TU Delft

Delft University of Technology

An exact optimization method for coordinating the arrival times of urban rail lines at a common corridor

Gkiotsalitis, K.; Cats, O.; Liu, T.; Bult, J. M.

DOI
10.1016/j.tre.2023.103265

Publication date
2023

Document Version
Final published version

Published in
Transportation Research Part E: Logistics and Transportation Review

Citation (APA)
Gkiotsalitis, K., Cats, O., Liu, T., & Bult, J. M. (2023). An exact optimization method for coordinating the arrival times of urban rail lines at a common corridor. *Transportation Research Part E: Logistics and Transportation Review*, *178*, Article 103265. https://doi.org/10.1016/j.tre.2023.103265

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Transportation Research Part E 178 (2023) 103265

Contents lists available at ScienceDirect

Transportation Research Part E

journal homepage: www.elsevier.com/locate/tre

An exact optimization method for coordinating the arrival times of urban rail lines at a common corridor

K. Gkiotsalitis [a,*], O. Cats [b], T. Liu [c], J.M. Bult [a]

[a] University of Twente, Department of Civil Engineering, 7500 AE Enschede, Netherlands
[b] Delft University of Technology, Department of Transport and Planning, Gebouw 23 Stevinweg 1 2628 CN Delft, Netherlands
[c] National Engineering Laboratory of Integrated Transportation Big Data Application Technology, School of Transportation and Logistics, Southwest Jiaotong University, Chengdu 611756, China

ARTICLE INFO

Keywords:
Train corridor
Coordination
Scheduling
Convex optimization

ABSTRACT

The trips of a high-frequency urban rail line are typically planned with the aim of achieving even time headways. This results in reliable services for each urban rail line, where successive trips have the same time headway. Maintaining even time headways for each service line has significant advantages for the passengers of the line, but it might result in safety issues, vehicle bunching, and increased transfer times at a common corridor served by multiple urban rail lines. This study investigates the problem of urban rail corridor coordination and develops an exact optimization method for coordinating the vehicle trips of different lines that serve stations along a joint corridor. The proposed formulation is a non-convex mathematical program which is reformulated as a mixed-integer quadratic program with a convex objective function. A branch-and-bound algorithm coupled with the Active-set method is proposed for solving the model to global optimality. Results from a toy network and a case study of the light rail service in The Hague, The Netherlands, demonstrate the improvement potential of time headways at a common corridor, while accounting for the effect on the variation of time headways at isolated segments of the individual service lines.

1. Introduction

Public transport services are planned at the strategic, tactical and operational level. At the strategic level, decisions are made about the location of stations and the routes of the service lines that can form broader mobility hubs (Murray, 2001; Arnold et al., 2004; Limbourg and Jourquin, 2009; Medina et al., 2013; Schöbel, 2012; Canca et al., 2019; Horjus et al., 2022; Wang and Liu, 2022; Geurs et al., 2023). At the tactical level, the pre-determined public transport network and passenger demand are used to set the frequencies of the service lines and provide a first estimation of the number of required vehicles to operate the service (Cadarso et al., 2013; Szeto and Jiang, 2014; Sun and Szeto, 2019; Gkiotsalitis and Cats, 2018). With the frequency settings stage, one determines the number of performed trips per line on an hourly basis. Determining the line frequencies, the next step is to schedule the dispatching time of each trip for each one of the service lines. These dispatching time decisions define the timetable of the service line (Parbo et al., 2014; Liu and Ceder, 2016; Gkiotsalitis and Alesiani, 2019). When determining the dispatching times, typical objectives are the even distribution of trips over time to maintain an even headway between successive trips (Gkiotsalitis and Cats, 2020) or the adherence to an even in-vehicle load at the maximum loading point of the line while considering time-dependent passenger arrival rate fluctuations (Ceder, 2016).

* Corresponding author.
E-mail address: k.gkiotsalitis@utwente.nl (K. Gkiotsalitis).

https://doi.org/10.1016/j.tre.2023.103265
Received 31 October 2022; Received in revised form 18 August 2023; Accepted 19 August 2023
Available online 30 August 2023
1366-5545/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

This study focuses on the timetabling phase of service lines, and in particular on the coordination of different service lines that share a common corridor. This service coordination problem often arises in light rail and metro services operating in urban regions, where different lines and line branches merge into a high-capacity public transport corridor serving a high demand area. There is limited knowledge on how to plan services for this prevalent network configuration as becomes apparent from our review of the literature in the following section (Sørensen and Longva, 2011; Liu and Ceder, 2016; Liu et al., 2021). Unlike problems that try to synchronize the arrival of vehicles of different lines at a common stop to facilitate passenger transfers, the problem of corridor coordination requires to maintain an even headway among trips that use the same corridor to mitigate bunching of vehicles from all lines operating along the corridor. This affects passengers traveling within the high-demand corridor as well as rail traffic (e.g., maintaining a minimum safety gap between successive trips).

The importance of managing the flow of vehicles in common public transport corridors becomes apparent when a delay of a vehicle results in a domino effect, where trains from other service lines have to wait for their turn to enter the corridor resulting in the propagation of delays in several service lines (Szymański et al., 2018). In these situations, it is important to maintain an even headway among successive trains traversing the corridor and thereby minimize the impacts of disturbances during operations. Unlike the case of transfer synchronization problems, the problem of corridor coordination does not seek to achieve as many simultaneous arrivals of trains to the corridor as possible. Conversely, we aim at maintaining a sufficient headway between successive passes of vehicles. This, however, is not a trivial task because vehicle-trips from different lines have different frequencies and there are line-specific restrictions related to rolling stock circulation and scheduling constraints at specific stations (Orth et al., 2012) that might conflict with the objectives of the corridor coordination process.

To this end, we develop a train corridor coordination approach which considers the constraints of individual service lines related to vehicle circulation and pre-determined dispatching time ranges. We formulate, solve and apply the urban rail corridor coordination problem by means of an exact method. The contributions of this study are fourfold:

- we develop a mathematical program for the corridor coordination problem
- we prove that the corridor coordination problem is non-convex and we introduce linearizations to reformulate it as a mixed-integer quadratic program
- we develop a branch-and-bound based algorithm coupled with the Active-set method and prove that our solution method guarantees the computation of a globally optimal solution
- we demonstrate the potential impact of the proposed corridor coordination using data from the light rail service in The Hague, The Netherlands

The remainder of the study is structured as follows. In Section 2 we provide a literature review of timetabling and transfer coordination/synchronization problems. In Section 3 we develop the mathematical program of the corridor coordination problem. In Section 4 we present the exact solution method, which is based on branch-and-bound and the Active-set method. Section 5 provides a demonstration of the model in a toy network to facilitate the reproduction of our model. Section 6 presents the results of our case study in The Hague. Finally, Section 7 concludes our study.

2. Literature review

2.1. Timetabling

2.1.1. Timetabling models for bus services

A distinct line of works in timetable optimization strives to ensure that the dispatching times of trips are evenly-spaced throughout the day. For instance, Ceder (2011), Ceder et al. (2013) strive to achieve a desired even-load level for all vehicles at their maximum loading point by determining trip dispatching times that do not deviate significantly from the desired even headway. Similarly, Daduna and Voß (1995) and Shafahi and Khani (2010) generated timetables with evenly-spaced dispatching times incorporating the additional objective of synchronizing passenger transfers.

Many works, especially on the scheduling of bus services, consider stochastic dwell times in the problem formulation during the optimization process. One example of such works is the work of Wu et al. (2015) which assumes stochastic bus travel times to calculate slack times at transfer stops and improve the synchronization of timetables. Studies on operational control have also considered stochastic travel times. For example, Xuan et al. (2011) used stochastic travel times to derive dynamic bus holding strategies and improve adherence to the planned schedule, and Hickman (2001) developed an analytical bus holding model that considers stochastic link travel times and boardings/alightings. These works use a typical probability distribution that approximates the stochastic nature of travel times and find a solution by solving a stochastic optimization problem.

2.1.2. Timetabling models for rail services

Unlike bus services, timetabling works of urban rail and train services consider typically fixed travel times and dwell times (Gkiotsalitis and Cats, 2020). This is especially the case in services that have a pre-defined time set for the opening and closing of doors. This does not mean, however, that all train services can maintain fixed dwell times (D'Acierno et al., 2017; Kuipers et al., 2021; Meloni et al., 2021).

From the works in traditional timetabling, there are several approaches that couple the timetabling problem with the frequency setting problem (Furth and Wilson, 1981; Peeters and Kroon, 2001; Ávila-Torres et al., 2018; Silva-Soto and Ibarra-Rojas, 2021) or the timetabling problem with the vehicle scheduling problem (Ceder, 2011; Ibarra-Rojas et al., 2014; Schmid and Ehmke, 2015;

Fonseca et al., 2018). Nevertheless, scheduling methods do not consider the impact of the determined timetables on vehicle flow in a common corridor of the public transport network. This shortcoming is the main focus of this study which develops a model for coordinating the trips of multiple urban rail lines at a common corridor.

2.2. Transfer synchronization

2.2.1. Transfer synchronization at the tactical planning stage

Public transport transfer synchronization aims to coordinate the arrival and departure times of vehicles from different lines at shared transfer stations so as to facilitate passenger transfers. It can be achieved through coordinated timetabling at the tactical planning level and real-time vehicle scheduling at the operational control level (see the survey papers of Liu et al. (2021) and Gkiotsalitis et al. (2022), respectively).

Coordinated timetabling at the tactical planning stage is one of the most cost-effective approaches in increasing public transport transfer synchronization because it only requires minor modifications of the service timetable without the need of performing infrastructure investments. It has been widely studied in the transportation research and operations research communities (e.g., Daduna et al., 1995; Ceder et al., 2001; Kwan and Chang, 2008; Shafahi and Khani, 2010; Ibarra-Rojas et al., 2014; Wu et al., 2016; Gkiotsalitis and Maslekar, 2018; Yin et al., 2021). Different solution approaches have been developed, including using heuristic rules, analytical models, mathematical programming models, and simulation approaches (Liu et al., 2021). Recent studies on public transport transfer synchronization mostly utilize mathematical programming approaches since they can generate more realistic and optimally coordinated timetables that can be directly applied into practice. Various optimization objectives have been considered in previous studies, such as minimizing the total transfer waiting time (Daduna et al., 1995; Wong et al., 2008; Shafahi and Khani, 2010), maximizing the number of simultaneous arrivals of vehicles at transfer stations (Ceder et al., 2001), minimizing a total passenger dissatisfaction index (Kwan and Chang, 2008), maximizing the number of passengers benefited from coordinated transfers (Ibarra-Rojas et al., 2014), maximizing the number of successful transfer connections (Tian and Niu, 2019), and maximizing the number of successful transfers for last/first train service (Yang et al., 2018; Huang et al., 2021). However, few studies have considered optimizing the arrival time gap between vehicles from different lines in a shared travel corridor (i.e., Canca et al., 2016), which is important for ensuring the safe operation of trains in a railway corridor and the evenly distributed passenger waiting times along that corridor. A relevant work in this direction is the approach of Canca et al. (2016), which proposed a mixed integer nonlinear programming (MINLP) model to determine optimal line frequencies and capacities in dense railway rapid transit networks in which typically several lines can run over the same open tracks. Given a certain demand matrix, the model determines the most appropriate frequency and train capacity for each line taking into account infrastructure capacity constraints, allocating lines to tracks while assigning passengers to lines.

2.2.2. Transfer synchronization at the operational control stage

At the operational control level, some near real-time control strategies, such as vehicle holding, stop-skipping, short-turning, speed change, rescheduling and rerouting, are employed to increase the actual occurrences of synchronized transfers (Gkiotsalitis et al., 2022). To the best of our knowledge, only a few real-time control studies have considered optimizing the public transport transfer synchronization in a shared travel corridor (e.g., Hadas and Ceder, 2008, 2010; Wu et al., 2021). However, these studies mainly focused on increasing the encounter probability between public transport vehicles, either normal buses (Hadas and Ceder, 2008, 2010) or autonomous modular vehicles (Wu et al., 2021), in order to facilitate passenger transfers. They did not consider the minimum time gap between vehicles required for the safe operation in a travel corridor and the possibility of maintaining evenly distributed headways between the trips of different lines operating in that corridor.

To bridge this research gap, this study proposes a new mathematical programming model to optimize the coordination of trains from different service lines in a shared travel corridor. Our model can help maintain an even headway among trains that use the same corridor so as to keep a safe operation gap as well as reducing vehicle bunching.

3. Urban rail corridor coordination problem

We define the Urban Rail Corridor Coordination problem (URCCP) as follows. Given a train corridor consisting of at least two consecutive stations which are served by at least two urban rail lines, modify the dispatching times of the trips of the lines so that the arrival times of trains from different lines at the stations of that corridor are as close as possible to pre-defined target time headway values.

From now on, the train corridor which meets this property will be referred to as *joint corridor* or *common corridor*. In this problem definition, the rail lines may operate also outside the joint corridor or serve only small parts of the corridor, as long as they do not intersect with other lines outside of the joint corridor. In case they do, an additional *weak* assumption stating that they do not need to synchronize their services at these exterior points is needed.

In addition, the pre-defined target headway values at the stations of the joint corridor may be defined in such a way that the time headways between trips of different lines that pass through the stations of the joint corridor are as evenly distributed as possible at each station of the corridor. This can provide an advantage compared to the case of only meeting a planned safety distance between successive trips at the stations of the joint corridor, because it ensures a relatively even distribution of vehicles passing at a station. This aspect is important if one wants to improve uneven vehicle arrivals and crowdedness levels at the platforms of train stations since applying real-time control measures to mitigate the uneven vehicle arrivals at stations (vehicle bunching) is a considerable

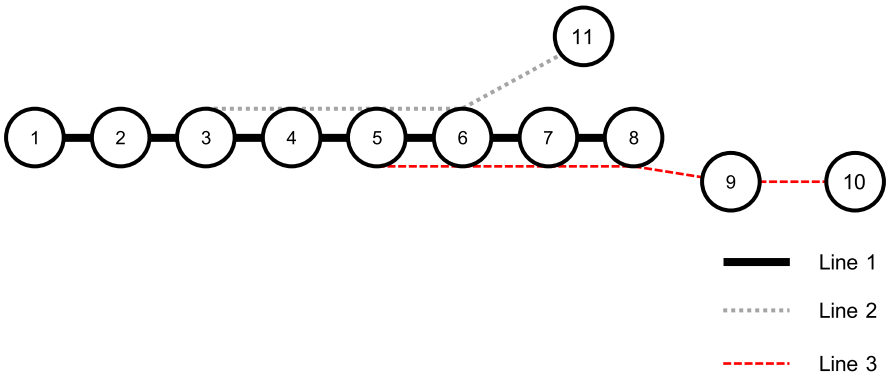

Fig. 1. Example network.

problem in public transport operations (Daganzo, 2009; Bartholdi III. and Eisenstein, 2012; Schmöcker et al., 2016; Gkiotsalitis and Van Berkum, 2020b). Finally, stations of the joint corridor with higher passenger boarding levels may receive priority compared to stations with lower passenger boarding levels because the coordination of train arrivals at the former stations might have higher importance.

An example of such network is presented in Fig. 1, where lines 1, 2, and 3 jointly serve stations 5 and 6 that form a common corridor.

In addition to the above, our model is based on the following assumptions:

1. there is a pre-determined time for door opening and closing at urban rail lines; thus, dwell times do not depend on the passenger demand levels.
2. as we operate at the tactical planning level, inter-station travel times are considered to be fixed and their values are estimated from historical data.
3. lines do not intersect with other lines that are not operating in the joint corridor.

At the remainder of this section, we present the sets, parameters, variables, the formulation, and the reformulation of our urban rail corridor coordination model.

3.1. Sets

We consider an unordered set of train lines $K = \{1, \dots, |K|\}$. These train lines have a set of train stations $S = \{1, \dots, |S|\}$. Each train line $k \in K$ has an ordered set of train stations S^k which is a subset of S, $S^k \subset S$. That is, S^k_1 is the first station of each train line $k \in K$. There is also a set of train trips $N = \{1, \dots, |N|\}$ that are assigned to the train lines.

3.2. Parameters

We consider column vector $\boldsymbol{\beta} = [\beta_1, \dots, \beta_{|N|}]^\intercal$, where $\beta_i \in K$ indicates the train line to which trip i is assigned to. In addition, we consider the following parameters (see Table 1). Note that because the total travel time of a trip τ_i is known (parameter), if we know the latest possible dispatching time of the next trip operated by the same train we can trivially compute the latest possible dispatching time of trip i (ω_i) by subtracting τ_i from the latest possible dispatching time of the next trip. In addition, the target headways at a station of the joint corridor may be defined such that the time headway between successive trips at that station are as evenly distributed as possible. For instance, if we have 4 lines where each one has a dispatching time headway of 10 min and they all pass by the same station, this station can have a target headway of 10 min/4 lines = 2.5 min. If another station of the joint corridor is served by 3 lines, then this station can have a target headway of 10 min/3 lines $\simeq$ 3.33 min.

3.3. Variables

Our variables include $\mathbf{A} = \{a_i^s\}$, where a_i^s indicates the arrival time of trip $i \in N$ at station $s \in S$. They also include $\mathbf{D} = \{d_i^s\}$, where d_i^s denotes the departure time of trip $i \in N$ from station $s \in S$. Finally, they include the $|N|$-valued column vector $\mathbf{x} = \{x_i\}$,

Table 1
Parameters.

$P = \{p_{ij}^k\}$	$\{0,1\}$ indicator parameters, where $p_{ij}^k = 1$ if trips i and j are successive trips of line $k \in K$.
$\boldsymbol{\Gamma} = \{\gamma_i^s\}$	$\{0,1\}$ indicator parameters, where $\gamma_i^k = 1$ if trip i serves train station $s \in S$.
$\boldsymbol{\Pi} = \{\pi_{ij}^k\}$	$\{0,1\}$ indicator parameters, where $\pi_{ij}^k = 1$ if trips i and j are trips of line $k \in K$ that are successively operated by the same train.
$T = \{t_i^s\}$	travel time of trip i from station $s-1$ to station s where $s-1 \in S^k, s \in S^k$ and $k = \beta_i$. If $s \notin S_k$ then t_i^s can take the value of any arbitrary finite number.
$\boldsymbol{\tau} = \{\tau_i\}$	total travel time of trip i, until its vehicle is ready to perform the next trip.
δ	safety distance among successive trains (a scalar).
$V = \{v_k\}$	minimum allowed dispatching time headway between successive train-trips of line $k \in K$
$U = \{u_k\}$	maximum allowed dispatching time headway between successive train-trips of line $k \in K$
$H = \{h_{ij}^s\}$	target headway between trips $i \in N$ and $j \in N \neq i$ at stop $s \in S$
$\boldsymbol{\eta} = \{\eta_i^s\}$	dwell time of trip i at stop $s \in S$
$\boldsymbol{\Phi} = \{\phi_{ij}^s\}$	$\{0,1\}$ indicator parameters, where $\phi_{ij}^s = 1$ if there is a target headway between trips $i \in N$ and $j \in N \neq i$ at station $s \in S$, and 0 otherwise.
$\boldsymbol{\Psi} = \{\psi_{ij}^s\}$	weight factor parameters with non-negative values indicating the importance of meeting the target headway between trips $i \in N$ and $j \in N \neq i$ at stop $s \in S$.
$\boldsymbol{\Omega} = \{\omega_i\}$	latest possible dispatching time for each trip $i \in N$ to avoid schedule sliding.
M	a very large positive real number (a scalar).

where $x_i \in \mathbb{R}_{\geq 0}$ is the dispatching time of train trip $i \in N$.

3.4. Feasible region

Reckon that S_1^k is the first station of all trips operating in line k. Because trip i is assigned to line β_i, we can set $s_i' \leftarrow S_1^{\beta_i}$, indicating the first station of each trip $i \in N$. The departure time of each trip i from its first station is then:

$$d_i^{s_i'} = x_i \quad \forall i \in N \tag{1}$$

The arrival time of each trip i at station s is given by:

$$a_i^s = \gamma_i^s (t_i^s + d_i^{s-1}) \quad \forall i \in N, \forall s \in S \setminus \{s_i'\} \tag{2}$$

and the departure time of each trip i from station $s \neq s_i'$ is:

$$d_i^s = \gamma_i^s (a_i^s + \eta_i^s) \quad \forall i \in N, \forall s \in S \setminus \{s_i'\} \tag{3}$$

Note that if trip i does not serve station s, then $a_i^s = d_i^s = 0$ because $\gamma_i^s = 0$. If we have two successive trips i, j of line k, then they should be dispatched within a lower and upper time headway $[v_k, u_k]$. This is enforced by the following inequality constraints:

$$p_{ij}^k v_k \leq p_{ij}^k (x_j - x_i) \leq u_k \quad \forall i, j \in N \mid i \neq j, \ \forall k \in K$$

(dispatching headway constraints) (4)

Note that if trips i, j are not successive or one of them is not assigned to line k, then the inequality constraint is satisfied because $p_{ij}^k = 0$ and $u_k \in \mathbb{R}_{\geq 0}$. The dispatching headway constraints are imposed to every line and guarantee that the regularity of lines will not be altered considerably, since headways will remain within a pre-determined range. This range is introduced in our model to attain a balance between the loss of regularity of individual lines and the improved coordination at common corridors.

In addition, we have safety gap constraints upon the arrival of vehicle trips at stations. For a safety gap δ, the following inequality constraints are enforced:

$$\gamma_i^s \gamma_j^s \| a_i^s - a_j^s \| \geq \gamma_i^s \gamma_j^s \delta \quad \forall i, j \in N \mid i \neq j, \ \forall s \in S$$

(safety gap constraints upon arrival at stations) (5)

Notice that if either trip i or j does not serve station s, the equality constraint is satisfied because $\gamma_i^s \gamma_j^s = 0$. Note that $\|\cdot\|_1$ is the l_1 norm which is equivalent to the absolute value $|a_i^s - a_j^s|$ because a_i^s, a_j^s are scalars. Similarly, we impose safety gap constraints upon the departure of vehicle trips from stations:

$$\gamma_i^s \gamma_j^s \|d_i^s - d_j^s\| \geq \gamma_i^s \gamma_j^s \delta \quad \forall i, j \in N \mid i \neq j, \ \forall s \in S$$

(safety gap constraints upon departure from stations) (6)

The definition of the safety gap constraints at the station level assumes that a station has a single platform and the incoming track of the successive trips is the same. In case of multiple tracks and platforms, the safety gap constraints may be applied at the platform level.

Trips i, j that are successive trips of line k operated by the same vehicle should also satisfy the vehicle circulation constraints expressed as:

$$x_j \geq \pi_{ij}^k (x_i + \tau_i) \quad \forall i \in N, j \in N \mid j \neq i, k \in K \text{ (vehicle circulation constraints)} \tag{7}$$

Finally, we cannot allow a trip to be dispatched with a considerable delay because this will result in schedule sliding, resulting in the delayed dispatching of subsequent trip operated by the same vehicle. This is enforced by the following constraints:

$$x_i \leq \omega_i \quad \forall i \in N \text{ (schedule sliding constraints)} \tag{8}$$

3.5. Objective function

The objective of our problem is to coordinate the vehicle trips when serving stations along the joint corridor. If there is a target (ideal) time headway h^s at every station then for every pair of trips that serve the station we seek to minimize the squared difference:

$$\left((a_i^s - a_j^s) - h^s\right)^2$$

Notice that we use the squared difference of the actual time headway $a_i^s - a_j^s$ and the target headway h^s instead of the absolute difference $\|(a_i^s - a_j^s) - h^s\|$ in order to overpenalize significant deviations from h^s since:

$$\left((a_i^s - a_j^s) - h^s\right)^2 > \|(a_i^s - a_j^s) - h^s\| \quad \text{for} \quad \|(a_i^s - a_j^s) - h^s\| > 1$$

One should strive to pursue the target headways between trips i, j that might belong to different lines, but serve the joint corridor station s in a successive order. This is achieved by using $\{0, 1\}$ indicator parameter ϕ_{ij}^s which is equal to 1 if there is a target headway between trips i and j at station s, and 0 if there is not (i.e., station s is not a joint corridor station, or trips i, j are not successive, or one of the two trips does not serve station s). Finally, meeting the target headway between a pair of trips i, j at station s might be more important due to higher passenger demand compared to another pair of trips i', j' at s or at another station s'. This is accounted for by using weight factors ψ_{ij}^s that attain a larger value when the importance of meeting that target headway due to higher passenger demand is higher. This yields the following objective function:

$$f(\mathbf{A}) \doteq \sum_{s \in S} \sum_{i \in N} \sum_{j \in N \setminus \{i\}} \phi_{ij}^s \psi_{ij}^s \left((a_i^s - a_j^s) - h_{ij}^s\right)^2 \tag{9}$$

which is a C^2 continuous and scalar-valued function, where $f : \mathbb{R}^{|N| \times |K|} \rightarrow \mathbb{R}_{\geq 0}$. Note that the objective function uses the squared deviation between the headways and their target values instead of the absolute deviation because the former overpenalizes large deviations from the target headway. This is a favorable property, because the absolute deviation focuses only on the deviation from the target headway and does not balance the headway deviation among trips (Newell, 1974; Trompet et al., 2011; Gkiotsalitis and Van Berkum, 2020a).

Closing, we note that by focusing only on the coordination of the arrival times of trains at the stations of the common corridor, we may neglect the transfer synchronization of lines that serve the common corridor at their exterior transfer points. If some of these exterior transfers are important and justify their synchronization in the expense of the coordination of trip arrival times at the joint corridor, one could add additional terms in the objective function to account for these transfer synchronizations. In more detail, if a train trip i that serves the joint corridor needs to have an arrival time as close as possible to the arrival time of another trip ρ at a transfer station s' which is outside of the common corridor, then one can add the term $(a_i^{s'} - a_\rho^{s'})^2$ to the objective function to account for a synchronized transfer. This can be performed for all transfers that need to be synchronized at exterior transfer stations and are not part of the joint corridor by simply adding similar terms to the objective function. The importance of each one of these transfer synchronizations can be regulated by multiplying the term $(a_i^{s'} - a_\rho^{s'})^2$ of each one of them by a weight factor which can take values in $[0, +\infty)$, where a value of 0 implies no importance and a very large value implies that this transfer is more important than the coordination of line i in the joint corridor.

3.6. Mathematical program

Before introducing the mathematical program, note that the schedule sliding constraints (8) and the vehicle circulation constraints (7) can potentially be in conflict, thereby resulting in an *empty* feasible set. Indeed, if we require from a vehicle-trip to be dispatched before a pre-defined time ω_i, there might be no vehicle to operate that trip. For this reason, the schedule sliding constraint is relaxed. From a *hard* constraint, it becomes a *soft* constraint which is included in the objective function by means of a penalty term:

$$\sum_{i\in N} M\max(x_i-\omega_i,0)$$

which adds a significant penalty each time a dispatching time x_i is greater than ω_i. The updated objective function which takes into consideration this penalty term is:

$$f(\mathbf{A},\mathbf{x}) \doteq \sum_{s\in S}\sum_{i\in N}\sum_{j\in N\setminus\{i\}} \phi_{ij}^s\psi_{ij}^s\left((a_i^s-a_j^s)-h_{ij}^s\right)^2 + \sum_{i\in N} M\max(x_i-\omega_i,0) \tag{10}$$

which is still C^2 and twice differentiable. The resulting mathematical program $\mathcal{Q}$ is:

$$(\mathcal{Q}):$$

$$\underset{\mathbf{A},\mathbf{D},\mathbf{x}}{\text{minimize}} \sum_{s\in S}\sum_{i\in N}\sum_{j\in N\setminus\{i\}} \phi_{ij}^s\psi_{ij}^s\left((a_i^s-a_j^s)-h_{ij}^s\right)^2 + \sum_{i\in N} M\max(x_i-\omega_i,0) \tag{11}$$

$$\text{subject to: } d_i^{s_i'} = x_i \qquad \forall i\in N \tag{12}$$

$$a_i^s = \gamma_i^s(t_i^s + d_i^{s-1}) \qquad \forall i\in N, \forall s\in S\setminus\{s_i'\} \tag{13}$$

$$d_i^s = \gamma_i^s(a_i^s + \eta_i^s) \qquad \forall i\in N, \forall s\in S\setminus\{s_i'\} \tag{14}$$

$$p_{ij}^k v_k \le p_{ij}^k(x_j-x_i) \le u_k \qquad \forall i\in N, j\in N \mid i\neq j,\ \forall k\in K \tag{15}$$

$$\gamma_i^s\gamma_j^s\|a_i^s-a_j^s\| \ge \gamma_i^s\gamma_j^s\delta \qquad \forall i\in N, j\in N \mid i\neq j,\ \forall s\in S \tag{16}$$

$$\gamma_i^s\gamma_j^s\|d_i^s-d_j^s\| \ge \gamma_i^s\gamma_j^s\delta \qquad \forall i\in N, j\in N \mid i\neq j,\ \forall s\in S \tag{17}$$

$$x_j \ge \pi_{ij}^k(x_i+\tau_i) \qquad \forall i\in N, j\in N \mid j\neq i, k\in K \tag{18}$$

$$\mathbf{x}\in\mathbb{R}_{\ge 0}^{|N|}, \mathbf{A}\in\mathbb{R}_{\ge 0}^{|N|\times|S|}, \mathbf{D}\in\mathbb{R}_{\ge 0}^{|N|\times|S|} \tag{19}$$

$\mathcal{Q}$ is a *compact* continuous program because its constraints do not increase exponentially with the size of the problem. It has $|N|+2|N||S|$ continuous variables, $|N|+2|N||S|$ equality constraints, and $2|N|^2|K|+2|N|^2|S|+|N|+2|N||S|$ inequality constraints. Using the Big-Oh asymptotic notation we can express these as $O(|N||S|)$, $O(|N||S|)$, and $O(|N|^2|S|+|N|^2|K|)$, respectively.

Theorem 3.1. *$\mathcal{Q}$ is a nonlinear and nonconvex minimization problem.*

Proof.

$$\gamma_i^s\gamma_j^s\|a_i^s-a_j^s\| \ge \gamma_i^s\gamma_j^s\delta \Rightarrow \begin{cases} \gamma_i^s\gamma_j^s(a_i^s-a_j^s) \ge \gamma_i^s\gamma_j^s\delta \\ \text{or} \\ -\gamma_i^s\gamma_j^s(a_i^s-a_j^s) \ge \gamma_i^s\gamma_j^s\delta \end{cases}$$

Thus, the feasible region has a gap and it is non-convex. The expressions also make it impossible for both to simultaneously hold true. Finally, $\mathcal{Q}$ is nonlinear because it has a quadratic and a max term in the objective function, and it also has the nonlinear constraints (16) and (17). ■

3.7. Linearizations

Because the minimization problem is nonlinear and noncovex, we proceed with a series of linearizations that make it possible to obtain a globally optimal solution of the problem. This comes at a price, however, because (i) we make the program less compact by introducing more variables and constraints, and (ii) we introduce binary variables that turn the continuous formulation of $\mathcal{Q}$ into a mixed-integer one. First, the absolute terms in Eqs. (16)–(17) are transformed into affine constraints by replacing

$$\|a_i^s-a_j^s\|$$

with

$$r_{ij}^s \ge a_i^s - a_j^s$$
$$r_{ij}^s \ge -(a_i^s-a_j^s)$$

$$r_{ij}^s \leq a_i^s - a_j^s + M\xi_{ij}^s$$
$$r_{ij}^s \leq -(a_i^s - a_j^s) + M(1 - \xi_{ij}^s)$$

and

$$\|d_i^s - d_j^s\|$$

with

$$\bar{r}_{ij}^s \geq d_i^s - d_j^s$$
$$\bar{r}_{ij}^s \geq -(d_i^s - d_j^s)$$
$$\bar{r}_{ij}^s \leq d_i^s - d_j^s + M\bar{\xi}_{ij}^s$$
$$\bar{r}_{ij}^s \leq -(d_i^s - d_j^s) + M(1 - \bar{\xi}_{ij}^s)$$

where $\xi_{ij}^s, \bar{\xi}_{ij}^s \in \{0,1\}$ and $r_{ij}^s, \bar{r}_{ij}^s \in \mathbb{R}$ are newly introduced variables. Note that we have to introduce the binary variables $\xi_{ij}^s, \bar{\xi}_{ij}^s$ because of the gap in the feasible region, even if this makes the problem mixed-integer resulting in a significant computational increase compared to its continuous version. At the same time, however, the continuous relaxation of the new mixed-integer formulation that is presented below guarantees global optimality, as we will later see in Theorem 4.1.

The linearized mathematical program $\tilde{\mathcal{Q}}$ is presented below. In this program we replace Eqs. (16)–(17) with (25)–(34). We also replace the nonlinear penalty term of the objective function:

$$\sum_{i \in N} M \max(x_i - \omega_i, 0)$$

by introducing new continuous variables $\mathbf{c} = [c_1, \ldots, c_{|N|}]^\intercal$ and inequality constraints (36)–(37):

$$(\tilde{\mathcal{Q}}): \quad \underset{\mathbf{A},\mathbf{D},\mathbf{x},\mathbf{r},\bar{\mathbf{r}},\boldsymbol{\xi},\bar{\boldsymbol{\xi}},\mathbf{c}}{\text{minimize}} \sum_{s \in S} \sum_{i \in N} \sum_{j \in N \setminus \{i\}} \phi_{ij}^s \psi_{ij}^s \left((a_i^s - a_j^s) - h_{ij}^s \right)^2 + \sum_{i \in N} c_i \tag{20}$$

$$\text{subject to: } d_i^{s_i'} = x_i \qquad \forall i \in N \tag{21}$$
$$a_i^s = \gamma_i^s (t_i^s + d_i^{s-1}) \qquad \forall i \in N, \forall s \in S \setminus \{s_i'\} \tag{22}$$
$$d_i^s = \gamma_i^s (a_i^s + \eta_i^s) \qquad \forall i \in N, \forall s \in S \setminus \{s_i'\} \tag{23}$$
$$p_{ij}^k v_k \leq p_{ij}^k (x_j - x_i) \leq u_k \qquad \forall i \in N, j \in N \mid i \neq j, \ \forall k \in K \tag{24}$$
$$\gamma_i^s \gamma_j^s r_{ij}^s \geq \gamma_i^s \gamma_j^s \delta \qquad \forall i, j \in N \mid i \neq j, \ \forall s \in S \tag{25}$$
$$r_{ij}^s \geq a_i^s - a_j^s \qquad \forall i, j \in N \mid i \neq j, \ \forall s \in S \tag{26}$$
$$r_{ij}^s \geq -(a_i^s - a_j^s) \qquad \forall i, j \in N \mid i \neq j, \ \forall s \in S \tag{27}$$
$$r_{ij}^s \leq a_i^s - a_j^s + M\xi_{ij}^s \qquad \forall i, j \in N \mid i \neq j, \ \forall s \in S \tag{28}$$
$$r_{ij}^s \leq -(a_i^s - a_j^s) + M(1 - \xi_{ij}^s) \qquad \forall i, j \in N \mid i \neq j, \ \forall s \in S \tag{29}$$
$$\gamma_i^s \gamma_j^s \bar{r}_{ij}^s \geq \gamma_i^s \gamma_j^s \delta \qquad \forall i, j \in N \mid i \neq j, \ \forall s \in S \tag{30}$$
$$\bar{r}_{ij}^s \geq d_i^s - d_j^s \qquad \forall i, j \in N \mid i \neq j, \ \forall s \in S \tag{31}$$
$$\bar{r}_{ij}^s \geq -(d_i^s - d_j^s) \qquad \forall i, j \in N \mid i \neq j, \ \forall s \in S \tag{32}$$
$$\bar{r}_{ij}^s \leq d_i^s - d_j^s + M\bar{\xi}_{ij}^s \qquad \forall i, j \in N \mid i \neq j, \ \forall s \in S \tag{33}$$
$$\bar{r}_{ij}^s \leq -(d_i^s - d_j^s) + M(1 - \bar{\xi}_{ij}^s) \qquad \forall i, j \in N \mid i \neq j, \ \forall s \in S \tag{34}$$
$$x_j \geq \pi_{ij}^k (x_i + \tau_i) \qquad \forall i \in N, j \in N \mid j \neq i, k \in K \tag{35}$$
$$c_i \geq x_i - \omega_i \qquad \forall i \in N \tag{36}$$
$$c_i \geq 0 \qquad \forall i \in N \tag{37}$$
$$\mathbf{x} \in \mathbb{R}_{\geq 0}^N, \mathbf{A} \in \mathbb{R}_{\geq 0}^{|N| \times |S|}, \mathbf{D} \in \mathbb{R}_{\geq 0}^{|N| \times |S|}$$
$$\mathbf{c} \in \mathbb{R}^{|N|}, \boldsymbol{\xi}, \bar{\boldsymbol{\xi}} \in \{0,1\}^{|N| \times |N| \times |S|}, \mathbf{r}, \bar{\mathbf{r}} \in \mathbb{R}^{|N| \times |N| \times |S|} \tag{38}$$

Program $(\tilde{\mathcal{Q}})$ is a linearized version of $(\mathcal{Q})$ because its feasible region consists of affine functions. This provides significant advantages and, as we will later see, allows us to find a *globally optimal* solution. To perform these linearizations we had to increase

the size of the problem as follows: $(\tilde{\mathcal{Q}})$ has $|N|+2|N|^2|S|$ additional continuous variables, $2|N|^2|S|$ new binary variables which make the formulation mixed-integer, and $8|N|^2|S|+2|N|$ additional inequality constraints. We can see this as the additional computational price we need to pay for guaranteeing global optimality. Summarizing, using the Big-Oh asymptotic notation the size of $(\tilde{\mathcal{Q}})$ is $O(|N|^2|S|)$ continuous variables, $O(|N|^2|S|)$ binary variables, $O(|N|^2|S| + |N|^2|K|)$ inequality constraints, and $O(|N||S|)$ equality constraints. Compared to program $(\mathcal{Q})$, the increase in size when $|N|, |S|, |K| \rightarrow +\infty$ comes from the binary and the continuous variables.

We note here that to avoid changing the order of trip arrivals at the joint corridor, one might also add the constraints:

$$\phi_{ij}^s a_i^s \leq \phi_{ij}^s a_j^s \qquad \forall i, j \in N \mid j \neq i, \forall s \in S \tag{39}$$

and

$$\phi_{ij}^s d_i^s \leq \phi_{ij}^s d_j^s \qquad \forall i, j \in N \mid j \neq i, \forall s \in S \tag{40}$$

4. Solution method

4.1. An active set method for the relaxed inequality QP

Consider the continuous relaxation of $(\tilde{\mathcal{Q}})$, renamed as $(\tilde{\mathcal{Q}}_{\text{rel.}})$. $(\tilde{\mathcal{Q}}_{\text{rel.}})$ has the same formulation as $(\tilde{\mathcal{Q}})$, but the integralities $\xi, \bar{\xi} \in \{0,1\}^{|N|\times|N|\times|S|}$ are now replaced by $\xi, \bar{\xi} \in \mathbb{R}^{|N|}$ and $0 \leq \xi \leq 1$, $0 \leq \bar{\xi} \leq 1$. This allows all variables to take real values. The solution of $(\tilde{\mathcal{Q}}_{\text{rel.}})$ provides a lower bound to the solution of the mixed-integer $(\tilde{\mathcal{Q}})$. This can be exploited by solving $(\tilde{\mathcal{Q}}_{\text{rel.}})$ iteratively with the addition of inequality constraints (i.e., with Branch and Bound (B&B)) resulting in finding the solution of $(\tilde{\mathcal{Q}})$. In fact, if $(\tilde{\mathcal{Q}}_{\text{rel.}})$ can be solved to global optimality, we can guarantee that we can find the globally optimal solution of $(\tilde{\mathcal{Q}})$ with B&B.

$(\tilde{\mathcal{Q}}_{\text{rel.}})$ is a continuous inequality QP. In the following theorem we examine the behavior of $(\tilde{\mathcal{Q}}_{\text{rel.}})$ and we prove its convexity.

Theorem 4.1. *$(\tilde{\mathcal{Q}}_{rel.})$ can be solved to global optimality.*

Proof. The feasible region $(\tilde{\mathcal{Q}}_{\text{rel.}})$ consists of affine equality and inequality constraints. Thus, it forms a polyhedron and it is a *convex set*. In addition, let us consider the scalar-valued function:

$$g_{ij}^s(a_i^s, a_j^s) := \phi_{ij}^s \psi_{ij}^s \left((a_i^s - a_j^s) - h^s \right)^2$$

The Hessian matrix of this function is:

$$\mathbf{H} = \begin{bmatrix} \frac{\partial^2 g_{ij}^s}{\partial (a_i^s)^2} & \frac{\partial^2 g_{ij}^s}{\partial a_i^s a_j^s} \\ \frac{\partial^2 g_{ij}^s}{\partial a_j^s a_i^s} & \frac{\partial^2 g_{ij}^s}{\partial (a_j^s)^2} \end{bmatrix} = \begin{bmatrix} 2\phi_{ij}^s \psi_{ij}^s & -2\phi_{ij}^s \psi_{ij}^s \\ -2\phi_{ij}^s \psi_{ij}^s & 2\phi_{ij}^s \psi_{ij}^s \end{bmatrix}$$

with eigenvalues $e_1 = 0$ and $e_2 = 4\phi_{ij}^s \psi_{ij}^s \geq 0$. Because the eigenvalues are non-negative, g_{ij}^s is convex and the objective function of $(\tilde{\mathcal{Q}}_{\text{rel.}})$ is convex as the sum of convex functions:

$$\sum_{s\in S} \sum_{i\in N} \sum_{j\in N\setminus\{i\}} g_{ij}^s(a_i^s, a_j^s) + \sum_{i\in N} c_i$$

Thus, $(\tilde{\mathcal{Q}}_{\text{rel.}})$ is convex and any locally optimal solution is a globally optimal solution. ■

Corollary 4.1.1. *Strong Duality holds for $(\tilde{\mathcal{Q}}_{rel.})$.*

Proof. $(\tilde{\mathcal{Q}}_{\text{rel.}})$ is convex and its equality and inequality constraints are affine functions. Thus, the revised Slater's constraint qualifications condition, which is sufficient for Strong Duality, holds (Slater, 2014). Ergo, a finite optimal solution of our primal problem has the same objective function value as the optimal solution of the dual problem:

$$\sup_{\nu\geq 0, \lambda} \inf_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \lambda, \nu)$$

where $\mathcal{L}(\mathbf{x}, \lambda, \nu)$ is the Lagrangian of the primal with *dual variables* λ associated to equality constraints and ν associated to inequality constraints. ■

The continuous relaxation can be solved in polynomial time with the active set method. The *active set* method for solving inequality QPs starts by finding a feasible point during an initial phase and then finds iterates that remain feasible while steadily

decreasing the objective function by solving a sequence of *equality-constrained QPs*. Without loss of generality, let us collect all variables $\mathbf{x} \in \mathbb{R}^{N}_{\geq 0}, \mathbf{A} \in \mathbb{R}^{|N|\times|S|}_{\geq 0}, \mathbf{N} \in \mathbb{R}^{|N|\times|S|}_{\geq 0}, \mathbf{c} \in \mathbb{R}^{|N|}, \boldsymbol{\xi}, \bar{\boldsymbol{\xi}} \in \{0,1\}^{|N|\times|N|\times|S|}, \mathbf{r}, \bar{\mathbf{r}} \in \mathbb{R}^{|N|\times|N|\times|S|}$ of $(\tilde{\mathcal{Q}}_{\text{rel.}})$ into column vector $\mathbf{y}$. Let also $\mathcal{F}$ be the feasible region of $(\tilde{\mathcal{Q}}_{\text{rel.}})$ which is comprised of equality and inequality constraints in the form of affine functions. Finally, let $\mathcal{L}$ be the set of equality constraints and $\mathcal{M}$ the set of inequality constraints of $(\tilde{\mathcal{Q}}_{\text{rel.}})$. Then, we have that:

$$c_i(\mathbf{y}) = 0 \quad \forall i \in \mathcal{L}$$

$$c_i(\mathbf{y}) \leq 0 \quad \forall i \in \mathcal{M}$$

where $c_i(\mathbf{y})$ are the affine functions of the problem's constraints.

The *active set* at any feasible point $\mathbf{y}^* \in \mathcal{F}$ of $(\tilde{\mathcal{Q}}_{\text{rel.}})$ is the set $\mathcal{A}(\mathbf{y}^*)$ which is the union of the set of all equality constraints $\mathcal{L}$ and the subset of inequality constraints $\mathcal{M}$ for which $c_i(\mathbf{y}^*) = 0$. That is, $\mathcal{A}(\mathbf{y}^*) = \mathcal{L} \cup \{i \in \mathcal{M} \mid c_i(\mathbf{y}_i^*) = 0\}$. It now follows that if an oracle informs us that the optimal solution $\mathbf{y}^*$ of $(\tilde{\mathcal{Q}}_{\text{rel.}})$ has active set $\mathcal{A}(\mathbf{y}^*)$, then we can solve the inequality QP of $(\tilde{\mathcal{Q}}_{\text{rel.}})$ as an equality-constrained QP by considering only the equality constraints in the active set since all other constraints would be inactive. This is the essence of this iterative numerical optimization method which is based on the following steps:

1. start from a feasible solution $\mathbf{y}_k$ of the inequality QP where $k = 0$ and find its active set $\mathcal{A}(\mathbf{y}_k)$. Note here that k refers to the iteration number of the active-set method and is not to be confused with the symbol used to for the train lines.
2. find the next iterate $\mathbf{y}_{k+1} = \mathbf{y}_k + \boldsymbol{\alpha}_k \mathbf{p}_k$ where $\mathbf{p}_k$ is the step direction and $\boldsymbol{\alpha}_k$ the step length
3. terminate when finding an optimal $\mathbf{y}_{k+1}$

The first question is how to find an initial feasible solution $\mathbf{y}_k$ for $k = 0$. Because this solution does not need to be optimal and all constraints are affine functions, if $(\tilde{\mathcal{Q}}_{\text{rel.}})$ is not infeasible we can easily find a feasible solution. To do so, we would have to turn $(\tilde{\mathcal{Q}}_{\text{rel.}})$ into a *Phase I* linear program by removing the nonlinear objective function of the problem and replacing it with a linear objective function of artificial slack variables. This Phase I linear program can be easily solved with simplex to obtain a basic feasible solution which will also be a feasible (but not necessarily optimal) solution of $(\tilde{\mathcal{Q}}_{\text{rel.}})$.

The second question is how to find the step direction $\mathbf{p}_k$ to update $\mathbf{y}_k$. This is achieved by solving $(\tilde{\mathcal{Q}}_{\text{rel.}})$ as an equality-constrained QP at the active set $\mathcal{A}(\mathbf{y}_k)$ after replacing $\mathbf{y}$ by $\mathbf{y}_k + \mathbf{p}$:

$$\begin{aligned} \min_{\mathbf{p}} \quad & f(\mathbf{y}_k + \mathbf{p}) \\ \text{s.t.} \quad & c_i(\mathbf{y}_k + \mathbf{p}) = 0 \quad \forall i \in \mathcal{L} \\ & c_i(\mathbf{y}_k + \mathbf{p}) = 0 \quad \forall i \in \mathcal{M} \cap \mathcal{A}(\mathbf{y}_k) \end{aligned}$$

where f is the quadratic objective function in Eq. (20). The solution $\mathbf{p}_k$ of this equality-constrained QP can be easily obtained by the null-space method (Gould et al., 2001).

Based on the optimal value of the step length $\mathbf{p}_k$, we have five cases.

Case 1: suppose that the optimal $\mathbf{p}_k$ is equal to 0 and the KKT multipliers of $\boldsymbol{\mu}_k \geq 0$. Then, we can terminate the active set algorithm because we have reached an optimal solution with $\mathbf{y}_{k+1} = \mathbf{y}_k$.

Case 2: suppose that the optimal $\mathbf{p}_k$ is equal to 0, but some components of $\boldsymbol{\mu}_k$ are negative. Then, we remove the most negative component from the active set $\mathcal{A}(\mathbf{y}_k)$ and we solve again the problem to obtain a new $\mathbf{p}_k, \lambda_k, \boldsymbol{\mu}_k$.

Case 3: suppose that the optimal $\mathbf{p}_k \neq 0$. We then need to decide how far we will move in the direction of $\mathbf{p}_k$ to maintain feasibility while reducing the objective function value. For this we choose step length $\boldsymbol{\alpha}_k \in [0,1]$ to be the largest value for which all constraints are satisfied for $\mathbf{y}_{k+1} = \mathbf{y}_k + \boldsymbol{\alpha}_k \mathbf{p}_k$. This $\boldsymbol{\alpha}_k$ is defined based on what happens to the inactive constraints at iteration k and it is obtained from:

$$\boldsymbol{\alpha}_k = \min\{1, \frac{\mathbf{b}_i - \mathbf{a}_i^{\intercal}\mathbf{y}_k}{\mathbf{a}_i^{\intercal}\mathbf{p}_k}, \quad \forall i \notin \mathcal{A}(\mathbf{y}_k) \text{ and } \mathbf{a}_i^{\intercal}\mathbf{p}_k < 0\}$$

Observe that if $\boldsymbol{\alpha}_k < 1$, then $\boldsymbol{\alpha}_k = \frac{\mathbf{b}_i - \mathbf{a}_i^{\intercal}\mathbf{y}_k}{\mathbf{a}_i^{\intercal}\mathbf{p}_k}$ for some $i \notin \mathcal{A}(\mathbf{y}_k)$ with $\mathbf{a}_i^{\intercal}\mathbf{p}_k < 0$. Then, we update the active set $\mathcal{A}(\mathbf{y}_{k+1}) = \mathcal{A}(\mathbf{y}_k) \cup i$ by making constraint i active.

Case 4: suppose that the optimal $\mathbf{p}_k \neq 0$ and $\boldsymbol{\alpha}_k = 1$. Then, there is no inequality constraint which becomes active in the direction of $\mathbf{p}_k$, and thus $\mathbf{y}_{k+1} = \mathbf{y}_k + \mathbf{p}_k$.

Case 5: suppose that the optimal $\mathbf{p}_k \neq 0$ and $\boldsymbol{\alpha}_k = 0$. Then, we have a constraint i which was active at $\mathbf{y}^k$ but we did not include it in the active set $\mathcal{A}(\mathbf{y}_k)$. Thus, we include it now and solve again the problem of determining $\mathbf{p}_k, \lambda_k, \boldsymbol{\mu}_k$.

The steps of the active set method are presented in Alg. 1.

Algorithm 1 Active Set method for inequality QP

set $k = 0$ and start from a feasible solution $\mathbf{y}_k$ of the inequality QP ($\tilde{Q}_{\text{rel.}}$) with the use of linear programming
find the active set $\mathcal{A}(\mathbf{y}_k)$ of the feasible solution $\mathbf{y}_k$
solve ($\tilde{Q}_{\text{rel.}}$) as an equality-constrained QP at the active set $\mathcal{A}(\mathbf{y}_k)$ after replacing $\mathbf{y}$ by $\mathbf{y}_k + \mathbf{p}$:

$$\begin{aligned} \min_{\mathbf{p}} \ & f(\mathbf{y}_k + \mathbf{p}) \\ \text{s.t. } & c_i(\mathbf{y}_k + \mathbf{p}) = 0 \quad \forall i \in \mathcal{L} \\ & c_i(\mathbf{y}_k + \mathbf{p}) = 0 \quad \forall i \in \mathcal{M} \cap \mathcal{A}(\mathbf{y}_k) \end{aligned}$$

if $\mathbf{p}_k = \mathbf{0}$ and the KKT multipliers $\boldsymbol{\mu}_k \geq \mathbf{0}$ **then**
  terminate because solution $\mathbf{y}_k$ is optimal.
end if
if $\mathbf{p}_k = \mathbf{0}$, but some components of $\boldsymbol{\mu}_k$ are negative **then**
  Remove the most negative component from the active set $\mathcal{A}(\mathbf{y}_k)$ and solve again the equality-constrained QP to obtain a new $\mathbf{p}_k, \lambda_k, \boldsymbol{\mu}_k$.
end if
if $\mathbf{p}_k \neq \mathbf{0}$ compute $\alpha_k = \min\{1, \frac{\mathbf{b}_i - \mathbf{a}_i^{\intercal}\mathbf{y}_k}{\mathbf{a}_i^{\intercal}\mathbf{p}_k}, \quad \forall i \notin \mathcal{A}(\mathbf{y}_k) \text{ and } \mathbf{a}_i^{\intercal}\mathbf{p}_k < 0\}$ **then**
  if $\alpha_k = 1$ **then**
    Set $\mathbf{y}_{k+1} = \mathbf{y}_k + \mathbf{p}_k$, $\mathcal{A}(\mathbf{y}_{k+1}) = \mathcal{A}(\mathbf{y}_k)$, and solve again the equality-constrained QP.
  end if
  if $\alpha_k = 0$ **then**
    Select inequality constraint $i \in M \setminus \mathcal{A}(\mathbf{y}_k)$ for which $\mathbf{a}_i^{\intercal}\mathbf{p}_k > 0$ and $b_i = \mathbf{a}_i^{\intercal}\mathbf{y}_k$
    Set $\mathbf{y}_{k+1} = \mathbf{y}_k$ and $\mathcal{A}(\mathbf{y}_{k+1}) = \mathcal{A}(\mathbf{y}_k) \cup \{i\}$ and solve again the equality-constrained QP.
  end if
  if $0 < \alpha_k < 1$ **then**
    Select $i \in M \setminus \mathcal{A}(\mathbf{y}_k)$ with $\mathbf{a}_i^{\intercal}\mathbf{p}_k > 0$ for which $\alpha_k = \frac{b_i - \mathbf{a}_i^{\intercal}\mathbf{y}_k}{\mathbf{a}_i^{\intercal}\mathbf{p}_k}$.
    Set $\mathbf{y}_{k+1} = \mathbf{y}_k + \alpha_k \mathbf{p}_k$ and $\mathcal{A}(\mathbf{y}_{k+1}) = \mathcal{A}(\mathbf{y}_k) \cup \{i\}$ and solve again the equality-constrained QP.
  end if
end if

4.2. Obtain an exact solution of ($\tilde{Q}$) with branch and bound

In the previous subsection we proved that the continuous relaxation of ($\tilde{Q}$), denoted as ($\tilde{Q}_{\text{rel.}}$), can be solved to global optimality and one can employ the active set method to compute a globally optimal solution. However, the globally optimal solution of ($\tilde{Q}_{\text{rel.}}$) is merely a lower bound of ($\tilde{Q}$) and not its globally optimal solution. Hereby, we propose a Branch and Bound (B&B) solution method to derive the globally optimal solution of ($\tilde{Q}$). B&B was proposed by Ailsa Land and Alison Doig in 1960 for mixed-integer linear programming (Land and Doig, 1960).

B&B is a more clever search technique compared to exhaustive enumeration (brute force) because it searches the solution space more efficiently, typically requiring to solve considerably fewer sub-problems than the brute force method. We should note, though, that the running time of B&B is still exponential at the worst-case scenario. Although B&B is not a polynomial time algorithm, it can perform well in practical applications bringing significant benefits compared to naive exact methods, such as the exhaustive enumeration.

($\tilde{Q}$) is a mixed-integer QP with binary and continuous variables. We start implementing B&B with iteration $k = 0$ by solving the continuous relaxation of ($\tilde{Q}$), denoted as ($\tilde{Q}_{\text{rel.}}$). Solving this continuous relaxation with the active-set method will give us solution $\mathbf{y}_0$ with performance $f(\mathbf{y}_0)$, where f is the quadratic objective function in Eq. (20). Solution $\mathbf{y}_0$ is a globally optimal solution of ($\tilde{Q}_{\text{rel.}}$) and a lower bound of ($\tilde{Q}$). Let $L_0 = f(\mathbf{y}_0)$ denote this lower bound. Let also $U_0 = +\infty$ denote the upper bound of ($\tilde{Q}$).

With B&B we perform iterations until the lower and upper bound of ($\tilde{Q}$) are equal, denoting that we have found the globally optimal solution of ($\tilde{Q}$). The upper bound value U_0 can be updated every time we find a feasible solution of ($\tilde{Q}$) with better performance. If, for instance, solution $\mathbf{y}_0$ happens to be a feasible solution of ($\tilde{Q}$), then $U_0 = f(\mathbf{y}_0)$, and because $L_0 = U_0$ the B&B algorithm terminates with $\mathbf{y}_0$ as the globally optimal solution. If $\mathbf{y}_0$ is not a feasible solution of ($\tilde{Q}$), then we proceed to the next iteration $k = k+1$. Each iteration populates the B&B tree where every leaf node of the tree represents an optimization problem. Initially, we have leaf node 0 representing problem ($\tilde{Q}_{\text{rel.}}$), which is the root of the B&B tree. From the B&B tree we have the *candidate list* $\mathcal{P} = \{0\}$ of the leaf nodes from which we have not performed branching.

At iteration k we select the leaf node from the candidate list $\mathcal{P}$ that has the smallest lower bound value and we generate two sub-problems (branching). Because at $k = 1$ we have only one problem in the candidate list, we select leaf node 0 representing ($\tilde{Q}_{\text{rel.}}$) and we generate two sub-problems by selecting one binary variable that had a continuous value in the solution $\mathbf{y}_0$. In the first sub-problem the value of this variable is set to 0 and in the second to 1. These two sub-problems (1 and 2) form two new leaf nodes and are subsequently solved with the active set method to derive their optimal solutions ($\mathbf{y}_1$ and $\mathbf{y}_2$), and their lower bound

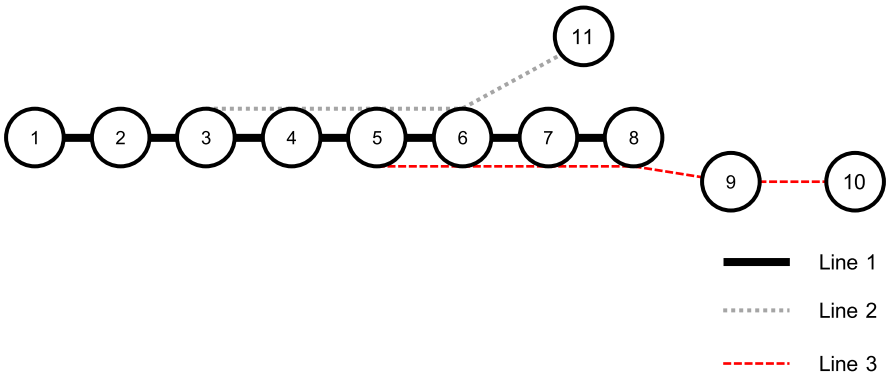

Fig. 2. Toy network.

values $L_1 = f(\mathbf{y}_1)$ and $L_2 = f(\mathbf{y}_2)$. If any of the solutions $\mathbf{y}_1$ and $\mathbf{y}_2$ is a feasible solution to $(\tilde{\mathcal{Q}})$, then we update the upper bound. If, for instance, $\mathbf{y}_1$ is a feasible solution to $(\tilde{\mathcal{Q}})$, then $U_k = f(\mathbf{y}_1)$. If this is the case, we can start the pruning phase where we remove all leaf nodes from $\mathcal{P}$ for which their lower bound values are greater than U_k since they cannot be globally optimal solutions. This iteration terminates by removing leaf node 0 from $\mathcal{P}$ since we branched out from this leaf node.

The iterations continue in a similar manner by finding the leaf node in $\mathcal{P}$ with the smallest lower bound value, creating two sub-problems, solving them with the active-set method to derive their solutions and lower bounds, checking if any of the two solutions is feasible to $(\tilde{\mathcal{Q}})$ with performance lower than U_k and updating U_k accordingly, and pruning leaf nodes from $\mathcal{P}$ that have a higher lower bound value than U_k. With this branching and pruning the algorithm terminates in an iteration where the smallest lower bound of all leaf nodes in $\mathcal{P}$ is equal to the upper bound value. The algorithmic steps are presented in Alg. 2.

Algorithm 2 B&B approach

1: solve the convex continuous relaxation $(\tilde{\mathcal{Q}}_{\text{rel.}})$ of the original problem $(\tilde{\mathcal{Q}})$ to compute L_0 with the active-set method. If this solution is feasible for the original problem, then $U_0 = L_0$. Otherwise, set $U_0 = +\infty$
2: Set $k = 0$ and $L_k = L_0$, $U_k = U_0$
3: initialize *candidate list* $\mathcal{P} = \{0\}$ which contains all leaf nodes from which we have not performed branching
4: **repeat**
5: from the leaf node in $\mathcal{P}$ with the smallest lower bound, generate two sub-problems by fixing the values of one of the binary variables $\boldsymbol{\xi}, \tilde{\boldsymbol{\xi}} \in \{0,1\}^{|N|\times|N|\times|S|}$ to 0 and 1 (branching)
6: for the two generated sub-problems, compute their lower bounds with the active-set method and their upper bounds in case of feasibility
7: add these two sub-problems to $\mathcal{P}$ and remove the parent leaf node of these two sub-problems from $\mathcal{P}$
8: set L_k as the smallest lower bound from the lower bounds of all sub-problems in $\mathcal{P}$
9: set U_k as the smallest upper bound of all sub-problems
10: remove all leaf nodes (sub-problems) from $\mathcal{P}$ for which their lower bound values are greater than U_k (pruning)
11: set $k = k + 1$
12: **until** $U_k = L_k$

5. Numerical experiments in a toy network

We first present the implementation of our urban rail corridor coordination model in a toy network (Fig. 2).

This toy network has $K = (1,2,3)$ train lines. It also has $(1,2,3,4,5,6,7,8,9,10,11)$ train stations. Trains have to operate in both directions, so we also have additional stations $(12,13,14,15,16,17,18,19,20,21,22)$ demonstrating the movement in the reverse direction, where station 12 symbolizes the opposite direction station of 11, 13 the opposite direction station of 10, and so on. In addition, we have a set of 6 train trips $N = (1,2,3,4,5,6)$. Trips 1 and 4 are successive trips of line 1, $P^1_{1,4} = 1$, trips 2 and 5 are successive trips of line 2, $P^2_{2,5} = 1$, and trips 3 and 6 are successive trips of line 3, $P^3_{3,6} = 1$. The ordered stations of line 1 are $S^1 = (1,2,3,4,5,6,7,8)$ in the main direction and $(15,16,17,18,19,20,21,22)$ in the opposite, of line 2 are $S^2 = (3,4,5,6,11)$ in the main direction and $(12,17,18,19,20)$ in the opposite, and of line 3 are $S^3 = (5,6,7,8,9,10)$ in the main direction and $(13,14,15,16,17,18)$ in the opposite. Trips 1 and 4 are assigned to line 1, trips 2 and 4 to line 2, and trips 3 and 6 to line 3. That is, $\beta_1 = \beta_4 = 1$, $\beta_2 = \beta_5 = 2$, and $\beta_3 = \beta_6 = 3$.

Trips 1 and 4 are successively operated by the same train ($\pi^1_{1,4} = 1$). The same applies for trips 2 and 5 ($\pi^2_{2,5} = 1$), and trips 3 and 6 ($\pi^3_{3,6} = 1$). The safety distance among successive trains is set to $\delta = 1$. The minimum allowed dispatching headway between

Table 2
Optimal trip arrival times at the stations of the joint corridor in minutes.

Station	Trip 1	Trip 2	Trip 3	Station	Trip 4	Trip 5	Trip 6
3	3	4	–	15	11.869	–	16.869
4	4	5	–	16	12.869	–	17.869
5	5	6	7	17	13.869	14.869	18.869
6	6	7	8	18	14.869	15.869	19.869
7	7	–	9	19	15.869	16.869	–
8	8	–	10	20	16.869	17.869	–

successive trips of any train line $k \in K$ is set to $\upsilon_k = 2$, and the maximum allowed dispatching headway is set to $u_k = 20$. The dwell time at stops is set to $\eta = 0.1$ minutes.

In addition, we initialize the target headways $\mathbf{H}$, the indicator parameters $\boldsymbol{\Phi}$ that indicate whether we should meet a target headway between two trips at a station, and the weight factor parameters $\boldsymbol{\Psi}$ as:

- $h^s_{ij} = \{(i, j, s) : 0 \text{ for } i \text{ in } N \text{ for } j \text{ in } N \text{ for } s \text{ in } S\}$
- $\phi^s_{ij} = \{(i, j, s) : 0 \text{ for } i \text{ in } N \text{ for } j \text{ in } N \text{ for } s \text{ in } S\}$
- $\psi^s_{ij} = \{(i, j, s) : 0 \text{ for } i \text{ in } N \text{ for } j \text{ in } N \text{ for } s \text{ in } S\}$

We then set their values as follows:
for i in N:
  for j in N:
    for s in S:
      if $\gamma^s_i = 1$ and $\gamma^s_j = 1$ and $i \neq j$: $h^s_{i,j} = 2$ minutes
      $\psi^s_{i,j} = 1.5$
      if $i = 1$ and $j = 2$: $\phi^s_{i,j} = 1$
      if $i = 2$ and $j = 3$: $\phi^s_{i,j} = 1$
      if $i = 4$ and $j = 5$: $\phi^s_{i,j} = 1$
      if $i = 5$ and $j = 6$: $\phi^s_{i,j} = 1$

In addition to the above, we set the latest possible dispatching time to ensure that the train trips in the opposite direction will not be delayed and thus avoid schedule sliding as $\omega_i = 14$ minutes for all $i \in N$, and $M = 10000$. Solving the corridor coordination problem in this toy network results in a model with 1572 variables and 432 constraints. The solution time of this problem implemented on a computer with Intel Core i7-9750H CPU @ 2.60 GHz was 0.04 s.

The optimal solution of this problem after implementing Alg. 2 results in the new trip dispatching times:

$$x_1 = 0, x_2 = 3, x_3 = 6, x_4 = 10.869, x_5 = 12.869, x_6 = 13.869$$

with

$$c_i = 0 \ \forall i \in N$$

resulting in a globally optimal solution with objective function value 591 and an optimality gap of 0%. The arrival times of the train trips at the stations of the joint corridor are presented in Table 2.

6. Case study

6.1. Description

The potential impact of urban rail corridor coordination is demonstrated on a light rail service in The Hague, The Netherlands. The Hague is a city on the North Sea coast of the western Netherlands, and it is the seat of the Dutch parliament. It has a population of more than half a million inhabitants and a public transport network consisting of buses and trams (light rail) operated by HTM. The data used for the case study is the network and timetable data (GTFS) and passenger demand data – processed from smart card data (see Yap et al., 2017 for information on the data processing and inferences performed) – from 2018. The network consists of 7 light rail lines which share tracks in the tram tunnel that runs through the central parts of The Hague. Three of the seven lines are supporting peak hour service lines that are operated but not communicated as separate lines. The network consists of a total of 99 stations. The characteristics of the 7 lines are presented in Table 3. Note that 3 of them (2K, 3K, and 4K) are peak-hour lines and do not operate throughout the day.

The network configuration is presented in Fig. 3. In this network, the supporting peak-hour service lines 2K, 3K, and 4K serve parts of the main lines 2, 3, and 4, respectively.

In our case study, we analyze the trips that serve the shared corridor between 16:50-17:50 because this is the peak afternoon period during which travel demand and service frequencies are at the highest level and all 7 lines are operational. Note that line 2K has only one departure per hour (providing a short-line addition for line 2). The basis for the analysis is the timetable as it was

Table 3
Characteristics of the service lines.

Line	Operational period	Frequency	Nr. of stations	Total trip time
2	04:45-00:30	6 trips/h	31	0:41:53
2K (peak-only)	07:30-08:45 and 16:45-17:15	5 trips/day	20	0:32:51
3	06:15-00:30	6 trips/h	41	1:08:35
3K (peak-only)	07:15-09:15 and 16:50-17:50	6 trips/h	15	0:21:37
4	06:00-01:00	6 trips/h	31	0:55:46
4K (peak-only)	07:30-09:15 and 16:10-18:10	6 trips/h	23	0:39:21
6	05:30-00:15	6 trips/h	28	0:37:22

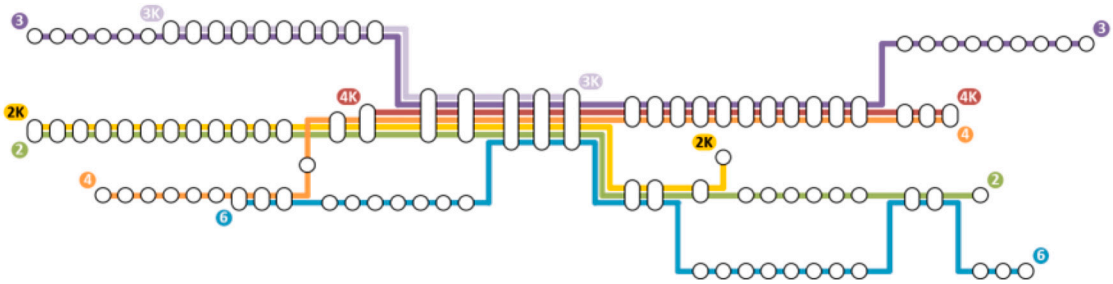

Fig. 3. Network topology of the urban rail lines.

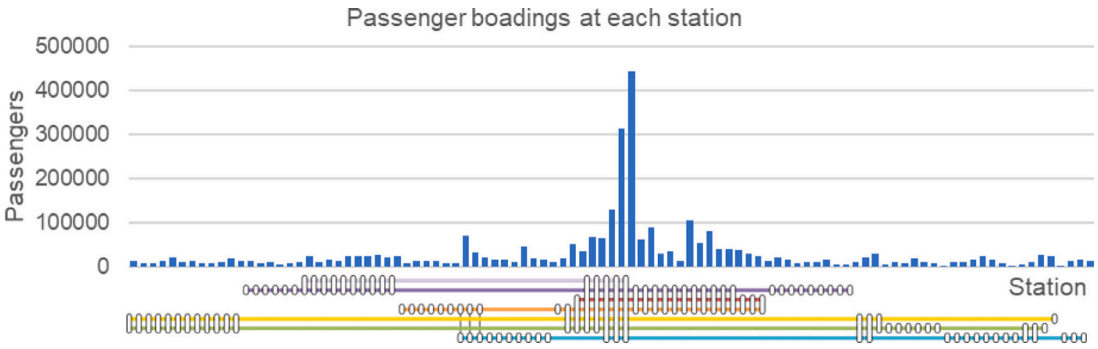

Fig. 4. Number of boarding passengers at each station.

implemented on 11-09-2018. The weight factor ψ_{ij}^{s} is calculated for each station based on the number of boarding passengers. The weight is computed as a normalized value of the number of boarding passengers as follows:

$$\psi_{ij}^{s} := \frac{B^{s}}{\sum_{s \in S} B^{s}} \quad \forall s \in S,\ \forall i, j \in N \tag{41}$$

where B^{s} is the number of passengers boarding at station s. The passenger demand data with the number of boardings at each station is presented in Fig. 4. This data is retrieved from the actual smart card data records available from 11-09-2018.

Note that the weight factor ψ_{ij}^{s} is calculated for each station based on the number of boarding passengers. That is, the weight factor value is affected from the number of boarding passengers, and not their arrival pattern; implying that the arrivals of passengers at stations are uniformly distributed (which is the most common assumption for urban public transport services operating in high frequencies (Gkiotsalitis, 2023)).

Based on the original schedule retrieved from the GTFS data, the safety distance parameter δ has a value of 2 s. The minimum dispatching time is $\upsilon_k = 120$ seconds and the maximum dispatching time $u_k = 1200$ seconds for all seven lines. The dwell time per station is retrieved from the original timetable.

Since dwell times are constant and the travel time for traversing common line segments is line-agnostic, it is possible to speed up the calculation by simplifying the network by means of aggregating stations served by the same lines by summing the respective weights ψ_{ij}^{s}. This reduces the number of stations from 99 to 22 resulting in the simplified network of Fig. 5.

6.2. As-is scenario

In the original timetable of 2018, all individual urban rail lines have a time headway of 10 min. In Fig. 5 we present how the 10-minute headways of the lines are positioned in relation to one another. At multiple stations, the headways are not equally divided

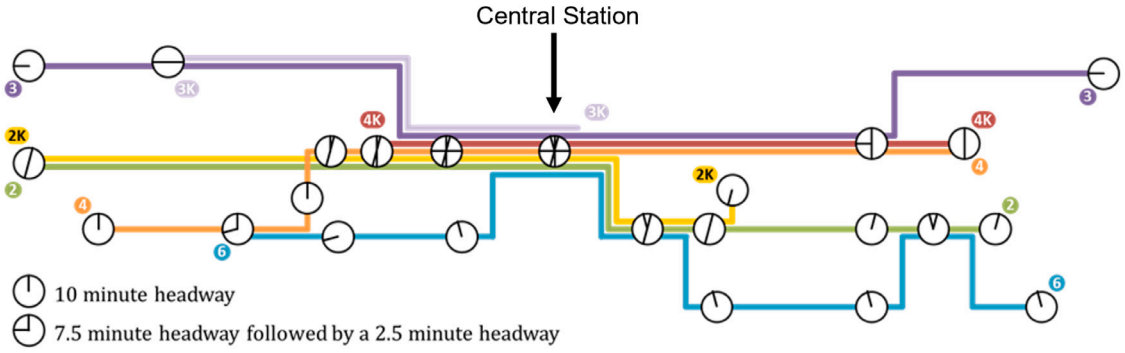

Fig. 5. Time headways in the original 2018 timetable.

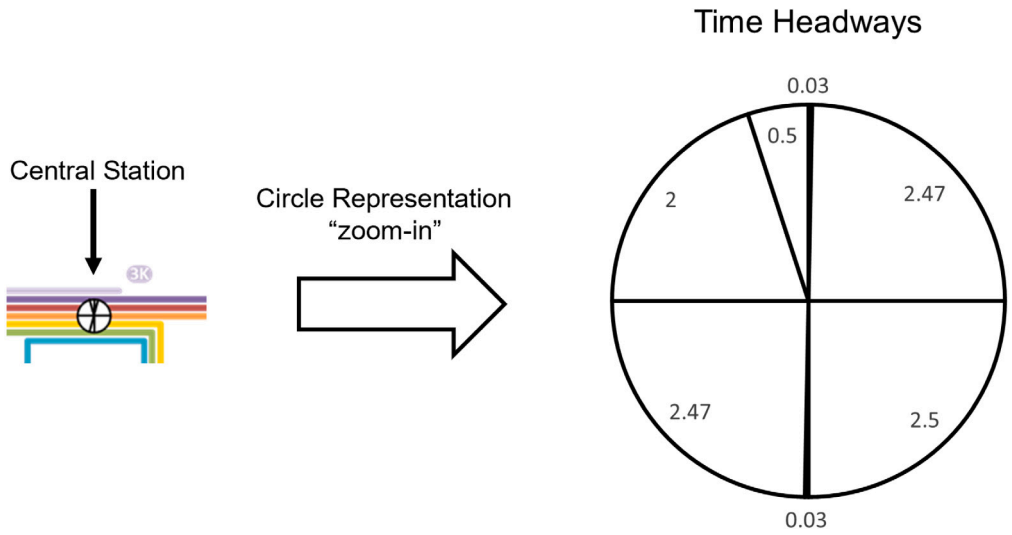

Fig. 6. Example of the full circle representation at the Central Station, where the time headways of the consecutive trips of the seven lines are expressed in minutes.

over the 10-minute cycles. For example, at the Central Station which is served by all seven lines, headways between successive trips are consecutively 0.03 min, 2.47 min, 2.50 min, 0.03 min, 2.47 min, 2.00 min, and 0.50 min; thereby, yielding a very uneven service arrival pattern with the expected ramifications for passengers' waiting times and uneven on-board crowding conditions due to bunching. This also causes safety issues because the gap between successive vehicles is as low as 0.03 min.

In Fig. 5, a full circle corresponds to a 10-minute cycle. The positions of lines within the circle represent the arrival of trips within the 10-minute interval. Note that all lines within our network have a 10-minute headway. Stations with small variations in headways have lines distributed evenly within their respective circles, whereas unevenly distributed circles represent stations with large variations in headways. For instance, the first station of line 4 has an evenly distributed headway of 10 min. In contrast, the Central Station has an uneven service arrival pattern. Note that the position of line 6 in the 10-minute cycle changes along its route, as it has different travel times than lines 4 and 2 between their shared stations.

To clarify the meaning of the full circles, consider the full circle at the Central Station with consecutive time headways between the trips of the seven lines being 0.03 min, 2.47 min, 2.50 min, 0.03 min, 2.47 min, 2.00 min, and 0.50 min. This is represented in Fig. 6 which provides more information about the circle used at the Central Station.

We note that at the Central Station, which is the most visited station, we have 10 platforms and our lines are served by platforms B and D.

6.3. Results of model implementation

The model formulated in Section 3 and the solution method of the B&B algorithm coupled with the Active Set method (Alg. 2) are implemented in Python on a computer with AMD Ryzen 7 5700U CPU @ 1.80 GHz processor and a 16 GB RAM. The globally optimal solution for this network with 7 lines and 22 stations is found in 2,928.26 s. The results of the model are presented in Table 4. The number of binary variables is 57,024.

Table 4
Computation time, variables and constraints of the exact urban rail corridor coordination model.

Number of binary variables	57,024
Number of constraints	351,452
Computation time	2,928.26 s
Optimality Gap	0% (optimal solution)
Objective function value	2,213,745 s^2
Objective function value of the original timetable	7,999,773 s^2

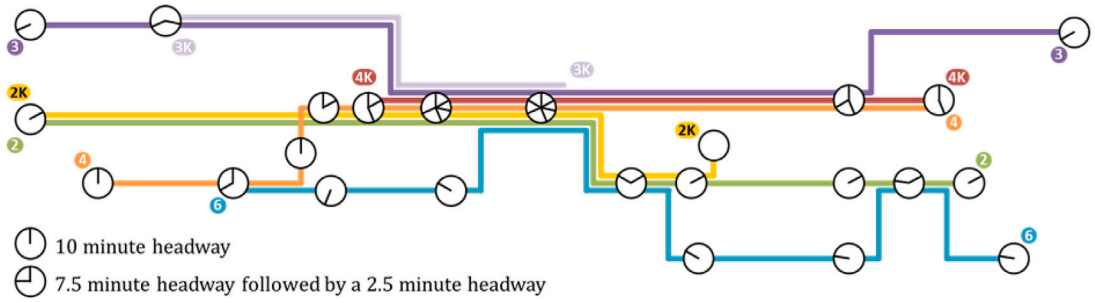

Fig. 7. Time headways in the optimized timetable.

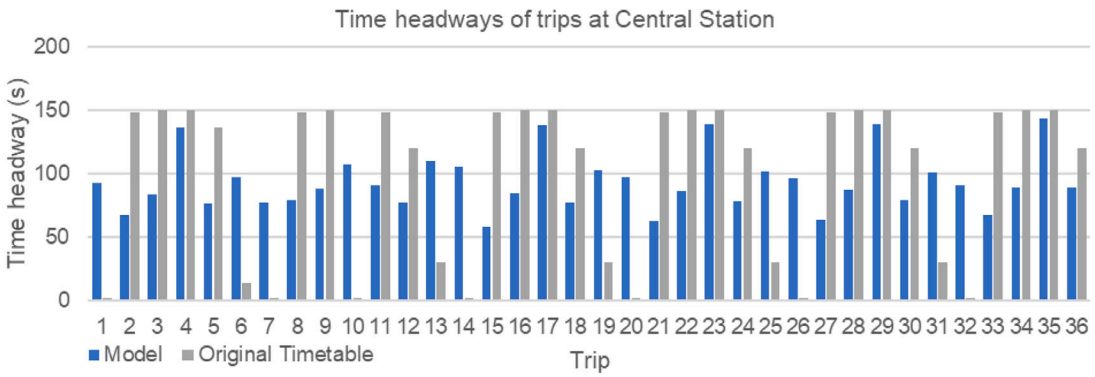

Fig. 8. Example of time headways of successive trips at Central Station within the hourly period 16:50-17:50, 11-09-2018.

The solution of the optimization process results in a new set of dispatching times for the vehicle trips of the urban rail line. Based on this solution, the time headways at the stations of the urban rail network are presented in Fig. 7.

In Fig. 7, a full circle is a 10-minute cycle. The positions of lines within the circle represent the arrival of trips within the 10-minute interval. Evenly distributed circles are stations with small variations in headways and unevenly distributed circles represent stations with large variations in headways. Note that the optimized solution improved the distribution of headways at the common corridor. For instance, the time headways of successive trips at the Central Station are more evenly distributed compared to the time headways of the original timetable presented in Fig. 5. As a result, there is no vehicle bunching in the common corridor.

In general, our corridor coordination model spaces the trips more evenly over time. It is notable that the model avoids very short time headways which are observed in the original timetable of Fig. 5. The minimum time headway between successive trips on the network increases from 2 s to 57 s, while maintaining the same average headway. This results in reduced bunching for the same service frequency (see Fig. 8). Note that by planning for more evenly distributed vehicle arrivals at the common corridor, there is lesser need for complex real-time control strategies aimed at reducing bunching during operations. In Fig. 8 we present the successive time headways between the 37 trips of all lines that serve the Central Station from 16:50 until 17:50.

From Fig. 8 one can note that several time headways of the original timetable are almost 0 (i.e., the time headways 1, 7, 14, 20, 26, and 32). On the contrary, the optimized time headways after the implementation of our model are never below 57 s, resulting in reduced bunching.

From the results it becomes evident that the large variations in time headways that were present at some stations of the joint corridor are reduced in our optimized timetable. This, however, might increase the headway variation at isolated network segments that are served only by a single line. This is presented in Fig. 9 where the coefficient of variation of time headways at the isolated stations 2, 3, 4, 6, 7, 13, 14, 17, 18, 20, 21, and 22 – which are served only by one line – is increased compared to the respective coefficient of variation of the original timetable. The coefficient of variation at each station is calculated as the standard deviation

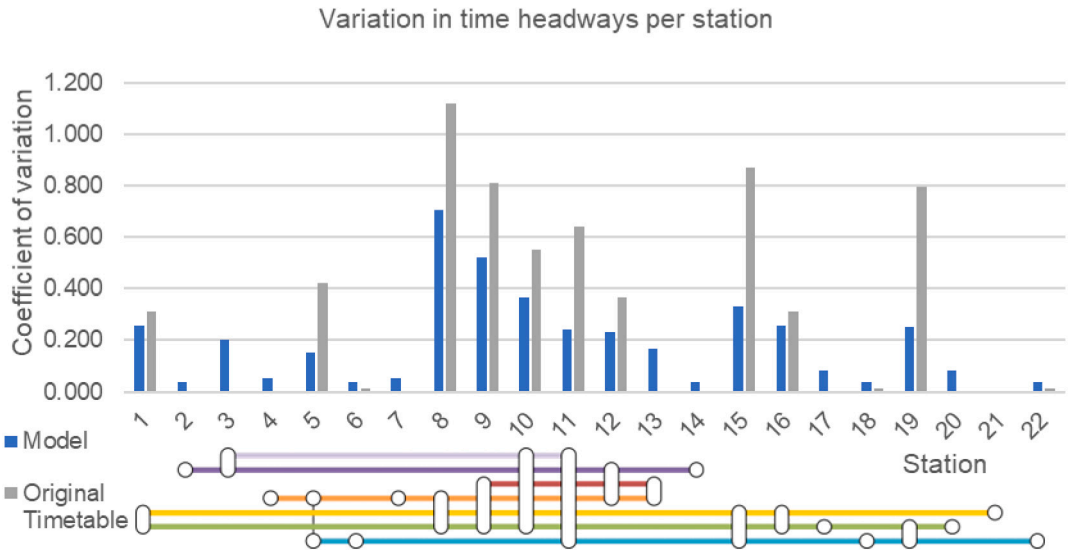

Fig. 9. Coefficient of variation of the time headways at each station.

Table 5
Mean time headways (in seconds) and coefficient of variation (in seconds) of time headways at the different stations.

Station	Model		Original timetable		Change of variation
	Mean	Coefficient of variation	Mean	Coefficient of variation	
1	480	0.255	500	0.310	−0.055
2	579	0.038	600	0.000	0.038
3	283	0.204	300	0.000	0.204
4	577	0.053	600	0.000	0.053
5	292	0.152	311	0.423	−0.272
6	581	0.036	597	0.012	0.024
7	577	0.053	600	0.000	0.053
8	248	0.703	250	1.117	−0.413
9	174	0.520	183	0.812	−0.291
10	109	0.367	115	0.551	−0.184
11	93	0.241	99	0.640	−0.399
12	193	0.234	203	0.364	−0.130
13	285	0.170	300	0.000	0.170
14	579	0.038	600	0.000	0.038
15	273	0.329	297	0.870	−0.541
16	480	0.255	500	0.310	−0.055
17	576	0.081	600	0.000	0.081
18	581	0.036	597	0.012	0.024
19	295	0.251	322	0.795	−0.544
20	576	0.081	600	0.000	0.081
21	0	0.000	0	0.000	−0.000
22	581	0.036	597	0.012	0.024

of the headways of consecutive trips arriving at this station divided by the mean value of the headway at this station. Figs. 8 and 9 illustrate the trade-off between attaining more even headways at the common corridor and increasing the variation of headways at isolated segments of the urban rail network.

The detailed change in the coefficient of variation at the time headway of each station before and after applying our urban rail corridor coordination model is presented in Table 5.

It is important to note that the coefficient of variation of time headways at stations that are served by more than one line (namely, stations 1, 5, 8, 9, 10, 11, 12, 13, 15, 16, 17) has been significantly reduced after the implementation of our proposed corridor coordination model. This has resulted in minor coefficient of variation increases at stations served by a single line (namely, stations 2, 3, 4, 6, 7, 14, 17, 18, 20, 21, 22). That is, by performing the same number of trips, there is a significant improvement of the distribution of time headways at the common corridor in the expense of a slight increase of headway variability at isolated stations with a small number of passengers.

Table 6
Computation time of the exact urban rail corridor coordination model.

Time period	No of trips	Binary variables	Computation time (h)
1 h	36	57,024	0.813
2 h	72	228,096	1.127
3 h	108	513,216	2.214
4 h	144	912,384	6.127
5 h	180	1,425,600	11.423

6.4. Computational times

We previously presented the improvement potential of the implementation of our method in a joint corridor. To understand the practicality of our approach in terms of computational complexity, we perform additional experiments. Because the computational complexity of our approach is mainly affected from the binary variables of our problem since the continuous relaxation of it can be solved in polynomial time, we concentrate on settings where the number of binary variables increases. We have already shown that the number of binary variables in our model is $2|N|^2|S|$, where $|N|$ is the total number of trips of all lines operating in the joint corridor during the time period of our selection, and $|S|$ the total number of stations.

In our case study, we have $|S| = 22$ stations. In addition, during the 1-hour period of our experiments we had 36 trips, resulting in $2 \cdot 36^2 \cdot 22 = 57{,}024$ binary variables. Given that the main driving force of the increase in the binary variables is the number of trips $|N|$ because its value is squared, we perform additional computational experiments considering longer time periods of the day. The results are presented in Table 6, where column 1 presents the time period, column 2 the number of trips, column 3 the number of binary variables, and column 4 the computational cost in hours. The experiments are performed in a conventional computer machine with an AMD Ryzen 7 5700U CPU @ 1.80 GHz processor and a 16 GB RAM.

7. Concluding remarks

This study developed a corridor coordination model for urban rail networks. It particularly focused on the pressing issue of vehicle bunching at the stations of a joint corridor which results in safety-related issues and increased waiting times for a high number of passengers that travel through these dense corridors. The developed model was a nonlinear and non-convex mathematical program which was reformulated as a mixed-integer quadratic program. The resulting mixed-integer quadratic program has a convex objective function and can be solved to global optimality. A branch-and-bound algorithm coupled with the active-set method was proposed as a solution method for this problem.

This model was implemented in the urban rail network of The Hague, focusing on 7 lines with a common corridor. Using actual passenger demand and vehicle trip data from 11-09-2018, we demonstrated the improvement potential in the main part of the urban rail network in The Hague. Because the original timetables of the urban rail network lines are produced with the objective of improving the regularity of each service line or synchronizing the vehicle arrivals at common stations (Liebchen, 2007; Liu et al., 2021), the vehicle trips at a common corridor have uneven headways resulting in bunching and long/uneven waiting times for passengers. Our approach alleviates this issue by producing more even arrivals between vehicle trips of different lines that serve a common corridor at the expense of increasing the irregularity at low-demand, isolated network segments.

In future research, our approach can be used to reduce vehicle bunching in conjunction with targeted, real-time control measures such as vehicle holding at stations and speed control. Although our approach is implemented at the tactical planning level where we use estimates of inter-station travel times from historical data, one can extend our model to cater for travel time uncertainties by adopting a stochastic optimization formulation. Closing, by incorporating constraints related to the required maintenance of tracks and trains, our model can also be applied in case of anticipated disruptions in the urban rail network.

CRediT authorship contribution statement

K. Gkiotsalitis: Conceptualization, Methodology, Software, Writing – original draft, Writing – review & editing. **O. Cats:** Conceptualization, Methodology, Writing – review & editing. **T. Liu:** Conceptualization, Methodology, Writing – review & editing. **J.M. Bult:** Data curation, Software.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

Arnold, P., Peeters, D., Thomas, I., 2004. Modelling a rail/road intermodal transportation system. Transp. Res. E Logist. Transp. Rev. 40 (3), 255–270.

Ávila-Torres, P., López-Irarragorri, F., Caballero, R., Ríos-Solís, Y., 2018. The multimodal and multiperiod urban transportation integrated timetable construction problem with demand uncertainty. J. Ind. Manag. Optim. 14 (2), 447.

Bartholdi III., J.J., Eisenstein, D.D., 2012. A self-coördinating bus route to resist bus bunching. Transp. Res. B 46 (4), 481–491.

Cadarso, L., Marín, Á., Maróti, G., 2013. Recovery of disruptions in rapid transit networks. Transp. Res. E Logist. Transp. Rev. 53, 15–33.

Canca, D., Barrena, E., De-Los-Santos, A., Andrade-Pineda, J.L., 2016. Setting lines frequency and capacity in dense railway rapid transit networks with simultaneous passenger assignment. Transp. Res. B 93, 251–267.

Canca, D., De-Los-Santos, A., Laporte, G., Mesa, J.A., 2019. Integrated railway rapid transit network design and line planning problem with maximum profit. Transp. Res. E Logist. Transp. Rev. 127, 1–30.

Ceder, A.A., 2011. Optimal multi-vehicle type transit timetabling and vehicle scheduling. Procedia-Soc. Behav. Sci. 20, 19–30.

Ceder, A., 2016. Public Transit Planning and Operation: Modeling, Practice and Behavior. CRC Press.

Ceder, A., Golany, B., Tal, O., 2001. Creating bus timetables with maximal synchronization. Transp. Res. A 35 (10), 913–928.

Ceder, A.A., Hassold, S., Dano, B., 2013. Approaching even-load and even-headway transit timetables using different bus sizes. Public Transp. 5 (3), 193–217.

D'Acierno, L., Botte, M., Placido, A., Caropreso, C., Montella, B., 2017. Methodology for determining dwell times consistent with passenger flows in the case of metro services. Urban Rail Transit 3 (2), 73–89.

Daduna, J.R., Branco, I., Paixao, J.M., 1995. Computer-Aided Transit Scheduling: Proceedings of the Sixth International Workshop on Computer-Aided Scheduling of Public Transport, Vol. 430. Springer Science & Business Media.

Daduna, J.R., Voß, S., 1995. Practical experiences in schedule synchronization. In: Computer-Aided Transit Scheduling. Springer, pp. 39–55.

Daganzo, C.F., 2009. A headway-based approach to eliminate bus bunching: Systematic analysis and comparisons. Transp. Res. B 43 (10), 913–921.

Fonseca, J.P., van der Hurk, E., Roberti, R., Larsen, A., 2018. A matheuristic for transfer synchronization through integrated timetabling and vehicle scheduling. Transp. Res. B 109, 128–149.

Furth, P.G., Wilson, N.H., 1981. Setting frequencies on bus routes: Theory and practice. Transp. Res. Rec. 818 (1981), 1–7.

Geurs, K., Grigolon, A., Münzel, K., Gkiotsalitis, K., Duran-Rodas, D., Büttner, B., Kirchberger, C., Pappers, J., Martinez Ramirez, L., Graf, A., et al., 2023. The smarthubs integration ladder: a conceptual model for the categorisation of shared mobility hubs. Transp. Rev. 1–28.

Gkiotsalitis, K., 2023. Public Transport Optimization. Springer Nature, Switzerland, http://dx.doi.org/10.1007/978-3-031-12444-0.

Gkiotsalitis, K., Alesiani, F., 2019. Robust timetable optimization for bus lines subject to resource and regulatory constraints. Transp. Res. E Logist. Transp. Rev. 128, 30–51.

Gkiotsalitis, K., Cats, O., 2018. Reliable frequency determination: Incorporating information on service uncertainty when setting dispatching headways. Transp. Res. C 88, 187–207.

Gkiotsalitis, K., Cats, O., 2020. Timetable recovery after disturbances in metro operations: an exact and efficient solution. IEEE Trans. Intell. Transp. Syst.

Gkiotsalitis, K., Cats, O., Liu, T., 2022. A review of public transport transfer synchronisation at the real-time control phase. Transport reviews 1–20.

Gkiotsalitis, K., Maslekar, N., 2018. Towards transfer synchronization of regularity-based bus operations with sequential hill-climbing. Public Transp. 10 (2), 335–361.

Gkiotsalitis, K., Van Berkum, E., 2020a. An analytic solution for real-time bus holding subject to vehicle capacity limits. Transp. Res. C 121, 102815.

Gkiotsalitis, K., Van Berkum, E., 2020b. An exact method for the bus dispatching problem in rolling horizons. Transp. Res. C 110, 143–165.

Gould, N.I., Hribar, M.E., Nocedal, J., 2001. On the solution of equality constrained quadratic programming problems arising in optimization. SIAM J. Sci. Comput. 23 (4), 1376–1395.

Hadas, Y., Ceder, A., 2008. Improving bus passenger transfers on road segments through online operational tactics. Transp. Res. Rec. 2072 (1), 101–109.

Hadas, Y., Ceder, A.A., 2010. Optimal coordination of public-transit vehicles using operational tactics examined by simulation. Transp. Res. C 18 (6), 879–895.

Hickman, M.D., 2001. An analytic stochastic model for the transit vehicle holding problem. Transp. Sci. 35 (3), 215–237.

Horjus, J., Gkiotsalitis, K., Nijënstein, S., Geurs, K., 2022. Integration of shared transport at a public transport stop: mode choice intentions of different user segments at a mobility hub. J. Urban Mobil. 2, 100026.

Huang, K., Wu, J., Liao, F., Sun, H., He, F., Gao, Z., 2021. Incorporating multimodal coordination into timetabling optimization of the last trains in an urban railway network. Transp. Res. C 124, 102889.

Ibarra-Rojas, O.J., Giesen, R., Rios-Solis, Y.A., 2014. An integrated approach for timetabling and vehicle scheduling problems to analyze the trade-off between level of service and operating costs of transit networks. Transp. Res. B 70, 35–46.

Kuipers, R.A., Palmqvist, C.-W., Olsson, N.O., Winslott Hiselius, L., 2021. The passenger's influence on dwell times at station platforms: a literature review. Transp. Rev. 41 (6), 721–741.

Kwan, C.M., Chang, C.S., 2008. Timetable synchronization of mass rapid transit system using multiobjective evolutionary approach. IEEE Trans. Syst. Man Cybern. C 38 (5), 636–648.

Land, A., Doig, A., 1960. An automatic method of solving discrete programming problems. econometrica. v28.

Liebchen, C., 2007. Periodic timetable optimization in public transport. In: Operations Research Proceedings 2006. Springer, pp. 29–36.

Limbourg, S., Jourquin, B., 2009. Optimal rail-road container terminal locations on the European network. Transp. Res. E Logist. Transp. Rev. 45 (4), 551–563.

Liu, T., Cats, O., Gkiotsalitis, K., 2021. A review of public transport transfer coordination at the tactical planning phase. Transp. Res. C 133, 103450.

Liu, T., Ceder, A., 2016. Synchronization of public transport timetabling with multiple vehicle types. Transp. Res. Rec. 2539 (1), 84–93.

Medina, M., Giesen, R., Muñoz, J.C., 2013. Model for the optimal location of bus stops and its application to a public transport corridor in Santiago, Chile. Transp. Res. Rec. 2352 (1), 84–93.

Meloni, C., Pranzo, M., Samà, M., 2021. Risk of delay evaluation in real-time train scheduling with uncertain dwell times. Transp. Res. E Logist. Transp. Rev. 152, 102366.

Murray, A.T., 2001. Strategic analysis of public transport coverage. Socio-Econ. Plan. Sci. 35 (3), 175–188.

Newell, G.F., 1974. Control of pairing of vehicles on a public transportation route, two vehicles, one control point. Transp. Sci. 8 (3), 248–264.

Orth, H., Weidmann, U., Dorbritz, R., 2012. Development of measurement system for public transport performance. Transp. Res. Rec. 2274 (1), 135–143.

Parbo, J., Nielsen, O.A., Prato, C.G., 2014. User perspectives in public transport timetable optimisation. Transp. Res. C 48, 269–284.

Peeters, L., Kroon, L., 2001. A cycle based optimization model for the cyclic railway timetabling problem. In: Computer-Aided Scheduling of Public Transport. Springer, pp. 275–296.

Schmid, V., Ehmke, J.F., 2015. Integrated timetabling and vehicle scheduling with balanced departure times. OR Spectr. 37 (4), 903–928.

Schmöcker, J.-D., Sun, W., Fonzone, A., Liu, R., 2016. Bus bunching along a corridor served by two lines. Transp. Res. B 93, 300–317.

Schöbel, A., 2012. Line planning in public transportation: models and methods. OR Spectr. 34 (3), 491–510.

Shafahi, Y., Khani, A., 2010. A practical model for transfer optimization in a transit network: Model formulations and solutions. Transp. Res. A 44 (6), 377–389.

Silva-Soto, Y., Ibarra-Rojas, O., 2021. Timetabling with flexible frequencies to synchronise groups of bus lines at common stops. Transp. A Transp. Sci. 17 (4), 978–1001.

Slater, M., 2014. Lagrange multipliers revisited. In: Traces and Emergence of Nonlinear Programming. Springer, pp. 293–306.

Sørensen, C.H., Longva, F., 2011. Increased coordination in public transport—which mechanisms are available? Transp. Policy 18 (1), 117–125.

Sun, S., Szeto, W., 2019. Optimal sectional fare and frequency settings for transit networks with elastic demand. Transp. Res. B 127, 147–177.
Szeto, W.Y., Jiang, Y., 2014. Transit route and frequency design: Bi-level modeling and hybrid artificial bee colony algorithm approach. Transp. Res. B 67, 235–263.
Szymański, P., Żołnieruk, M., Oleszczyk, P., Gisterek, I., Kajdanowicz, T., 2018. Spatio-temporal profiling of public transport delays based on large-scale vehicle positioning data from GPS in Wrocław. IEEE Trans. Intell. Transp. Syst. 19 (11), 3652–3661.
Tian, X., Niu, H., 2019. A bi-objective model with sequential search algorithm for optimizing network-wide train timetables. Comput. Ind. Eng. 127, 1259–1272.
Trompet, M., Liu, X., Graham, D.J., 2011. Development of key performance indicator to compare regularity of service between urban bus operators. Transp. Res. Rec. 2216 (1), 33–41.
Wang, Z., Liu, X., 2022. Cyber security of railway cyber-physical system (CPS)–A risk management methodology. Commun. Transp. Res. 2, 100078.
Wong, R.C., Yuen, T.W., Fung, K.W., Leung, J.M., 2008. Optimizing timetable synchronization for rail mass transit. Transp. Sci. 42 (1), 57–69.
Wu, J., Kulcsár, B., Qu, X., et al., 2021. A modular, adaptive, and autonomous transit system (MAATS): An in-motion transfer strategy and performance evaluation in urban grid transit networks. Transp. Res. A 151, 81–98.
Wu, W., Liu, R., Jin, W., 2016. Designing robust schedule coordination scheme for transit networks with safety control margins. Transp. Res. B 93, 495–519.
Wu, Y., Tang, J., Yu, Y., Pan, Z., 2015. A stochastic optimization model for transit network timetable design to mitigate the randomness of traveling time by adding slack time. Transp. Res. C 52, 15–31.
Xuan, Y., Argote, J., Daganzo, C.F., 2011. Dynamic bus holding strategies for schedule reliability: Optimal linear control and performance analysis. Transp. Res. B 45 (10), 1831–1845.
Yang, S., Yang, K., Yang, L., Gao, Z., 2018. MILP formulations and a TS algorithm for reliable last train timetabling with uncertain transfer flows. J. Oper. Res. Soc. 69 (8), 1318–1334.
Yap, M., Cats, O., van Oort, N., Hoogendoorn, S., 2017. A robust transfer inference algorithm for public transport journeys during disruptions. Transp. Res. Procedia 27, 1042–1049.
Yin, J., D'Ariano, A., Wang, Y., Yang, L., Tang, T., 2021. Timetable coordination in a rail transit network with time-dependent passenger demand. European J. Oper. Res. 295 (1), 183–202.