A Benders Decomposition Algorithm for Demand-Driven Metro Scheduling

Tommaso Schettini[a,*], Ola Jabali[a], Federico Malucelli[a]

[a] *Politecnico di Milano, Milan, Italy*

Abstract

Metro timetables are usually planned with a top-down approach. After dividing the day into different periods, the trains are scheduled between the terminals of the line with a fixed frequency per period. In this paper we adopt an alternative paradigm where trains are scheduled individually. The schedule is developed so as to best match the passenger demand, and trains may short-turn at intermediate stations, thus reversing their direction before reaching the line terminal. This type of approach is particularly suited for automated metro lines, since it has a limited impact on personnel management. Considering the objective of minimizing the passenger waiting times on a two-directional metro corridor, we make two operating assumptions when designing the train schedule. Specifically, we assume the presence of a root station, which cannot be skipped by short-turning, and we assume that idling is only allowed immediately after a short-turn, and for a maximum amount of time. We present a path-based formulation for the problem and develop an efficient exact algorithm for it using a Benders-based branch-and-cut algorithm. We evaluate the proposed formulation and algorithm on a number of test instances. Through our computational experiments, we demonstrate the effectiveness of the developed formulation and algorithm.

Keywords: Metro scheduling, Short-turning, Demand-driven, Benders decomposition

1. Introduction

Tactical planning for metro lines is typically performed through a strictly hierarchical approach (see Farahani et al. (2013), and Ceder (2016)). This planning is based on categorizing each day, e.g., winter-workday, winter-festive. Each day is then further divided into several time periods, such as morning peak, off-peak. Based on historical data, a timetable, predominately based on fixed frequencies for each time period, is determined to guarantee a certain level of service while minimizing costs. Lastly, based on this timetable, the operational planning phase establishes the train and crew schedules.

In practice, metro timetables are frequently constructed in a periodic manner, e.g., Liebchen (2008), Kroon et al. (2009), Nachtigall & Voget (1996), Liebchen et al. (2010). Caprara et al. (2002) define the train timetabling problem (TTP), which consists of determining a periodic timetable (i.e., a timetable that repeats at every period) for a set of trains operating in a single one-way track, to minimize the deviations from a given ideal timetable. The TTP has attracted much attention in the literature, e.g., Cacchiani et al. (2008), Cacchiani et al. (2012).

More recently, a growing number of studies have started considering the train timetabling issue from a passenger perspective. These optimize timetables by directly minimizing the passenger waiting times. Sun et al. (2014) develop a demand-driven optimization model to determine the optimal headways of a one-directional metro corridor. Canca et al. (2014) develop a non-linear integer programming model to optimize demand-driven timetables for a one-directional metro corridor. Barrena et al. (2014a) and Barrena et al. (2014b) investigate relaxing the periodic assumption in the timetable of a rail corridor under given passenger

*Corresponding author

Email addresses: `tommaso.schettini@polimi.it` (Tommaso Schettini), `ola.jabali@polimi.it` (Ola Jabali), `federico.malucelli@polimi.it` (Federico Malucelli)

demands. Yin et al. (2017) develop two models for the joint minimization of the passenger waiting time and the operating costs of the line in a bi-directional metro corridor.

Other recent research strands focused on the use of acceleration strategies in railway optimization, such as stop skipping and short-turning. The latter entails that trains are not obliged to operate from terminal to terminal, but may turn and reverse their direction in certain stations of a line. The use of acceleration strategies has been extensively studied in transit networks, in particular in a bus context, e.g., Furth (1987), Tirachini et al. (2011) and Cortés et al. (2011). Comparatively, their study in rail transit applications has been rather limited.

Canca et al. (2012) discuss the introduction of additional shuttle services alongside the regular service on overloaded segments of a railway line using short-turning, while allowing trains to run empty throughout portions of the line. Canca et al. (2016) propose a model for handling disruptions in the demand of a transit system through the optimization of short-turning services introduced alongside the pre-planned timetable. To cope with overcrowding at stations of a metro line after a service disruption, Gao et al. (2016) propose using stop-skipping patterns in the timetable of the trains. The objective of the model is minimizing the delays with respect to the original timetable as well as the passengers' waiting time, while guaranteeing a minimum service level during the recovery period. Yang et al. (2020) propose a demand-driven train timetabling model while employing a flexible short-turning strategy. The model is solved through the use of a Lagrangian Relaxation. This was applied to a set of small artificially generated instances and on a realistic testbed taken from the Beijing metro network.

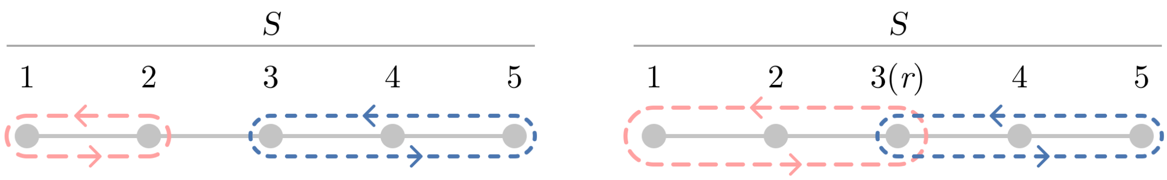
Schettini et al. (2021b) optimize the schedule of a metro line during a large event, e.g., a football match or a concert. This causes a surge of demand at a single station, where the event venue is located. The demand not associated with the event is ignored. To better serve the line, each train is scheduled individually, and trains are allowed to short-turn. The authors develop an iterated local search metaheursitic to optimize the service. Four objective functions representing different service measures are evaluated.

Schettini et al. (2021a) define a demand-driven control paradigm on a discretized planning horizon called *direct timetabling*. According to this paradigm, given the passenger demands over a fixed planning horizon, trains are scheduled individually, with the possibility of short-turning, and with no predetermined periodic structure on the train schedule. Moreover, trains are allowed to idle at any station. The resulting optimization problem is called the direct timetabling problem (DTP), and is studied considering the objective of minimizing passenger waiting times. The DTP allows a great deal of flexibility in the control decisions of the line, which is used to adapt to variations in passenger demand. However, this flexibility results in a complex problem setting, since each train is individually scheduled. Considering a bi-directional metro, Schettini et al. (2021a) propose a cut generation algorithm developed for a flow-based formulation for the DTP. Their procedure is capable of exactly solving medium sized instances.

In this paper, we consider a similar problem setting as the DTP proposed by Schettini et al. (2021a). However, we impose the following two operating assumptions on the structure of the train schedule.

1) We assume the existence of a *root station*, which is a station that cannot be skipped by short-turning (i.e., all trains are required to reach r before short-turning). This ensures the connectivity of the line. Indeed, allowing trains to short-turn in a number of stations throughout the line may cause the schedule to be disaggregated into several services. For example, considering a five-station line as in Figure 1, there could be trains that loop predominately between stations one and two, and trains that predominately loop between stations three and five. Such a situation is undesirable from an operational perspective. Designating station three as the root station mitigates this behavior.
2) We impose that trains are only allowed to idle right after short-turning and for a maximum amount time. As passengers will always alight before a train short-turns, this assumption entails that passengers will not wait inside an idling train, which is desirable from a passenger perspective.

We refer to the DTP with the two above operating assumptions as the direct timetabling problem with a root station (DTPR). We analyze the theoretical structure of DTPR solutions. Based on this analysis, we define a path-based formulation to the DTPR. This is in contrast to the flow-based formulation defined in Schettini et al. (2021a) for the DTP. We develop a Benders-based branch-and-cut algorithm for the DTPR path-based formulation. We then demonstrate the effectiveness of this algorithm.



(a) A schedule without the root assumption.

(b) A schedule with the root assumption.

Figure 1: Comparison of possible train schedules with and without the root assumption.

The contributions of this paper are as follows:

1) defining the DTPR and proposing a path-based formulation for it.
2) developing an efficient exact Benders-based branch-and-cut algorithm for the DTPR.
3) demonstrating, through extensive experiments, the effectiveness of the developed algorithm.

The remainder of the paper is organized as follows: In Section 2, we describe the DTPR and provide a path-based formulation for the problem, when short-turning is allowed in each station. Then, in Section 3, we develop a Benders-based branch-and-cut algorithm for the DTPR, based on its path-based formulation. In Section 4, we develop a compression approach to reduce the size of the problem when short-turning is not allowed at all stations. In Section 5, we present our computational results. Lastly, our conclusions are presented in Section 6. For the sake of conciseness, proofs are omitted from the main sections of the paper and are presented in Appendix B.

2. The Direct Timetabling Problem with a Root Station

The DTPR consists of constructing the train schedule of a bi-directional metro over a given discretized time horizon. Given the passenger demands over the time horizon, the objective of the DTPR is minimizing the total passenger waiting time.

Let $S = \{1, ..., m\}$ denote the ordered set of stations on the line, and $\Gamma = \{-1, 1\}$ is the set of directions, where 1 denotes the upstream direction, (i.e., $1, 2, \dots, m$), and -1 denotes the downstream direction, (i.e., $m, m-1, \dots, 1$). Let F_γ be the first station in direction $\gamma \in \Gamma$. We denote by $S_{\gamma,i}$ the set of stations that lie in direction γ relative to station i (See Fig. 2). Lastly, we use the short-hand $\upsilon(i, j)$ to denote the direction implied by the pair of stations i and j (e.g., $\upsilon(4, 1) = -1$).

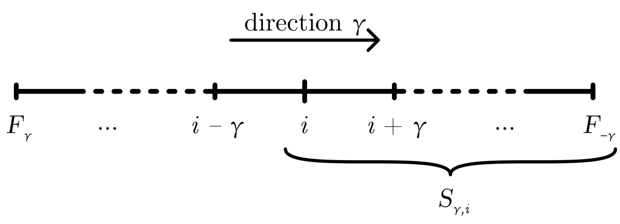


Figure 2: Example of the stations on the line w.r.t. a given station i and a direction γ.

The planning horizon is divided in time-steps of length δ. Let $T = \{1, ..., h\}$ be the set of time-steps. Thus, time-step $t \in T$ corresponds to the time interval $(\delta(t-1), \delta t]$. The line is operated by Ψ identical trains. We assume that the time to board at any station is fixed and constant for all stations. We denote by $\tau_{i,j}$ the number of time-steps required for a train to go from station i to station j without short-turning, including dwelling time and time to board. For the sake of brevity, we will not make the distiction between a turning operation performed at a terminal station and a short-turning operation performed at some other station on the line. Therefore, in what follows we will use the term *short-turn* to refer to all turning operations regardless of the station. We denote by ρ the number of time-steps required to perform a short-turn at any station. Similar to Barrena et al. (2014a,b), the capacity of the trains is assumed to be sufficient to accommodate all the passengers without overcrowding.

We denote by D the set of Origin-Destination pairs, i.e., $D = \{(i, j) : i, j \in S, i \neq j\}$. The number of passengers that arrive at station $i \in S$ with destination $j \in S$ at time-step $t \in T$ is denoted by $a^t_{i,j}$. We impose an upper bound g on the total number of time-steps that a passenger may wait. We note that the parameter g is introduced for modeling purposes. In practice, g can be set to a rather high value, so that a waiting time of g time-steps is not attained by any passenger. We have done so in the computational experiments.

In the case of the DTP, no assumption is placed upon the schedule of the trains. As stated in the introduction, in the DTPR we make the following two operating assumptions on the structure of the schedule. Firstly, we assume the existence of a *root station* on the line. The root station is a designated station r, which acts as the anchoring point of the line. All trains are required to reach r before short-turning. Thus, a train going in direction $\gamma \in \Gamma$ has to be moving away from r before short-turning, i.e., a short-turn can be performed in station i if $\gamma(i - r) \geq 0$. We refer to this assumption as the *root assumption*. The root station is a parameter of the problem and should be selected by the line operator considering the distribution of the passenger demand and exogenous factors, such as the position of relevant points of interest. We note that the root station can conceivably be one of the terminal stations of the line, e.g., a metro line linking an airport to the city. For the ease of notation, we use S_γ as a shorthand for $S_{\gamma,r}$. Secondly, we assume trains are only permitted to idle right after having short-turned for a maximum of β time-steps. Since idling is only permitted immediately after a short-turn, any idle time in the schedule is performed while no passengers are on board the trains. Thus, the trains will dwell for the minimum required amount of time at each station. We refer to this assumption as the *no-idling assumption*.

The solution to the DTPR consists of the sequence of actions performed by the trains over the planning horizon, describing the schedule of the line that minimizes the total waiting time of the passengers. At any given time-step, each train is either 1) moving to the next station in its current direction, or 2) performing a short-turn at the current station, or 3) idling for one time-step.

We now develop a path-based formulation (PF0) for the DTPR. This formulation builds upon the basic formulation of Schettini et al. (2021b) which was defined for a simpler problem setting. In the path-based formulation, we represent the train movements using the paths they take relative to the root station. The remainder of this section is organized as follows. In Subsection 2.1, we present a formal definition of a train path, and discuss how paths can be used to represent a train schedule. In Subsection 2.2, we describe the concept of equivalent time, which presents several modeling advantages in the formulation of the problem. In Subsection 2.3, we present a partitioning of D which is used in the formulation of the problem. Lastly, in Subsection 2.4, we present the PF0.

2.1. Train Paths

We define a *path* as the sequence of actions of a train between two consecutive visits of the root station. Since the train only visits the root station at the start and the end of the path, a train will short-turn exactly once while operating a path. Additionally, as for the no-idling assumption, the trains can only idle immediately after short-turning. Therefore, all paths adhere the following structure: 1) the train departs from the root station, 2) reaches a given destination station where it short-turns, 3) idles at the destination station, and 4) returns to the root station.

Formally, a path is described by the tuple (t, γ, d, n), where t denotes the path starting time-step, γ denotes the initial direction of the path, d is the destination station, and n is the path duration. As such, path (t, γ, d, n) uniquely describes a train that departs from the root station in direction γ at time-step t, then proceeds to its destination station d, arriving at d at time-step $t + \tau_{d,r}$, short-turns at d, and finally departs from d in direction $-\gamma$ at time-step $t + n - \tau_{d,r}$, arriving at r at time-step $t + n$, and thus effectively

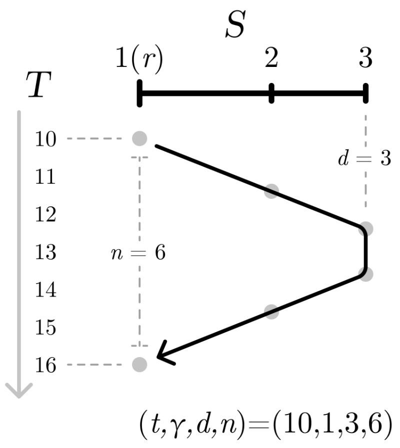


Figure 3: Example of path on a three-station line.

waiting $n-\rho-\tau_{r,d}-\tau_{d,r}$. Figure 3 shows an example of path on a five-station line. Since a train can idle for at most β time-steps, a path with destination d has a maximum duration of $\beta+\tau_{r,d}+\tau_{d,r}+\rho$. We denote by $Q_d=\{\tau_{r,d}+\tau_{d,r}+\rho,\ldots,\beta+\tau_{r,d}+\tau_{d,r}+\rho\}$ the set of possible path durations in time-steps associated with destination d.

A feasible train schedule to the DTPR can be fully represented through paths. To do so, we express the schedule of each train as a sequence of paths in alternating directions. Furthermore, we are able to model any possible starting position of the trains by artificially dispatching trains in time-steps preceding the start of the planning horizon. Let $T_e=\{-sup_{i\in S}(Q_i),\ldots,0,\ldots,h\}$ be the set of possible starting times of the paths. The value $-sup_{i\in S}(Q_i)$ is selected to allow for all possible initial positions of the trains on the line. Ending conditions are directly implied by the paths starting near the end of the planning horizon.

In the path-based formulation we describe the train schedule using the binary path variable $x^t_{\gamma,d,n}$, which takes the value of one if the path described by the tuple (t,γ,d,n) is used in the schedule, and zero otherwise. We denote by x the path vector, which aggregates the path variables $x^t_{\gamma,d,n}$ $\forall\gamma\in\Gamma,d\in S_\gamma,n\in Q_d,t\in T$.

2.2. Equivalent Time

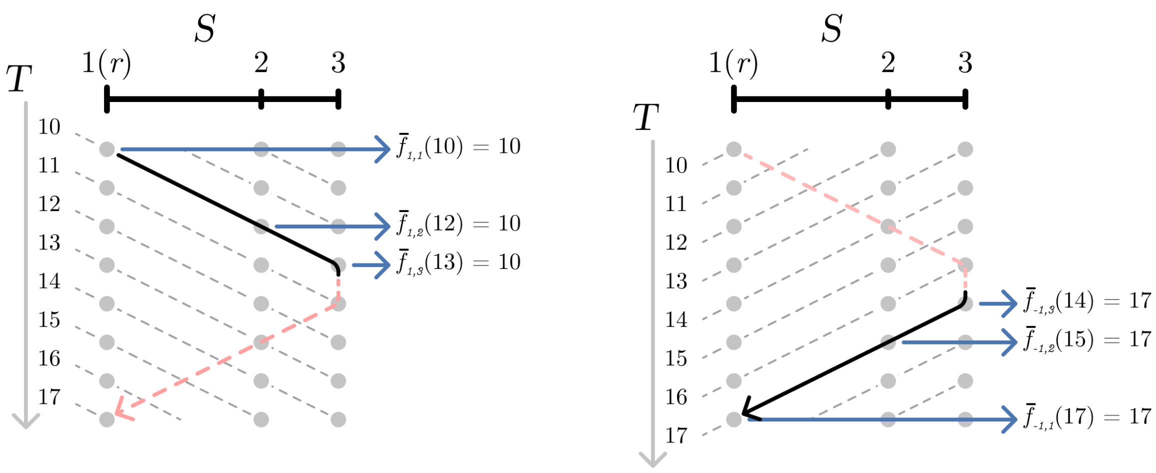
In this section, we discuss a remapping of the time axis, which presents several modeling advantages. In principle, for each station, we will offset the time coordinate to account for the train travel time (Newell (1993)). We refer to this local time coordinate as "Equivalent Time". We note that Schettini et al. (2021b) first applied this idea in the context of a two directional metro corridor. We further adapt this concept to the DPTR problem setting.

The equivalent time relative to station i in direction γ at time t is defined with respect to the root station r as:

$$\bar{f}_{\gamma,i}(t)=\begin{cases}t-\tau_{r,i} & \text{if } (i-r)\gamma\geq 0\\ t+\tau_{i,r} & \text{if } (i-r)\gamma<0.\end{cases}\qquad \forall\gamma\in\Gamma,i\in S,t\in T \tag{1}$$

The equivalent time at a station i corresponds to the time t scaled by the travel time between r and i, when moving in direction γ. By definition, the equivalent time at the root station will apply no offset to the time axis, i.e., $\bar{f}_{\gamma,r}(t)=t$ $\forall t\in T,\gamma\in\Gamma$.

To demonstrate the value of this conversion, let us consider a train that departs from station i in direction γ at time-step t, and reaches $i+\gamma$ at time-step $t+\tau_{i,i+\gamma}$. Converting those times to equivalent time, we obtain $\bar{f}_{\gamma,i}(t)=\bar{f}_{\gamma,i+\gamma}(t+\tau_{i,i+\gamma})$. Thus, a train serving the line from terminal to terminal without idling would stop at all stations at the same equivalent time-step (relative to each station). See Fig. 4 for an example of the conversion to equivalent time. Note that the definition of equivalent time is also valid in the case when the travel time between stations is a fractional number of time-steps.



(a) Moving away from the root station.

(b) Approaching the root station.

Figure 4: Example of equivalent time on a three station line with non integer travel time between the stations. $r = 1$, $\tau_{1,2} = 2$, $\tau_{2,3} = 1$.

For modeling convenience, we define $\bar{T}_{\gamma,i}$ as the equivalent time set for station i in direction γ. The elements of $\bar{T}_{\gamma,i}$ correspond to the elements of T scaled according to $\bar{f}_{\gamma,i}$. As an example, consider the line depicted in Figure 4, let $T = \{1, 2, \dots, 20\}$, then the equivalent time set at station 3 in the upstream direction is $\bar{T}_{1,3} = \{4, 5, \dots, 23\}$; and in the downstream direction, the equivalent time set is $\bar{T}_{-1,3} = \{-1, 0, \dots, 18\}$.

We redefine the passenger demand accounting for the equivalent time. We denote by $\bar{a}^{\bar{t}}_{i,j}$ the number of passengers with origin i and destination j that enter the line at equivalent time $\bar{t}$. Parameter $a^t_{i,j}$ relates to the redefined passenger demand $\bar{a}^{\bar{t}}_{i,j}$ as follows:

$$\bar{a}^{\bar{t}}_{i,j} = a^t_{i,j} \qquad \forall (i,j) \in D, t \in T, \bar{t} \in \bar{T}_{\upsilon(i,j),i} \ : \ \bar{t} = \bar{f}_{\upsilon(i,j),i}(t).$$

For the sake of brevity, in what follows, we will use the tuple (i, j, t) to denote the passengers that arrives at station i with destination j at equivalent time-step t relative to station i.

2.3. Partitioning of the OD pairs

We partition the set of OD pairs in D into three mutually exclusive sub-sets, depending on how they relate to r. We refer to Fig. 5 for examples of the various OD pairs that belong to each partition.

The set D_f is the set of OD pairs (i, j) in which i and j lie on the same side of the line, relative to the root station, and i is closer to r than j, i.e., $i, j \in S_\gamma$ and i is closer to r than j is. Formally:

$$D_f = \{i, j \in S \ : \ |i - r| < |j - r| \ \wedge \ (\exists \gamma \in \Gamma \ : \ i, j \in S_\gamma)\}. \tag{2}$$

Similarly, D_b is the set of OD pairs (i, j) in which i and j lie on the same side of the line, relative to the root station, and i is farther from r than j, i.e., $i, j \in S_\gamma$ and j is closer to r than i is. Formally:

$$D_b = \{i, j \in S \ : \ |i - r| > |j - r| \ \wedge \ (\exists \gamma \in \Gamma \ : \ i, j \in S_\gamma)\}. \tag{3}$$

Lastly, D_c is the set of OD pairs that lie on opposite sides of the line, relative to the root station, thus representing passengers that need to pass through the root station. Formally:

$$D_c = \{i, j \in S \ : \ (\exists \gamma \in \Gamma \ : \ i \in S_\gamma \ \wedge \ j \notin S_\gamma)\}. \tag{4}$$

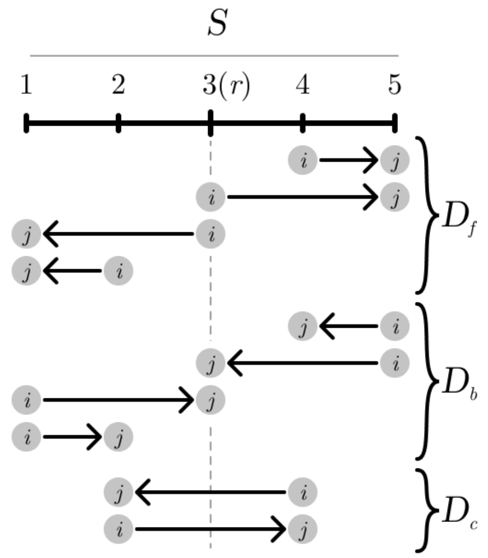


Figure 5: Examples of OD pairs belonging to D_f, D_b and D_c.

We define the *trip* of passengers (i, j, t) as the sequence of paths that the passengers take to reach their destination. We note that a train operating a path in direction γ is not limited to serving passengers headed in that direction. Passengers headed in direction $-\gamma$ may also board the train while it is returning towards the root station. Considering passengers (i, j, t), we now state the conditions under which they are able to board a given train operating path (t', γ, d, n). We distinguish three cases.

1) passengers (i, j, t) where $(i, j) \in D_f$, are able to board a train operating path (t', γ, d, n) if

$$\gamma = \upsilon(i, j),\ d \in S_{i,\gamma},\ t' \geq t.$$

Additionally, if $d \in S_{j,\gamma}$, then path (t, γ, d, n) reaches station j before short-turning and the passengers arrive at their destination at equivalent time-step t. Conversely, if path (t, γ, d, n) does not reach j before short-turning, the passengers will need to alight at station d. After alighting, the passengers will have to wait for a future train to reach their destination as if they just arrived at station d at equivalent time-step t'.

2) passengers (i, j, t) where $(i, j) \in D_b$, are able to board a train operating path (t', γ, d, n) if

$$\gamma = -\upsilon(i, j),\ d \in S_{\gamma,i},\ t' + n \geq t.$$

We note that in this case, all paths that reach station i also pass by station j at equivalent time-step $t' + n$, since j lies between i and r.

3) passengers (i, j, t) where $(i, j) \in D_c$. No single path is able to reach both stations i and j, since those stations lie on opposite sides of r. In these cases, the trajectory of the passengers has to be split into two trip legs: i to r and r to j. On the first leg, the passengers are able to board a train operating path (t', γ, d, n) if

$$\gamma = -\upsilon(i, j),\ d \in S_{\gamma,i},\ t' + n \geq t.$$

We note that this is the same boarding condition for OD pairs $(i, j) \in D_b$. If these conditions hold, the passengers are able to reach r at time-step $t' + n$. From r, they are then able to board a train that is operating path (t'', γ', d', n') if

$$\gamma' = -\gamma,\ t'' \geq t' + n.$$

If the passengers board the train, they depart from the root station at time-step t and, if $d' \in S_{j,\gamma'}$, arrive at their destination at equivalent time-step t. If $d' \notin S_{j,\gamma'}$ the same considerations we drew for case 1) may be applied. We note that accounting for trip legs i to r and r to j in separation does not imply that the passengers alight at the root station. Indeed, the same train may be operating the path that delivers the passengers from i to r as well as the path that delivers the passengers from r to j.

The number of possible trips that passengers can take may grow exponentially with the number of stations. Thus, in general, it is not possible to fully enumerate them. In what follows, we establish that only a limited number of trips need to be considered. Specifically, we show that the trip with the least waiting time for passengers (i, j, t) can be completed by boarding at most two properly selected trains. Lemmas 1 and 2, show that for $(i, j) \in D_f$, $(i, j) \in D_b$ the trip with the least waiting time is established by boarding a single train, whereas Lemma 3 shows that for $(i, j) \in D_c$ the trip with the least waiting time is established by boarding at most two trains.

Lemma 1. *Given a feasible path vector x, the trip with the least waiting time for passengers (i, j, t) where $(i, j) \in D_f$ is established by only boarding the train stating at time-step:*

$$\min_{t' \in T} \Big(t' \ : \sum_{d \in S_{\upsilon(i,j),j}} \sum_{n \in Q_d} x^{t'}_{\upsilon(i,j),d,n} = 1 \ \wedge \ t' \geq t \Big),$$

i.e., the first train passing by station i headed to station j seen by passengers (i, j, t).

Lemma 2. *Given a feasible path vector x, the trip with the least waiting time for passengers (i, j, t) where $(i, j) \in D_b$ is established by only boarding the train starting at time-step:*

$$\min_{t' \in T} \Big(t' \ : \sum_{d \in S_{-\upsilon(i,j),i}} \sum_{n \in Q_d} x^{t'-n}_{-\upsilon(i,j),d,n} = 1 \ \wedge \ t' \geq t \Big),$$

i.e., the first train headed towards r seen by passengers (i, j, t).

Lemma 3. *Given a feasible path vector x, the trip with the least waiting time for passengers (i, j, t) where $(i, j) \in D_c$ is established by boarding at most two trains, specifically boarding the first train passing by station i headed towards r or j at time-step:*

$$\min_{t' \in T} \Big(t' \ : \sum_{d \in S_{-\upsilon(i,j),i}} \sum_{n \in Q_d} x^{t'-n}_{-\upsilon(i,j),d,n} = 1 \ \wedge \ t' \geq t \Big),$$

then departing from the root station at time-step:

$$\min_{t'' \in T} \Big(t'' \ : \sum_{d \in S_{\upsilon(i,j),j}} \sum_{n \in Q_d} x^{t'}_{\upsilon(i,j),d,n} = 1 \ \wedge \ t'' \geq t' \Big).$$

As a result of Lemmas 1–3, we conclude that the passengers can be assumed to never wait at any station other than their origin station or the root station. Recalling that passengers may spend at most g total time-steps waiting for trains to arrive, the number of possible trips of passengers (i, j, t) where $(i, j) \in D_f \cup D_b$ is at most g, since the passengers will board exactly one train to reach their destination and therefore may only wait at their origin station. Similarly, the theoretical number of possible trips available to passengers (i, j, t) where $(i, j) \in D_c$ is at most $\binom{g+2}{2} = \frac{(g+2)(g+1)}{2}$, accounting for the fact that the passengers can wait for a train both at their origin station as well as at r, waiting at most g time-steps in total.

Given that g is at most equal to the length of the planning horizon, those trips can be enumerated and explicitly addressed in the model. We denote by $\Delta = \{0, ..., g\}$ the set of possible waiting times experienced by the passengers at a station.

We define the binary variable $y_{i,j}^{t,\delta}$, describing the trajectory of the passengers (i, j, t) where $(i, j) \in D_f \cup D_b$. The variable $y_{i,j}^{t,\delta}$ takes the value of one if passengers (i, j, t) board a train, headed to j after having waited $\delta \in \Delta$ time-steps (i.e., the passengers board a train headed to their destination at equivalent time-step $t + \delta$), and zero otherwise. Similarly, the binary variable $y_{i,j}^{t,\delta,\delta'}$ describes the trajectory on the line of the passengers (i, j, t) where $(i, j) \in D_c$. The variable $y_{i,j}^{t,\delta,\delta'}$ is one if passengers (i, j, t) board a train headed to the root station after having waited $\delta \in \Delta$ time-steps, once at the root station the passengers wait for $\delta' - \delta$ time-steps for the train that will bring them to their destination (with $\delta' \in \Delta$), and zero otherwise. Note that $\delta' = \delta$ implies that the passengers will not alight the first train they boarded and will move directly from i to j.

2.4. Path-Based Formulation

Using the concepts and notation described thus far, the DTPR is formulated as follows.

$$\textbf{[PF0]} \quad \min \sum_{(i,j)\in D_f} \sum_{t\in \bar{T}_{\upsilon(i,j),i}} \sum_{\delta\in\Delta} \left(\delta \bar{a}_{i,j}^{t} y_{i,j}^{t,\delta}\right) + \sum_{(i,j)\in D_b} \sum_{t\in \bar{T}_{\upsilon(i,j),i}} \sum_{\delta\in\Delta} \left(\delta \bar{a}_{i,j}^{t} y_{i,j}^{t,\delta}\right) + \sum_{(i,j)\in D_c} \sum_{t\in \bar{T}_{\upsilon(i,j),i}} \sum_{\delta\in\Delta} \sum_{\delta'\in\Delta:\delta'\geq\delta} \left(\delta' \bar{a}_{i,j}^{t} y_{i,j}^{t,\delta,\delta'}\right) \tag{5}$$

$$\sum_{t\in T_e\setminus T} \sum_{\gamma\in\Gamma} \sum_{d\in S_\gamma} \sum_{n\in Q_d} x_{\gamma,d,n}^{t} \leq \Psi \tag{6}$$

$$\sum_{d\in S_{-\gamma}} \sum_{n\in Q_d} x_{-\gamma,d,n}^{t-n} = \sum_{d\in S_\gamma} \sum_{n\in Q_d} x_{\gamma,d,n}^{t} \qquad \forall t \in T, \gamma \in \Gamma \tag{7}$$

$$\sum_{d\in S_\gamma} \sum_{n\in Q_d} x_{\gamma,d,n}^{t} \leq 1 \qquad \forall t \in T, \gamma \in \Gamma \tag{8}$$

$$x_{\gamma,d,n}^{t} + \sum_{d'\in S_{\gamma,d+\gamma}} \sum_{n'\in Q_{d'}} x_{\gamma,d',n'}^{t'-\tau_{r,d}-n'} \leq 1 \qquad \forall \gamma \in \Gamma, d \in S_\gamma, n \in Q_d, t \in T, t' \in \mathbb{T}_d^t,$$
$$\mathbb{T}_d^t = \{t' \in T_e \ : \ \tau_{r,d} + 1 + t \leq t' \leq t + n - \tau_{r,d}\} \qquad \forall \gamma \in \Gamma, t, t' \in T_e, d \in S_\gamma \tag{9}$$

$$\sum_{\delta\in\Delta} y_{i,j}^{t,\delta} = 1 \qquad \forall (i, j) \in D_f \cup D_b, t \in \bar{T}_{\upsilon(i,j),i} \tag{10}$$

$$\sum_{\delta\in\Delta} \sum_{\delta'\in\Delta:\delta'\geq\delta} y_{i,j}^{t,\delta,\delta'} = 1 \qquad \forall (i, j) \in D_c, t \in \bar{T}_{\upsilon(i,j),i} \tag{11}$$

$$y_{i,j}^{t,\delta} \leq \sum_{d\in S_{-\upsilon(i,j),i}} \sum_{n\in Q_d} x_{\upsilon(i,j),d,n}^{t+\delta} \qquad \forall (i, j) \in D_f, t \in \bar{T}_{\upsilon(i,j),i}, \delta \in \Delta \tag{12}$$

$$y_{i,j}^{t,\delta} \leq \sum_{d\in S_{\upsilon(i,j),j}} \sum_{n\in Q_d} x_{-\upsilon(i,j),d,n}^{t+\delta-n} \qquad \forall (i, j) \in D_b, t \in \bar{T}_{\upsilon(i,j),i}, \delta \in \Delta \tag{13}$$

$$\sum_{\delta'\in\Delta:\delta'\geq\delta} y_{i,j}^{t,\delta,\delta'} \leq \sum_{d\in S_{-\upsilon(i,j),i}} \sum_{n\in Q_d} x_{-\upsilon(i,j),d,n}^{t+\delta-n} \qquad \forall \delta \in \Delta, (i, j) \in D_c, t \in \bar{T}_{\upsilon(i,j),i} \tag{14}$$

$$\sum_{\delta\in\Delta:\delta\leq\delta'} y_{i,j}^{t,\delta,\delta'} \leq \sum_{d\in S_{\upsilon(i,j),j}} \sum_{n\in Q_d} x_{\upsilon(i,j),d,n}^{t+\delta'} \qquad \forall \delta' \in \Delta, (i, j) \in D_c, t \in \bar{T}_{\upsilon(i,j),i} \tag{15}$$

$$x_{\gamma,d,n}^{t} \in \{0, 1\} \qquad \forall \gamma \in \Gamma, d \in S_\gamma, n \in Q_d, t \in T_e \tag{16}$$

$$y_{i,j}^{t,\delta} \in \{0, 1\} \qquad \forall (i, j) \in D_f \cup D_b, t \in \bar{T}_{\upsilon(i,j),i}, \delta \in \Delta \tag{17}$$

$$y_{i,j}^{t,\delta,\delta'} \in \{0, 1\} \qquad \forall (i, j) \in D_c, t \in \bar{T}_{\upsilon(i,j),i}, \delta, \delta' \in \Delta \ : \ \delta \leq \delta' \tag{18}$$

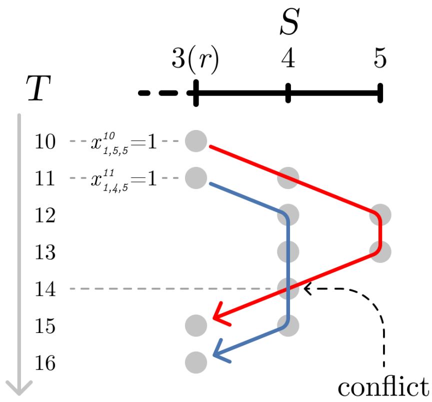


Figure 6: An example of two unfeasible paths that occur when $\beta > 0$.

The objective function (5) minimizes the total waiting time experienced by the passengers. Constraint (6) enforces the limit on the number of trains. It limits the number of paths that can be dispatched before the start of the planning horizon, i.e., paths with starting time $t \in T_e \backslash T$. Constraints (7) balance the flow of trains at the root station. At each time-step $t \in T$, a path can start in direction $\gamma \in \Gamma$ if another path with direction $-\gamma$ ends at the same time. Constraints (8) prevent more than one path from starting at the same time-step in the same direction. In combination with (7), the constraints also prevent more than one path from ending at the same time-step. We note that this is not sufficient to fully prevent scheduling conflicts on the line, as it fails to avoid conflicts that may occur when $\beta > 0$ (see Fig. 6 for an example). To cover these cases, constraints (9) prevent the idle period of a path from overlapping with another one, at each short-turning station, and at each time-step.

Constraints (10) and (11) ensure that all passengers select a trip to their destination. Constraints (12) and (13) enforce the consistency of the selected trips with the selected paths for passengers with OD $(i,j) \in D_f \cup D_b$. Similarly, constraints (14) and (15) enforce the consistency of the selected trips with the selected paths for passengers with OD $(i,j) \in D_c$ with the train paths, considering separately the two trip legs (i,r) and (r,j). Lastly, constraints (16), (17), and (18) enforce the domain of the variables. We now show that constraints (17)–(18) can be relaxed.

Lemma 4. *Given a feasible path vector $\bar{x}$, the constraint matrix of formulation PF0 is totally unimodular.*

Assume that path vector $\bar{x}$ satisfies (6)–(9), (16), i.e., yields a feasible schedule. Given that the path vector $\bar{x}$ is binary, the right-hand side of the constraints (10)–(15) is integer. Thus, as for Lemma 4, the binary domain constraints (17)–(18) can be replaced with:

$$y_{i,j}^{t,\delta} \geq 0 \qquad \forall (i,j) \in D_f \cup D_b, t \in \bar{T}_{\upsilon(i,j),i}, \delta \in \Delta \tag{19}$$

$$y_{i,j}^{t,\delta} \leq 1 \qquad \forall (i,j) \in D_f \cup D_b, t \in \bar{T}_{\upsilon(i,j),i}, \delta \in \Delta \tag{20}$$

$$y_{i,j}^{t,\delta,\delta'} \geq 0 \qquad \forall (i,j) \in D_c, t \in \bar{T}_{\upsilon(i,j),i}, \delta, \delta' \in \Delta \ : \ \delta \leq \delta'. \tag{21}$$

$$y_{i,j}^{t,\delta,\delta'} \leq 1 \qquad \forall (i,j) \in D_c, t \in \bar{T}_{\upsilon(i,j),i}, \delta, \delta' \in \Delta \ : \ \delta \leq \delta'. \tag{22}$$

The upper bound constraints (20) and (22) are dominated by constraints (10) and (11) respectively, and thus can be omitted from the formulation. Therefore, replacing constraints (17)–(18) with (19) and (21) results in a formulation equivalent to PF0, we refer to this new formulation as PF.

3. A Benders-Based Branch-and-Cut Algorithm for the DTPR

In this section, we develop a Benders-based branch-and-cut algorithm (BD) based on formulation PF. An overview of Benders decomposition for linear and integer programming is provided by Rahmaniani et al.

(2017). The basic idea of the decomposition is to simplify the solution of a complex problem by isolating its complicating variables, i.e., variables which, once fixed, make the resulting sub-problem significantly easier to solve. Using Benders decomposition, problems are solved in an iterative way. At each iteration, a relaxed version of the problem with only the complicating variables is solved, then its solution is linearized and projected onto the space of the remaining variables. From the projection, valid inequalities on the complicating variables are derived, and the process is iterated until converging to optimality.

In practice, it is convenient to implement the Benders decomposition in a branch-and-cut framework. This approach is commonly referred to as Benders-based branch-and-cut (Naoum-Sawaya & Elhedhli, 2013; Rahmaniani et al., 2017). In this framework, the separation of violated Benders cuts is performed within the search through the use of solver callback functions. This has the advantage not having to solve a MILP from scratch at every iteration, thus, considerably speeding up the solution. Therefore, we implement the Benders decomposition through a Benders-based branch-and-cut framework.

Additionally, Benders based methods can be further accelerated through the use of tailored implementations that exploit the structure of the problem. Tailored methods have been successfully applied in several fields such as production-inventory planning (Golari et al., 2017), layout planning (Sudermann-Merx et al., 2021), and stochastic network design (Crainic et al., 2021). In our proposed algorithm, we derived closed-form solutions for the Benders sub-problems. This allows us to solve the sub-problems analytically, without solving the associated LP.

For presentation convenience, in Subsection 3.1, we first show the structure of the Benders sub-problems utilized in the decomposition. In Subsection 3.2, we show the formulation of the master problem. In Subsection 3.3, we discuss the properties of Benders cuts added to the master problem and derive closed-form solutions for the Benders sub-problems. Lastly, in Subsection 3.4, we present an efficient procedure for the separation of the Benders cuts, given an incumbent solution. We recall that formal proofs are presented in Appendix B.

3.1. Benders Sub-Problems

We designate the path variables, i.e., $x^t_{\gamma,d,n}$ $\forall \gamma \in \Gamma, d \in S_\gamma, n \in Q_d, t \in T$, as the complicating variables. By fixing the path vector x to $\bar{x}$, constraints (6)–(9), (16) can be dropped from formulation PF0. Therefore, PF reduces to constraints (10)–(15), (19), (21), and decomposes into several smaller sub-problems. Specifically, one for each passenger tuple (i, j, t).

Each Benders sub-problem (i, j, t), addresses the minimization of the total waiting time experienced by the corresponding passenger, denoted as $\eta^t_{i,j}$. The Benders sub-problems have different structures depending on the partition of OD pairs defined in Section 2.3. We recall that $\upsilon(i, j)$ denotes the direction implied by the pair of stations i and j. For the ease of notation, in the rest of this section, let $\upsilon = \upsilon(i, j)$.

3.1.1. Case D_f and Case D_b

The sub-problem associated with passengers (i, j, t) where $(i, j) \in D_f$ is as follows.

$$[PD_f] \quad \min \quad \sum_{\delta\in\Delta}(\delta \bar{a}^t_{i,j} y^{t,\delta}_{i,j}) \tag{23}$$

$$\sum_{\delta\in\Delta} y^{t,\delta}_{i,j} = 1 \tag{24}$$

$$y^{t,\delta}_{i,j} \le \sum_{d\in S_{\gamma,j}} \sum_{n\in Q_d} \bar{x}^{t+\delta}_{\upsilon,d,n} \qquad \forall \delta \in \Delta \tag{25}$$

$$y^{t,\delta}_{i,j} \ge 0 \qquad \forall \delta \in \Delta \tag{26}$$

We write the dual of the sub-problem PD_f using $\lambda^t_{i,j}$ as the dual variable corresponding to constraint (24), and $\pi^{t,\delta}_{i,j}$ as the dual variables corresponding to constraints (25).

$$[BD_f] \quad \max \quad \eta^t_{i,j} = \lambda^t_{i,j} - \sum_{\delta\in\Delta}\Big(\pi^{t,\delta}_{i,j} \sum_{d\in S_{\gamma,j}} \sum_{n\in Q_d} \bar{x}^{t+\delta}_{\upsilon,d,n}\Big) \tag{27}$$

$$\lambda_{i,j}^{t} - \pi_{i,j}^{t,\delta} \leq \delta \bar{a}_{i,j}^{t} \qquad \forall \delta \in \Delta \tag{28}$$

$$\lambda_{i,j}^{t} \geq 0 \tag{29}$$

$$\pi_{i,j}^{t,\delta} \geq 0 \qquad \forall \delta \in \Delta \tag{30}$$

We derive the sub-problem associated with case D_b and dualize it following the same procedure. The formulation of sub-problem PD_b and its dual BD_b are reported in Appendix A.

3.1.2. Case D_c

The Benders sub-problem associated with passengers (i, j, t) where $(i, j) \in D_c$ is as follows.

$$[PD_c] \quad \min \quad \sum_{\delta \in \Delta} \sum_{\delta' \in \Delta : \delta' \geq \delta} \left(\delta' \bar{a}_{i,j}^{t} y_{i,j}^{t,\delta,\delta'} \right) \tag{31}$$

$$\sum_{\delta \in \Delta} \sum_{\delta' \in \Delta : \delta' \geq \delta} y_{i,j}^{t,\delta,\delta'} = 1 \tag{32}$$

$$\sum_{\delta' \in \Delta : \delta' \geq \delta} y_{i,j}^{t,\delta,\delta'} \leq \sum_{d \in S_{-\gamma,i}} \sum_{n \in Q_d} \bar{x}_{\upsilon,d,n}^{t+\delta-n} \qquad \forall \delta \in \Delta \tag{33}$$

$$\sum_{\delta \in \Delta : \delta \leq \delta'} y_{i,j}^{t,\delta,\delta'} \leq \sum_{d \in S_{\gamma,j}} \sum_{n \in Q_d} \bar{x}_{-\upsilon,d,n}^{t+\delta'} \qquad \forall \delta' \in \Delta \tag{34}$$

$$y_{i,j}^{t,\delta,\delta'} \geq 0 \qquad \forall \delta, \delta' \in \Delta : \delta \leq \delta' \tag{35}$$

We use $\xi_{i,j}^{t}$ as the dual variable corresponding to constraint (32), $\theta_{i,j}^{t,\delta}$ as the dual variables corresponding to constraints (33), and $\phi_{i,j}^{t,\delta}$ as the dual variables corresponding to (34). We dualize the linear relaxation of sub-problem PD_c as follows.

$$[BD_c] \quad \max \quad \eta_{i,j}^{t} = \xi_{i,j}^{t} - \sum_{\delta \in \Delta} \left(\theta_{i,j}^{t,\delta} \sum_{d \in S_{-\gamma,i}} \sum_{n \in Q_d} x_{\upsilon,d,n}^{t+\delta-n} + \phi_{i,j}^{t,\delta} \sum_{d \in S_{\gamma,j}} \sum_{n \in Q_d} x_{-\upsilon,d,n}^{t+\delta} \right) \tag{36}$$

$$\xi_{i,j}^{t} - \theta_{i,j}^{t,\delta} - \phi_{i,j}^{t,\delta'} \leq \delta' \bar{a}_{i,j}^{t} \qquad \forall \delta, \delta' \in \Delta \ : \ \delta \leq \delta' \tag{37}$$

$$\xi_{i,j}^{t} \geq 0 \tag{38}$$

$$\theta_{i,j}^{t,\delta} \geq 0 \qquad \forall \delta \in \Delta \tag{39}$$

$$\phi_{i,j}^{t,\delta} \geq 0 \qquad \forall \delta \in \Delta \tag{40}$$

3.2. Benders Master Problem

In the Benders decomposition, we use variable $\eta_{i,j}^{t}$ to denote the waiting time experienced by all passengers (i, j, t). We note that $\eta_{i,j}^{t}$ corresponds to the objective value of the Benders sub-problem associated with passengers (i, j, t). We denote by ω the iteration number of the Benders-based branch-and-cut. The set $K_{(i,j,t)}^{(\omega)}$ denotes the set of Benders optimality cuts added to the sub-problem associated with passengers (i, j, t) at iteration ω. For a given sub-problem (i, j, t) (where $(i, j) \in D_f \cup D_b$) and index $k \in K_{(i,j,t)}^{(\omega)}$, the parameters $\lambda_{i,j}^{t,(k)}$ and $\pi_{i,j}^{t,\delta,(k)}$ store the values of the dual solution that was generated in the k-th optimality cut for the sub-problem. The parameters $\xi_{i,j}^{t,(k)}, \theta_{i,j}^{t,\delta,(k)}$ and $\phi_{i,j}^{t,\delta,(k)}$ do the same in the case of $(i, j) \in D_c$. We maintain cuts over the iterations. At iteration ω the Benders master problem is as follows.

$$[\mathbf{BMP}]^{(\omega)} \quad \min \sum_{(i,j) \in D} \sum_{t \in \bar{T}_{\upsilon(i,j),i}} \eta_{i,j}^{t} \tag{41}$$

$$(6)\text{–}(9),\ (16)$$

$$\eta_{i,j}^{t} \geq \lambda_{i,j}^{t,(k)} - \sum_{\delta\in\Delta} \pi_{i,j}^{t,\delta,(k)} \sum_{d\in S_{\upsilon(i,j),j}} \sum_{n\in Q_d} x_{\upsilon(i,j),d,n}^{t+\delta} \qquad \forall (i,j)\in D_f, t\in \bar{T}_{\upsilon(i,j),i}, k\in K_{(i,j,t)}^{(\omega)} \tag{42}$$

$$\eta_{i,j}^{t} \geq \lambda_{i,j}^{t,(k)} - \sum_{\delta\in\Delta} \pi_{i,j}^{t,\delta,(k)} \sum_{d\in S_{-\upsilon(i,j),i}} \sum_{n\in Q_d} x_{-\upsilon(i,j),d,n}^{t+\delta-n} \qquad \forall (i,j)\in D_b, t\in \bar{T}_{\upsilon(i,j),i}, k\in K_{(i,j,t)}^{(\omega)} \tag{43}$$

$$\eta_{i,j}^{t} \geq \xi_{i,j}^{t,(k)} - \sum_{\delta\in\Delta} \Big(\theta_{i,j}^{t,\delta,(k)} \sum_{d\in S_{-\upsilon(i,j),i}} \sum_{n\in Q_d} x_{-\upsilon(i,j),d,n}^{t'+\delta-n} + \phi_{i,j}^{t,\delta,(k)} \sum_{d\in S_{\upsilon(i,j),j}} \sum_{n\in Q_d} x_{\upsilon(i,j),d,n}^{t'+\delta}\Big) \qquad \forall (i,j)\in D_c, t\in \bar{T}_{\upsilon(i,j),i}, k\in K_{(i,j,t)}^{(\omega)} \tag{44}$$

The objective function (41) minimizes the total passenger waiting time. Constraints (42)–(44) impose the optimality cuts generated by the Benders sub-problems. Note that, by construction, sub-problems BD_f, BD_b, and BD_c are always feasible and bounded from above. Therefore, since the sub-problems cannot be unbounded, no feasibility cuts are required.

As previously mentioned, PF is solved using a Benders-based branch-and-cut algorithm through the use of solver callbacks which are called at integer nodes of the search tree. Additionally, in the implemented algorithm, constraints (9) are treated as lazy cuts and verified when no violated Benders cuts are found. For the sake of convenience, we refer to the relaxed master problem formulation, i.e., (41), (6)–(8), (16), as RMP, which is solved using a state-of-the-art MILP solver. At integer nodes, the separation procedure is invoked. If no violated cut is found, the current solution is feasible and can be accepted as the new incumbent. If a cut is found, the node has to be solved again. Afterwards, the solver resumes the search. For simplicity, in what follows, we will be using "cuts" to refer both to the solution of the dual sub-problem and to the actual cut that is introduced in the master problem.

3.3. The Benders Cuts

In this section, we discuss several theoretical properties of the Benders cuts (42)–(44). Namely, we develop parametric closed-form solutions for each of the dual sub-problems BD_f, BD_b, and BD_c. We also show the choices of the parameters that yield an optimal solution and prove them to be Pareto-optimal according to Magnanti & Wong (1981). For the ease of notation, in the rest of this section, let $\upsilon = \upsilon(i,j)$.

3.3.1. Case D_f and Case D_b

In the subsequent discussion, we focus only on case D_f. We note that all the properties and Lemmas described for case D_f also hold for case D_b with minimal adjustments.

For (i,j,t), where $(i,j)\in D_f$, for a given $\sigma\in\mathbb{N}_0$, a special case of (42) is the following.

$$\eta_{i,j}^{t} \geq \bar{a}_{i,j}^{t}\Big(\sigma - \sum_{\delta\in\Delta:\delta\leq\sigma} (\sigma-\delta) \sum_{d\in S_{\upsilon,j}} \sum_{n\in Q_d} x_{\upsilon,d,n}^{t+\delta}\Big) \tag{45}$$

Lemma 5. *Cut (45) is valid for the DTPR, and dominates all other cuts (42) with $\lambda_{i,j}^{t} = \bar{a}_{i,j}^{t}\sigma$, $\forall\sigma\in\mathbb{N}_0$.*

Lemma 6. *Given a path vector $\bar{x}$, for a given $\sigma\in\mathbb{N}_0$, Benders cut (45) is generated by an optimal solution of the BD_f sub-problem if exactly one train is available to serve the customers (i,j) in the equivalent time interval $[t, t+\sigma]$, i.e., it holds:*

$$\sigma \geq 0 \ :\ \sum_{\delta\in\Delta:\delta\leq\sigma} \sum_{d\in S_{\upsilon,j}} \sum_{n\in Q_d} \bar{x}_{\upsilon,d,n}^{t+k} = 1 \tag{46}$$

Proposition 1. *Benders cuts of the form (45) are Pareto-optimal $\forall\sigma\in\mathbb{N}_0$.*

3.3.2. Case D_c

For (i,j,t), where $(i,j) \in D_c$, the optimality cut associated with sub-problem BD_c is (44). Given $\sigma, \mu \in \mathbb{N}_0 : \mu \leq \sigma$, a special case of cut (44) is the following.

$$\eta^t_{i,j} \geq \bar{a}^t_{i,j}\Big(\sigma - \sum_{\delta\in\Delta:\delta<\mu}(\sigma-\delta)\sum_{d\in S_{-\upsilon,i}}\sum_{n\in Q_d} x^{t+\delta-n}_{-\upsilon,d,n} - \sum_{\delta\in\Delta:\mu\leq\delta<\sigma}(\sigma-\delta)\sum_{d\in S_{\upsilon,j}}\sum_{n\in Q_d} x^{t+\delta}_{\upsilon,d,n}\Big) \tag{47}$$

Lemma 7. *Let $\sigma, \mu \in \mathbb{N}_0 \ : \ \mu \leq \sigma$, cut (47) is valid for DTPR.*

Given a path vector $\bar{x}$, let us consider the trajectory of passengers (i,j,t). As for Lemma 3, passengers always reach their destination boarding at most two trains. Thus, we can distinguish two cases: 1) passengers (i,j,t) board exactly one train to reach their destination, or 2) passengers (i,j,t) board two different trains, alighting at station r.

We now show how $\bar{x}$ relates to these two cases. Let us denote by t_1 the earliest equivalent time-step passengers (i,j,t) can depart from i to reach station r. Similarly, let t_2 be the earliest equivalent time-step passengers (i,j,t) can depart from r to reach station j. If $t_1 = t_2$ passengers (i,j,t) will board one train to reach their destination, i.e., case 1). Conversely, if $t_1 < t_2$ they will instead board two trains, i.e., case 2). Formally, case 1) implies the following.

$$\begin{aligned}\exists p \geq 0 \ : \ &\sum_{d\in S_{-\upsilon,i}}\sum_{n\in Q_d} \bar{x}^{t+p-n}_{-\upsilon,d,n} = 1 \quad\wedge\quad \sum_{\delta\in\Delta:\delta\leq p}\sum_{d\in S_{-\upsilon,i}}\sum_{n\in Q_d} \bar{x}^{t+\delta-n}_{-\upsilon,d,n} = 1 \\ \wedge\quad &\sum_{d\in S_{\upsilon,j}}\sum_{n\in Q_d} \bar{x}^{t+p}_{\upsilon,d,n} = 1\end{aligned} \tag{48}$$

Case 2) implies the following.

$$\begin{aligned}\exists p' > p \geq 0 \ : \ &\sum_{d\in S_{-\upsilon,i}}\sum_{n\in Q_d} \bar{x}^{t+p-n}_{-\upsilon,d,n} = 1 \quad\wedge\quad \sum_{\delta\in\Delta:\delta<p}\sum_{d\in S_{-\upsilon,i}}\sum_{n\in Q_d} \bar{x}^{t+\delta-n}_{-\upsilon,d,n} = 0 \\ \wedge\quad &\sum_{d\in S_{\upsilon,j}}\sum_{n\in Q_d} \bar{x}^{t+p'}_{\upsilon,d,n} = 1 \quad\wedge\quad \sum_{\delta\in\Delta:p<\delta<p'}\sum_{d\in S_{\upsilon,j}}\sum_{n\in Q_d} \bar{x}^{t+\delta}_{\upsilon,d,n} = 0.\end{aligned} \tag{49}$$

Lemma 8. *Given a path vector $\bar{x}$, if $\bar{x}$ satisfies (48), there exist $\sigma, \mu \in \mathbb{N}_0$ such that: 1) in the equivalent time interval $[t, t+\mu]$ exactly one train is available to bring the passengers from station i to station j at equivalent time-step $t+k$, with $k \in [0,\mu]$, 2) in the equivalent time interval $[t, t+\mu]$ no other trains are available to bring the passengers at station i to r, and 3) in the equivalent time interval $[t+\mu, t+\sigma]$ no trains depart from r headed to station j. Formally:*

$$1)\quad \exists k \in [0,\mu] \ : \sum_{d\in S_{-\upsilon,i}}\sum_{n\in Q_d} \bar{x}^{t+k-n}_{-\upsilon,d,n} = 1 \ \wedge \sum_{d\in S_{\upsilon,j}}\sum_{n\in Q_d} \bar{x}^{t+k}_{\upsilon,d,n} = 1 \tag{50}$$

$$2)\quad \sum_{\delta\in\Delta:\delta\leq\mu}\sum_{d\in S_{-\upsilon,i}}\sum_{n\in Q_d} \bar{x}^{t+\delta-n}_{-\upsilon,d,n} = 1 \tag{51}$$

$$3)\quad \sum_{\delta\in\Delta:\delta>\mu}\sum_{d\in S_{\upsilon,j}}\sum_{n\in Q_d} \bar{x}^{t+\delta}_{\upsilon,d,n} = 0 \tag{52}$$

Then an optimal solution of BD_c generates cut (47).

Lemma 9. *Given a path vector $\bar{x}$, if $\bar{x}$ satisfies (49), there exist $\sigma, \mu \in \mathbb{N}_0 : \sigma \geq \mu$ such that: 1) in the equivalent time interval $[t+\mu, t+\sigma]$ exactly one train departs from station r headed to station j, at equivalent time-step $t+k$, with $k \in [\mu, \sigma]$, 2) in the equivalent time interval $[t, t+\mu)$ no train is available to bring* *the passengers from station i to r, and 3) in the equivalent time interval $[t+\mu, t+k]$ at least one train is available to bring the passengers from station i to station r. Formally:*

$$1) \quad \exists k \in [\mu, \sigma] \; : \sum_{\delta \in \Delta: \mu < h \leq k} \sum_{d \in S_{-\upsilon,i}} \sum_{n \in Q_d} \bar{x}^{t+\delta-n}_{-\upsilon,d,n} = 1 \;\wedge\; \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+k}_{\upsilon,d,n} = 1 \tag{53}$$

$$2) \quad \sum_{\delta \in \Delta: \delta < \mu} \sum_{d \in S_{-\upsilon,i}} \sum_{n \in Q_d} \bar{x}^{t+\delta-n}_{-\upsilon,d,n} = 0 \tag{54}$$

$$3) \quad \sum_{\delta \in \Delta: \mu < h \leq \sigma} \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+\delta}_{\upsilon,d,n} = 1 \tag{55}$$

Then an optimal solution of BD_c generates cut (47).

Proposition 2. *Benders cuts of the form (47), are Pareto-optimal $\forall \sigma, \mu \in \mathbb{N}_0 : \mu \leq \sigma$.*

3.4. Separation of the Benders Cut

The implemented separation procedure is invoked at integer nodes of the search tree. When invoked, the incumbent solution is projected on all sub-problems, to search for violated Benders cuts. Sub-problems BD_f, BD_b, and BD_c are always feasible and bounded from above. Therefore, for any sub-problem, substituting the optimal dual values of the Benders sub-problem in constraints (42), (43) and (44) leads to a valid optimality cut for the master problem. If no violated Benders cuts are found, we verify if constraints (9) are satisfied; if not, then the violated constraint (9) is added to the master problem. To summarize, the pseudo-code of the entire solution procedure is reported in Algorithm 1.

In the rest of this Section, we present the theoretical framework for separating the Benders cuts, given an incumbent solution satisfying (6)–(8), (16). We specify the necessary and sufficient conditions to verify if a cut needs to be added to the BMP and the rules for the selection of violated cuts given an incumbent solution. These results allow us to separate Pareto optimal violated cuts for all Benders sub-problems in polynomial time, without having to solve the dual problems.

3.4.1. Case D_f and Case D_b

We now discuss the separation of Benders cuts associated with sub-problems (i, j, t), where $(i, j) \in D_f$. Note that all presented results and Lemmas can be adapted to case D_b with minimal adjustments.

Proposition 3. *Given path vector $\bar{x}$ associated with the current BMP solution, a cut associated with sub-problem (i, j, t), where $(i, j) \in D_f$, needs to be generated if and only if either of the following conditions hold:*

$$1) \quad \zeta = \frac{\eta^t_{i,j}}{\bar{a}^t_{i,j}} \text{ is not integer,} \tag{56}$$

$$2) \quad \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+\zeta}_{\upsilon,d,n} = 0. \tag{57}$$

Lemma 10. *Given path vector $\bar{x}$ associated with the current BMP solution, if all the Benders cuts associated with sub-problem (i, j, t), where $(i, j) \in D_f$, are of the form (45) with $\sigma \in \mathbb{N}_0$, the quantity $\zeta = \eta^t_{i,j}/\bar{a}^t_{i,j}$ is integer.*

Proposition 4. *Given path vector $\bar{x}$ associated with the current BMP solution, if all the Benders cuts associated with sub-problem (i, j, t) where $(i, j) \in D_f$ are of the form (45) with $\sigma \in \mathbb{N}_0$, and a cut needs to be generated for the sub-problem; a cut of the form (45) with $\sigma = \zeta + 1$ where $\zeta = \eta^t_{i,j}/a^t_{i,j}$, is violated by the incumbent solution.*

Algorithm 1: The implemented Benders-based branch-and-cut.

Input: initial solution $\bar{x}_0$

$UB \leftarrow$ the total passenger waiting time resulting from $\bar{x}_0$, $LB \leftarrow 0$;
initialize Branch-and-bound;
while $UB + \epsilon < LB \ \wedge \ elapsed\_time < time\_limit$ **do**
  $added \leftarrow False$, solve RMP;
  $\bar{x} \leftarrow$ the solution to RMP;
  if $\bar{x}$ *is integer* **then**
    for $(i, j) \in D_f \cup D_b$ **do**
      for $t \in \bar{T}_{\upsilon,i}$ **do**
        if (57) *holds* **then** add cut (45) with $\sigma = \zeta$ for (i, j, t) to RMP, $added \leftarrow True$;
    for $(i, j) \in D_c$ **do**
      $t \leftarrow \bar{t}^1_{i,\gamma} - 1$, $\upsilon \leftarrow \upsilon(i, j)$, $k \leftarrow t$;
      while $t \in \bar{T}_{\upsilon,i}$ **do**
        $t \leftarrow t + 1$;
        if $t > T$ **then**
          $k \leftarrow t$;
          while $k \in \bar{T}_{\upsilon,i} \wedge \sum_{d \in S_{-\upsilon,i}} \sum_{n \in Q_d} \bar{x}^{t'}_{-\upsilon,d,n} = 0$ **do** $k \leftarrow k + 1$;
        $\zeta \leftarrow \frac{\eta^t_{i,j}}{\bar{a}^t_{i,j}}$;
        if $k > t + \zeta$ **then**
          add cuts (47) with $\sigma = \zeta$, $\forall \mu \in [0, \sigma]$ for (i, j, t) to RMP, $added \leftarrow True$;
        else if $\sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+\zeta}_{\upsilon,d,n} = 0$ **then**
          add cuts (47) with $\sigma = \zeta$, $\forall \mu \in [0, \sigma]$ for (i, j, t) to RMP, $added \leftarrow True$;
    if $added = False$ **then**
      for $\forall \gamma \in \Gamma, d \in S_\gamma, n \in Q_d, t \in T, t' \in \mathbb{T}^t_d$ **do**
        verify if (9) holds, if not add corresponding constraint to RMP, $added \leftarrow True$;
    if $added = True$ **then goto** line 4 ;
  update bounds UB and LB;
  branch;

To separate violated cuts, given $\bar{x}$, for each sub-problem (i, j, t), where $(i, j) \in D_f$, we verify condition (57). If condition (57) holds, a cut needs to be generated. As for Proposition 3, a cut should be generated if either (56) or (57) hold. However, in the algorithm, all Benders cuts associated with the sub-problem follow the structure of (45) with $\sigma \in \mathbb{N}_0$. Therefore, as for Lemma 10, condition (56) is never verified. If a cut needs to be generated, cut (45) with $\sigma = \eta^t_{i,j}/a^t_{i,j} + 1$ is violated by the incumbent solution, as for Proposition 4. Cuts generated following this procedure are violated by exactly $\bar{a}^t_{i,j}$.

We note that $|S_{\upsilon,j}| \le |S|$, $\forall \upsilon, j$, and that $|Q_d| = 1 + \beta$, $\forall d$. Therefore, in the worst case scenario, verifying condition (57) involves checking $(\beta + 1)|S|$ path variables. By extension, the separation of violated cuts associated with sub-problem (i, j, t) has a computational complexity $\mathcal{O}(\beta|S|)$.

3.4.2. Case D_c

We now discuss the separation of Benders cuts associated with sub-problems (i, j, t), where $(i, j) \in D_c$.

Proposition 5. *Given path vector $\bar{x}$ associated with the current BMP solution, a cut associated with sub-problem (i, j, t) where $(i, j) \in D_f$ needs to be generated if and only if at least one the following conditions* *hold:*

$$1)\quad \zeta = \frac{\eta^t_{i,j}}{\bar{a}^t_{i,j}} \text{ is not integer,} \tag{58}$$

$$2)\quad \zeta \text{ is integer } \wedge \sum_{d\in S_{v,j}} \sum_{n\in Q_d} \bar{x}^{t+\zeta}_{v,d,n} = 0 \tag{59}$$

$$3)\quad \nexists k' \;:\; 0 \le k' \le \zeta \;\wedge\; \sum_{d\in S_{-v,i}} \sum_{n\in Q_d} \bar{x}^{t+k'-n}_{-v,d,n} = 1 \tag{60}$$

Lemma 11. *Given path vector $\bar{x}$ associated with the current BMP solution, if all the Benders cuts associated with sub-problem (i,j,t), where $(i,j)\in D_c$, are of the form (47) with $\sigma \le \mu \in \mathbb{N}_0$; the quantity $\zeta = \eta^t_{i,j}/\bar{a}^t_{i,j}$ is integer.*

Proposition 6. *Given path vector $\bar{x}$ associated with the current BMP solution, if all the Benders cuts associated with sub-problem (i,j,t), where $(i,j)\in D_c$, are of the form (47) with $\sigma \le \mu \in \mathbb{N}_0$, and a cut needs to be generated for the sub-problem; there exists a cut of the form (47) with $\sigma = \zeta + 1$ where $\zeta = \eta^t_{i,j}/a^t_{i,j}$ that is violated by the incumbent solution.*

To separate violated Benders cuts, given an incumbent solution, for each sub-problem (i,j,t), where $(i,j)\in D_c$, we verify conditions (59) and (60). If either condition holds, a cut needs to be generated. Similar to case D_f, Proposition 5 states that a cut needs to be generated if (58) or (59) or (60) hold. However, in the algorithm, all Benders cuts associated with the sub-problem follow the structure of (47) with $\sigma \le \mu \in \mathbb{N}_0$. Therefore, as for Lemma 11, condition (58) is never verified. If a cut needs to be generated, given an incumbent solution, Proposition 6 guarantees that there exists a cut (47) with $\sigma = \lambda^t_{i,j}/a^t_{i,j}$ and $\mu \le \sigma$ that is violated. All cuts with $\sigma = \lambda^t_{i,j}/a^t_{i,j}$ are added simultaneously, i.e., when a cut needs to be generated, a cut for each value of $\mu \in [0,\sigma]$ is added to the master problem.

To accelerate the separation, we aggregate sub-problems $\{(i,j,t)\ \forall t \in \bar{T}_{v,i}\}$. Let us denote by $P(t)$ the first equivalent time-step after t at which a train is available to bring passengers waiting at station i to r, i.e.,

$$P(t) = \min_{t'\in T}\Big(t' \;:\; t' \ge t \;\wedge\; \sum_{d\in S_{v,j}} \sum_{n\in Q_d} \bar{x}^{t'}_{v,d,n} = 1\Big).$$

We use $P(t)$ to rewrite (60) as follows:

$$P(t) > t + \zeta. \tag{61}$$

Thus, it is possible to evaluate condition (60) in constant time, provided we are able to compute $P(t)$. The value of $P(t)$ can be efficiently computed incrementally. Given $P(t-1)$, if $t-1 \ne P(t-1)$ the following holds:

$$P(t) = \begin{cases} P(t) = P(t-1) & \text{if } t-1 \ne P(t-1), \\ \min_{t'\in T}\Big(t' \;:\; \sum_{d\in S_{v,j}} \sum_{n\in Q_d} \bar{x}^{t'}_{v,d,n} = 1 \;\wedge\; t' \ge t\Big) & \text{if } t-1 = P(t-1). \end{cases} \tag{62}$$

Using this result, for a given $(i,j)\in D_c$, we perform the separation of violated cuts on all sub-problems $\{(i,j,t)\ \forall t \in \bar{T}_{v,i}\}$ sequentially, iterating over the planning horizon (lines 10–21 in Algorithm1). Overall, computing $P(t)\ \forall t \in \bar{T}_{v,i}$ has a complexity of $\mathcal{O}(\beta|S||T|)$. Therefore, separating violated cuts on sub-problems $\{(i,j,t)\ \forall t \in \bar{T}_{v,i}\}$ has a complexity of $\mathcal{O}(\beta|S||T|)$. In total, the number of sub-problems is $|S|^2|T|$. Thus, the computational complexity associated with the separation of all sub-problems is $\mathcal{O}(\beta|S|^3|T|)$.

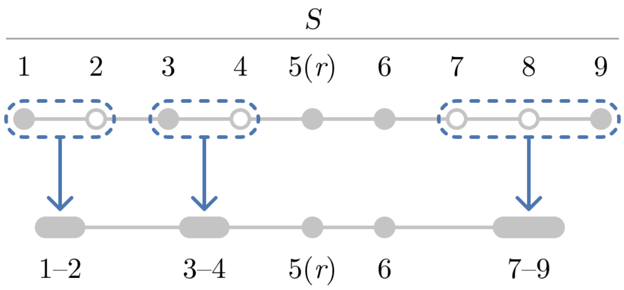


Figure 7: Example of the merging of the non-short-turning stations with the short-turning stations. Stations drawn as a gray dot are short-turning stations, stations drawn as a white circle are non-short-turning.

4. Compression of the Line

In most real lines, short-turning can be performed at a few stations, rather than at all stations of the line, as it was assumed in the model. The methods we presented thus far can be adapted to accommodate such a case by fixing to zero any path variable whose destination is a non-short-turning station. In this section, we present an alternative strategy for enforcing non-short-turning stations, which yields a more compact model.

We denote by $\bar{S}$ the set of stations where trains are allowed to short-turn, with $\bar{S} \subseteq S$ When $\bar{S} \subset S$ we apply a compression to the line, to remove unnecessary non-short-turning stations. See Figure 7 for an illustration of this concept. By definition, the destination of a feasible path is a short-turning station. We enforce this by redefining the set $S_{\gamma,i}$ to exclude any non-short-turning station. We denote by $\bar{S}_{\gamma,i}$ the set of short-turning stations that lie in direction γ relative to i. Formally:

$$\bar{S}_{\gamma,i} \equiv S_{\gamma,i} \cap \bar{S} \qquad \forall i \in S, \gamma \in \Gamma$$

For a given station i, the set $\bar{S}_{\upsilon(r,i),i}$ denotes the set of possible destinations of a path that visits station i. To reduce the size of the model, we merge various stations into station groups, ensuring that all non-short-turning stations are combined with exactly one short-turning station. Specifically, each non-short-turning station $j \in S \backslash \bar{S}$ is merged with its closest short-turning station in direction $\upsilon(r,j)$. We denote by h_j the short-turning station that is merged with station j. Formally:

$$h_j = \min_{i \in \bar{S}_{\upsilon(r,j),j}} |i - r| \qquad \forall j \in S$$

A station group is a set of stations that are merged with the same short-turning station. We denote by $\mathbb{S}_i$ the station group associated with $i \in \bar{S}$. We note that the root station is never merged with any other station, i.e., $\mathbb{S}_r \equiv \{r\}$. We distinguish two categories of passengers: 1) passengers with $(i,j) \in D$ such that $h_i \neq h_j$, and 2) passengers with $(i,j) \in D$ such that $h_i = h_j$.

Let us consider passengers (i,j,t) such that $h_i \neq h_j$, i.e., passengers whose origin and destination belong to different station groups. From a modeling perspective, passengers (i,j,t) are equivalent to passengers (h_i,h_j,t). Indeed, by the definition of h_i, any train path that reaches station i also reaches station h_i, at the same equivalent time-step. Similarly, any train path that reaches station j also reaches station h_j at the same equivalent time-step. Thus, passengers (h_i,h_j,t) and (i,j,t), are served by the same train paths and thus experience the same delay reaching their destination. We recall that the arrival time of the passengers is expressed in equivalent time. Therefore, the actual arrival time of passengers (i,j,t), and passengers (h_i,h_j,t) does not coincide if $i \neq h_i$ (see Section 2.2). Additionally, the travel time of the passengers is not accounted for in the computation of the delay; therefore, the total trip duration of the passengers may differ since passengers (h_i,h_j,t) and (i,j,t) may require different travel times on the line.

Since the passengers (h_i,h_j,t) and (i,j,t) can be modeled equivalently, we combine all passengers (i',j',t) such that $h_i = h_{i'}, h_j = h_{j'}$, and treat them as a single set of passengers with (h_i,h_j), arriving at equivalent time-step t. Given station groups $S_i \in \bar{S}$ and $S_j \in \bar{S}$, we denote by $\alpha^t_{i,j}$ the total number of passengers (i', j', t) such that $h_i = h_{i'},\ h_j = h_{j'}$, i.e.,

$$\alpha^t_{i,j} = \sum_{i' \in S_i} \sum_{j' \in S_j} \bar{a}^t_{i',j'}.$$

We now consider the case of passengers (i, j, t) such that $h_i = h_j$, i.e., passengers whose origin and destination belong to the same station group. In this context, we further distinguish two cases, either $(i, j) \in D_f$, or $(i, j) \in D_b$. We note that the case where $(i, j) \in D_c$ cannot occur since $\mathbb{S}_r \equiv \{r\}$. Focusing on the first case, $(i, j) \in D_f$, let us consider two passengers (i, j, t) and (i', j', t), with $h_i = h_j = h_{i'} = h_{j'}$ and $(i, j), (i', j') \in D_f$. From a modeling perspective, passengers (i, j, t) are equivalent to passengers (i', j', t). Indeed, any path that serves station i in direction $v(i, j)$, also serves all stations in $\mathbb{S}_i$ in direction $v(i, j)$ at the same equivalent time-step. Thus, for a given time-step t and short-turning station i, all passengers (i', j', t) such that $(i', j') \in D_f$ and $h_{i'} = h_{j'} = h_i$ can board the same trains and experience the same delay. Similarly, for a given time-step t and short-turning station i, all passengers (i', j', t) such that $(i', j') \in D_b$ and $h_{i'} = h_{j'} = h_i$ are also equivalent to one another. Thus, similar to the previously discussed case (i.e., $h_i \neq h_j$), those passengers can be merged and modeled as a group. We will call this portion of passenger demand the *self-demand* of a station group.

We introduce some additional notation to allow us to distinguish the two cases $(i, j) \in D_f$, and $(i, j) \in D_b$. We denote as (i, i^+) a fictional OD pair that represents the self-demand of a short-turning station i associated with passengers whose OD pairs belong to D_f. Similarly, we denote as (i, i^-) the fictional OD pair that represents the self-demand of a short-turning station i associated with passengers whose OD pairs belong to D_b.

Given a station group $S_i \in \bar{S}$ we denote by α^t_{i,i^+} the total number of passengers (i', j', t) such that $h_i = h_{i'} = h_{j'},\ (i, j) \in D_f$, and by α^t_{i,i^-} the total number of passengers (i', j', t) such that $h_i = h_{i'} = h_{j'},\ (i, j) \in D_b$. Formally:

$$\alpha^t_{i,i^-} = \sum_{i' \in S_i} \sum_{j \in S_i : (i', j) \in D_b} \bar{a}^t_{i',j}$$

$$\alpha^t_{i,i^+} = \sum_{i' \in S_i} \sum_{j \in S_i : (i', j) \in D_f} \bar{a}^t_{i',j}$$

Finally, To solve the the problem, we use the same models and algorithms used to solve the DTPR, with the following adjustments. The set of stations S is replaced by the set of short-turning stations $\bar{S}$. The set of OD pairs D is extended to include the fictional OD pairs (i, i^+) and (i, i^-) $\forall i \in \bar{S}$. OD pairs (i, i^+) and (i, i^-) are to be treated as any other OD pair, using i as both the origin and destination stations of the passenger, and considering $(i, i^+) \in D_f$ and $(i, i^-) \in D_b$. The demand $a^t_{i,j}$ is replaced by the parameter $\alpha^t_{i,j}$.

5. Experimental Results

We carried out a series of computational tests to evaluate the effectiveness of the PF model and the Benders-based branch-and-cut algorithm, as well as the effectiveness of the compression approach presented in Section 4. In Subsection 5.1, we discuss the implementation details. Initially, we assume that trains can short-turn at each station. In Subsection 5.2, we compare the computational performance of the PF, the Benders-based branch-and-cut algorithm, and a naive implementation of the Benders decomposition. In Subsection 5.3, we compare the DTPR solutions to the frequently used regular timetables. In Subsection 5.4, we consider that trains may short-turn in a subset of the stations; using these instances, we demonstrate the effectiveness of the compression procedure presented in Section 4.

5.1. Implementation

All experiments have been performed on a 3.20Ghz AMD Ryzen 5 1600, with 16 GB of RAM and 12 cores, running Windows 10. We implemented formulation PF for the DTPR (Section 2), and the Benders-based branch-and-cut algorithm (BD) for the DTPR (Section 3). Additionally, we consider a second implementation of the Benders decomposition algorithm which uses CPLEX's built-in automated Benders decomposition (Bonami et al., 2020), we denote this by BDA. In this case, the path variables $x^t_{\gamma,d,n}$ are assigned to the master problem, and the passenger variables $y^{t,\delta}_{i,j}$ and $y^{t,\delta,\delta'}_{i,j}$ are set to the Benders sub-problem. This allows us to compare our BD to an efficient non-tailored implementation of a Benders decomposition.

PF, BDA, and BD are solved using the CPLEX solver version 12.10.0.0 and have been implemented in Julia 1.4.2 using the JuMP modeling interface version 0.18.5 (Dunning et al., 2017). In the solution of formulation PF, constraints (9) are added to the problem as lazy cuts. As previously stated, in the BD, those cuts are evaluated in case no Benders cuts are found. Each experiment was run on a single thread, the search is stopped when the optimality gap is smaller or equal to 10^{-5} or when the time limit of 3600 seconds is reached.

PF, BDA and BD are provided with an initial feasible solution computed through the construction of a regular timetable for the line. The regular timetable is constructed scheduling trains from the root station at regular intervals. Afterwards, the solution is completed by computing the passenger waiting time induced by the regular timetable.

We adapted the instances of Schettini et al. (2021a) to the DTPR. Specifically, we considered a set of artificial instances with $|S| = \{5, 10, 15, 20\}$ stations, and planning horizons of $h = \{10, 20, \ldots, 100\}$ time-steps. The maximum number of trains operating on the line is set to $\Psi = |S| - 1$. The maximum number of time-steps passengers may wait g is set to 10, and the maximum number of time-steps a train may idle β is set to 5. Additionally, the root station was set either to station 1 or to the centermost station of the line, i.e., $\lfloor |S|/2 \rfloor$. Overall, the number of instances considered was 80. The passenger demand was generated by sampling a Poisson distribution at each time-step for each OD pair. Each station is assigned a popularity score. The mean of the Poisson distribution for a given OD pair is computed as the product of the popularity score of the origin and the popularity score of the destination station. The test instances are available at Schettini (2020).

5.2. Algorithmic Comparison

In this section, we compare the performance of PF, BDA, and BD. Table 1 reports a summary of this comparison. Detailed results are found in Tables 4 and 5 in Appendix C. Additionally, for each algorithm (PF, BDA, BD) and root position (i.e., $r = 1$, $r = \lfloor |S|/2 \rfloor$) we plotted the number of instances solved to optimality as a function of the computational time. We report the plots in Figure 8. We observe that BDA outperforms PF, when the root station is located in $r = \lfloor |S|/2 \rfloor$. Conversely, PF outperforms BDA, when the root station is located in $r = 1$. This may be explained by the fact that PF contains significantly more variables when the root station is located in the middle of the line. Indeed, in this case, passengers (i, j, t) with $(i, j) \notin D_c$ require g variables to be modeled. Conversely, passengers (i, j, t) with $(i, j) \in D_c$ require $\frac{(g+2)(g+1)}{2}$ variables. In general, we note that the performance of both PF and BDA significantly deteriorates on large instances. On some instances, CPLEX is not able to solve the linear relaxation of the problem within the time limit, resulting in a 100% optimality gap.

Overall, the results show that on average our developed BD outperforms all other methods, solving more instances to optimality, achieving the best optimality gaps, and requiring the lowest average computational time. Using BD, we are able to solve all instances with up to 80 time-steps within an optimality gap of 2.2% or less.

5.3. Comparison with Regular Timetables

In this section, we compare the DTPR solutions to the regular timetables to evaluate the benefits of the optimization. In the analysis, horizons of 10 and 20 time-steps are ignored as they would bias the results in favor of the DTPR.

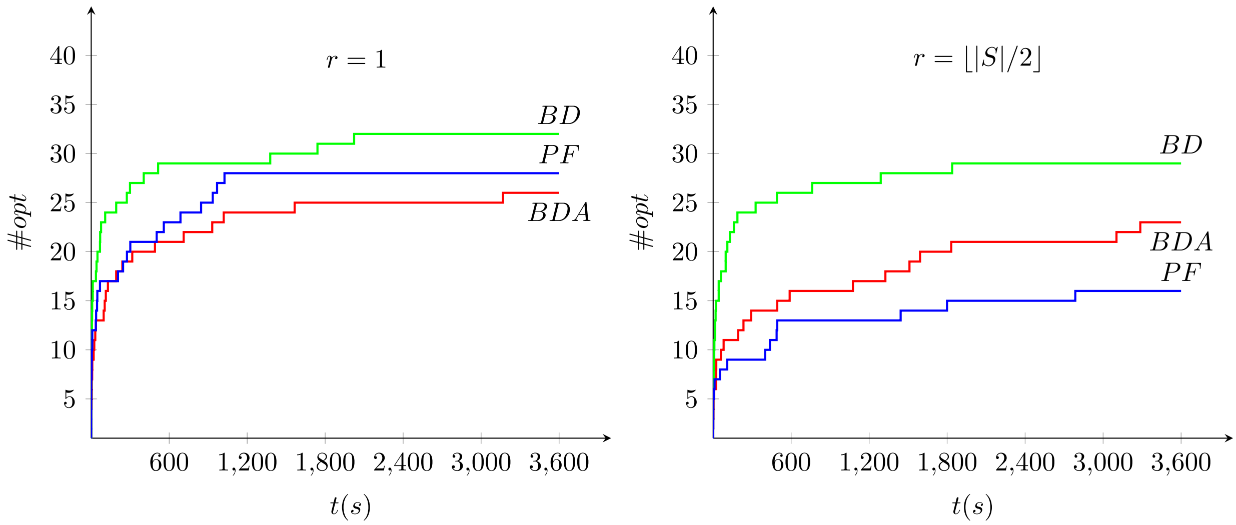


Figure 8: Number of instances solved (out of 40) for the various solution algorithms as a function of computational time. On the right side, instances with $r = 1$ are reported, on the left side, instances with $r = \lfloor |S|/2 \rfloor$. Reported algorithms are BD (Benders Decomposition), PF (Path Formulation), and BDA (Automated Benders Decomposition).

Table 1: Results for PF, BDA, and BD, average on 10 instances.

		PF				BDA				BD			
$\|S\|$	r	#s	t(s)	TWT	gap(%)	#s	t(s)	TWT	gap(%)	#s	t(s)	TWT	gap(%)
5	1	10	160	2111	0.01	10	137	2111	0.00	10	45	2111	0.00
10	1	8	969	3054	0.28	7	1329	3169	2.63	10	384	3050	0.01
15	1	5	1982	4229	1.18	5	1988	4544	5.48	6	1646	4248	1.16
20	1	5	1890	4751	0.81	4	2507	5163	6.16	6	1538	4747	0.59
all		**28**	**1250**	**3536**	**0.57**	**26**	**1490**	**3747**	**3.57**	**32**	**903**	**3539**	**0.44**
		PF				BDA				BD			
$\|S\|$	r	#s	t(s)	TWT	gap(%)	#s	t(s)	TWT	gap(%)	#s	t(s)	TWT	gap(%)
5	3	7	1325	2116	0.84	9	883	2111	0.22	8	817	2114	0.44
10	5	4	2345	3127	3.10	7	1684	3124	2.01	8	807	3052	0.20
15	8	3	2843	4752	7.77	4	2436	4717	15.06	6	1498	4388	0.57
20	10	2	2930	5319	24.50	3	2790	5347	15.61	7	1411	5034	0.84
all		**16**	**2361**	**3828**	**9.05**	**23**	**1948**	**3825**	**8.23**	**29**	**1133**	**3647**	**0.51**

Table 2 reports the comparison of the DTPR with the regular timetables. We compare solutions w.r.t. the total waiting time of the passengers (TWT), the maximum number of passengers on board a train ($\max(B)$), the mean value of the number of passengers on board a train ($\text{avg}(B)$) and its variance ($\sigma^2(B)$). We denote by D the measures associated with the DTPR solution and by R measures associated with the regular timetable.

From Table 2, we observe that the DTPR solutions tend to reduce the maximum number of boarded passengers, the variance observed in the number of boarded passengers. Conversely, the average number of passengers on board a train tends to increase. This implies that DTPR timetables typically achieve a better utilization of the available trains, with fewer trains that are completely empty or carry a very small number of passengers. Additionally, the maximum number of boarded passengers tends to be reduced, thus reducing the probability of observing overcrowded trains.

5.4. Application to Lines with a Subset of Short-turning Stations

Now, we discuss the impact of reducing the number of short-turning stations on the DTPR. We modify the previously presented instances so that a percentage of the stations are non-short-turning. Specifically,

Table 2: Comparison of the DTPR solutions to the regular timetables, average on 8 instances with $h \in \{30, 40, \dots, 100\}$.

$\|S\|$	r	$\frac{(\text{TWT}_R - \text{TWT}_D)}{\text{TWT}_R}(\%)$ all	optimal	$\frac{(\max(B_R) - \max(B_D))}{\max(B_R)}(\%)$ all	optimal	$\frac{(\text{avg}(B_R) - \text{avg}(B_D))}{\text{avg}(B_R)}(\%)$ all	optimal	$\frac{(\sigma^2(B_R) - \sigma^2(B_D))}{\sigma^2(B_R)}(\%)$ all	optimal
5	1	3.3	3.3	0.7	0.7	-3.9	-3.9	2.7	2.7
10	1	14.2	14.2	21.4	21.4	-3.6	-3.6	3.3	3.3
15	1	13.9	15.8	18.1	18.4	-4.5	-4.4	4.2	4.0
20	1	14.4	16.3	12.5	11.5	-4.6	-4.5	4.4	4.3
all		**11.4**	**12.4**	**13.2**	**13.0**	**-4.1**	**-4.1**	**3.6**	**3.6**

$\|S\|$	r	$\frac{(\text{TWT}_R - \text{TWT}_D)}{\text{TWT}_R}(\%)$ all	optimal	$\frac{(\max(B_R) - \max(B_D))}{\max(B_R)}(\%)$ all	optimal	$\frac{(\text{avg}(B_R) - \text{avg}(B_D))}{\text{avg}(B_R)}(\%)$ all	optimal	$\frac{(\sigma^2(B_R) - \sigma^2(B_D))}{\sigma^2(B_R)}(\%)$ all	optimal
5	3	3.2	3.5	0.3	-0.3	-3.9	-4.2	2.6	2.8
10	5	14.1	15.0	22.0	23.3	-3.6	-3.9	3.3	3.6
15	8	10.8	12.0	20.7	18.0	-2.5	-2.6	2.3	2.4
20	10	9.4	10.9	19.9	18.8	-1.9	-2.1	1.8	2.0
all		**9.4**	**10.3**	**15.8**	**15.0**	**-3.0**	**-3.2**	**2.5**	**2.7**

with the exclusion of the extreme stations of the line, we set one out of four stations or one out of two stations as non-short-turning. When presenting the results, we denote the case where short-turning is allowed at all stations as S0, the case where one out of four stations are non-short-turning as S1, and the case where one out of two stations are non-short-turning as S2.

Table 3 reports the performance achieved by the BD in cases S0, S1 and S2. For case S1 and S2, two versions of the BD are presented: 1) by fixing the relevant variables to zero in the BD (No Compression), 2) by utilizing the compression procedure presented in Section 4. The table reports the average solution times in seconds (t(s)), the average optimality gaps, and the number of instances solved to optimality out of 10 (#s), the average objective value (TWT). Detailed results are found in Appendix C, in Tables 8 and 9.

We observe that reducing the number of short-turning stations allows BD to solve a larger number of instances. Overall, utilizing the compression procedure improves the performance of the algorithm. Utilizing compression, we are able to solve more instances in less time and achieve better optimality gaps. In both cases S1 and S2, we are able to solve to optimality 77 out of 80 test instances when using compression.

6. Concluding Remarks

In this paper we propose the DTPR problem, which is a particular case of the more general DTP. In the DTPR, we utilize a flexible scheduling strategy to minimize the total waiting time of the passengers. Unlike traditional approaches, in the DTPR, the trains are scheduled individually, and are allowed to short-turn. Compared to the DTP, in the DTPR we assume the presence of a designated root station, i.e., a station that cannot be skipped by short-turning, and we assume that trains are not allowed to idle with passengers on board. We developed a path-based formulation for the DTPR and an efficient Benders-based branch-and-cut algorithm. We derived closed-form solutions of the Benders sub-problems used in the decomposition, proved their optimality, and specify an efficient algorithm to generate violated cuts for a given incumbent solution. Furthermore, we developed compression approach to reduce the size of the problem when short-turning is not allowed at all stations of the line.

We showed the effectiveness of the developed formulation and the Benders-based branch-and-cut algorithm. Then, we compared the DTPR solutions to the commonly used regular timetables. We observed the DTPR solutions tend to reduce the maximum number of boarded passengers, the variance of the occupancy, and the distance traveled by the trains. Lastly, we tested the developed compression approach on instances where short-turning is not allowed at all stations of the line, and evaluated the impact of such limitation. We tested two short-turning configurations, with a different number of allowed short-turning stations. In both configurations, we observed that the developed compression yields significant computational benefits, allowing us to solve 77 out of 80 instances.

Table 3: Results of the Benders-based branch-and-cut algorithm in configurations S0 and S1, with and without compression of the graph.

		S0				S1				S1-compression			
$\|S\|$	r	#S	t(s)	TWT	gap(%)	#S	t(s)	TWT	gap(%)	#S	t(s)	TWT	gap(%)
5	1	10	45	2111	0.0	10	1	2127	0.0	10	1	2127	0.0
10	1	10	384	3050	0.0	9	692	3051	0.1	10	298	3050	0.0
15	1	6	1646	4248	1.2	9	786	4227	0.2	9	602	4222	0.0
20	1	6	1538	4747	0.6	9	665	4770	0.0	10	604	4770	0.0
all		**32**	**903**	**3539**	**0.4**	**37**	**536**	**3544**	**0.1**	**39**	**376**	**3542**	**0.0**

		S2				S2-compression			
$\|S\|$	r	#S	t(s)	TWT	gap(%)	#S	t(s)	TWT	gap(%)
5	1	10	1	2127	0.0	10	1	2127	0.0
10	1	8	782	3082	0.1	10	501	3081	0.0
15	1	8	817	4251	0.2	8	771	4249	0.1
20	1	10	716	4778	0.0	10	298	4778	0.0
all		**36**	**579**	**3559**	**0.1**	**38**	**392**	**3559**	**0.0**

		S0				S1				S1-compression			
$\|S\|$	r	#S	t(s)	TWT	gap(%)	#S	t(s)	TWT	gap(%)	#S	t(s)	TWT	gap(%)
5	3	8	817	2114	0.4	10	3	2127	0.0	10	3	2127	0.0
10	5	8	807	3052	0.2	9	842	3050	0.1	10	494	3050	0.0
15	8	6	1498	4388	0.6	10	693	4424	0.0	10	382	4424	0.0
20	10	7	1411	5034	0.8	8	849	5045	0.1	8	875	5047	0.1
all		**29**	**1133**	**3647**	**0.5**	**37**	**597**	**3661**	**0.0**	**38**	**438**	**3662**	**0.0**

		S2				S2-compression			
$\|S\|$	r	#S	t(s)	TWT	gap(%)	#S	t(s)	TWT	gap(%)
5	3	10	1	2127	0.0	10	1	2127	0.0
10	5	8	800	3085	0.2	9	748	3082	0.1
15	8	10	382	4425	0.0	10	179	4425	0.0
20	10	7	1232	5076	0.2	10	610	5075	0.0
all		**35**	**604**	**3678**	**0.1**	**39**	**384**	**3677**	**0.0**

The exact solution of the DTPR provides a system optimum for the direct timetabling paradigm. Heuristics capable of handling large-sized lines or networks within a reasonable computational time will render the paradigm more applicable. Moreover, such heuristics may enable incorporating further operational considerations and constraints.

References

Barrena, E., Canca, D., Coelho, L. C., & Laporte, G. (2014a). Exact formulations and algorithm for the train timetabling problem with dynamic demand. *Computers & Operations Research*, *44*, 66–74.

Barrena, E., Canca, D., Coelho, L. C., & Laporte, G. (2014b). Single-line rail rapid transit timetabling under dynamic passenger demand. *Transportation Research Part B: Methodological*, *70*, 134–150.

Bonami, P., Salvagnin, D., & Tramontani, A. (2020). Implementing automatic benders decomposition in a modern mip solver. In *IPCO* (pp. 78–90).

Cacchiani, V., Caprara, A., & Fischetti, M. (2012). A lagrangian heuristic for robustness, with an application to train timetabling. *Transportation Science*, *46*, 124–133.

Cacchiani, V., Caprara, A., & Toth, P. (2008). A column generation approach to train timetabling on a corridor. *4OR*, *6*, 125–142.

Canca, D., Barrena, E., Algaba, E., & Zarzo, A. (2014). Design and analysis of demand-adapted railway timetables. *Journal of Advanced Transportation*, *48*, 119–137.

Canca, D., Barrena, E., Laporte, G., & Ortega, F. A. (2016). A short-turning policy for the management of demand disruptions in rapid transit systems. *Annals of Operations Research*, *246*, 145–166.

Canca, D., Barrena, E., Zarzo, A., Ortega, F., & Algaba, E. (2012). Optimal train reallocation strategies under service disruptions. *Procedia-Social and Behavioral Sciences*, *54*, 402–413.

Caprara, A., Fischetti, M., & Toth, P. (2002). Modeling and solving the train timetabling problem. *Operations research*, *50*, 851–861.

Ceder, A. (2016). *Public transit planning and operation: Modeling, practice and behavior*. Boca Raton, USA: CRC press.

Cortés, C. E., Jara-Díaz, S., & Tirachini, A. (2011). Integrating short turning and deadheading in the optimization of transit services. *Transportation Research Part A: Policy and Practice*, *45*, 419–434.

Crainic, T. G., Hewitt, M., Maggioni, F., & Rei, W. (2021). Partial benders decomposition: general methodology and application to stochastic network design. *Transportation Science*, *55*, 414–435.

Dunning, I., Huchette, J., & Lubin, M. (2017). Jump: A modeling language for mathematical optimization. *SIAM Review*, *59*, 295–320. doi:10.1137/15M1020575.

Farahani, R. Z., Miandoabchi, E., Szeto, W. Y., & Rashidi, H. (2013). A review of urban transportation network design problems. *European Journal of Operational Research*, *229*, 281–302.

Furth, P. G. (1987). Short turning on transit routes. *Transportation Research Record*, *1108*, 42–52.

Gao, Y., Kroon, L., Schmidt, M., & Yang, L. (2016). Rescheduling a metro line in an over-crowded situation after disruptions. *Transportation Research Part B: Methodological*, *93*, 425–449.

Golari, M., Fan, N., & Jin, T. (2017). Multistage stochastic optimization for production-inventory planning with intermittent renewable energy. *Production and Operations Management*, *26*, 409–425.

Kroon, L., Huisman, D., Abbink, E., Fioole, P.-J., Fischetti, M., Maróti, G., Schrijver, A., Steenbeek, A., & Ybema, R. (2009). The new dutch timetable: The OR revolution. *Interfaces*, *39*, 6–17.

Liebchen, C. (2008). The first optimized railway timetable in practice. *Transportation Science*, *42*, 420–435.

Liebchen, C., Schachtebeck, M., Schöbel, A., Stiller, S., & Prigge, A. (2010). Computing delay resistant railway timetables. *Computers & Operations Research*, *37*, 857–868.

Magnanti, T. L., & Wong, R. T. (1981). Accelerating benders decomposition: Algorithmic enhancement and model selection criteria. *Operations research*, *29*, 464–484.

Nachtigall, K., & Voget, S. (1996). A genetic algorithm approach to periodic railway synchronization. *Computers & Operations Research*, *23*, 453–463.

Naoum-Sawaya, J., & Elhedhli, S. (2013). An interior-point benders based branch-and-cut algorithm for mixed integer programs. *Annals of Operations Research*, *210*, 33–55.

Newell, G. F. (1993). A simplified theory of kinematic waves in highway traffic, part ii: Queueing at freeway bottlenecks. *Transportation Research Part B: Methodological*, *27*, 289–303.

Rahmaniani, R., Crainic, T. G., Gendreau, M., & Rei, W. (2017). The benders decomposition algorithm: A literature review. *European Journal of Operational Research*, *259*, 801–817.

Schettini, T. (2020). DTP-instances. `https://github.com/tommaso-schettini-polimi/DTP-Instances`.

Schettini, T., Jabali, O., & Malucelli, F. (2021a). Demand-driven scheduling for automated metro lines. Preprint on webpage at `https://tschettini.files.wordpress.com/2021/04/demand-driven-metro.pdf`.

Schettini, T., Jabali, O., & Malucelli, F. (2021b). Metro scheduling for special events. *Transportation Research Procedia*, *52*, 147–154.

Sudermann-Merx, N., Rebennack, S., & Timpe, C. (2021). Crossing minimal edge-constrained layout planning using benders decomposition. *Production and Operations Management*, available at webpage `https://onlinelibrary.wiley.com/doi/abs/10.1111/poms.13441`.

Sun, L., Jin, J. G., Lee, D.-H., Axhausen, K. W., & Erath, A. (2014). Demand-driven timetable design for metro services. *Transportation Research Part C: Emerging Technologies*, *46*, 284–299.

Tirachini, A., Cortés, C. E., & Jara-Díaz, S. R. (2011). Optimal design and benefits of a short turning strategy for a bus corridor. *Transportation*, *38*, 169–189.

Yang, L., Yao, Y., Shi, H., & Shang, P. (2020). Dynamic passenger demand-oriented train scheduling optimization considering flexible short-turning strategy. *Journal of the Operational Research Society*, *Advance online publication*. doi:10.1080/01605682.2020.1806745.

Yin, J., Yang, L., Tang, T., Gao, Z., & Ran, B. (2017). Dynamic passenger demand oriented metro train scheduling with energy-efficiency and waiting time minimization: Mixed-integer linear programming approaches. *Transportation Research Part B: Methodological*, *97*, 182–213.

Appendix A: The Benders dual sub-problem for case D_b

The sub-problem associated with passengers (i, j, t) where $(i, j) \in D_b$, is as follows.

$$\min \quad \sum_{\delta \in \Delta} (\delta \bar{a}^t_{i,j} y^{t,\delta}_{i,j}) \tag{63}$$

$$\sum_{\delta \in \Delta} y^{t,\delta}_{i,j} = 1 \tag{64}$$

$$y^{t,\delta}_{i,j} \leq \sum_{d \in S_{-\gamma,i}} \sum_{n \in Q_d} \bar{x}^{t+\delta-n}_{\upsilon,d,n} \qquad \forall \delta \in \Delta \tag{65}$$

$$y^{t,\delta}_{i,j} \geq 0 \qquad \forall \delta \in \Delta. \tag{66}$$

Let $\lambda^t_{i,j}$ be the dual variables corresponding to constraint (64), and $\pi^{t,\delta}_{i,j}$ be the dual variables corresponding to constraints (65). The dual sub-problem is as follows.

$$[BD_b] \quad \max \quad \eta^t_{i,j} = \lambda^t_{i,j} - \sum_{\delta \in \Delta} \left(\pi^{t,\delta}_{i,j} \sum_{d \in S_{-\gamma,i}} \sum_{n \in Q_d} \bar{x}^{t+\delta-n}_{\upsilon,d,n} \right) \tag{67}$$

$$\lambda^t_{i,j} - \pi^{t,\delta}_{i,j} \leq \delta \bar{a}^t_{i,j} \qquad \forall \delta \in \Delta \tag{68}$$

$$\lambda^t_{i,j} \geq 0 \tag{69}$$

$$\pi^{t,\delta}_{i,j} \geq 0 \qquad \forall \delta \in \Delta \tag{70}$$

We note that sub-problems BD_f and BD_b have a similar structure. The difference between the two cases lies in the components of the path vector $\bar{x}$ in the objective function. This is due to the fact that the primal sub-problems associated with case D_f consists of constraints (10) and (12). Conversely, in case D_b the sub-problem consists of constraints (10) and (13).

Appendix B: Proofs

Lemma 1. *Given a feasible path vector x, the trip with the least waiting time for passengers (i, j, t) where $(i, j) \in D_f$ is established by only boarding the train stating at time-step:*

$$\min_{t' \in T} \left(t' \; : \; \sum_{d \in S_{\upsilon(i,j),j}} \sum_{n \in Q_d} x^{t'}_{\upsilon(i,j),d,n} = 1 \;\wedge\; t' \geq t \right),$$

i.e., the first train passing by station i headed to station j seen by passengers (i, j, t).

Proof. Passengers that intend to reach station j from station i can do so either by boarding one or multiple trains. If they board exactly one train, they must board a train that is associated with a path that reaches station j. In this case, the passengers will board the first train that passes by station i, headed to station j.

Consider passengers boarding two trains to reach their destination. Thus, the passengers board a train at station i, reach an intermediate station k, alight, and wait for the next train to perform the last leg of their trip to station j. Any path that reaches station j from station k must also pass by station i. Trains cannot overtake each other, therefore no reduction in waiting time can be established by taking trains that short-turn before reaching j as the passengers would end up eventually boarding the same train they would have boarded from station i. The same argument can be used in the case of more than two legs. □

Lemma 2. *Given a feasible path vector x, the trip with the least waiting time for passengers (i,j,t) where $(i,j) \in D_b$ is established by only boarding the train starting at time-step:*

$$\min_{t' \in T} \Big(t' \ : \sum_{d \in S_{-\upsilon(i,j),i}} \sum_{n \in Q_d} x^{t'-n}_{-\upsilon(i,j),d,n} = 1 \ \wedge \ t' \geq t \Big),$$

i.e., the first train headed towards r seen by passengers (i,j,t).

Proof. Let us consider a station $k \in S$ that lies between station i and j, i.e., $i < k < j \leq r$ or $r \leq i < k < j$. Any path that connects station i to station k, also connects i to j. Since trains cannot overtake each other, no reduction in waiting time can be established by having passengers (i,j,t) alight before reaching their destination station to take another train. □

Lemma 3. *Given a feasible path vector x, the trip with the least waiting time for passengers (i,j,t) where $(i,j) \in D_c$ is established by boarding at most two trains, specifically boarding the first train passing by station i headed towards r or j at time-step:*

$$\min_{t' \in T} \Big(t' \ : \sum_{d \in S_{-\upsilon(i,j),i}} \sum_{n \in Q_d} x^{t'-n}_{-\upsilon(i,j),d,n} = 1 \ \wedge \ t' \geq t \Big),$$

then departing from the root station at time-step:

$$\min_{t'' \in T} \Big(t'' \ : \sum_{d \in S_{\upsilon(i,j),j}} \sum_{n \in Q_d} x^{t'}_{\upsilon(i,j),d,n} = 1 \ \wedge \ t'' \geq t' \Big).$$

Proof. Trains are not allowed to skip stops, thus the trip with the least waiting time that connects two stations i,j can be separated into the two sub-trips connecting i to an intermediate station i', and i' to the destination station j. Where i' that lies between i and j. By construction, the ending time of the trip between i and i' corresponds with the starting time of the trip between i' and j. Since trains cannot overtake each other, the trip with the least waiting time can be established by independently minimizing the waiting time in each of the two sub-trips.

Considering r as the intermediate station, it holds $(i,r) \in D_b$ and $(r,j) \in D_f$, each of the two legs of the trip connecting i to j can be performed with the least waiting time without alighting at any intermediate stations as per Lemmas 1 and 2. Thus, the trip with the least waiting time between (i,j) can always be completed by changing trains at most once at the root station. □

Lemma 4. *Given a feasible path vector $\bar{x}$, the constraint matrix of formulation PF0 is totally unimodular.*

Proof. If the path vector is fixed to $\bar{x}$, formulation PF0 reduces to constraints (10)–(15), (17), (18). Furthermore, Constraints (14) and (15) can be rewritten as:

$$y^{t,\delta,\delta'}_{i,j} \leq \Big(\sum_{d \in S_{-\upsilon,i}} \sum_{n \in Q_d} \bar{x}^{t+\delta-n}_{-\upsilon,d,n} \Big) \Big(\sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+\delta'}_{\upsilon,d,n} \Big) \quad \forall (i,j) \in D_c, t \in \bar{T}_{\upsilon(i,j),i}, \delta, \delta' \in \Delta \ : \ \delta \leq \delta'. \tag{71}$$

The resulting formulation can be decomposed into a sub-problem for each passenger tuple (i,j,t). Each sub-problem belongs to one out of three cases. We now detail these cases.

In the case where $(i,j) \in D_f$, the sub-problems are structured as follows.

$$[SD_f] \quad \min \sum_{\delta \in \Delta} (\delta \bar{a}^t_{i,j} y^{t,\delta}_{i,j}) \tag{72}$$

$$\sum_{\delta \in \Delta} y^{t,\delta}_{i,j} = 1 \tag{73}$$

$$y_{i,j}^{t,\delta} \leq \sum_{d \in S_{\gamma,j}} \sum_{n \in Q_d} \bar{x}_{\upsilon,d,n}^{t+\delta} \qquad \forall \delta \in \Delta \tag{74}$$

$$y_{i,j}^{t,\delta} \in \{0,1\} \qquad \forall \delta \in \Delta \tag{75}$$

In the case where $(i,j) \in D_b$, the sub-problems are structured as follows.

$$[SD_b] \quad \min \quad \sum_{\delta \in \Delta} (\delta \bar{a}_{i,j}^t y_{i,j}^{t,\delta}) \tag{76}$$

$$\sum_{\delta \in \Delta} y_{i,j}^{t,\delta} = 1 \tag{77}$$

$$y_{i,j}^{t,\delta} \leq \sum_{d \in S_{\upsilon(i,j),j}} \sum_{n \in Q_d} \bar{x}_{-\upsilon(i,j),d,n}^{t+\delta-n} \qquad \forall \delta \in \Delta \tag{78}$$

$$y_{i,j}^{t,\delta} \in \{0,1\} \qquad \forall \delta \in \Delta \tag{79}$$

In the case where $(i,j) \in D_c$, the sub-problems are structured as follows.

$$[SD_c] \quad \min \quad \sum_{\delta \in \Delta} \sum_{\delta' \in \Delta : \delta' \geq \delta} \left(\delta' \bar{a}_{i,j}^t y_{i,j}^{t,\delta,\delta'} \right) \tag{80}$$

$$\sum_{\delta \in \Delta} \sum_{\delta' \in \Delta : \delta' \geq \delta} y_{i,j}^{t,\delta,\delta'} = 1 \tag{81}$$

$$y_{i,j}^{t,\delta,\delta'} \leq \Big(\sum_{d \in S_{-\upsilon,i}} \sum_{n \in Q_d} \bar{x}_{-\upsilon,d,n}^{t+\delta-n} \Big) \Big(\sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}_{\upsilon,d,n}^{t+\delta'} \Big) \qquad \forall \delta, \delta' \in \Delta \ : \ \delta \leq \delta' \tag{82}$$

$$y_{i,j}^{t,\delta,\delta'} \in \{0,1\} \qquad \forall \delta, \delta' \in \Delta : \delta \leq \delta' \tag{83}$$

In the case of sub-problem SD_f, we note that constraint (73) is a special case of the knapsack constraint, where all items have a weight of one, and the knapsack has a capacity of one. Additionally, for a fixed value of $\bar{x}$ constraints (74) can either be $y_{i,j}^{t,\delta} \leq 1$ or $y_{i,j}^{t,\delta} \leq 0$. In the latter case, the involved variable is always zero and can be removed from the formulation. Thus, formulation SD_f is a special case of the knapsack problem, where $\{\delta \in \Delta \ : \ \sum_{d \in S_{\gamma,j}} \sum_{n \in Q_d} \bar{x}_{\upsilon,d,n}^{t+\delta} \geq 1\}$ is the set of items. In this case, sub-problem SD_f is totally unimodular.

Following the same logic, it is possible to prove that SD_b and SD_c, are also totally unimodular. Therefore, given a feasible path vector $\bar{x}$, formulation PF0 is totally unimodular.

□

Lemma 5. *Cut (45) is valid for the DTPR, and dominates all other cuts (42) with $\lambda_{i,j}^t = \bar{a}_{i,j}^t \sigma$, $\forall \sigma \in \mathbb{N}_0$.*

Proof. For a given $\sigma \in \mathbb{N}_0$, cut (45) corresponds to having the dual variables set as follows.

$$\lambda_{i,j}^t = \bar{a}_{i,j}^t \sigma, \quad \pi_{i,j}^{t,\delta} = \bar{a}_{i,j}^t (\sigma - \delta) \ \forall \ \delta \in \Delta \ : \ \delta \leq \sigma, \quad \pi_{i,j}^{t,\delta} = 0 \ \forall \ \delta \in \Delta \ : \ \delta > \sigma \tag{84}$$

The configuration described in (84) is feasible for BD_f by inspection. Thus, the inequality (45) is valid for DTPR.

Every dual variable $\pi_{i,j}^{t,\delta}$ only appears in constraint set (28), i.e., in constraint $\lambda_{i,j}^t - \pi_{i,j}^{t,\delta} \leq \delta \bar{a}_{i,j}^t$. The variables $\pi_{i,j}^{t,\delta}$ are all associated with non-positive coefficients in the objective function (27). Therefore, each $\pi_{i,j}^{t,\delta}$ can be minimized independently with respect to the objective function. The optimal value of $\pi_{i,j}^{t,\delta}$, $\forall \delta \in \Delta$ is $\bar{a}_{i,j}^t \max(0, \sigma - \delta)$, which is independent of $\bar{x}$. Thus, cut (45) dominates all other cuts with $\lambda_{i,j}^t = \bar{a}_{i,j}^t \sigma$. □

Lemma 6. *Given a path vector $\bar{x}$, for a given $\sigma \in \mathbb{N}_0$, Benders cut (45) is generated by an optimal solution of the BD_f sub-problem if exactly one train is available to serve the customers (i,j) in the equivalent time interval $[t, t+\sigma]$, i.e., it holds:*

$$\sigma \geq 0 \ : \sum_{\delta \in \Delta : \delta \leq \sigma} \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+k}_{\upsilon,d,n} = 1 \tag{46}$$

Proof. When the condition (46) is verified, the following holds.

$$\exists! k \in \Delta \ : \ \pi^{t,k}_{i,j} > 0 \ \wedge \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+k}_{\upsilon,d,n} = 1.$$

Therefore, the passengers will wait for k time-steps before boarding the train that will lead them to their destination, i.e., the primal objective function is $\bar{a}^t_{i,j}k$. The dual objective function, associated with cut (45) can be developed to:

$$\bar{a}^t_{i,j}\sigma - \bar{a}^t_{i,j}(\sigma - k) = \bar{a}^t_{i,j}k.$$

By strong duality, we conclude that the cut (45) is optimal for the dual sub-problem. □

Proposition 1. *Benders cuts of the form (45) are Pareto-optimal $\forall \sigma \in \mathbb{N}_0$.*

Proof. We conduct this proof by contradiction. Let us assume that cut (45) with $\sigma \in \mathbb{N}_0$ is not Pareto-optimal. Therefore, there exists another cut that dominates it with properly selected parameters:

$$\eta^t_{i,j} \geq \bar{\lambda}^t_{i,j} - \sum_{\delta \in \Delta} \bar{\pi}^{t,\delta}_{i,j} \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} x^{t+\delta}_{\upsilon,d,n} \tag{85}$$

For the ease of notation, let us denote by $\mathbb{X}$ the polytope described by constraints (6)–(9) and (16). For cut (85) to dominate (45) the following must hold.

$$\lambda^t_{i,j} - \sum_{\delta \in \Delta} \pi^{t,\delta}_{i,j} \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+\delta}_{\upsilon,d,n} \leq \bar{\lambda}^t_{i,j} - \sum_{\delta \in \Delta} \bar{\pi}^{t,\delta}_{i,j} \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+\delta}_{\upsilon,d,n} \qquad \forall \bar{x} \in \mathbb{X} \tag{86}$$

Where the parameters of cut (45) are specified as follows.

$$\lambda^t_{i,j} = \bar{a}^t_{i,j}\sigma, \quad \pi^{t,\delta}_{i,j} = \bar{a}^t_{i,j}(\sigma - \delta) \quad \forall \delta \in \Delta \ : \ \delta \leq \sigma, \quad \pi^{t,\delta}_{i,j} = 0 \quad \forall \delta \in \Delta \ : \ \delta > \sigma.$$

Additionally, at least one of the inequalities (86) must strictly hold. We examine four cases of $\bar{x}$, and utilize (86) to infer conditions on the parameters of cut (85).

Case 1:. First, let us consider a path vector $\bar{x}$ such that $\bar{x}^{t'}_{\gamma,d,n} = 0 \ \ \forall t' \in T, \gamma \in \Gamma, d \in S_\gamma, n \in Q_d$. In this case, inequality (86) reduces to:

$$\bar{\lambda}^t_{i,j} \geq \lambda^t_{i,j} = \bar{a}^t_{i,j}\sigma \tag{87}$$

Case 2:. We consider path vectors $\bar{x}$ such that:

1) a train is available to connect stations i and j at time-step $t+k$, with $k \leq \sigma$,
2) cut (45) is optimal,
3) no other trains connect stations i and j during the rest of the planning horizon, $\forall k \in \mathbb{N}_0 : k \leq \sigma$.

Formally, the following holds:

$$
\begin{aligned}
&1) \quad \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+k}_{\upsilon,d,n} = 1,\\
&2) \quad \sum_{\delta \in \Delta : \delta \leq \sigma} \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+\delta}_{\upsilon,d,n} = 1,\\
&3) \quad \sum_{\delta \in \Delta : \delta > \sigma} \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+\delta}_{\upsilon,d,n} = 0.
\end{aligned}
$$

Replacing such a solution into inequality (86) it reduces to:

$$
\bar{\lambda}^t_{i,j} - \bar{\pi}^{t,\delta}_{i,j} \geq \lambda^t_{i,j} - \pi^{t,\delta}_{i,j} = \bar{a}^t_{i,j} k \qquad \forall k \in \Delta \ : \ k \leq \sigma.
$$

Combining this result with the dual constraint (28), we conclude the following.

$$
\bar{\lambda}^t_{i,j} - \bar{\pi}^{t,k}_{i,j} = \lambda^t_{i,j} - \pi^{t,k}_{i,j} = \bar{a}^t_{i,j} k \qquad \forall k \in \Delta \ : \ k \leq \sigma \tag{88}
$$

Therefore, it holds:

$$
\bar{\pi}^{t,\delta}_{i,j} = \bar{\lambda}^t_{i,j} - \bar{a}^t_{i,j} k \qquad \forall k \in \Delta \ : \ k \leq \sigma \tag{89}
$$

Case 3:. We consider path vector $\bar{x}$ such that:

1) a train is available to connect stations i and j at time-step $t+k$, with $k \leq \sigma$,
2) cut (45) is optimal,
3) exactly one train connects stations i to j after time-step $t+\sigma$ at time-step $t+k'$, $\forall k, k' \in \Delta \ : \ k \leq \sigma < k'$

Formally, the following holds:

$$
\begin{aligned}
&1) \quad \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+k}_{\upsilon,d,n} = 1,\\
&2) \quad \sum_{\delta \in \Delta : \delta \leq \sigma} \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+\delta}_{\upsilon,d,n} = 1,\\
&3) \quad \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+k'}_{\upsilon,d,n} = 1, \ \wedge \ \sum_{\delta \in \Delta : \delta > \sigma} \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+\delta}_{\upsilon,d,n} = 1.
\end{aligned}
$$

We recall that $\pi^{t,k'}_{i,j} = 0$; therefore, we can write inequality (86) in this case as follows:

$$
\bar{\lambda}^t_{i,j} - \bar{\pi}^{t,k}_{i,j} - \bar{\pi}^{t,k'}_{i,j} \geq \lambda^t_{i,j} - \pi^{t,k}_{i,j} \qquad \forall \ k, k' \in \Delta \ : \ k \leq \sigma < k'.
$$

However, based on (88), it holds: $\bar{\lambda}^t_{i,j} - \bar{\pi}^{t,k}_{i,j} = \lambda^t_{i,j} - \pi^{t,k}_{i,j}$, therefore:

$$
-\bar{\pi}^{t,k'}_{i,j} \geq 0 \qquad \forall \ k' \in \Delta \ : \ k' > \sigma.
$$

Which, as for dual constraint (30), becomes:

$$
\bar{\pi}^{t,k'}_{i,j} = 0 \qquad \forall \ k' \in \Delta \ : \ k' > \sigma \tag{90}
$$

Case 4:. Lastly, considering $\bar{x}$ such that exactly two trains are available to connect stations i and j at time-steps $t+k$ and $t+k'$, $\forall k, k' \in \mathbb{N}_0 : k < k' \leq \sigma$, i.e.,

$$\sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+k}_{\upsilon,d,n} = 1 \ \wedge \ \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+k'}_{\upsilon,d,n} = 1 \ \wedge \ \sum_{\delta \in \Delta : \delta \leq \sigma} \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+\delta}_{\upsilon,d,n} = 2.$$

In this case, inequality (86) reduces to the following:

$$\bar{\lambda}^t_{i,j} - \bar{\pi}^{t,k}_{i,j} - \bar{\pi}^{t,k'}_{i,j} \geq \lambda^t_{i,j} - \pi^{t,k}_{i,j} - \pi^{t,k'}_{i,j}.$$

We substitute (89) in the condition, as follows:

$$-\bar{\lambda}^t_{i,j} + \bar{a}^t_{i,j}(k+k') \geq -\lambda^t_{i,j} + \bar{a}^t_{i,j}(k+k').$$

Which, along with condition (87) implies:

$$\bar{\lambda}^t_{i,j} = \lambda^t_{i,j} \tag{91}$$

As for conditions (89), (90), and (91) we conclude that (85) must correspond to (45), violating the initial hypotheses that there exists a dominating cut. □

Lemma 7. *Let $\sigma, \mu \in \mathbb{N}_0 \ : \ \mu \leq \sigma$, cut (47) is valid for DTPR.*

Proof. For given $\sigma, \mu \in \mathbb{N}_0$, cut (44) corresponds to having the dual variables set as follows:

$$\xi^t_{i,j} = \bar{a}^t_{i,j}\sigma, \qquad \phi^{t,\delta}_{i,j} = 0 \ \forall \delta \in \Delta \ : \ \delta < \mu,$$
$$\theta^{t,\delta}_{i,j} = \bar{a}^t_{i,j}(\sigma - \delta) \ \forall \delta \in \Delta \ : \ \delta < \mu, \qquad \phi^{t,\delta}_{i,j} = \bar{a}^t_{i,j}(\sigma - \delta) \ \forall \delta \in \Delta \ : \ \mu \leq \delta < \sigma,$$
$$\theta^{t,\delta}_{i,j} = 0 \ \forall \delta \in \Delta \ : \ \delta \geq \mu, \qquad \phi^{t,\delta}_{i,j} = 0 \ \forall \delta \in \Delta \ : \ \delta \geq \sigma.$$

Which can be verified to be feasible for BD_c by inspection. Thus, inequality (47) is valid for DTPR. □

Lemma 8. *Given a path vector $\bar{x}$, if $\bar{x}$ satisfies (48), there exist $\sigma, \mu \in \mathbb{N}_0$ such that: 1) in the equivalent time interval $[t, t+\mu]$ exactly one train is available to bring the passengers from station i to station j at equivalent time-step $t+k$, with $k \in [0, \mu]$, 2) in the equivalent time interval $[t, t+\mu]$ no other trains are available to bring the passengers at station i to r, and 3) in the equivalent time interval $[t+\mu, t+\sigma]$ no trains depart from r headed to station j. Formally:*

$$1) \quad \exists k \in [0, \mu] \ : \ \sum_{d \in S_{-\upsilon,i}} \sum_{n \in Q_d} \bar{x}^{t+k-n}_{-\upsilon,d,n} = 1 \ \wedge \ \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+k}_{\upsilon,d,n} = 1 \tag{50}$$

$$2) \quad \sum_{\delta \in \Delta : \delta \leq \mu} \sum_{d \in S_{-\upsilon,i}} \sum_{n \in Q_d} \bar{x}^{t+\delta-n}_{-\upsilon,d,n} = 1 \tag{51}$$

$$3) \quad \sum_{\delta \in \Delta : \delta > \mu} \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+\delta}_{\upsilon,d,n} = 0 \tag{52}$$

Then an optimal solution of BD_c generates cut (47).

Proof. If (48) holds, setting $\sigma = p$ and $\mu = 0$ satisfies conditions (50)–(52). Cut (47) along with the conditions (50)–(52) yields: $\eta^t_{i,j} \geq \bar{a}^t_{i,j}\sigma - \bar{a}^t_{i,j}(\sigma - k) = \bar{a}^t_{i,j}k$. Condition (50) implies that there exists one trip that is able to bring the passengers from their origin station i to their destination j where the passengers wait for a total of k time-steps. Condition (51) implies that in the time interval $[t, t+\mu]$ exactly one train is available to bring the passengers from i to r. Condition (52) implies that no train is available to bring the passengers from r to j in the equivalent time interval $[t+\mu, t+\sigma]$. Overall, these conditions imply that in the best trip available to the passengers (i, j, t) is the one associated with condition (50). Thus, the optimal objective value of the primal problem corresponds to $\bar{a}^t_{i,j}k$. Therefore, by strong duality we can conclude that the dual variables associated with cut (47) are optimal for the sub-problem. □

Lemma 9. *Given a path vector $\bar{x}$, if $\bar{x}$ satisfies (49), there exist $\sigma, \mu \in \mathbb{N}_0 : \sigma \geq \mu$ such that: 1) in the equivalent time interval $[t+\mu, t+\sigma]$ exactly one train departs from station r headed to station j, at equivalent time-step $t+k$, with $k \in [\mu, \sigma]$, 2) in the equivalent time interval $[t, t+\mu)$ no train is available to bring the passengers from station i to r, and 3) in the equivalent time interval $[t+\mu, t+k]$ at least one train is available to bring the passengers from station i to station r. Formally:*

$$1) \quad \exists k \in [\mu, \sigma] \ : \sum_{\delta \in \Delta: \mu < h \leq k} \sum_{d \in S_{-\upsilon,i}} \sum_{n \in Q_d} \bar{x}^{t+\delta-n}_{-\upsilon,d,n} = 1 \ \wedge \ \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+k}_{\upsilon,d,n} = 1 \tag{53}$$

$$2) \quad \sum_{\delta \in \Delta: \delta < \mu} \sum_{d \in S_{-\upsilon,i}} \sum_{n \in Q_d} \bar{x}^{t+\delta-n}_{-\upsilon,d,n} = 0 \tag{54}$$

$$3) \quad \sum_{\delta \in \Delta: \mu < h \leq \sigma} \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+\delta}_{\upsilon,d,n} = 1 \tag{55}$$

Then an optimal solution of BD_c generates cut (47).

Proof. If (49) holds, setting $\sigma = p'$ and $\mu = p$ satisfies conditions (54)–(55). Cut (47) along with the conditions (54)–(55) yields: $\eta^t_{i,j} \geq \bar{a}^t_{i,j}\sigma - \bar{a}^t_{i,j}(\sigma - k) = \bar{a}^t_{i,j}k$. Condition (54) implies that the passengers will not leave their origin station in the time interval $[t, t+\mu-1]$. Condition (53) requires that there exists one trip that is able to bring the passengers from i to j where the passengers wait for a total of k time-steps. Condition (55) implies that exactly one train in the equivalent time interval $[t+\mu, t+\sigma]$ is able to bring the passengers from r to j. Overall, these conditions imply that in the best trip available to passengers (i, j, t) is the one associated with condition (53). Thus, the optimal objective value of the primal problem corresponds to $\bar{a}^t_{i,j}k$. Therefore, by strong duality we can conclude that the dual variables associated with cut (47) are optimal for the sub-problem. □

Proposition 2. *Benders cuts of the form (47), are Pareto-optimal $\forall \sigma, \mu \in \mathbb{N}_0 : \mu \leq \sigma$.*

Proof. We conduct this proof by contradiction. Let us assume that cut (47) with $\sigma, \mu \in \mathbb{N}_0$ is not Pareto-optimal. Therefore, there exists another cut that dominates it with properly chosen parameters:

$$\eta^t_{i,j} \geq \bar{\xi}^t_{i,j} - \sum_{\delta \in \Delta} \left(\bar{\theta}^{t,\delta}_{i,j} \sum_{d \in S_{-\upsilon,i}} \sum_{n \in Q_d} x^{t'+\delta-n}_{-\upsilon,d,n} + \bar{\phi}^{t,\delta}_{i,j} \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} x^{t'+\delta}_{\upsilon,d,n} \right) \tag{92}$$

For the ease of notation, let us denote by $\mathbb{X}$ the polytope described by constraints (6)–(9) and (16). For cut (92) to dominate (47) the following must hold.

$$\begin{aligned} &\bar{\xi}^t_{i,j} - \sum_{\delta \in \Delta} \left(\bar{\theta}^{t,\delta}_{i,j} \sum_{d \in S_{-\upsilon,i}} \sum_{n \in Q_d} \bar{x}^{t'+\delta-n}_{-\upsilon,d,n} + \bar{\phi}^{t,\delta}_{i,j} \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t'+\delta}_{\upsilon,d,n} \right) \\ &\leq \xi^t_{i,j} - \sum_{\delta \in \Delta} \left(\theta^{t,\delta}_{i,j} \sum_{d \in S_{\upsilon,i}} \sum_{n \in Q_d} \bar{x}^{t'+\delta-n}_{-\upsilon,d,n} + \phi^{t,\delta}_{i,j} \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t'+\delta}_{\upsilon,d,n} \right) \qquad \forall \bar{x} \in \mathbb{X} \end{aligned} \tag{93}$$

Where the parameters of cut (47) are specified as follows.

$$\begin{aligned} &\xi^t_{i,j} = \bar{a}^t_{i,j}\sigma, && \phi^{t,\delta}_{i,j} = 0 \ \forall \delta \in \Delta \ : \ \delta < \mu, \\ &\theta^{t,\delta}_{i,j} = \bar{a}^t_{i,j}(\sigma - \delta) \ \forall \delta \in \Delta \ : \ \delta < \mu, && \phi^{t,\delta}_{i,j} = \bar{a}^t_{i,j}(\sigma - \delta) \ \forall \delta \in \Delta \ : \ \mu \leq \delta < \sigma, \\ &\theta^{t,\delta}_{i,j} = 0 \ \forall \delta \in \Delta \ : \ \delta \geq \mu, && \phi^{t,\delta}_{i,j} = 0 \ \forall \delta \in \Delta \ : \ \delta \geq \sigma. \end{aligned}$$

Additionally, at least one of the inequalities (93) must strictly hold. We examine five cases of $\bar{x}$, and utilize (93) to infer conditions on the parameters of cut (92).

Case 1:. Consider a path vector $\bar{x}$ such that $\bar{x}^{t'}_{\gamma,d,n} = 0 \ \ \forall t' \in T, \gamma \in \Gamma, d \in S_\gamma, n \in Q_d$; inequality (93) is reduced to:

$$\bar{\xi}^t_{i,j} \leq \xi^t_{i,j} = \bar{a}^t_{i,j}\sigma \tag{94}$$

Case 2:. We consider path vectors $\bar{x}$ such that:

1) there is exactly one train that connects station i to station j at equivalent time-step $t+\delta$, with $\delta \leq \sigma$,
2) there are no other trains that connect station i to station r,
3) there are no other trains that connect station r to station j, i.e., $\forall \delta \in \mathbb{N}_0 : \delta \leq \sigma$,

Formally, the following holds.

$$\begin{aligned}
&1) \quad \sum_{d\in S_{\upsilon,i}} \sum_{n\in Q_d} \bar{x}^{t+\delta-n}_{-\upsilon,d,n} = 1 \ \wedge \ \sum_{d\in S_{\upsilon,j}} \sum_{n\in Q_d} \bar{x}^{t+\delta}_{\upsilon,d,n} = 1,\\
&2) \quad \sum_{t'\in T} \sum_{d\in S_{\upsilon,i}} \sum_{n\in Q_d} \bar{x}^{t'-n}_{-\upsilon,d,n} = 1,\\
&3) \quad \sum_{t'\in T} \sum_{d\in S_{\upsilon,j}} \sum_{n\in Q_d} \bar{x}^{t'}_{\upsilon,d,n} = 1.
\end{aligned}$$

Replacing $\bar{x}$ into inequality (93) it reduces to:

$$\bar{\xi}^t_{i,j} - \bar{\phi}^{t,k}_{i,j} - \bar{\psi}^{t,k}_{i,j} \leq \xi^t_{i,j} - \phi^{t,k}_{i,j} - \psi^{t,k}_{i,j} = \bar{a}^t_{i,j}\sigma - \bar{a}^t_{i,j}(\sigma - k) = \bar{a}^t_{i,j}k.$$

Combing this result with dual constraints (37) we obtain the following.

$$\bar{\xi}^t_{i,j} - \bar{\phi}^{t,k}_{i,j} - \bar{\psi}^{t,k}_{i,j} = \xi^t_{i,j} - \phi^{t,k}_{i,j} - \psi^{t,k}_{i,j} = \bar{a}^t_{i,j}k \qquad \forall k \in [0,\sigma] \tag{95}$$

Case 3:. We consider path vectors $\bar{x}$ such that:

1) there is exactly one train that connects station i to station j at time-step $t+k$, with $t+k \in [t, t+\mu]$,
2) there exists an additional train that connects station i to r at time-step $t+k'$, with $k \neq k : t+k' \in [t, t+\mu]$,
3) there are no other trains that connect station r to station j for the entire planning horizon,
4) there are no other trains that connect station i to station r in the rest of the planning horizon, $\forall k, k' \in [0,\mu]$.

Formally, the following holds.

$$\begin{aligned}
&1) \quad \sum_{d\in S_{\upsilon,i}} \sum_{n\in Q_d} \bar{x}^{t+k-n}_{-\upsilon,d,n} = 1 \ \wedge \ \sum_{d\in S_{\upsilon,j}} \sum_{n\in Q_d} \bar{x}^{t+\delta}_{\upsilon,d,n} = 1\\
&2) \quad \sum_{d\in S_{\upsilon,i}} \sum_{n\in Q_d} \bar{x}^{t+k'-n}_{-\upsilon,d,n} = 1\\
&3) \quad \sum_{t'\in T} \sum_{d\in S_{\upsilon,i}} \sum_{n\in Q_d} \bar{x}^{t'-n}_{-\upsilon,d,n} = 2\\
&4. \quad \sum_{t'\in T} \sum_{d\in S_{\upsilon,j}} \sum_{n\in Q_d} \bar{x}^{t'}_{\upsilon,d,n} = 1
\end{aligned}$$

Replacing $\bar{x}$ into inequality (93) it reduces to:

$$\bar{\xi}^t_{i,j} - \bar{\phi}^{t,k}_{i,j} - \bar{\phi}^{t,k'}_{i,j} - \bar{\psi}^{t,k}_{i,j} \geq \xi^t_{i,j} - \phi^{t,k}_{i,j} - \phi^{t,k'}_{i,j} - \psi^{t,k}_{i,j} = \xi^t_{i,j} - \phi^{t,k}_{i,j} - \psi^{t,k}_{i,j} = \bar{a}^t_{i,j}k.$$

This, combined with (95), implies the following:

$$-\bar{\phi}^{t,k'}_{i,j} \geq 0.$$

As for the non-negativity constraint of the dual variables (40), it results:

$$\bar{\phi}^{t,k'}_{i,j} = 0 \qquad \forall k' \in [0,\mu] \tag{96}$$

Case 4:. We consider path vectors $\bar{x}$ such that:

1) there is exactly one train that connects station i to station j at time-step $t+k$, with $k \in (\mu, \sigma]$,
2) there exists an additional train that connects station r to j at time-step $k' \neq k \in (\mu, \sigma]$,
3) there are no other trains that connect station i to station r for the entire planning horizon,
4) there are no other trains that connect station r to station j in the rest of the planning horizon, $\forall k, k' \in (\mu, \sigma]$.

Formally, the following holds:

$$\begin{aligned}
&1) \sum_{d \in S_{\upsilon,i}} \sum_{n \in Q_d} \bar{x}^{t+k-n}_{-\upsilon,d,n} = 1 \ \wedge \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+k}_{\upsilon,d,n} = 1, \\
&2) \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+k'}_{\upsilon,d,n} = 1, \\
&3) \sum_{t' \in T} \sum_{d \in S_{\upsilon,i}} \sum_{n \in Q_d} \bar{x}^{t'-n}_{-\upsilon,d,n} = 1, \\
&4) \sum_{t' \in T} \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t'}_{\upsilon,d,n} = 2.
\end{aligned}$$

Replacing $\bar{x}$ into inequality (93) and applying the same reasoning as Case 3, we obtain the following:

$$\bar{\theta}^{t,k'}_{i,j} = 0 \qquad \forall k' \in (t+\mu, t+\sigma]. \tag{97}$$

Case 5:. We consider path vectors $\bar{x}$ such that:

1) there is exactly one train connecting station i to station j in equivalent time interval $[t, t+\sigma]$, at time-step $t+k$ with $k \in [0, \sigma]$,
2) there is exactly one train connecting station i to station j after equivalent time interval $[t, t+\sigma]$, at time-step $t+k'$ with $k' > \sigma$,
3) there are no other trains that connect station i to station r for the entire planning horizon, and
4) there are no other trains that connect station r to station j for the entire planning horizon, $\forall k, k' \in \Delta : k \leq \sigma, k' > \sigma$, i.e.,

Formally, the following holds:

$$\begin{aligned}
&1) \sum_{d \in S_{\upsilon,i}} \sum_{n \in Q_d} \bar{x}^{t+\delta-n}_{-\upsilon,d,n} = 1 \ \wedge \sum_{d \in S_{\upsilon,i}} \sum_{n \in Q_d} \bar{x}^{t+\delta'-n}_{-\upsilon,d,n} = 1, \\
&2) \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+\delta}_{\upsilon,d,n} = 1 \ \wedge \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+\delta'}_{\upsilon,d,n} = 1, \\
&3) \sum_{t' \in T} \sum_{d \in S_{\upsilon,i}} \sum_{n \in Q_d} \bar{x}^{t'-n}_{-\upsilon,d,n} = 2, \\
&4) \sum_{t' \in T} \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t'}_{\upsilon,d,n} = 2.
\end{aligned}$$

Replacing $\bar{x}$ into inequality (93) and applying the same reasoning as Case 3, we obtain the following:

$$\bar{\phi}^{t,k'}_{i,j} = 0 \qquad \forall k' \geq \sigma, \tag{98}$$

$$\bar{\theta}^{t,k'}_{i,j} = 0 \qquad \forall k' \geq \sigma. \tag{99}$$

Lastly, let us consider inequality (95) for $k = \sigma$. We incorporate inequalities (98) and (99) to obtain the following:

$$\bar{\xi}^{t}_{i,j} = \bar{a}^{t}_{i,j}(\sigma + \mu). \tag{100}$$

As for inequalities (95), (96), (97), and (100) we conclude that cut (92) corresponds to cut (47), violating the initial hypothesis that there exists a dominating cut. □

Proposition 3. *Given path vector $\bar{x}$ associated with the current BMP solution, a cut associated with sub-problem (i, j, t), where $(i, j) \in D_f$, needs to be generated if and only if either of the following conditions hold:*

$$1) \quad \zeta = \frac{\eta^t_{i,j}}{\bar{a}^t_{i,j}} \text{ is not integer,} \tag{56}$$

$$2) \quad \sum_{d \in S_{v,j}} \sum_{n \in Q_d} \bar{x}^{t+\zeta}_{v,d,n} = 0. \tag{57}$$

Proof. $(A) \Rightarrow (B)$: Additional cuts have to be added to sub-problem (i, j, t) if the current value of $\eta^t_{i,j}$ cannot be projected onto a feasible solution in the original space of variables. Such projection must result in a sub-problem with objective value which equates $\eta^t_{i,j}$. The binary nature of the passenger variables, in conjunction with constraint (10), imply that this can only be achieved through the following projection:

$$y^{t,\zeta}_{i,j} = 1 \tag{101}$$

$$y^{t,\delta}_{i,j} = 0 \qquad \forall \delta \in \Delta \ : \ \delta \neq \zeta \tag{102}$$

If this projection is feasible for the primal sub-problem, then no cut needs to be added, otherwise, additional optimality cuts need to be generated. Requiring the feasibility of the projection implies that ζ has to be integer, otherwise the variable $y^{t,\zeta}_{i,j}$ would not exist in the primal model. Furthermore, the chosen variables must also satisfy the constraints of the sub-problem involving the passenger variables, i.e., constraints (10) and (12). Constraint (10) is directly satisfied by the chosen projection. Constraints (12) for passengers (i, j, t) and $\delta = \zeta$, correspond to:

$$\sum_{d \in S_{v,j}} \sum_{n \in Q_d} x^{t+\zeta}_{v,d,n} \geq y^{t,\zeta}_{i,j} = 1,$$

i.e., the negation of condition (57).

$(A) \Leftarrow (B)$: If (56) and (57) do not hold, it implies that:

$$\sum_{d \in S_{v,j}} \sum_{n \in Q_d} x^{t+\zeta}_{v,d,n} \geq 1.$$

Thus, the projection defined by (101) and (102) is feasible for the primal problem, and therefore no cut needs to be added. □

Lemma 10. *Given path vector $\bar{x}$ associated with the current BMP solution, if all the Benders cuts associated with sub-problem (i, j, t), where $(i, j) \in D_f$, are of the form (45) with $\sigma \in \mathbb{N}_0$, the quantity $\zeta = \eta^t_{i,j}/\bar{a}^t_{i,j}$ is integer.*

Proof. The dual variables η^t_{ij} are minimized in the master problem BMP and are only bounded from below by cuts (42).

$$\eta^t_{i,j} \geq \bar{a}^t_{i,j} \Big(\sigma - \sum_{\delta \in \Delta : \delta \leq \sigma} (\sigma - \delta) \sum_{d \in S_{v,j}} \sum_{n \in Q_d} x^{t+\delta}_{v,d,n} \Big)$$

If $\sigma \in \mathbb{N}_0$, all the parameters of the cut are integer multiples of $\bar{a}^t_{ij}$, therefore, the bound imposed by the cut will always be such that $\eta^t_{i,j}/\bar{a}^t_{i,j} \in \mathbb{N}_0$. □

Proposition 4. *Given path vector $\bar{x}$ associated with the current BMP solution, if all the Benders cuts associated with sub-problem (i, j, t) where $(i, j) \in D_f$ are of the form (45) with $\sigma \in \mathbb{N}_0$, and a cut needs to be generated for the sub-problem; a cut of the form (45) with $\sigma = \zeta + 1$ where $\zeta = \eta^t_{i,j}/a^t_{i,j}$, is violated by the incumbent solution.*

PROOF. Lemma 10 guarantees that $\zeta = \eta^t_{i,j}/\bar{a}^t_{i,j}$ is integer. Additionally, cut (45) is valid as for Lemma 7. Thus, the total waiting time of passengers (i, j, t) cannot be smaller than $\eta^t_{i,j}$, i.e.,

$$\sum_{\delta \in \Delta : \delta \leq \zeta} \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+\delta}_{\gamma,d,n} = 0.$$

Substituting this information in cut (45) with $\sigma = \zeta + 1$, we obtain:

$$\eta^t_{i,j} \geq \bar{a}^t_{i,j}(\zeta + 1)$$

Which is trivially violated. □

Proposition 5. *Given path vector $\bar{x}$ associated with the current BMP solution, a cut associated with sub-problem (i, j, t) where $(i, j) \in D_f$ needs to be generated if and only if at least one the following conditions hold:*

$$1) \quad \zeta = \frac{\eta^t_{i,j}}{\bar{a}^t_{i,j}} \text{ is not integer,} \tag{58}$$

$$2) \quad \zeta \text{ is integer} \wedge \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+\zeta}_{\upsilon,d,n} = 0 \tag{59}$$

$$3) \quad \nexists k' \ : \ 0 \leq k' \leq \zeta \ \wedge \sum_{d \in S_{-\upsilon,i}} \sum_{n \in Q_d} \bar{x}^{t+k'-n}_{-\upsilon,d,n} = 1 \tag{60}$$

PROOF. $(A) \Rightarrow (B)$: Additional cuts have to be added to sub-problem (i, j, t) if the current value of $\eta^t_{i,j}$ cannot be projected onto a feasible solution in the original space of variables. Such projection must result in a sub-problem with objective value which equates $\eta^t_{i,j}$. The binary nature of the passenger variables, in conjunction with constraint (11), imply that this can only be achieved through the following projection:

$$y^{t,k,\zeta}_{i,j} = 1 \tag{103}$$

$$y^{t,\delta,\delta'}_{i,j} = 0 \qquad \forall \delta, \delta' \in \Delta \ : \ \delta \neq k \wedge \delta' \neq \zeta \tag{104}$$

With $k \in \mathbb{N}_0 : k \leq \zeta$. If the projection described by (103) and (104) is feasible for the primal sub-problem, then no cut needs to be added. Otherwise, additional optimality cuts need to be generated. Requiring the feasibility of the projection implies that ζ has to be integer, otherwise no variable $y^{t,k,\zeta}_{i,j}$ would exist in the primal model $\forall k \in \mathbb{N}_0$.
Secondly, the chosen projection must satisfy the constraints of the sub-problem involving the passenger variables, i.e., constraints (11)–(15). Constraint (11) is directly satisfied by the chosen projection. For passengers (i, j, t) constraints (14) are:

$$\sum_{d \in S_{-\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+\delta-n}_{-\upsilon,d,n} \geq \sum_{\delta' \in \Delta : \delta' \geq \delta} y^{t,\delta,\delta'}_{i,j} \qquad \forall \delta \in \Delta.$$

Clearly, if $y^{t,\delta,\delta'}_{i,j} = 0$ the constraint is always satisfied. Conversely, for the case $y^{t,k,\zeta}_{i,j} = 1$ it must hold:

$$\sum_{d \in S_{-\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+k-n}_{-\upsilon,d,n} \geq 1.$$

Thus, a necessary condition for the projection to be feasible is:

$$\exists k \in \mathbb{N}_0 \ : \ k \leq \zeta \ \wedge \sum_{d \in S_{-\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+k-n}_{-\upsilon,d,n} \geq 1,$$

i.e., the negation of condition (60).

For passengers (i,j,t) constraints (15) are:

$$\sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+\delta}_{\upsilon,d,n} \geq \sum_{\delta' \in \Delta : \delta' \leq \delta} y^{t,\delta',\delta}_{i,j} \qquad \forall \delta \in \Delta,$$

If $y^{t,\delta,\delta'}_{i,j} = 0$ the constraint is always satisfied. Conversely, for the case $y^{t,k,\zeta}_{i,j} = 1$ it must hold:

$$\sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+\zeta}_{\upsilon,d,n} \geq 1,$$

i.e., the negation of condition (59).

$(A) \Leftarrow \ (B)$: If (58), (59) and (60) it implies that:

$$\sum_{d \in S_{-\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+k-n}_{-\upsilon,d,n} \geq 1 \ \wedge \ \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+\zeta}_{\upsilon,d,n} \geq 1,$$

with $k \in \mathbb{N}_0 : k \leq \zeta$. Thus making the projection defined by (103) and (104) feasible for the primal problem, and therefore no cut needs to be added. □

Lemma 11. *Given path vector $\bar{x}$ associated with the current BMP solution, if all the Benders cuts associated with sub-problem (i,j,t), where $(i,j) \in D_c$, are of the form (47) with $\sigma \leq \mu \in \mathbb{N}_0$; the quantity $\zeta = \eta^t_{i,j}/\bar{a}^t_{i,j}$ is integer.*

Proof. The proof is identical to the proof of Lemma 10. The dual variables η^t_{ij} are minimized in the master problem **BMP** and are bounded from below by the Benders cuts (47). Recalling the structure of the cut is the following:

$$\eta^t_{i,j} \geq \bar{a}^t_{i,j}(\sigma - \sum_{\delta \in \Delta : \delta < \mu} (\sigma - \delta) \sum_{d \in S_{-\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+\delta-n}_{-\upsilon,d,n} - \sum_{\delta \in \Delta : \mu \leq \delta < \sigma} (\sigma - \delta) \sum_{d \in S_{\upsilon,j}} \sum_{n \in Q_d} \bar{x}^{t+\delta}_{\upsilon,d,n}).$$

If $\sigma, \mu \in \mathbb{N}_0$, all the parameters of the cut are integer multiples of a^t_{ij}, therefore, regardless of the selection of the train variables, which are binary, the bound imposed by the cut will always be such that $\eta^t_{i,j}/a^t_{i,j} \in \mathbb{N}_0$. □

Proposition 6. *Given path vector $\bar{x}$ associated with the current BMP solution, if all the Benders cuts associated with sub-problem (i,j,t), where $(i,j) \in D_c$, are of the form (47) with $\sigma \leq \mu \in \mathbb{N}_0$, and a cut needs to be generated for the sub-problem; there exists a cut of the form (47) with $\sigma = \zeta + 1$ where $\zeta = \eta^t_{i,j}/a^t_{i,j}$ that is violated by the incumbent solution.*

Proof. As for Lemma 11, $\zeta = \eta^t_{i,j}/\bar{a}^t_{i,j}$ is integer. Additionally, cut (47) is valid as for Lemma 7. Thus, the total waiting time of passengers (i,j,t) cannot be smaller than $\eta^t_{i,j}$, i.e., $\bar{x}$ is such that passengers (i,j,t) wait more than $\zeta = \eta^t_{i,j}/\bar{a}^t_{i,j}$ to reach their destination. Therefore, one of the following statements holds:

1) there are no trains in the time interval $[t, t+\zeta]$ that connect station r to j,
2) there are no trains in the time interval $[t, t+k]$ that connect station i to r, with $k \leq \zeta$ such that there is at least a train in the time interval $[t+k, t+\zeta]$ that connect station r to j.

Formally:

$$1)\qquad \sum_{\delta\in\Delta:\delta\le\zeta}\ \sum_{d\in S_{\upsilon,j}}\ \sum_{n\in Q_d} \bar{x}^{t+\delta}_{\upsilon,d,n} = 0 \tag{105}$$

$$2)\qquad \nexists k\in[0,\zeta]\ :\ \sum_{d\in S_{-\upsilon,j}}\ \sum_{n\in Q_d} \bar{x}^{t+k-n}_{-\upsilon,d,n} = 1\ \wedge \sum_{\delta\in\Delta:k\le\delta\le\zeta}\ \sum_{d\in S_{\upsilon,j}}\ \sum_{n\in Q_d} \bar{x}^{t+\delta}_{\upsilon,d,n} = 1 \tag{106}$$

If (105) holds, setting $\mu = 0$ leads to the inequality:

$$\eta^t_{i,j} \ge \bar{a}^t_{i,j}(\sigma - \sum_{\delta\in\Delta:\delta<\sigma}(\sigma-\delta)\sum_{d\in S_{\upsilon,j}}\ \sum_{n\in Q_d}\bar{x}^{t+\delta}_{\upsilon,d,n}),$$

which becomes:

$$\eta^t_{i,j} \ge \bar{a}^t_{i,j}(\zeta+1),$$

which is trivially violated.

If (106) holds, then:

$$\exists k\in[0,\zeta]\ :\ \sum_{\delta\in\Delta:\delta<k}\ \sum_{d\in S_{-\upsilon,j}}\ \sum_{n\in Q_d} \bar{x}^{t-n+\delta}_{-\upsilon,d,n} = 0\ \wedge \sum_{\delta\in\Delta:k\le\delta\le\zeta}\ \sum_{d\in S_{\upsilon,j}}\ \sum_{n\in Q_d} \bar{x}^{t+\delta}_{\upsilon,d,n} = 0.$$

Setting $\mu = k$ leads to inequality:

$$\eta^t_{i,j} \ge \bar{a}^t_{i,j}\big(\zeta - \sum_{\delta\in\Delta:\delta<\mu}(\zeta-\delta)\sum_{d\in S_{-\upsilon,j}}\ \sum_{n\in Q_d}\bar{x}^{t+\delta-n}_{-\upsilon,d,n} - \sum_{\delta\in\Delta:k\le\delta<\zeta}(\zeta-\delta)\sum_{d\in S_{\upsilon,j}}\ \sum_{n\in Q_d}\bar{x}^{t+\delta}_{\upsilon,d,n}\big),$$

which becomes:

$$\eta^t_{i,j} \ge \bar{a}^t_{i,j}(\zeta+1),$$

which is trivially violated. □

Appendix C: Detailed Tables

In this section, we report the comprehensive tables of results for the computational experiments.

Tables 4 and 5 report the detailed results of PF, BDA, and BD on instances with $r = 1$ and $r = \lfloor |S|/2 \rfloor$ respectively. For each instance and each method, we report the computational time, the objective function value of the incumbent solution at the end of the computation (TWT), and the final optimality gap (gap(%)).

Tables 6 and 7 report the detailed comparison of the DTPR solutions to the regular timetables and the DTP solutions. The comparison is conducted w.r.t. total waiting time of the passengers (TWT), maximum number of passengers on board a train (max(B)), mean value of the number of passengers on board a train (avg(B)) and its variance ($\sigma^2(B)$) We denote by D the measures associated with the DTPR solution (i.e., TWT$_D$, max(B_D), avg(B_D), and $\sigma^2(B_D)$), by R the measures associated with the regular timetable (i.e., TWT$_R$, max(B_R), avg(B_R), and $\sigma^2(B_R)$), and by N the measures associated with the regular timetable (i.e., TWT$_N$, max(B_N), avg(B_N), and $\sigma^2(B_N)$). Instances marked with * are not solved to optimality by the DTPR, instances marked with $\diamond$ are not solved to optimality by the DTP.

Tables 8 and 9 report the detailed results of the BD in configurations S0, S1, and S2, with and without compression on instances with $r = 1$ and $r = \lfloor |S|/2 \rfloor$ respectively. For each instance and each configuration, we report the computational time achieved, the objective value of the incumbent solution at the end of the computation (TWT), and the final optimality gap (gap(%)).

Lastly, Tables 10–13 report the detailed comparison of the DTPR solutions in configurations S1 and S2, to the regular timetables and to the DTPR solutions of configuration S0. We denote by SN the measures associated with the DTPR solution in configuration N (i.e., TWT_{SN}, $\max(B_{SN})$, $\text{avg}(B_{SN})$, $\sigma^2(B_{SN})$, and DIST_{SN}). Instances marked with * are not solved to optimality by the DTPR in configurations S1 or S2, instances marked with $\diamond$ are not solved to optimality by the DTPR in configuration S0.

Table 4: Comprehensive results of the flow formulation, cut generation algorithm, path formulation, and Benders-based branch-and-cut algorithm with $r = 1$.

			PF			BDA			BD		
$\|S\|$	$\|T\|$	r	t(s)	TWT	gap(%)	t(s)	TWT	gap(%)	t(s)	TWT	gap(%)
5	10	1	1	366	0.0	1	366	0.0	1	366	0.0
	20		1	696	0.0	2	696	0.0	1	696	0.0
	30		1	1113	0.0	3	1113	0.0	1	1113	0.0
	40		5	1544	0.0	9	1544	0.0	1	1544	0.0
	50		3	1966	0.0	7	1966	0.0	1	1966	0.0
	60		47	2308	0.0	31	2308	0.0	7	2308	0.0
	70		68	2646	0.0	37	2646	0.0	5	2646	0.0
	80		245	3041	0.0	112	3041	0.0	69	3041	0.0
	90		206	3533	0.0	240	3533	0.0	68	3533	0.0
	100		1026	3898	0.0	931	3898	0.0	299	3898	0.0
10	10	1	1	442	0.0	1	442	0.0	1	442	0.0
	20		1	1049	0.0	5	1049	0.0	1	1049	0.0
	30		8	1645	0.0	19	1645	0.0	3	1645	0.0
	40		45	2214	0.0	96	2214	0.0	8	2214	0.0
	50		275	2723	0.0	316	2723	0.0	37	2723	0.0
	60		687	3335	0.0	1565	3335	0.0	76	3335	0.0
	70		504	3893	0.0	491	3893	0.0	41	3893	0.0
	80		969	4507	0.0	3600	4547	2.2	274	4507	0.0
	90		3600	5042	1.4	3600	5621	13.1	1378	5020	0.0
	100		3600	5689	1.3	3600	6221	10.9	2023	5675	0.0
15	10	1	1	495	0.0	1	495	0.0	1	495	0.0
	20		3	1383	0.0	24	1383	0.0	2	1383	0.0
	30		38	2205	0.0	129	2205	0.0	12	2205	0.0
	40		846	3029	0.0	711	3029	0.0	108	3029	0.0
	50		935	3805	0.0	1019	3805	0.0	192	3805	0.0
	60		3600	4532	1.6	3600	4536	2.6	1741	4508	0.0
	70		3600	5494	2.1	3600	5819	8.4	3600	5495	1.8
	80		3600	6155	1.6	3600	7239	16.9	3600	6145	1.1
	90		3600	7159	2.2	3600	7919	12.0	3600	7303	4.0
	100		3600	8033	4.2	3600	9011	14.8	3600	8107	4.8
20	10	1	1	554	0.0	5	554	0.0	1	552	0.0
	20		9	1446	0.0	104	1446	0.0	3	1446	0.0
	30		35	2504	0.0	192	2504	0.0	12	2504	0.0
	40		301	3391	0.0	3168	3391	0.0	48	3391	0.0
	50		3600	4240	0.2	3600	4352	4.0	404	4239	0.0
	60		559	5206	0.0	3600	5456	5.7	515	5206	0.0
	70		3600	6254	1.9	3600	6686	8.8	3600	6211	0.2
	80		3600	6955	1.4	3600	8042	15.2	3600	6994	1.4
	90		3600	8048	1.6	3600	9095	13.2	3600	8065	1.5
	100		3600	8911	3.0	3600	10099	14.7	3600	8864	2.7

Table 5: Comprehensive results of the flow formulation, cut generation algorithm, path formulation, and Benders-based branch-and-cut algorithm with $r = \lfloor |S|/2 \rfloor$.

			PF			BDA			BD		
$\|S\|$	$\|T\|$	r	t(s)	TWT	gap(%)	t(s)	TWT	gap(%)	t(s)	TWT	gap(%)
5	10	3	1	366	0.0	1	366	0.0	1	366	0.0
	20		1	687	0.0	1	687	0.0	1	687	0.0
	30		5	1113	0.0	3	1113	0.0	1	1113	0.0
	40		108	1539	0.0	22	1539	0.0	4	1539	0.0
	50		52	1966	0.0	23	1966	0.0	8	1966	0.0
	60		488	2308	0.0	193	2308	0.0	96	2308	0.0
	70		1799	2646	0.0	292	2646	0.0	97	2646	0.0
	80		3600	3041	1.1	1592	3041	0.0	762	3041	0.0
	90		3600	3539	2.6	3103	3533	0.0	3600	3534	0.7
	100		3600	3954	4.7	3600	3906	2.1	3600	3940	3.7
10	10	5	1	442	0.0	1	442	0.0	15	442	0.0
	20		11	1049	0.0	24	1049	0.0	19	1049	0.0
	30		400	1645	0.0	81	1645	0.0	5	1645	0.0
	40		1443	2214	0.0	233	2214	0.0	15	2214	0.0
	50		3600	2723	1.3	1325	2723	0.0	60	2723	0.0
	60		3600	3361	3.2	3286	3335	0.0	159	3335	0.0
	70		3600	4042	5.7	1075	3890	0.0	109	3890	0.0
	80		3600	4611	4.8	3600	4529	2.6	491	4507	0.0
	90		3600	5268	8.3	3600	5131	4.8	3600	5022	0.9
	100		3600	5915	7.6	3600	6279	12.7	3600	5688	1.1
15	10	8	1	516	0.0	2	516	0.0	1	516	0.0
	20		436	1468	0.0	59	1468	0.0	3	1468	0.0
	30		2786	2310	0.0	589	2310	0.0	22	2310	0.0
	40		3600	3262	6.6	1832	3155	0.0	43	3155	0.0
	50		3600	4166	7.5	3600	4114	5.2	187	3991	0.0
	60		3600	4816	5.7	3600	4875	6.5	327	4701	0.0
	70		3600	6393	14.8	3600	6393	14.3	3600	5739	1.6
	80		3600	7225	14.7	3600	7225	14.3	3600	6489	1.7
	90		3600	8200	13.8	3600	7953	10.3	3600	7387	1.5
	100		3600	9163	14.6	3600	9163	100.0	3600	8121	0.9
20	10	10	2	572	0.0	8	572	0.0	1	570	0.0
	20		493	1600	0.0	493	1600	0.0	3	1600	0.0
	30		3600	2643	1.0	1510	2643	0.0	9	2643	0.0
	40		3600	3637	3.4	3600	3697	4.2	43	3591	0.0
	50		3600	4544	4.3	3600	4844	10.3	129	4467	0.0
	60		3600	6057	11.9	3600	6043	11.1	1289	5528	0.0
	70		3600	7065	10.8	3600	7001	9.4	1839	6551	0.0
	80		3600	8011	13.7	3600	8011	11.3	3600	7397	2.2
	90		3600	9049	100.0	3600	9049	9.7	3600	8656	4.1
	100		3600	10012	100.0	3600	10012	100.0	3600	9339	2.0

Table 6: Comprehensive comparison of the DTPR schedules w.r.t. the regular timetables (denoted with R) and the DTP solutions (denoted with N) with $r = 1$. Instances marked with * are not solved to optimality by the DTPR, instances marked with ⋄ are not solved to optimality by the DTP.

$\|S\|$	$\|T\|$	r	$\frac{(\text{TWT}_R - \text{TWT}_D)}{\text{TWT}_R}(\%)$	$\frac{(\max(B_R) - \max(B_D))}{\max(B_R)}(\%)$	$\frac{(\text{avg}(B_R) - \text{avg}(B_D))}{\text{avg}(B_R)}(\%)$	$\frac{(\sigma^2(B_R) - \sigma^2(B_D))}{\sigma^2(B_R)}(\%)$	$\frac{(\text{TWT}_N - \text{TWT}_D)}{\text{TWT}_N}(\%)$
5	10	1	4.7	1.9	-16.3	11.1	0.0
	20		2.1	-11.5	-4.4	2.9	-1.3
	30		4.2	0.0	-6.4	4.0	0.0
	40		4.3	-1.6	-3.9	2.3	-0.3
	50		2.5	0.0	-5.0	3.0	0.0
	60		0.9	0.0	-2.7	2.0	0.0
	70		4.2	0.0	-3.9	2.6	0.0
	80		4.4	0.0	-3.8	3.1	0.0
	90⋄		3.9	0.0	-2.2	1.7	0.1
	100⋄		1.6	7.4	-3.3	2.5	0.0
10	10	1	25.1	0.0	-3.1	2.2	0.0
	20		22.1	23.6	-4.2	3.5	0.0
	30		16.2	26.2	-5.4	4.8	0.0
	40		18.8	26.0	-5.8	5.1	0.0
	50		14.6	23.4	-3.9	3.6	0.0
	60		14.0	21.4	-2.1	2.0	0.0
	70		14.5	28.8	-1.7	1.6	-0.1
	80⋄		12.2	8.9	-4.5	4.1	0.2
	90⋄		11.6	19.6	-2.8	2.7	0.7
	100⋄		11.4	16.8	-2.4	2.2	1.2
15	10	1	41.9	10.2	-6.9	4.2	0.0
	20		22.5	24.2	-2.8	2.9	0.0
	30		15.8	15.3	-4.2	4.2	0.0
	40⋄		16.2	15.6	-5.3	4.5	-0.1
	50		14.8	20.0	-4.5	4.2	-0.2
	60⋄		16.3	22.7	-3.5	3.1	0.2
	70*⋄		12.9	13.1	-5.4	5.0	1.0
	80*⋄		15.1	20.0	-4.6	4.3	2.3
	90*⋄		10.1	26.1	-4.2	4.0	2.2
	100*⋄		10.0	12.0	-4.5	4.2	0.5
20	10	1	45.3	31.6	-7.2	4.5	0.0
	20		26.4	-5.3	-4.8	5.5	0.0
	30		19.4	13.5	-4.7	4.7	0.0
	40		19.1	9.0	-3.4	3.4	0.0
	50⋄		13.5	8.3	-5.4	5.1	0.8
	60⋄		13.0	15.2	-4.3	4.1	0.4
	70*⋄		13.5	6.0	-5.3	5.1	0.8
	80*⋄		13.0	19.5	-5.1	4.8	1.2
	90*⋄		11.3	8.8	-3.9	3.7	4.1
	100*⋄		12.2	20.1	-4.6	4.4	12.2

Table 7: Comprehensive comparison of the DTPR schedules w.r.t. the regular timetables (denoted with R) and the DTP solutions (denoted with N) with $r = \lfloor |S|/2 \rfloor$. Instances marked with * are not solved to optimality by the DTPR, instances marked with ⋄ are not solved to optimality by the DTP.

$\|S\|$	$\|T\|$	r	$\frac{(\text{TWT}_R - \text{TWT}_D)}{\text{TWT}_R}(\%)$	$\frac{(\max(B_R) - \max(B_D))}{\max(B_R)}(\%)$	$\frac{(\text{avg}(B_R) - \text{avg}(B_D))}{\text{avg}(B_R)}(\%)$	$\frac{(\sigma^2(B_R) - \sigma^2(B_D))}{\sigma^2(B_R)}(\%)$	$\frac{(\text{TWT}_N - \text{TWT}_D)}{\text{TWT}_N}(\%)$
5	10	3	4.7	1.9	-16.3	11.1	0.0
	20		3.4	-17.3	-7.7	4.4	0.0
	30		4.2	0.0	-6.4	4.0	0.0
	40		4.6	-1.6	-4.7	3.0	0.0
	50		2.5	0.0	-5.0	3.0	0.0
	60		0.9	0.0	-2.1	1.5	0.0
	70		4.2	0.0	-3.9	2.6	0.0
	80		4.4	0.0	-3.3	2.7	0.0
	90*⋄		3.9	-3.1	-2.6	2.1	0.1
	100*⋄		0.6	7.4	-3.1	2.2	-1.1
10	10	5	25.1	-5.6	-0.5	0.0	0.0
	20		22.1	23.6	-4.2	3.5	0.0
	30		16.2	26.2	-5.0	4.4	0.0
	40		18.8	26.0	-5.5	4.8	0.0
	50		14.6	23.4	-3.9	3.6	0.0
	60		14.0	21.4	-2.1	2.0	0.0
	70		14.5	33.9	-2.2	2.1	0.0
	80⋄		12.2	8.9	-5.0	4.6	0.2
	90*⋄		11.6	19.6	-2.8	2.7	0.7
	100*⋄		11.2	16.8	-2.1	2.0	1.0
15	10	8	39.4	10.2	-5.0	2.1	-4.2
	20		17.7	19.7	-1.9	1.8	-6.1
	30		11.8	4.4	-2.0	2.0	-4.8
	40⋄		12.7	17.0	-2.7	2.2	-4.2
	50		10.6	27.1	-2.1	2.0	-5.1
	60⋄		12.7	23.4	-3.6	3.3	-4.1
	70*⋄		9.1	21.4	-4.1	3.8	-3.4
	80*⋄		10.4	18.5	-2.1	2.1	-3.2
	90*⋄		9.1	24.8	-2.0	1.9	1.1
	100*⋄		9.9	29.1	-1.4	1.4	0.3
20	10	10	43.5	31.6	-5.9	3.0	-3.3
	20		18.6	12.8	-1.3	1.8	-10.7
	30		14.9	17.3	-1.8	1.6	-5.6
	40		14.4	9.0	-2.2	2.2	-5.9
	50⋄		8.9	21.8	-1.7	1.5	-4.6
	60⋄		7.6	19.6	-2.5	2.5	-5.8
	70⋄		8.8	26.3	-2.2	2.2	-4.6
	80*⋄		8.0	21.3	-1.6	1.5	-4.5
	90*⋄		4.8	15.0	-1.5	1.6	-2.9
	100*⋄		7.5	29.1	-1.7	1.7	7.5

Table 8: Comprehensive results of BD in configurations S0, S1, and S2, with and without compression, with $r = 1$.

			S0				S1				S1-compression				S2				S2-compression			
$\lvert S\rvert$	$\lvert T\rvert$	r	t(s)	TWT	gap(%)	cuts	t(s)	TWT	gap(%)	cuts	t(s)	TWT	gap(%)	cuts	t(s)	TWT	gap(%)	cuts	t(s)	TWT	gap(%)	cuts
	10		1	366	0.0	361	1	366	0.0	314	1	366	0.0	250	1	366	0.0	314	1	366	0.0	179
	20		1	696	0.0	656	1	709	0.0	607	1	709	0.0	478	1	709	0.0	607	1	709	0.0	343
	30		1	1113	0.0	1067	1	1138	0.0	945	1	1138	0.0	762	1	1138	0.0	889	1	1138	0.0	543
	40		1	1544	0.0	1512	1	1544	0.0	1218	1	1544	0.0	959	1	1544	0.0	1218	1	1544	0.0	672
5	50	1	1	1966	0.0	1789	1	1986	0.0	1519	1	1986	0.0	1227	1	1986	0.0	1470	1	1986	0.0	925
	60		7	2308	0.0	2211	1	2328	0.0	1873	1	2328	0.0	1520	1	2328	0.0	1971	1	2328	0.0	1095
	70		5	2646	0.0	2539	1	2673	0.0	2204	1	2673	0.0	1743	1	2673	0.0	2167	1	2673	0.0	1178
	80		69	3041	0.0	3266	1	3073	0.0	2549	1	3073	0.0	1974	1	3073	0.0	2549	1	3073	0.0	1380
	90		68	3533	0.0	3484	1	3546	0.0	2785	1	3546	0.0	2236	1	3546	0.0	2785	1	3546	0.0	1557
	100		299	3898	0.0	3950	1	3905	0.0	3084	1	3905	0.0	2431	1	3905	0.0	3193	1	3905	0.0	1711
	10		1	442	0.0	620	1	442	0.0	620	1	442	0.0	493	1	442	0.0	620	1	442	0.0	396
	20		1	1049	0.0	1473	1	1049	0.0	1507	1	1049	0.0	1138	1	1070	0.0	1538	1	1070	0.0	955
	30		3	1645	0.0	2819	3	1645	0.0	2870	2	1645	0.0	2235	2	1651	0.0	2899	1	1651	0.0	1629
	40		8	2214	0.0	4116	9	2214	0.0	4329	5	2214	0.0	3063	8	2226	0.0	3980	4	2226	0.0	2579
10	50	1	37	2723	0.0	5178	43	2723	0.0	4678	19	2723	0.0	3868	10	2723	0.0	4793	9	2723	0.0	2935
	60		76	3335	0.0	5893	179	3335	0.0	6067	54	3335	0.0	4740	114	3370	0.0	5644	43	3370	0.0	3955
	70		41	3893	0.0	6629	60	3893	0.0	7007	31	3893	0.0	5261	196	3969	0.0	6690	113	3969	0.0	4074
	80		274	4507	0.0	7802	330	4507	0.0	7712	168	4507	0.0	6342	286	4542	0.0	8049	144	4542	0.0	5252
	90		1378	5020	0.0	7044	2697	5020	0.0	9444	1378	5020	0.0	7044	3600	5070	0.4	8996	2121	5069	0.0	5849
	100		2023	5675	0.0	10141	3600	5685	1.0	10684	1325	5675	0.0	7881	3600	5753	0.9	10365	2571	5750	0.0	6550
	10		1	495	0.0	1090	1	495	0.0	1102	1	495	0.0	923	1	500	0.0	1037	1	500	0.0	693
	20		2	1383	0.0	2875	1	1384	0.0	2423	1	1384	0.0	2297	2	1389	0.0	2493	1	1389	0.0	1733
	30		12	2205	0.0	4624	4	2211	0.0	4137	2	2211	0.0	3570	3	2224	0.0	4323	2	2224	0.0	2528
	40		108	3029	0.0	6190	22	3052	0.0	6854	11	3052	0.0	5352	25	3071	0.0	6357	14	3071	0.0	3606
15	50	1	192	3805	0.0	8043	32	3847	0.0	7325	21	3847	0.0	6172	27	3864	0.0	7041	17	3864	0.0	4408
	60		1741	4508	0.0	9866	133	4550	0.0	8929	61	4550	0.0	7381	50	4555	0.0	8984	30	4555	0.0	5423
	70		3600	5495	1.8	11790	445	5476	0.0	11102	123	5476	0.0	9111	249	5489	0.0	11159	101	5489	0.0	7005
	80		3600	6145	1.1	14052	548	6157	0.0	13430	179	6157	0.0	10801	608	6213	0.0	12394	345	6213	0.0	8114
	90		3600	7303	4.0	15472	3077	7156	0.0	15104	2022	7156	0.0	11766	3600	7228	1.0	14695	3600	7216	0.6	9291
	100		3600	8107	4.8	16782	3600	7943	1.5	16350	3600	7896	0.4	12706	3600	7974	1.2	16754	3600	7968	0.7	9936
	10		1	552	0.0	1482	1	554	0.0	1439	1	554	0.0	1160	1	554	0.0	1439	1	554	0.0	880
	20		3	1446	0.0	3420	2	1455	0.0	3397	2	1455	0.0	2826	2	1455	0.0	3513	2	1455	0.0	2067
	30		12	2504	0.0	6087	6	2529	0.0	6086	6	2529	0.0	4715	19	2563	0.0	6107	8	2563	0.0	3326
	40		48	3391	0.0	9025	24	3415	0.0	7975	22	3415	0.0	6758	31	3423	0.0	8641	12	3423	0.0	4769
20	50	1	404	4239	0.0	10106	71	4272	0.0	10084	50	4272	0.0	7482	81	4284	0.0	9863	29	4284	0.0	5786
	60		515	5206	0.0	11881	69	5255	0.0	11663	56	5255	0.0	9313	58	5255	0.0	11670	33	5255	0.0	6631
	70		3600	6211	0.2	15014	1476	6276	0.0	14758	1194	6276	0.0	11485	1404	6282	0.0	14925	413	6282	0.0	8454
	80		3600	6994	1.4	16568	1050	7016	0.0	15864	2246	7016	0.0	13477	1747	7016	0.0	16106	399	7016	0.0	8958
	90		3600	8065	1.5	19214	354	8060	0.0	18810	570	8060	0.0	13654	322	8060	0.0	18268	122	8060	0.0	9814
	100		3600	8864	2.7	23254	3600	8867	0.2	20396	1891	8867	0.0	16661	3497	8883	0.0	20706	1961	8883	0.0	11673

Table 9: Comprehensive results of BD in configurations S0, S1, and S2, with and without compression, with $r = \lfloor |S|/2 \rfloor$.

			S0				S1				S1-compression				S2				S2-compression			
$\lvert S\rvert$	$\lvert T\rvert$	r	t(s)	TWT	gap(%)	cuts	t(s)	TWT	gap(%)	cuts	t(s)	TWT	gap(%)	cuts	t(s)	TWT	gap(%)	cuts	t(s)	TWT	gap(%)	cuts
	10		1	366	0.0	439	1	366	0.0	368	1	366	0.0	287	1	366	0.0	333	1	366	0.0	195
	20		1	687	0.0	777	1	709	0.0	678	1	709	0.0	555	1	709	0.0	675	1	709	0.0	404
	30		1	1113	0.0	1298	1	1138	0.0	1172	1	1138	0.0	922	1	1138	0.0	995	1	1138	0.0	619
	40		4	1539	0.0	1948	1	1544	0.0	1584	1	1544	0.0	1260	1	1544	0.0	1315	1	1544	0.0	738
5	50	3	8	1966	0.0	2305	1	1986	0.0	1850	1	1986	0.0	1506	1	1986	0.0	1803	1	1986	0.0	1017
	60		96	2308	0.0	2967	3	2328	0.0	2563	2	2328	0.0	1962	1	2328	0.0	2159	1	2328	0.0	1264
	70		97	2646	0.0	3517	3	2673	0.0	2823	3	2673	0.0	2217	1	2673	0.0	2501	1	2673	0.0	1437
	80		762	3041	0.0	4139	4	3073	0.0	3206	3	3073	0.0	2529	1	3073	0.0	2954	1	3073	0.0	1647
	90		3600	3534	0.7	4465	9	3546	0.0	3521	4	3546	0.0	2914	1	3546	0.0	3198	1	3546	0.0	1774
	100		3600	3940	3.7	4875	12	3905	0.0	4236	10	3905	0.0	3081	2	3905	0.0	3527	1	3905	0.0	2149
	10		15	442	0.0	917	1	442	0.0	767	1	442	0.0	786	1	442	0.0	777	1	442	0.0	513
	20		19	1049	0.0	1856	1	1049	0.0	1680	1	1049	0.0	1387	1	1070	0.0	2291	1	1070	0.0	1257
	30		5	1645	0.0	3894	4	1645	0.0	3974	2	1645	0.0	2973	3	1651	0.0	3800	2	1651	0.0	2390
	40		15	2214	0.0	5597	17	2214	0.0	4996	6	2214	0.0	4308	7	2226	0.0	5591	4	2226	0.0	3238
10	50	5	60	2723	0.0	6156	45	2723	0.0	7070	26	2723	0.0	5481	13	2723	0.0	6825	13	2723	0.0	4265
	60		159	3335	0.0	7734	139	3335	0.0	8698	85	3335	0.0	7169	56	3370	0.0	8353	68	3370	0.0	5489
	70		109	3890	0.0	10291	119	3893	0.0	8127	56	3893	0.0	7487	509	3969	0.0	9348	202	3969	0.0	6903
	80		491	4507	0.0	11606	1021	4507	0.0	11547	687	4507	0.0	9310	214	4542	0.0	11060	339	4542	0.0	7024
	90		3600	5022	0.9	13882	3600	5020	0.7	13551	2027	5020	0.0	10641	3600	5074	0.5	13145	3253	5069	0.0	8248
	100		3600	5688	1.1	14782	3471	5675	0.0	14464	2050	5675	0.0	11043	3600	5782	1.9	15345	3600	5757	1.2	9732
	10		1	516	0.0	1101	1	516	0.0	1089	1	516	0.0	925	1	516	0.0	1089	1	516	0.0	698
	20		3	1468	0.0	2978	1	1474	0.0	2697	1	1474	0.0	2347	3	1474	0.0	2741	1	1474	0.0	1665
	30		22	2310	0.0	5909	8	2342	0.0	5715	3	2342	0.0	4541	6	2347	0.0	5606	5	2347	0.0	3433
	40		43	3155	0.0	8379	20	3200	0.0	8106	13	3200	0.0	6699	13	3200	0.0	8647	12	3200	0.0	5182
15	50	8	187	3991	0.0	9608	103	4067	0.0	11134	71	4067	0.0	8641	99	4070	0.0	10905	66	4070	0.0	6887
	60		327	4701	0.0	14320	88	4784	0.0	11749	111	4784	0.0	10453	85	4787	0.0	11372	59	4787	0.0	7590
	70		3600	5739	1.6	16454	1486	5779	0.0	16105	1403	5779	0.0	13839	950	5779	0.0	15496	559	5779	0.0	9315
	80		3600	6489	1.7	18533	2895	6508	0.0	18263	644	6508	0.0	15389	613	6508	0.0	18407	414	6508	0.0	11553
	90		3600	7387	1.5	20748	1706	7398	0.0	20686	585	7398	0.0	16881	1732	7398	0.0	20280	454	7398	0.0	11806
	100		3600	8121	0.9	22804	619	8169	0.0	19499	990	8169	0.0	19012	316	8169	0.0	21667	219	8169	0.0	13067
	10		1	570	0.0	1518	1	572	0.0	1611	1	572	0.0	1223	1	572	0.0	1511	1	572	0.0	959
	20		3	1600	0.0	3910	2	1601	0.0	3495	1	1601	0.0	2946	1	1601	0.0	3578	1	1601	0.0	2197
	30		9	2643	0.0	7008	6	2670	0.0	7149	4	2670	0.0	5482	6	2694	0.0	6783	5	2694	0.0	4242
	40		43	3591	0.0	11196	13	3606	0.0	9955	9	3606	0.0	7676	21	3626	0.0	10992	10	3626	0.0	5691
20	50	10	129	4467	0.0	13850	44	4484	0.0	12070	23	4484	0.0	9365	56	4533	0.0	12019	27	4533	0.0	6770
	60		1289	5528	0.0	16667	374	5594	0.0	15894	238	5594	0.0	12364	404	5621	0.0	16796	159	5621	0.0	9725
	70		1839	6551	0.0	19954	402	6609	0.0	19390	547	6609	0.0	15333	1033	6644	0.0	19961	335	6644	0.0	11435
	80		3600	7397	2.2	23428	445	7401	0.0	20017	731	7401	0.0	16618	3600	7489	0.7	20292	1322	7485	0.0	12421
	90		3600	8656	4.1	26663	3600	8532	0.3	24920	3600	8532	0.5	19553	3600	8564	0.5	26013	1793	8564	0.0	13613
	100		3600	9339	2.0	28178	3600	9377	0.5	26956	3600	9402	0.6	20405	3600	9412	0.7	27538	2443	9412	0.0	15816

Table 10: Comprehensive comparison of the DTPR schedules with configuration S1 w.r.t. the regular timetables (denoted with R) and DTPR solutions with configuration S0, with $r = 1$. Instances marked with * are not solved to optimality by the DTPR in configuration S1, instances marked with $\diamond$ are not solved to optimality by the DTPR in configuration S0.

$\|S\|$	$\|T\|$	r	$\frac{(\text{TWT}_R - \text{TWT}_D)}{\text{TWT}_R}(\%)$	$\frac{(\max(B_R) - \max(B_D))}{\max(B_R)}(\%)$	$\frac{(\text{avg}(B_R) - \text{avg}(B_D))}{\text{avg}(B_R)}(\%)$	$\frac{(\sigma^2(B_R) - \sigma^2(B_D))}{\sigma^2(B_R)}(\%)$	$\frac{(\text{TWT}_N - \text{TWT}_D)}{\text{TWT}_N}(\%)$
5	10	1	4.7	1.9	-16.3	11.1	0.0
	20		0.3	-17.3	-6.1	2.9	-1.9
	30		2.1	0.0	-2.8	2.0	-2.2
	40		4.3	-1.6	-3.1	1.5	0.0
	50		1.5	0.0	-5.8	4.3	-1.0
	60		0.0	0.0	0.0	0.0	-0.9
	70		3.2	0.0	-2.4	1.3	-1.0
	80		3.4	0.0	-3.4	2.7	-1.1
	90		3.6	0.0	-2.2	1.7	-0.4
	100		1.4	7.4	-2.4	1.5	-0.2
10	10	1	25.1	0.0	-3.1	2.2	0.0
	20		22.1	23.6	-4.2	3.5	0.0
	30		16.2	26.2	-5.4	4.8	0.0
	40		18.8	26.0	-5.5	4.8	0.0
	50		14.6	23.4	-3.9	3.6	0.0
	60		14.0	21.4	-2.1	2.0	0.0
	70		14.5	28.8	-1.8	1.7	0.0
	80		12.2	8.9	-4.7	4.3	0.0
	90		11.6	19.6	-2.8	2.7	0.0
	100		11.4	16.8	-2.5	2.3	0.0
15	10	1	41.9	10.2	-6.9	4.2	0.0
	20		22.4	24.2	-3.6	3.6	-0.1
	30		15.6	15.3	-4.3	4.2	-0.3
	40		15.6	15.6	-5.0	4.3	-0.8
	50		13.8	20.7	-4.1	3.9	-1.1
	60		15.5	19.1	-4.6	4.3	-0.9
	70$\diamond$		13.2	13.1	-4.7	4.4	0.3
	80$\diamond$		14.9	20.0	-3.8	3.6	-0.2
	90$\diamond$		11.9	26.1	-3.9	3.8	2.0
	100*$\diamond$		12.4	29.7	-3.5	3.3	2.6
20	10	1	45.1	31.6	-7.0	4.5	-0.4
	20		26.0	-5.3	-3.7	4.0	-0.6
	30		18.6	-21.8	-6.6	7.7	-1.0
	40		18.6	14.1	-3.9	3.8	-0.7
	50		12.9	19.9	-5.5	5.2	-0.8
	60		12.2	15.2	-3.6	3.5	-0.9
	70$\diamond$		12.6	18.0	-4.2	4.0	-1.0
	80$\diamond$		12.8	7.3	-5.2	4.9	-0.3
	90$\diamond$		11.4	8.8	-4.2	4.0	0.1
	100$\diamond$		12.2	14.0	-4.8	4.6	0.0

Table 11: Comprehensive comparison of the DTPR schedules with configuration S1 w.r.t. the regular timetables (denoted with R) and DTPR solutions with configuration S0, with $r = \lfloor |S|/2 \rfloor$. Instances marked with * are not solved to optimality by the DTPR in configuration S1, instances marked with ⋄ are not solved to optimality by the DTPR in configuration S0.

$\|S\|$	$\|T\|$	r	$\frac{(\mathrm{TWT}_R - \mathrm{TWT}_D)}{\mathrm{TWT}_R}(\%)$	$\frac{(\max(B_R) - \max(B_D))}{\max(B_R)}(\%)$	$\frac{(\mathrm{avg}(B_R) - \mathrm{avg}(B_D))}{\mathrm{avg}(B_R)}(\%)$	$\frac{(\sigma^2(B_R) - \sigma^2(B_D))}{\sigma^2(B_R)}(\%)$	$\frac{(\mathrm{TWT}_N - \mathrm{TWT}_D)}{\mathrm{TWT}_N}(\%)$
5	10	3	4.7	1.9	-16.3	11.1	0.0
	20		0.3	-17.3	-7.7	4.4	-3.2
	30		2.1	0.0	-2.8	2.0	-2.2
	40		4.3	-1.6	-3.9	2.3	-0.3
	50		1.5	0.0	-5.8	4.3	-1.0
	60		0.0	0.0	-1.0	1.0	-0.9
	70		3.2	0.0	-2.4	1.3	-1.0
	80		3.4	0.0	-3.8	3.1	-1.1
	90⋄		3.6	0.0	-2.2	1.7	-0.3
	100⋄		1.4	7.4	-2.7	1.9	0.9
10	10	5	25.1	0.0	-3.1	2.2	0.0
	20		22.1	23.6	-4.2	3.5	0.0
	30		16.2	26.2	-5.4	4.8	0.0
	40		18.8	26.0	-5.5	4.8	0.0
	50		14.6	23.4	-3.3	3.1	0.0
	60		14.0	21.4	-2.3	2.2	0.0
	70		14.5	28.8	-1.8	1.7	-0.1
	80		12.2	8.9	-4.5	4.1	0.0
	90⋄		11.6	19.6	-2.8	2.7	0.0
	100⋄		11.4	16.8	-2.1	2.0	0.2
15	10	8	39.4	10.2	-5.0	2.1	0.0
	20		17.4	19.7	-1.4	1.5	-0.4
	30		10.6	9.5	-1.4	1.5	-1.4
	40		11.5	17.0	-3.5	3.0	-1.4
	50		8.9	23.6	-2.0	1.8	-1.9
	60		11.1	22.7	-2.0	1.9	-1.8
	70⋄		8.4	24.8	-2.0	1.8	-0.7
	80⋄		10.1	2.2	-1.4	1.3	-0.3
	90⋄		8.9	33.8	-1.9	1.8	-0.1
	100⋄		9.3	29.1	-2.1	2.0	-0.6
20	10	10	43.3	25.4	-6.6	3.5	-0.4
	20		18.5	12.8	-1.0	1.3	-0.1
	30		14.0	21.2	-1.2	1.3	-1.0
	40		14.0	13.5	-2.1	2.2	-0.4
	50		8.5	21.8	-1.7	1.6	-0.4
	60		6.5	10.1	-2.4	2.4	-1.2
	70		8.0	14.4	-2.4	2.3	-0.9
	80⋄		8.0	15.9	-1.2	1.2	-0.1
	90*⋄		6.2	12.5	-1.2	1.0	1.4
	100*⋄		6.9	27.4	-1.3	1.4	-0.7

Table 12: Comprehensive comparison of the DTPR schedules with configuration S2 w.r.t. the regular timetables (denoted with R) and w.r.t. the DTPR solutions with configuration S0, with $r = 1$. Instances marked with * are not solved to optimality by the DTPR in configuration S1, instances marked with ⋄ are not solved to optimality by the DTPR in configuration S0.

$\|S\|$	$\|T\|$	r	$\frac{(\text{TWT}_R - \text{TWT}_D)}{\text{TWT}_R}(\%)$	$\frac{(\max(B_R) - \max(B_D))}{\max(B_R)}(\%)$	$\frac{(\text{avg}(B_R) - \text{avg}(B_D))}{\text{avg}(B_R)}(\%)$	$\frac{(\sigma^2(B_R) - \sigma^2(B_D))}{\sigma^2(B_R)}(\%)$	$\frac{(\text{TWT}_N - \text{TWT}_D)}{\text{TWT}_N}(\%)$
5	10	1	4.7	1.9	-16.3	11.1	0.0
	20		0.3	-17.3	-6.1	2.9	-1.9
	30		2.1	0.0	-2.8	2.0	-2.2
	40		4.3	-1.6	-3.1	1.5	0.0
	50		1.5	0.0	-5.8	4.3	-1.0
	60		0.0	0.0	0.0	0.0	-0.9
	70		3.2	0.0	-2.4	1.3	-1.0
	80		3.4	0.0	-3.4	2.7	-1.1
	90		3.6	0.0	-2.2	1.7	-0.4
	100		1.4	7.4	-2.4	1.5	-0.2
10	10	1	25.1	0.0	-3.1	2.2	0.0
	20		20.6	23.6	-4.5	3.5	-2.0
	30		15.9	15.5	-4.8	4.0	-0.4
	40		18.3	19.0	-4.5	3.9	-0.5
	50		14.6	23.4	-3.9	3.6	0.0
	60		13.1	18.4	-3.3	3.2	-1.0
	70		12.8	31.4	-2.6	2.4	-2.0
	80		11.5	5.0	-4.7	4.3	-0.8
	90		10.8	19.6	-2.8	2.7	-1.0
	100		10.2	21.5	-1.8	1.6	-1.3
15	10	1	41.3	10.2	-6.1	3.5	-1.0
	20		22.1	24.2	-3.6	3.6	-0.4
	30		15.1	15.3	-4.3	4.2	-0.9
	40		15.1	15.6	-5.2	4.5	-1.4
	50		13.4	12.9	-4.0	3.8	-1.6
	60		15.4	19.9	-5.2	4.9	-1.0
	70⋄		13.0	13.1	-5.3	5.0	0.1
	80⋄		14.2	3.7	-4.2	4.0	-1.1
	90*⋄		11.2	20.4	-3.7	3.6	1.2
	100*⋄		11.6	29.7	-3.7	3.6	1.7
20	10	1	45.1	25.4	-8.1	5.6	-0.4
	20		26.0	-5.3	-3.7	4.0	-0.6
	30		17.5	13.5	-4.0	3.9	-2.4
	40		18.4	14.1	-3.0	3.0	-0.9
	50		12.6	19.9	-5.7	5.4	-1.1
	60		12.2	15.2	-3.8	3.7	-0.9
	70⋄		12.6	15.6	-4.9	4.7	-1.1
	80⋄		12.8	7.3	-4.7	4.5	-0.3
	90⋄		11.4	8.8	-4.1	3.9	0.1
	100⋄		12.0	14.0	-4.6	4.4	-0.2

Table 13: Comprehensive comparison of the DTPR schedules with configuration S2 w.r.t. the regular timetables (denoted with R) and w.r.t. the DTPR solutions with configuration S0, with $r = \lfloor |S|/2 \rfloor$. Instances marked with * are not solved to optimality by the DTPR in configuration S1, instances marked with ⋄ are not solved to optimality by the DTPR in configuration S0.

$\|S\|$	$\|T\|$	r	$\frac{(\text{TWT}_R - \text{TWT}_D)}{\text{TWT}_R}(\%)$	$\frac{(\max(B_R) - \max(B_D))}{\max(B_R)}(\%)$	$\frac{(\text{avg}(B_R) - \text{avg}(B_D))}{\text{avg}(B_R)}(\%)$	$\frac{(\sigma^2(B_R) - \sigma^2(B_D))}{\sigma^2(B_R)}(\%)$	$\frac{(\text{TWT}_N - \text{TWT}_D)}{\text{TWT}_N}(\%)$
5	10	3	4.7	1.9	-16.3	11.1	0.0
	20		0.3	-17.3	-6.1	2.9	-3.2
	30		2.1	0.0	-2.8	2.0	-2.2
	40		4.3	-1.6	-3.1	1.5	-0.3
	50		1.5	0.0	-5.8	4.3	-1.0
	60		0.0	0.0	-1.0	1.0	-0.9
	70		3.2	0.0	-2.4	1.3	-1.0
	80		3.4	0.0	-3.4	2.7	-1.1
	90⋄		3.6	0.0	-1.5	1.0	-0.3
	100⋄		1.4	7.4	-2.7	1.9	0.9
10	10	5	25.1	0.0	-3.1	2.2	0.0
	20		20.6	23.6	-4.5	3.5	-2.0
	30		15.9	15.5	-4.8	4.0	-0.4
	40		18.3	19.0	-4.1	3.6	-0.5
	50		14.6	23.4	-3.9	3.6	0.0
	60		13.1	18.4	-3.5	3.4	-1.0
	70		12.8	31.4	-3.1	3.0	-2.0
	80		11.5	5.0	-4.7	4.3	-0.8
	90⋄		10.8	19.6	-2.8	2.7	-0.9
	100*⋄		10.1	14.0	-1.6	1.5	-1.2
15	10	8	39.4	10.2	-5.8	2.8	0.0
	20		17.4	19.7	-1.4	1.5	-0.4
	30		10.4	19.0	-2.2	2.2	-1.6
	40		11.5	17.0	-3.3	2.8	-1.4
	50		8.8	25.7	-1.4	1.2	-2.0
	60		11.1	22.7	-1.8	1.6	-1.8
	70⋄		8.4	25.5	-1.1	1.0	-0.7
	80⋄		10.1	20.0	-0.9	0.9	-0.3
	90⋄		8.9	35.0	-1.5	1.4	-0.1
	100⋄		9.3	29.1	-1.6	1.5	-0.6
20	10	10	43.3	25.4	-6.9	4.0	-0.4
	20		18.5	12.8	-0.8	1.1	-0.1
	30		13.2	18.6	-1.9	2.0	-1.9
	40		13.5	14.7	-2.1	2.2	-1.0
	50		7.5	16.0	-2.0	1.8	-1.5
	60		6.1	21.5	-1.7	1.7	-1.7
	70		7.5	14.4	-2.7	2.7	-1.4
	80⋄		6.9	15.9	-1.3	1.2	-1.2
	90⋄		5.8	15.0	-1.4	1.3	1.1
	100⋄		6.8	28.5	-1.2	1.2	-0.8