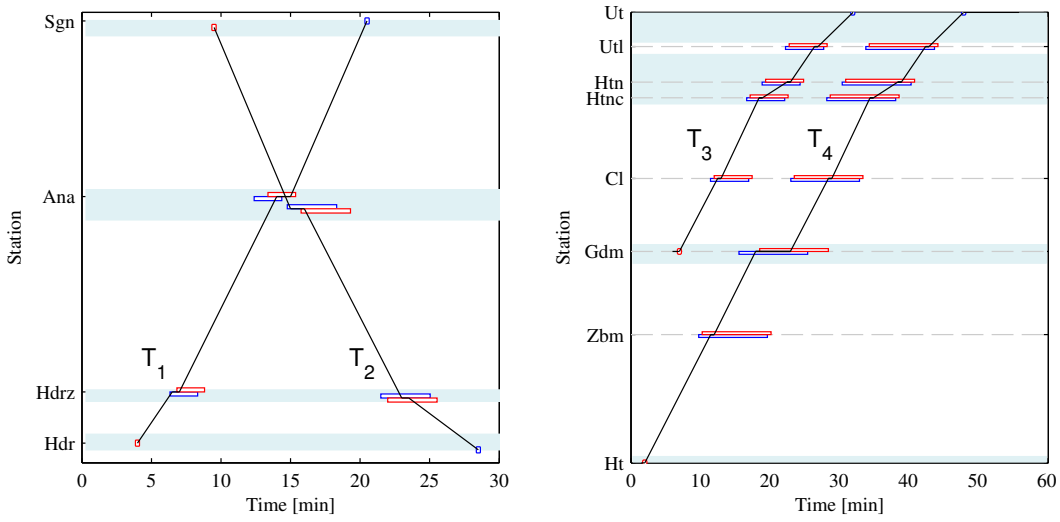

Figure 8: TCS time windows of train T_1, T_2, T_3 and T_4 (red rectangles refer to departure time windows, the blue rectangles refer to arrival time windows).

5.2.1. Timetable analysis

Table 2 and 3 present the arrival and departure times of the original timetables and the optimized timetables. The optimized arrival and departure times are slightly different from the ones indicated by the

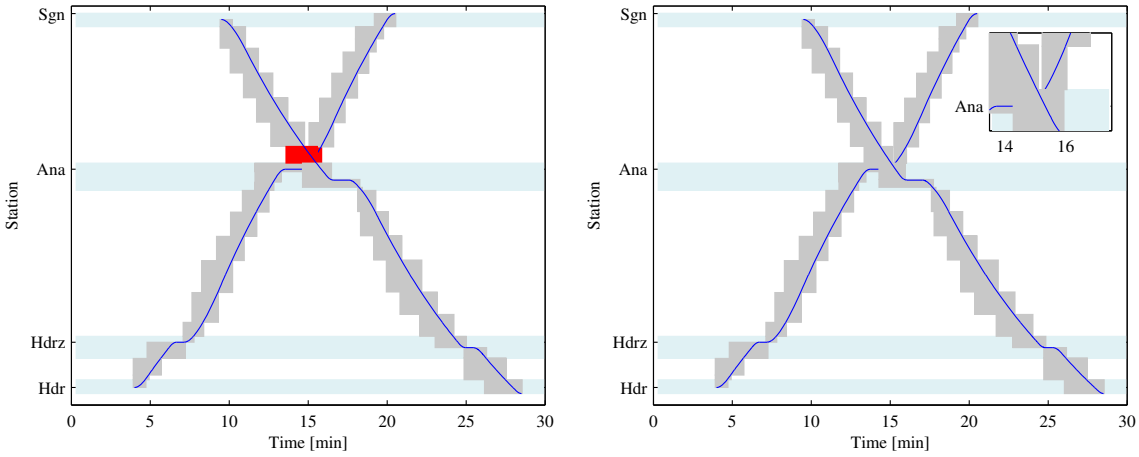

Figure 9: Blocking time diagrams of train T_1 and T_2 computed by STTO (left) and MTTO (right) (red refers to conflicts).

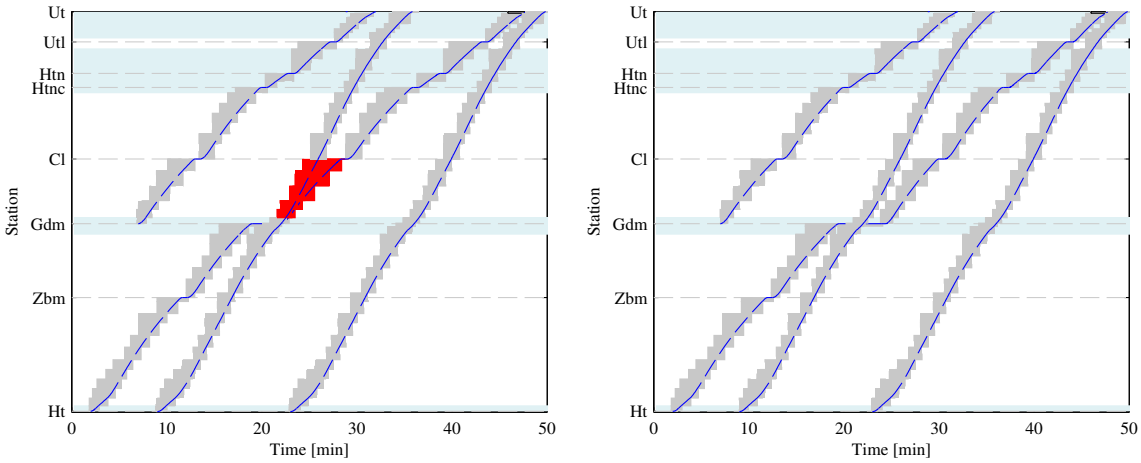

Figure 10: Blocking time diagrams of train T_3, T_4, T_5 and T_6 computed by STTO (left) and MTTO (right) (red refers to conflicts).

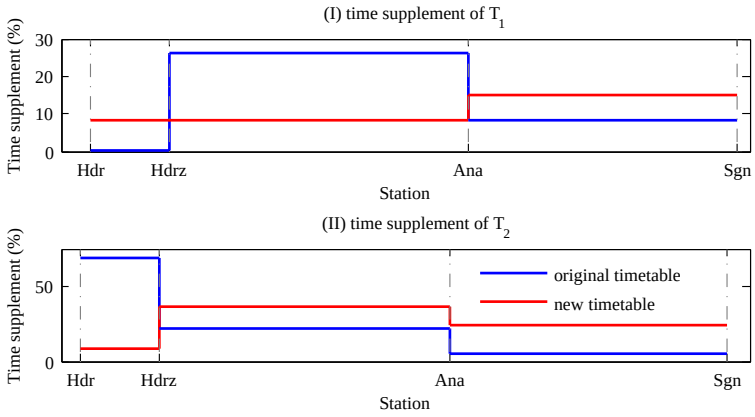

Figure 11: Time supplements of the original timetable and optimized timetable for trains on the single-track corridor.

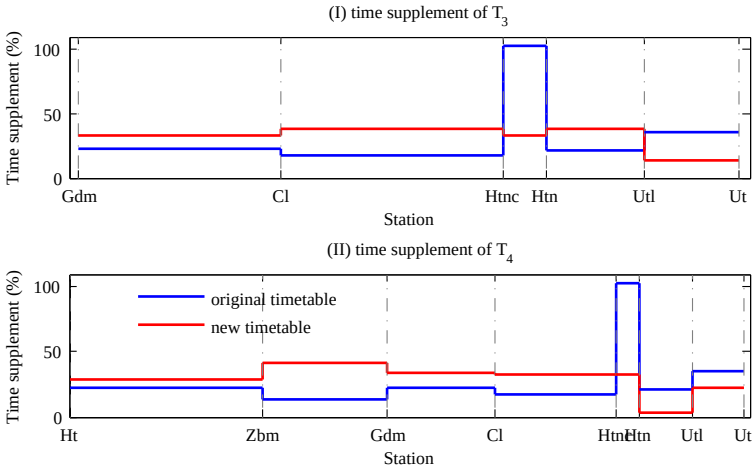

Figure 12: Time supplements of the original timetable and optimized timetable for trains on the double-track corridor.

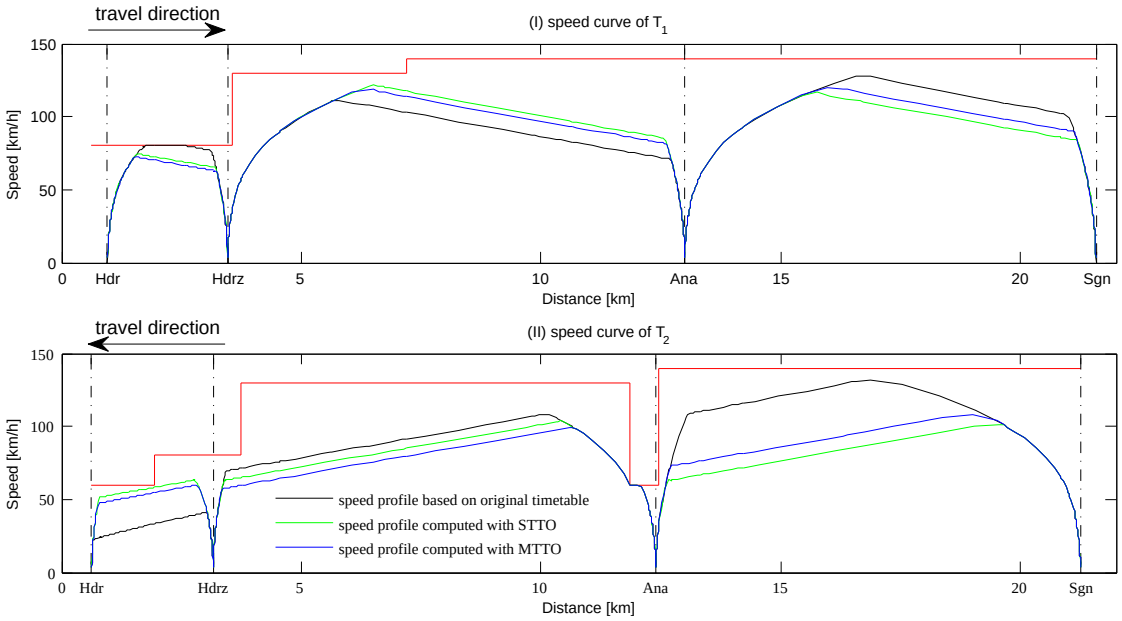

Figure 13: Speed profiles for trains on the single-track corridor (T_1 runs from Hdr to Sgn and T_2 goes from Sgn to Hdr.).

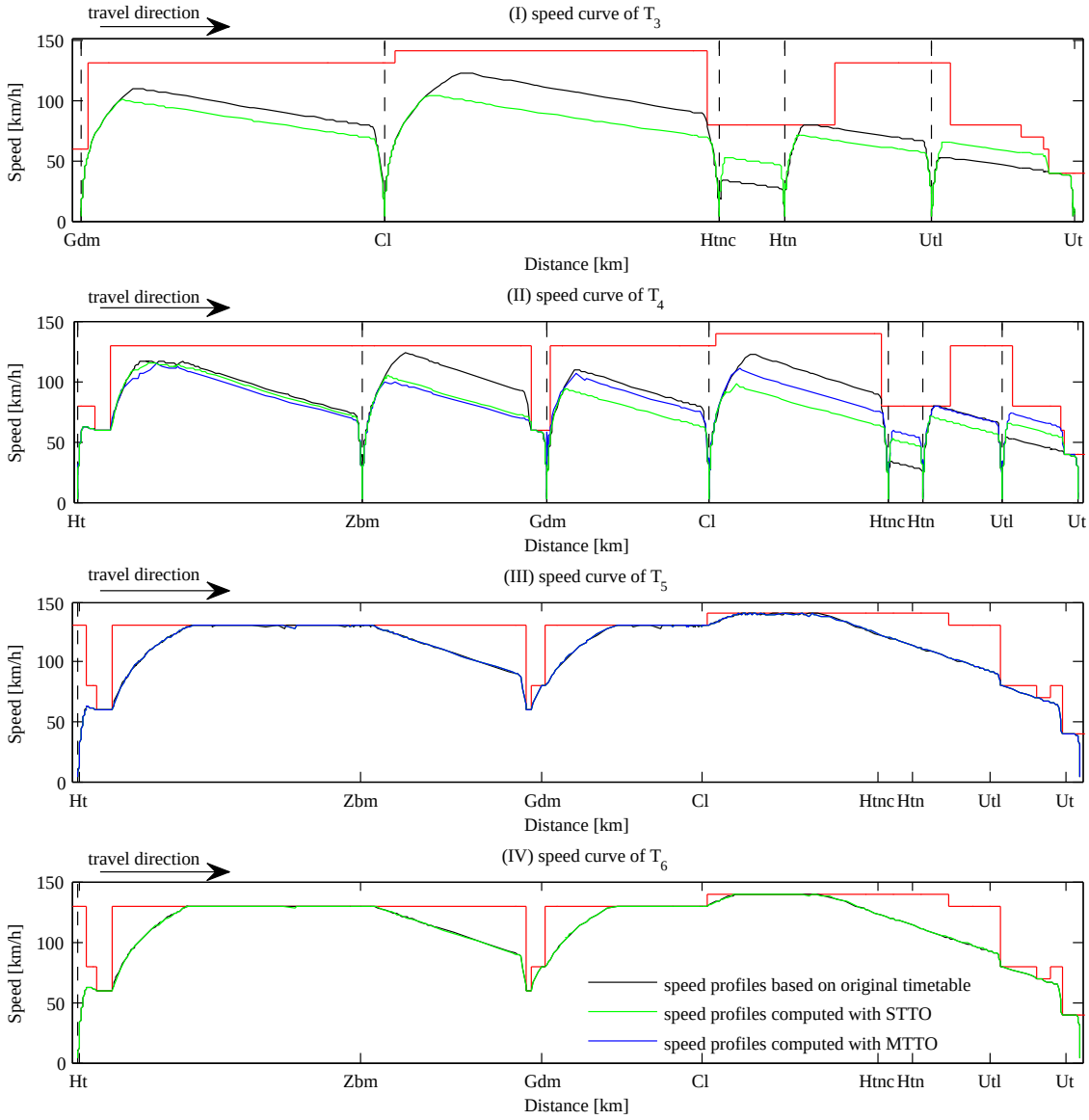

Figure 14: Speed profiles for trains on the double-track corridor.

Table 2: The original and optimized timetables, and their computation times and energy consumptions for the single-track corridor.

	Train	Event	Hdr	Hdrz	Ana	Sgn	Computation time [s]	Number of Phases/ Variables/Constraints	Energy consumption [J]	Energy saving	Conflict
Original timetable	T_1	A	–	06:30	14:00	20:30	–	20/925/701	6.9378×10^5	–	no
		D	04:00	07:00	15:00	–					
	T_2	A	28:30	23:00	15:00	–	–	22/1077/815	5.4670×10^5	–	no
		D	–	23:30	16:00	09:30					
STTO	T_1	A	–	06:42	13:30	20:30	28.61	20/925/701	6.4663×10^5	6.80%	yes
		D	04:00	07:12	14:30	–					
	T_2	A	28:30	25:00	16:36	–	13.93	22/1077/815	4.1349×10^5	24.37%	yes
		D	–	25:30	17:36	09:30					
MTTO	T_1	A	–	06:42	13:42	20:30	81.12	25/1800/1353	6.4364×10^5	7.23%	no
		D	04:00	07:12	14:42	–					
	T_2	A	28:30	24:48	16:00	–			4.1366×10^5	24.34%	no
		D	–	25:18	17:00	09:30					

*A: arrival time, [mm:ss]; D: departure time, [mm:ss].

Table 3: The original and optimized timetables, and their computation times and energy consumptions for the double-track corridor.

	Train	Event	Ht	Zbm	Gdm	Cl	Htnc	Htn	Utl	Ut	Computation time [s]	Number of Phases/ Variables/Constraints	Energy consumption [J]	Energy saving
Original timetable	T_3	A	–	–	–	12:30	18:30	22:30	26:30	32:00	–	76/740/558	3.9658×10^5	–
		D	–	–	07:00	13:00	19:00	23:00	27:00	–				
	T_4	A	–	11:30	18:00	28:30	34:30	38:30	42:30	48:00	–	128/999/754	7.3417×10^5	–
		D	02:00	12:00	23:00	29:00	35:00	39:00	43:00	–				
	T_5	A	–	–	–	–	–	–	–	37:00	–	135/1295/978	1.2394×10^6	–
		D	09:00	–	–	–	–	–	–	–				
	T_6	A	–	–	–	–	–	–	–	51:00	–	135/1295/978	1.2391×10^6	–
		D	23:00	–	–	–	–	–	–	–				
STTO	T_3	A	–	–	–	13:00	20:00	22:48	27:18	32:00	18.79	76/740/558	3.2810×10^5	17.27%
		D	–	–	07:00	13:30	20:30	23:18	27:48	–				
	T_4	A	–	11:42	19:00	28:30	36:00	38:48	43:12	48:00	49.76	128/999/754	5.7409×10^5	21.80%
		D	02:00	12:12	22:00	29:00	36:30	39:18	43:42	–				
	T_5	A	–	–	–	–	–	–	–	37:00	31.31	135/1295/978	1.2394×10^6	0%
		D	09:00	–	–	–	–	–	–	–				
	T_6	A	–	–	–	–	–	–	–	51:00	32.83	135/1295/978	1.2391×10^6	0%
		D	23:00	–	–	–	–	–	–	–				
MTTO	T_4	A	–	12:00	20:00	30:00	36:42	39:30	43:00	48:00	256.13	156/6811/5914	6.3550×10^5	13.44%
		D	02:00	12:30	24:00	30:30	37:12	40:00	43:30	–				
	T_5	A	–	–	–	–	–	–	–	37:00			1.2394×10^6	0%
		D	09:00	–	–	–	–	–	–	–				

*A: arrival time, [mm:ss]; D: departure time, [mm:ss].

original timetables. The departure times are integral multiples of 6 seconds because of the restriction of Equation (7). The optimized dwell times of train T_1 and T_2 at Ana are 1 minute. The optimized dwell times of T_4 at Gdm are respectively 3 minutes by using STTO, and 4 minutes by using MTTO. In addition to that, the optimized dwell times at the other stations of all trains are 0.5 minutes. The results show that short dwell times are beneficial for energy-efficiency. The MTTO results 1 minute longer dwell time of tain T_4 at Gdm to satisfy the constraint (8) and to eliminate conflicts between T_4 and T_5, see Fig. 10.

5.2.2. Blocking time diagram analysis

Fig. 9 and 10 present the blocking time diagrams to check conflicts between trains. The blocking time diagrams were computed according to the optimized results using STTO (left) or MTTO (right). Both in Fig. 9 and 10, the light-blue rectangles represent station areas with multiple tracks. The blocking time diagrams are only allowed to intersect within the station areas.

In the left plot of Fig. 9, T_1 and T_2's blocking times intersect at the single-track line between Sgn and Ana, which is not feasible. Therefore the MTTO method is adopted to compute new conflict-free time paths. The right plot in Fig. 9 shows that the blocking time diagrams for T_1 and T_2 overlap at station Ana, which has two parallel tracks. T_1 uses track 1 while the opposing train uses track 2, so the conflict is eliminated with the optimized results. This proves that the timetable produced by the MTTO method is conflict-free.

In Fig. 10, the left plot presents an overlap of the blocking times of train T_4 and T_5 on the section between Gdm and Cl. Both T_4 and T_5 leave Gdm at 00:22:00, thus there is no sufficient headway between the moments that the trains start using the section between Gdm and Cl. The results by using the MTTO method are shown in the right plot of Fig. 10, where train T_4 leaves Gdm at 00:24:00, causing no conflicts with T_5.

5.2.3. Time supplement and energy consumption analysis

Fig. 11 and 12 present the time supplements of the original and optimized timetables for T_1, T_2, T_3 and T_4. The time supplements of T_5 and T_6 are not discussed herein, because their running times between Ht and Ut are the same as the ones provided by the original timetable. It can be seen that the time supplements are quite high. In general, the time supplement is 5%-20% depending on timetable constraints. In practice, it is possible to have high time supplements, especially on short distances where a minute more is large in percentage, such as nearly 100% time supplements of T_4 between Htnc and Htn in the original timetable. In practice there will be less time supplement, since in this case study we did not consider the impact of gradients (because we do not have the data). This is fine for demonstrating the model of this paper but for practical results the gradient profile needs to be considered as well. (Note that varying gradients can be included in the method of multiple-phase optimal control model, see Wang and Goverde (2016a).). Most importantly, it can be observed that the running time supplements are re-allocated. The optimized results show more reasonable time supplements along the routes.

The re-allocation of time supplements has a direction influence on speed profiles and energy consumptions. Trains coast more on the sections with more running time supplements. As reported in Table 2 and 3, the energy consumption of trains T_1-T_4 are decreased by 7.23%, 24.34%, 17.27% and 13.44% (The energy savings are calculated by dividing the saved energy of the optimized timetable by the energy consumption of the original timetable.). In addition, the optimized dwell time of train T_4 at Gdm is 4 minutes, so that T_4 got 1 minute more running time compared to the original timetable. The extra running time is beneficial to energy saving. The STTO is able to save more energy consumption than the MTTO model, however, computing the trajectories separately may cause conflicts. The MTTO saves less energy consumption but eliminates the conflicts.

5.2.4. Speed profile analysis

Fig. 13 shows the optimized speed profiles of train T_1 from left to right, and train T_2 from right to left. Fig. 14 shows the speed profiles of T_3, T_4, T_5 and T_6. In Fig. 13 and 14, the red horizontal lines represent the static speed limits, the black lines refer to the energy-efficient speed profiles based on the original timetables, the green lines represent the optimized speed profiles within TCS windows computed with the STTO model, and the blue lines represent the optimized speed profiles computed with the MTTO model.

The trains stop at every planned stop point, which includes not only the first and last stations of the corridor, but also the intermediate stops, such as Hdrz, Ana (for train T_1 and T_2), Zbm, Gdm, Cl, Htnc,

Htn, and Utl (for train T_3, T_4, T_5 and T_6). It shows that both STTO and MTTO can handle time and speed constraints at intermediate stations. The MTTO method is capable of optimizing the trajectories of trains on different directions and different routes. The optimal speed profiles, computed by both STTO and MTTO (as shown in Figs. 13 and 14), satisfy the energy-efficient train control theory by the application of PMP (Howlett and Pudney, 1995): the control regimes between every two stops include using maximum traction during the outbound processes, maximum braking force for the inbound processes, cruising at maximum speeds, and coasting before braking for energy saving. Train speeds do not exceed the static varying speed limits. The trains are capable of accelerating after entering high speed regions and decelerating to low speeds before getting in low speed areas. Fig. 14 (II) shows an exception: the cruising regimes are approximated by alternating traction and coasting phases. Cruising is a singular arc in the optimal control structure where a unique control solution does not exist. GPOPS does not see the difference between cruising at constant speed and the approximation with traction-coasting pairs, where it favours the latter. The average speed and energy consumption are however the same.

5.2.5. Computation time analysis

The computation times, as well as the number of phases, variables and constraints are reported in Tables 2 and 3. The STTO is able to compute trajectories for each train within a short time. For instance, it takes less than 1 minute to compute a trajectory for T_3-T_6 running on a 50 km long corridor between Ht and Ut. The MTTO causes longer computation times compared to the STTO. It is because the MTTO model includes more phases, variables, and constraints than the STTO model since the MTTO model takes into account more than one train. To apply the MTTO method in even larger problems (more trains, longer corridors), our method can provide solutions with longer computation times. If this leads to computational problems some intermediate event times can be fixed, which then decomposes the problem in parts. If no natural fixing point exists this can be embedded into an iterative scheme.

6. Conclusions

This paper proposed a novel energy-efficient timetabling strategy. This strategy is achieved by adjusting the arrival and departure times of an existing timetable to improve its energy efficiency using train trajectory optimization methods. The train trajectory optimization methods proposed by us integrate two types, a single-train trajectory optimization, and a multi-train trajectory optimization.

The STTO method optimizes every single train's timetable by re-allocating their time supplements separately. It improves the energy efficiency of the timetable however it cannot guarantee that the train trajectories will be conflict-free. A MTTO method is applied if the STTO generates conflicts. The MTTO model optimizes multiple trains' trajectories simultaneously with constraints to avoid conflicts between trains. The STTO and MTTO models are re-formulated as a multiple-phase optimal control problem and solved by a pseudospectral method. Both STTO and MTTO are adopted in finding the optimal speed profiles and conflict-free time paths, which in turn changes the arrival and departure times of the existing timetable, and improves the timetable's energy efficiency. Case studies of two half-hour timetables on a Dutch single-track corridor and a double-track corridor suggested that our method is able to produce energy-efficient timetables.

Future research aims at integrating the train trajectory optimization in timetable design. Along with the proposed MTTO model, the following extensions should be addressed: (1) a scheduling model to find optimal train orders; and (2) dynamic headway constraints to ensure a feasible timetable.

References

Albrecht, A., Howlett, P., Pudney, P., Vu, X., Zhou, P., 2016. The key principles of optimal train controlpart 1: Formulation of the model, strategies of optimal type, evolutionary lines, location of optimal switching points. Transportation Research Part B: Methodological 94, 482–508.

Albrecht, T., Binder, A., Gassel, C., 2013. Applications of real-time speed control in rail-bound public transportation systems. IET Intelligent Transport Systems 7 (3), 305–314.

Bešinović, N., Goverde, R. M. P., Quaglietta, E., Roberti, R., 2016. An integrated micro–macro approach to robust railway timetabling. Transportation Research Part B: Methodological 87, 14–32.

Brännlund, U., Lindberg, P. O., Nou, A., Nilsson, J.-E., 1998. Railway timetabling using lagrangian relaxation. Transportation science 32 (4), 358–369.

Cacchiani, V., Caprara, A., Toth, P., 2008. A column generation approach to train timetabling on a corridor. 4OR: A Quarterly Journal of Operations Research 6 (2), 125–142.

Cacchiani, V., Toth, P., 2012. Nominal and robust train timetabling problems. European Journal of Operational Research 219 (3), 727–737.

Caimi, G., Fuchsberger, M., Laumanns, M., Schüpbach, K., 2011. A multi-level framework for generating train schedules in highly utilised networks. Public Transport 3 (1), 3–24.

Canca, D., Zarzo, A., 2017. Design of energy-efficient timetables in two-way railway rapid transit lines. Transportation Research Part B: Methodological 102, 142–161.

Garg, D., 2011. Advances in global pseudospectral methods for optimal control. Ph.D. thesis, University of Florida.

Gill, P. E., Murray, W., Saunders, M. A., 2005. SNOPT: An SQP algorithm for large-scale constrained optimization. SIAM Review 47 (1), 99–131.

Goverde, R. M. P., Bešinović, N., Binder, A., Cacchiani, V., Quaglietta, E., Roberti, R., Toth, P., 2016. A three-level framework for performance-based railway timetabling. Transportation Research Part C: Emerging Technologies 67, 62–83.

Goverde, R. M. P., Corman, F., DAriano, A., 2013. Railway line capacity consumption of different railway signalling systems under scheduled and disturbed conditions. Journal of Rail Transport Planning & Management 3 (3), 78–94.

Gupta, S. D., Tobin, J. K., Pavel, L., 2016. A two-step linear programming model for energy-efficient timetables in metro railway networks. Transportation Research Part B: Methodological 93, 57–74.

Hansen, I. A., Pachl, J., 2014. Railway timetabling and operations. Eurailpress.

Howlett, P., 2000. The optimal control of a train. Annals of Operations Research 98 (1), 65–87.

Howlett, P., Pudney, P., 1995. Energy-efficient train control. Springer, london, UK.

Huang, Y., Yang, L., Tang, T., Cao, F., Gao, Z., 2016. Saving energy and improving service quality: bicriteria train scheduling in urban rail transit systems. IEEE Transactions on Intelligent Transportation Systems 17 (12), 3364–3379.

Huang, Y., Yang, L., Tang, T., Gao, Z., Cao, F., 2017. Joint train scheduling optimization with service quality and energy efficiency in urban rail transit networks. Energy 138, 1124–1147.

Khmelnitsky, E., 2000. On an optimal control problem of train operation. IEEE Transactions on Automatic Control 45 (7), 1257–1266.

Kraay, D., Harker, P. T., Chen, B., 1991. Optimal pacing of trains in freight railroads: model formulation and solution. Operations Research 39 (1), 82–99.

Li, X., Lo, H. K., 2014a. An energy-efficient scheduling and speed control approach for metro rail operations. Transportation Research Part B: Methodological 64, 73–89.

Li, X., Lo, H. K., 2014b. Energy minimization in dynamic train scheduling and control for metro rail operations. Transportation Research Part B: Methodological 70, 269–284.

Liu, R., Golovitcher, I. M., 2003. Energy-efficient operation of rail vehicles. Transportation Research Part A: Policy and Practice 37 (10), 917–932.

Lusby, R. M., Larsen, J., Bull, S., 2017. A survey on robustness in railway planning. European Journal of Operational Research, in press.

Peeters, L., 2003. Cyclic railway timetable optimization. Ph.D. thesis, Erasmus Universiteit Rotterdam, Rotterdam, the Netherlands.

ProRail, 2017. Network statement. Https: //www.prorail.nl /sites /default /files.

Rao, A., Benson, D., Darby, C., Mahon, B., Francolin, C., Patterson, M., Sanders, I., Huntington, G., 2011. User's manual for gpops version 4. x: a matlab software for solving multiple-phase optimal control problems using hp-adaptive pseudospectral methods. University of Florida, Gainesville.

Rao, A. V., 2003. Extension of a pseudospectral legendre method to non-sequential multiple-phase optimal control problems. In: 2003 AIAA Guidance, Navigation, and Control Conference and Exhibit. Austin, USA, pp. 11–14.

Rao, A. V., Benson, D. A., Darby, C., Patterson, M. A., Francolin, C., Sanders, I., Huntington, G. T., 2010. Algorithm 902: GPOPS, a matlab software for solving multiple-phase optimal control problems using the gauss pseudospectral method. ACM Transactions on Mathematical Software 37 (2), 22.

Scheepmaker, G. M., Goverde, R. M. P., 2015. The interplay between energy-efficient train control and scheduled running time supplements. Journal of Rail Transport Planning & Management 5 (4), 225–239.

Scheepmaker, G. M., Goverde, R. M. P., Kroon, L. G., 2017. Review of energy-efficient train control and timetabling. European Journal of Operational Research 257 (2), 355–376.

Serafini, P., Ukovich, W., 1989. A mathematical model for periodic scheduling problems. SIAM Journal on Discrete Mathematics 2 (4), 550–581.

Su, S., Li, X., Tang, T., Gao, Z., 2013. A subway train timetable optimization approach based on energy-efficient operation strategy. IEEE Transactions on Intelligent Transportation Systems 14 (2), 883–893.

Su, S., Tang, T., Li, X., Gao, Z., 2014. Optimization of multitrain operations in a subway system. IEEE transactions on intelligent transportation systems 15 (2), 673–684.

Wang, P., Goverde, R. M. P., 2016a. Multiple-phase train trajectory optimization with signalling and operational constraints. Transportation Research Part C: Emerging Technologies 69, 255–275.

Wang, P., Goverde, R. M. P., 2016b. Train trajectory optimization of opposite trains on single-track railway lines. In: 2016 IEEE International Conference on Intelligent Rail Transportation (ICIRT). pp. 23–31, birmingham, UK.

Wang, P., Goverde, R. M. P., 2016c. Two-train trajectory optimization with a green-wave policy. Transportation Research Record: Journal of the Transportation Research Board (2546), 112–120.

Wang, P., Goverde, R. M. P., 2017. Multi-train trajectory optimization for energy efficiency and delay recovery on single-track railway lines. Transportation Research Part B: Methodological 105, 340–361.

Wang, Y., Ning, B., Cao, F., De Schutter, B., van den Boom, T. J., 2011. A survey on optimal trajectory planning for train operations. In: 2011 IEEE International Conference on Service Operations, Logistics, and Informatics (SOLI). pp. 589–594, beijing, China.

Yang, L., Li, K., Gao, Z., 2009. Train timetable problem on a single-line railway with fuzzy passenger demand. IEEE Transactions on fuzzy systems 17 (3), 617–629.

Yang, L., Li, K., Gao, Z., Li, X., 2012. Optimizing trains movement on a railway network. Omega 40 (5), 619–633.

Yang, X., Li, X., Ning, B., Tang, T., 2016. A survey on energy-efficient train operation for urban rail transit. IEEE Transactions on Intelligent Transportation Systems 17 (1), 2–13.

Ye, H., Liu, R., 2016. A multiphase optimal control method for multi-train control and scheduling on railway lines. Transportation Research Part B: Methodological 93, 377–393.

Yin, J., Tang, T., Yang, L., Gao, Z., Ran, B., 2016. Energy-efficient metro train rescheduling with uncertain time-variant passenger demands: An approximate dynamic programming approach. Transportation Research Part B: Methodological 91, 178–210.

Zhou, L., Tong, L. C., Chen, J., Tang, J., Zhou, X., 2017. Joint optimization of high-speed train timetables and speed profiles: A unified modeling approach using space-time-speed grid networks. Transportation Research Part B: Methodological 97, 157–181.