Supplemental Material
Integrating Public Transport in Sustainable Last Mile Delivery

Diego Delle Donne, Alberto Santini and Claudia Archetti

1 Proofs of the theorems

1.1 Proof of Theorem 1

We begin by showing that, without valid inequalities (1i) and (1j), the optimal solution of the continuous relaxation of the 3T-DPPT can be arbitrarily small. Consider an instance with a single customer, $\mathcal{C} = \{c\}$, a single in-stop $\mathcal{S}^{\text{in}} = \{s_1\}$ and a single out-stop $\mathcal{S}^{\text{out}} = \{s_2\}$. The instance consists of one truck and one courier. Assume that both the truck and the courier have enough capacity to carry c's parcel, that c's time window corresponds to the entire planning horizon, and that $W^{\max}$ and $L^{\max}$ are equal to the planning horizon's length. Consider a bus line running through s_1 and s_2, with n buses scheduled during the time horizon, i.e., $\mathcal{P} = \{p_1, \dots, p_n\}$. Because of the assumptions above, $\mathcal{P}^{\text{in}}_{s_1 c} = \mathcal{P}^{\text{out}}_{s_2 c} = \mathcal{P}$.

Then set $\mathcal{R}^{\text{D}}$ consists of routes of type (o, s_1, o), with one route for each feasible start time. Analogously, $\mathcal{R}^{\text{F}}$ consists of routes of type (s_2, c, s_2), with one route for each feasible start time. All the routes in $\mathcal{R}^{\text{D}}$ have the same cost c_{D}, and all the routes in $\mathcal{R}^{\text{F}}$ have the same cost c_{F}.

Consider a route $r_d \in \mathcal{R}^{\text{D}}$ such that $t_{r_d s_1} \leq t^{p_1}_{s_1}$ and a route $r_f \in \mathcal{R}^{\text{F}}$ such that $t_{r_f s_2} \geq t^{p_n}_{s_2}$. An optimal solution of LP(1^-) is: $x_{r_d} = y_{r_f} = \frac{1}{n}$ and $z^{\text{in}}_{s_1 pc} = z^{\text{out}}_{s_2 pc} = \frac{1}{n}$ for all $p \in \mathcal{P}$. This solution satisfies constraints (1b)–(1h), and results in an objective value of $\frac{1}{n}(c_{\text{D}} + c_{\text{F}})$. Considering an instance with a sufficiently large number of buses n, then, one can make the objective value as small as desired.

We now consider the optimal solution of LP(1) for the same instance, i.e., the optimal solution of formulation 1 where (1i) and (1j) are included. Whatever the value of variables x and y in the solution of the relaxation, because of constraints (1i) and (1j), their respective sum is equal to 1. Thus the solution value is $c_{\text{D}} + c_{\text{F}}$ and $\frac{\text{LP}(1)}{\text{LP}(1^-)} = \frac{c_{\text{D}}+c_{\text{F}}}{\frac{1}{n}(c_{\text{D}}+c_{\text{F}})} = n$ which tends to infinity for $n \to \infty$. $\square$

More in general, constraints (1i) and (1j) guarantee that the value of LP(1) is always positive with a lower bound corresponding to the sum of the least cost route in $\mathcal{R}^{\text{D}}$ and $\mathcal{R}^{\text{F}}$. This lower bound is invalid in case constraints (1i) and (1j) are removed.

1.2 Proof of Proposition 1

To prove that L_{P_1} dominates L_{P_2} we shall prove that: (i) L_{P_1} can be extended to any *complete* route (i.e., ending at the CDC) to which L_{P_2} can be extended, and (ii) the corresponding extension of L_{P_1} has a cost better or equal than the extension of L_{P_2}.

Assume that L_{P_2} is extended by a sequence of nodes $(s_1, c_1), \dots, (s_k, c_k)$. This implies that $s_i \in \mathcal{S}_{P_2} \subseteq \mathcal{S}_{P_1}$, $c_i \in \mathcal{C}_{P_2} \subseteq \mathcal{C}_{P_1}$, for $i = 1, \dots, k$, and $\sum_{i=1}^{k} q_{c_i} \leq Q_{P_2} \leq Q_{P_1}$. Therefore, L_{P_1} can also be extended by the same sequence, thus proving (i). Now, the cost function for the extension of L_{P_1} to a node $w = (s, c) \in V$ is

$$\bar{C}_{P_1+w}(t) = \nu_{sc}(t) + \bar{C}_{P_1}(t - t_{vw}) + c_{vw} \leq \nu_{sc}(t) + \bar{C}_{P_2}(t - t_{vw}) + c_{vw} = \bar{C}_{P_2+w}(t),$$

where $v = v_{P_1} = v_{P_2}$. Hence, any sequence of extensions applied to both labels L_{P1} and L_{P2} will give routes in which the cost of the former is always better or equal to the cost of the latter, thus proving (ii). $\square$

As mentioned in Section 5.1.1, a key factor for the correctness of the dominance rule presented above is that the reduced cost associated with the label is calculated with respect to the ending vertex of the path, i.e., $\bar{C}_P(t)$, and not with respect to the starting time from the CDC, $C_P(t)$, as done by Tagmouti, Gendreau and Potvin (2007). Indeed, consider the example depicted in Figure 1 where, for

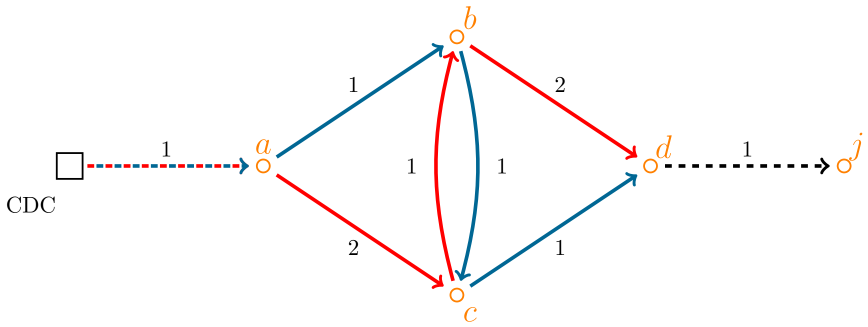

Figure 1: Example graph highlighting the importance of using $\bar{C}_P(t)$ over $C_P(t)$ in the dominance rule.

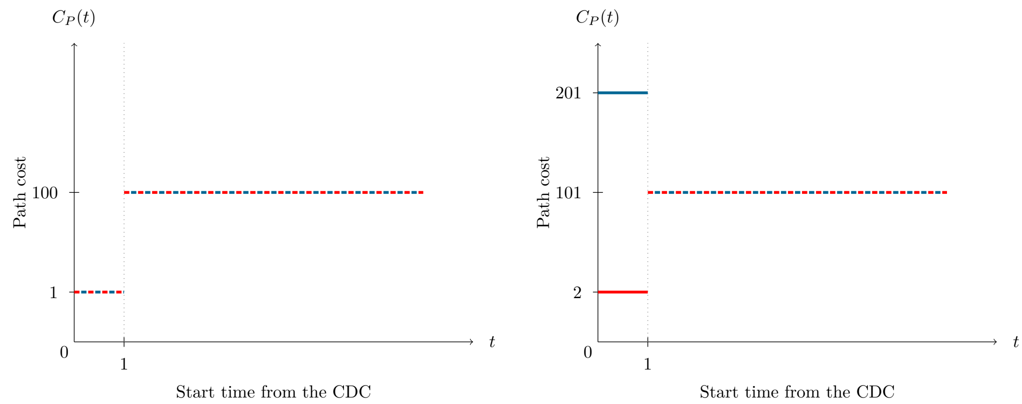

Figure 2: Cost functions of paths ending at vertex d (left) and paths extended to vertex j (right). Costs are in function to starting time from the CDC (as in Tagmouti, Gendreau and Potvin 2007).

ease of exposition, we do not consider the capacity constraint and where the numbers close to each arc are the corresponding travelling times.

The red and blue paths both end at vertex d. The total travel time is four for the blue path and six for the red one. We now extend the paths to vertex j. In our instance, the cost functions associated with each vertex of the graph are constant zero, except for vertices a and j where they are

$$C_a(t) = \begin{cases} 1 & \text{for } t \leq 2 \\ 100 & \text{for } t > 2 \end{cases} \qquad C_j(t) = \begin{cases} 200 & \text{for } t \leq 6 \\ 1 & \text{for } t > 6. \end{cases}$$

If we used the rule proposed by Tagmouti, Gendreau and Potvin 2007, calculating the costs with respect to the starting time from vertex the CDC, i.e., using $C_P(t)$, we would obtain the costs depicted in Figure 2 (left). The blue and red paths have the same cost for all values of t. Because the total travel time of the blue path is lower, it would dominate the red one. However, extending both labels to vertex j we get the costs depicted in Figure 2 (right). Here we see that the red path is no longer dominated because it has a lower cost for start times $t \leq 1$. In this example, in fact, it is preferable to arrive soon to node a but late to node d. Since trucks cannot wait at nodes, it is better to take a longer route between nodes a and d. Figure 3 shows the costs of the blue and red paths at vertex d, when computing costs using the arrival times at the nodes, i.e., using cost function $\bar{C}_P(t)$.

This example shows that we cannot apply the rule proposed by Tagmouti, Gendreau and Potvin (2007) to our case because trucks cannot wait along the route, and cost functions associated with nodes might be decreasing.

As a side note, we remark that there exists some ambiguity in the problem setting presented in Tagmouti, Gendreau and Potvin 2007. On one side, the compact problem formulation presented in

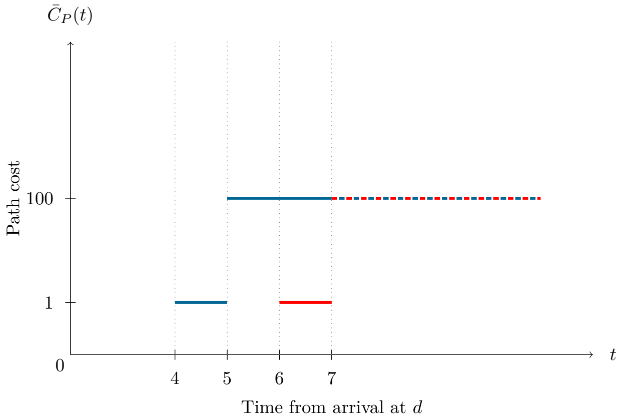

Figure 3: Cost functions of paths ending at vertex d, in function to arrival times (i.e., $\bar{C}_P(t)$).

the paper allows for waiting at customers. However, the problem description states that waiting is not allowed. Moreover, the way in which the cost function associated with each partial path is calculated is such that waiting at customers is not allowed.

2 A MIP approach to the pricing subproblem

The pricing problem SP_x can be tackled as a MIP and solved using a black-box solver. On small instances for which we can solve SP_x to optimality, the MIP formulation usually performs worse than the specialised approaches presented in Sections 5.1.1 and 5.1.2. Even when solving the pricing problem heuristically, the methods of Section 6.1 produce columns with lower reduced cost in a shorter time. On the other hand, the advantage of the MIP approach is that it promptly provides us with a lower (dual) bound on the reduced cost of a truck column. Such a bound can be exploited to obtain a dual bound for the entire 3T-DPPT, as in Section 3.1.

Consider variables $w_{ij} \in \{0,1\}$ defined for $i, j \in \{o\} \cup \mathcal{S}^{\text{in}}$ $(i \neq j)$ and taking value 1 iff j is visited immediately after i. Let $\pi_s \geq 0$ be a variable representing the departing time from $s \in \mathcal{S}^{\text{in}}$, if s is visited. Let $\gamma_{sc} \in \{0,1\}$ be a binary variable taking value 1 iff the route delivers the parcel of $c \in \mathcal{C}$ at in-stop $s \in \mathcal{S}^{\text{in}}$. Finally, let $\delta_{spc} \in \{0,1\}$ be a binary variable taking value 1 iff the route delivers the parcel of $c \in \mathcal{C}$ at in-stop $s \in \mathcal{S}^{\text{in}}$ at a time compatible with pick-up by bus $p \in \mathcal{P}_{cs}$. Then a MIP model to solve SP_x reads as follows:

$$
\begin{aligned}
\min \quad & \sum_{i,j \in \{o\} \cup \mathcal{S}^{\text{in}}} c_{ij} w_{ij} + \\
& \sum_{c \in \mathcal{C}} \sum_{s \in \mathcal{S}^{\text{in}}_c} \lambda^{(1i)}_c \gamma_{sc} + \\
& \sum_{c \in \mathcal{C}} \sum_{s \in \mathcal{S}^{\text{in}}_c} \sum_{p \in \mathcal{P}_{sc}} \lambda^{(1g)}_{spc} \delta_{spc} && && \text{(1a)} \\
\text{s.t.} \quad & \sum_{s \in \mathcal{S}^{\text{in}}} w_{os} = \sum_{s \in \mathcal{S}^{\text{in}}} w_{so} = 1 && && \text{(1b)} \\
& \sum_{\substack{s' \in \{o\} \cup \mathcal{S}^{\text{in}} \\ s' \neq s}} w_{ss'} = \sum_{\substack{s' \in \{o\} \cup \mathcal{S}^{\text{in}} \\ s' \neq s}} w_{s's} && \forall s \in \mathcal{S}^{\text{in}} && \text{(1c)} \\
& \pi_i + t_{ij} + T_j - M(1 - w_{ij}) \leq \pi_j && \forall i, j \in \mathcal{S}^{\text{in}} \cup \{0\},\ i \neq j, j \neq 0 && \text{(1d)} \\
& \pi_j \leq \pi_i + t_{ij} + T_j + M(1 - w_{ij}) && \forall i, j \in \mathcal{S}^{\text{in}} \cup \{0\},\ i \neq j, j \neq 0 && \text{(1e)} \\
& t^s_p - W^{\max} - M(1 - \delta_{spc}) \leq \pi_s && \forall c \in \mathcal{C},\ \forall s \in \mathcal{S}^{\text{in}}_c,\ \forall p \in \mathcal{P}^{\text{in}}_{sc} && \text{(1f)} \\
& \pi_s \leq t^s_p + M(1 - \delta_{spc}) && \forall c \in \mathcal{C},\ \forall s \in \mathcal{S}^{\text{in}}_c,\ \forall p \in \mathcal{P}^{\text{in}}_{sc} && \text{(1g)}
\end{aligned}
$$

$$\delta_{spc} \leq \gamma_{sc} \qquad \forall c \in \mathcal{C},\ \forall s \in \mathcal{S}_c^{\text{in}},\ \forall p \in \mathcal{P}_{sc}^{\text{in}} \tag{1h}$$

$$\gamma_{sc} \leq \sum_{p \in \mathcal{P}_{sc}} \delta_{spc} \qquad \forall c \in \mathcal{C},\ \forall s \in \mathcal{S}_c^{\text{in}} \tag{1i}$$

$$\gamma_{sc} \leq \sum_{i \in \{o\} \cup \mathcal{S}^{\text{in}}} w_{is} \qquad \forall c \in \mathcal{C},\ \forall s \in \mathcal{S}_c^{\text{in}} \tag{1j}$$

$$w_{ij} \leq \sum_{c \in \mathcal{C}} \gamma_{jc} \qquad \forall i \in \{o\} \cup \mathcal{S}^{\text{in}},\ \forall j \in \mathcal{S}^{\text{in}} \setminus \{i\} \tag{1k}$$

$$\sum_{s \in \mathcal{S}_c^{\text{in}}} \gamma_{sc} \leq 1 \qquad \forall c \in \mathcal{C} \tag{1l}$$

$$\sum_{c \in \mathcal{C}} \sum_{s \in \mathcal{S}_c} q_c \gamma_{sc} \leq Q \tag{1m}$$

$$w_{ij} \in \{0,1\} \qquad \forall i,j \in \{o\} \cup \mathcal{S}^{\text{in}},\ i \neq j \tag{1n}$$

$$\pi_s \geq 0 \qquad \forall s \in \mathcal{S}^{\text{in}} \tag{1o}$$

$$\gamma_{sc} \in \{0,1\} \qquad \forall c \in \mathcal{C},\ \forall s \in \mathcal{S}_c^{\text{in}} \tag{1p}$$

$$\delta_{spc} \in \{0,1\} \qquad \forall c \in \mathcal{C},\ \forall s \in \mathcal{S}_c^{\text{in}},\ \forall p \in \mathcal{P}_{sc}, \tag{1q}$$

where $M > 0$ is a sufficiently large number. The objective function (1a) minimises the reduced cost of the route (but for constant term $\lambda^{(1b)}$). Constraints (1b) and (1c) are classical arc-based formulation constraints ensuring flow, elementarity, and starting and ending at the CDC. Constraints (1d) are MTZ-like constraints used to set the value of variables π_s while Constraints (1e) forbids the truck to delay its route at a stop. Variables π and δ are linked through constraints (1f) and (1g); γ and δ through constraints (1h) and (1i); w and γ through constraints (1j) and (1k). Finally, constraints (1l) ensure that each parcel is delivered at most once, and (1m) make sure that the truck's capacity is respected.

3 Details of the bounding techniques

3.1 Lagrangean bound

Consider an optimal solution to the RRMP when potentially not all negative reduced cost columns have been generated. The corresponding dual solution, which we denote with $\bar{\lambda}$, can be unfeasible for the dual of MP_{CONT}. This is because a missing column in RRMP may correspond to a violated constraint in the dual. Therefore, its objective value $\bar{Z}$ is not a valid dual bound for MP_{CONT}. The key idea behind the first bounding technique is to provide a way to restore the dual feasibility of $\bar{\lambda}$. In this way, we obtain a new dual-feasible solution $\tilde{\lambda}$ with dual objective value $\tilde{Z}$. By weak duality, $\tilde{Z}$ is a valid dual bound for MP_{CONT} and, therefore, for the 3T-DPPT.

Let $\bar{Z}^{\text{D}}$ be a lower bound for the minimum reduced cost of a truck route, and $\bar{Z}_s^{\text{F}}$ be a lower bound for the minimum reduced cost of a courier route starting from $s \in \mathcal{S}^{\text{out}}$. In other words, the following inequalities hold for $\bar{Z}^{\text{D}}$ and $\bar{Z}_s^{\text{F}}$:

$$\bar{Z}^{\text{D}} \leq \min_{r \in \mathcal{R}^{\text{D}}} \Bigg\{ c_r + \bar{\lambda}^{(1b)} - \sum_{c \in \mathcal{C}_r} \bar{\lambda}_c^{(1i)} - \sum_{c \in \mathcal{C}_r} \sum_{p \in \mathcal{P}_c} \sum_{\substack{s \in \mathcal{S}_{pc}^{\text{out}} \\ \text{s.t. } r \in \mathcal{R}_{spc}^{\text{D}}}} \bar{\lambda}_{spc}^{(1g)} \Bigg\} \tag{2}$$

$$\bar{Z}_s^{\text{F}} \leq \min_{r \in \mathcal{R}_s^{\text{F}}} \Bigg\{ c_r + \bar{\lambda}_s^{(1c)} - \sum_{c \in \mathcal{C}_r} \bar{\lambda}_c^{(1j)} - \sum_{c \in \mathcal{C}_r} \sum_{\substack{p \in \mathcal{P}_c \\ \text{s.t. } r \in \mathcal{R}_{spc}^{\text{F}}}} \bar{\lambda}_{spc}^{(1h)} \Bigg\}. \tag{3}$$

Consider now the dual solution $\tilde{\lambda}$ obtained from $\bar{\lambda}$ modifying the following components:

$$\tilde{\lambda}^{(1b)} = \bar{\lambda}^{(1b)} - \bar{Z}^{\text{D}} \tag{4}$$

$$\tilde{\lambda}_s^{(1c)} = \bar{\lambda}_s^{(1c)} - \bar{Z}_s^{\mathrm{F}} \quad \forall s \in \mathcal{S}^{\mathrm{out}}. \tag{5}$$

Theorem SM.1 establishes that $\tilde{\lambda}$ is, indeed, feasible for the dual of $\mathrm{MP}_{\mathrm{CONT}}$. The corresponding dual bound $\tilde{Z}$, takes value

$$\tilde{Z} = \bar{Z} - \left(\tilde{\lambda}^{(1b)} - \bar{\lambda}^{(1b)}\right) \cdot |D| - \sum_{s \in \mathcal{S}^{\mathrm{out}}} \left(\tilde{\lambda}_s^{(1c)} - \bar{\lambda}_s^{(1c)}\right) \cdot n_s.$$

Theorem SM.1. *Dual solution $\tilde{\lambda}$ obtained modifying $\bar{\lambda}$ according to* (4) *and* (5)*, and keeping all other components equal, is feasible for the dual of* $\mathrm{MP}_{\mathrm{CONT}}$.

Proof. Violated dual constraints correspond to missing primal variables x and y. There are two families of such constraints:

$$-\lambda^{(1b)} + \sum_{c \in \mathcal{C}_r} \lambda_c^{(1i)} + \sum_{c \in \mathcal{C}_r} \sum_{p \in \mathcal{P}_c} \sum_{\substack{s \in \mathcal{S}_{pc}^{\mathrm{out}} \\ \text{s.t. } r \in \mathcal{R}_{spc}^{\mathrm{D}}}} \lambda_{spc}^{(1g)} \leq c_r \qquad \forall r \in \mathcal{R}^{\mathrm{D}} \tag{6}$$

$$-\lambda_s^{(1c)} + \sum_{c \in \mathcal{C}_r} \lambda_c^{(1j)} + \sum_{c \in \mathcal{C}_r} \sum_{\substack{p \in \mathcal{P}_c \\ \text{s.t. } r \in \mathcal{R}_{spc}^{\mathrm{F}}}} \lambda_{spc}^{(1h)} \leq c_r \qquad \forall s \in \mathcal{S}^{\mathrm{out}},\ \forall r \in \mathcal{R}_s^{\mathrm{F}}. \tag{7}$$

Furthermore, $\lambda^{(1b)}$ only appears in (6) and $\lambda^{(1c)}$ only appears in (7). It follows that we shall prove that $\tilde{\lambda}^{(1b)}$ satisfies (6) and non-negativity constraints and that $\tilde{\lambda}^{(1c)}$ satisfies (7) and non-negativity constraints.

Indeed, (2) and (4) imply that (6) is satisfied. Analogously, (7) follows from (3) and (5). Finally, we remark that the RHS of (2) and (3) are non-positive if negative reduced cost columns are missing from the reduced sets, and thus the non-negativity of $\tilde{\lambda}^{(1b)}$ and $\tilde{\lambda}^{(1c)}$ follows. In case any of the two RHS is strictly positive, then one can always set the corresponding LHS ($\bar{Z}^{\mathrm{D}}$ or $\bar{Z}_s^{\mathrm{F}}$) equal to zero. In this case, the corresponding component of $\tilde{\lambda}$ is equal to the original component in $\bar{\lambda}$, and this component satisfies both the dual constraint and the non-negativity condition. □

Finally, we explain how to obtain lower bounds $\bar{Z}^{\mathrm{D}}$ and $\bar{Z}_s^{\mathrm{F}}$. Because, in practice, we can always solve SP_y to optimality in fractions of a second, we simply use the RHS of (3) as the value of $\bar{Z}_s^{\mathrm{F}}$. We then focus on $\bar{Z}^{\mathrm{D}}$. A straightforward approach to bound the reduced cost of a truck route involves relaxing some of the constraints of SP_x. In particular, one can relax the “hard” in-stop and customer elementarity constraints by applying the state-space relaxation technique (Christofides, Mingozzi and Toth 1981) to the labelling algorithms proposed in Sections 5.1.1 and 5.1.2. In practice, however, such relaxed subproblems are still time-consuming and produce loose bounds. Therefore, we decide to use a dual bound from the MIP model introduced in Section 2. We solve the model with a black-box solver and a short time limit and use the best dual bound returned by the solver.

3.2 Decomposition bound

We solve the GVRPTW introduced in Section 5.3 via branch-price-and-cut, adapting the algorithm presented by Pessoa et al. (2023). We make two modifications to their model. First, we add support for time windows, introducing a time resource and appropriate bounds in the pricing problem. Second, we change the VrpSolver (Pessoa et al. 2020) edge mapping to consider a directed graph and an asymmetric vehicle routing problem. The reason is that once time windows are introduced, the direction in which each route is traversed becomes essential.

4 Generating an initial set of first-tier columns

To initialise the column generation algorithm, we populate $\mathcal{R}^{\mathrm{D}}$ and $\mathcal{R}^{\mathrm{F}}$ with dummy columns which have: (i) a very high cost c_r; (ii) coefficient zero in all inequalities (1b) and (1c); (iii) coefficient one

Algorithm 1 Procedure to generate the initial columns of $\mathcal{R}^{\mathrm{D}}$.

$\bar{\mathcal{C}} \leftarrow \emptyset$ ▷ set of covered customers
$\bar{\mathcal{R}} \leftarrow \emptyset$ ▷ set of routes without start time
Phase 1: greedy creation
while $\bar{\mathcal{C}} \neq \mathcal{C}$ **do**
  $r \leftarrow (o)$ ▷ create an empty route
  $Q_r \leftarrow 0$ ▷ initialise truck used capacity
  $\underline{S} \leftarrow \mathcal{S}^{\mathrm{in}}$ ▷ available in-stops
  while $\underline{S} \neq \emptyset$ and $Q_r \leq Q^{\mathrm{D}}$ **do**
    Let s be the stop in $\underline{S}$ closest to the endpoint of r
    Append s to r
    $\underline{S} \leftarrow \underline{S} \setminus \{s\}$ ▷ update available in-stops
    for $c \in \left(\mathcal{C} \setminus \bar{\mathcal{C}}\right) \cap \mathcal{C}_s$ **do**
      if $Q_r + q_c \leq Q^{\mathrm{D}}$ **then**
        Route r will deliver c at s
        $Q_r \leftarrow Q_r + q_c$ ▷ update truck capacity
        $\bar{\mathcal{C}} \leftarrow \bar{\mathcal{C}} \cup \{c\}$ ▷ update covered customers
  $\bar{\mathcal{R}} \leftarrow \bar{\mathcal{R}} \cup \{r\}$ ▷ Add new route to initial set
Phase 2: greedy augmentation
for $r \in \bar{\mathcal{R}}$ **do**
  for in-stop s visited by r **do**
    for customer c not covered by r **do** ▷ c is covered by another route
      if $Q_r + q_c \leq Q^{\mathrm{D}}$ **then**
        Route r will deliver c at s
        $Q_r \leftarrow Q_r + q_c$
Phase 3: time assignment
$\bar{T} \leftarrow$ {possible start times}
for $r \in \bar{\mathcal{R}}$ **do**
  for start time $t \in \bar{T}$ **do**
    $r_t \leftarrow$ a copy of r with truck start time t
    for customer c covered by r_t **do**
      $s \leftarrow$ in-stop at which r_t delivers c's parcel
      $t' \leftarrow$ time at which c's parcel is ready for bus pick-up at s, according to r_t
      if $t' \notin \Theta_{sc}$ **then** ▷ no bus can pick up c's parcel
        Remove c from route r_t
    $\mathcal{R}^{\mathrm{D}} \leftarrow \mathcal{R}^{\mathrm{D}} \cup \{r_t\}$

Labelling algorithm	Heuristic	Solved instances	Avg. gap
cost-function	No heuristic	46	14.33 %
	PATH	46	13.72 %
	BEST_1	48	15.04 %
	BEST_2	47	14.61 %
	BEST_3	49	15.30 %
scalar-cost	No heuristic	46	14.52 %
	PATH	49	14.56 %
	BEST_1	49	15.38 %
	BEST_2	48	14.88 %
	BEST_3	46	14.27 %

Table 1: Number of instances with a feasible solution (denoted here as "solved") and average solution gaps for different configurations.

in all inequalities (1g), (1h), (1i) and (1j). Any solution of MP which selects a dummy column is infeasible for the 3T-DPPT.

To speed up the convergence of the pricing algorithm, in addition to the dummy column, we populate $\mathcal{R}^{\text{D}}$ with feasible columns generated through Algorithm 1. In the first phase, the algorithm creates a set of truck routes without specifying their start time. It greedily adds new routes until all customers are served. In the second phase, for each route, it tries to fill the truck capacity greedily, adding more parcels to the route. This means that some parcels might be present in more than one route. This, however, is not a problem due to set-covering constraints (1i) and the fact that constraints (1g) ensure that exactly one parcel per client will be loaded onto a bus. In the last phase, each route is copied multiple times, changing its start time. When assigning times, it can happen that a route delivers c's parcel at an in-stop s at time t', but there is no bus which can pick up the parcel at a compatible time, i.e., $t' \notin \Theta_{sc}$, where Θ_{sc} is defined in A. In this case, such a route would lead to a worse continuous relaxation because it would cover row (1i) for customer c, but not row (1g) associated with c. Therefore, in the third phase, we remove from each route all parcels with no compatible bus.

5 Detailed numerical results

Table 1 shows the results obtained by different configurations for the pricing algorithm with respect to the pricing of truck routes. In particular, the table shows the number of instances (over a set of 50) for which a feasible solution is found (column "Solved instances") and the average solution gaps with respect to the best bound known for each instance. For each configuration, the average gap is calculated considering only the set of instances for which the configuration found a feasible solution, thus one should be careful comparing these figures, as each of them is average over a potentially different set. Three configurations tie in the first place, finding feasible solutions for 49 out of the 50 instances. From these, we choose the two achieving the best solution gaps: the BEST_k heuristic with $k = 3$ for the algorithm using cost functions and the PATH heuristic for the algorithm using the scalar costs.

Table 2 shows the solutions obtained by our methods on the instances from Mandal and Archetti 2023, and a comparison of these with the results for the compact formulation (CF) and with the best results obtained by the decomposition heuristics (DH) proposed in Mandal and Archetti 2023.

		Best setting from		Improvement vs		
Ins	CF	DH	This work	CMP	DH	Best
1	2295.02	2295.02	2295.02	0.00 %	0.00 %	0.00 %
2	1460.23	1544.11	1460.23	0.00 %	5.43 %	0.00 %
3	1250.02	1364.22	1250.02	0.00 %	8.37 %	0.00 %
4	2182.70	2300.73	2182.70	0.00 %	5.13 %	0.00 %
5	1826.87	1921.52	1754.81	3.94 %	8.68 %	**3.94 %**
6	1929.18	2201.15	1929.18	0.00 %	12.36 %	0.00 %
7	2876.60	2992.42	2667.57	7.27 %	10.86 %	**7.27 %**
8	3402.97	3720.52	3403.64	-0.02 %	8.52 %	-0.02 %
9	2630.94	2748.94	2473.11	6.00 %	10.03 %	**6.00 %**
10	–	3303.32	3225.65	–	2.35 %	**2.35 %**
11	3294.88	3502.31	3137.74	4.77 %	10.41 %	**4.77 %**
12	4852.02	4540.10	4310.31	11.16 %	5.06 %	**5.06 %**
13	–	3957.66	4203.86	–	-6.22 %	-6.22 %
14	–	3532.50	3424.44	–	3.06 %	**3.06 %**
15	–	4113.17	3930.21	–	4.45 %	**4.45 %**
16	–	6491.98	5831.86	–	10.17 %	**10.17 %**
17	–	5690.91	5130.44	–	9.85 %	**9.85 %**
18	–	4817.65	4631.55	–	3.86 %	**3.86 %**
19	–	8102.66	7536.34	–	6.99 %	**6.99 %**
20	–	6941.22	6239.99	–	10.10 %	**10.10 %**
21	–	6315.74	5589.22	–	11.50 %	**11.50 %**
22	–	7184.62	6122.47	–	14.78 %	**14.78 %**
23	–	6388.33	6009.89	–	5.92 %	**5.92 %**
24	–	8350.63	7903.83	–	5.35 %	**5.35 %**
			# Improves:	18	22	17
			# Deteriorates:	1	1	2
			# Ties:	5	1	5

Table 2: Solutions obtained by our methods on the instances from Mandal and Archetti 2023, compared with the results for the compact formulation (CF) and with the best results obtained by the decomposition heuristics (DH) proposed in the mentioned literature.

6 The compact formulation of (Mandal and Archetti 2023) and our improvement

In the following, we detail the modifications required to forbid trucks from waiting at public transit stops to the formulation presented by Mandal and Archetti (2023). The notation used in that paper is different from the one used in our paper. The reader should refer to Mandal and Archetti's Table 1 for the list of sets and parameters and Table 2 for the list of decision variables. The formulation is presented in (3.1)–(3.35), in Section 1 of (Mandal and Archetti 2023). For the sake of completeness, we recall here that decision variable t^1_{ud} represents the time at which truck $d \in \mathcal{D}$ leaves stop u (or the CDC for $u = 0$) and binary decision variable w_{uvd} equals 1 if truck $d \in \mathcal{D}$ traverses arc (u, v). Also, parameters T^1_{uvd} and T'_v represent the time needed by for truck $d \in \mathcal{D}$ to traverse arc (u, v) and the service time at stop v, respectively. The following constraints are added to the formulation to prevent trucks from waiting at public transit stops:

$$t^1_{vd} \leq t^1_{ud} + T^1_{uvd} + T'_v + M(1 - w_{uvd}), \qquad \forall i, j \in \mathcal{S}^{\text{in}} \cup \{0\},\ i \neq j,\ j \neq 0,\ \forall d \in \mathcal{D} \tag{8}$$

where $M > 0$ is a sufficiently large number.

7 On stop capacity in (Mandal and Archetti 2023) and in our formulation

In the 3T-DPPT formulation introduced by Mandal and Archetti (2023), the storage capacity at in- and out-stops is unlimited. To some extent, storing too many parcels at stops is discouraged by the parameter $W^{\max}$ (i.e., the maximum time for a package to be at a stop). Hard storage capacities, however, are an important aspect of the freight-on-transit (Delle Donne et al. 2023) system, suggesting that a 3T-DPPT formulation should take them into account. The following theorem shows that the compact formulation cannot model explicit stop capacities without introducing extra variables.

Theorem SM.2. *The compact formulation (CF) cannot be used to model storage capacities at in-stops with linear constraints unless extra variables are included in the model.*

Proof. Formulation CF uses several sets of variables, associated with the first, second and third tiers of 3T-DPPT. For the sake of this proof, we give a brief description of those associated to the first tier: binary variable r_{csd} equals 1 iff the package for customer c is delivered by truck d to the drop-in stop s; binary variable w_{uvd} equals 1 iff truck d traverses arc (u, v); continuous variable t_{ud} indicates the time when truck d finishes its drops at (or start from) node $u \in \mathcal{S}^{\text{in}} \cup \{o\}$, where o is the CDC; and binary variable y_{csp} equals 1 iff the package for customer c is picked up by public vehicle p from drop-in stop s.

Consider an instance of 3T-DPPT with 3 trucks $\mathcal{D} = \{d_1, d_2, d_3\}$, 3 customers $\mathcal{C} = \{c_1, c_2, c_3\}$ with parcel sizes $q_{c_i} = 1$ ($i \in \{1, 2, 3\}$), and a single drop-in stop $\mathcal{S}^{\text{in}} = \{s\}$. Assume that the storage capacity of stop s is $Q_s = 2$, hence any 2 parcels can be stored at s simultaneously, but the 3 parcels cannot. Consider a feasible solution S_1 for CF in which parcels c_1, c_2 and c_3 are delivered at s by trucks d_1, d_2 and d_3 at times 1, 2 and 8 and picked up at times 6, 9 and 9. Figure 4 illustrates this scenario on a timeline. This solution does not violate the storage capacity of s, as the 3 parcels are never simultaneously at the stop. Consider now another feasible solution S_2 almost equal to S_1 but in which the delivery times of parcels c_2 and c_3 are switched (see Figure 4). Note that S_1 and S_2 are encoded identically into the model variables except for the continuous variables representing:

- The starting times of the trucks d_2 and d_3, i.e., t_{od_2} and t_{od_3}.
- And their departure times from stop s, i.e., t_{sd_2} and t_{sd_3}, after delivering parcels c_2 and c_3.

Table 3 shows the values associated with these variables, assuming a travel time from o to s of 1, including the service time. Consider now the vectors x^1 and x^2 representing solutions S_1 and S_2,

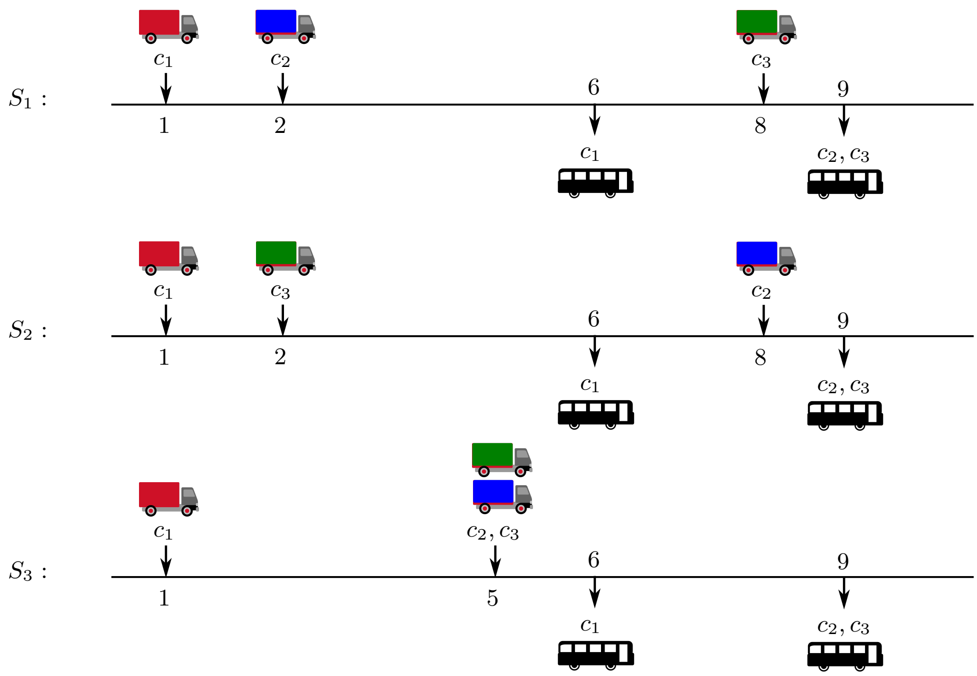

Figure 4: Example proving that it is not possible to model stop capacities in CF with linear constraints and without adding extra variables.

	t_{od_2}	t_{od_3}	t_{sd_2}	t_{sd_3}
S_1	1	7	2	8
S_2	7	1	8	2
S_3	4	4	5	5

Table 3: Values of the continuous variables of example solutions S_1, S_2 and S_3.

respectively. Let $x^3 = (x^1 + x^2)/2$ be their midpoint. This point is identical to x^1 and x^2, except for variables t_{od_2}, t_{od_3}, t_{sd_2} and t_{sd_3}. In particular, x^3 satisfies all integrality constraints because the associated variables are identical to those of x^1 and x^2. Being a convex combination of feasible solutions, x^3 also satisfies all linear constraints of CF, and therefore it is a feasible solution. Let S_3 be the solution represented by x^3. In this new solution both parcels c_1 and c_2 are delivered at s at time $(2+8)/2 = 5$ (see Table 3 and Figure 4). Hence, x^3 violates the storage capacity of s because all three parcels are in s from time 5 to 6. To summarise, we produced an infeasible solution x^3 that, being a convex combination of two feasible solutions and satisfying all integrality constraints, cannot be eliminated by means of linear inequalities. $\square$

We conclude this section by noting that stop capacities can be easily modelled in our extended formulation. Since the storage load at an in-stop only decreases when a bus picks some parcels up, we need only to check that capacities are not exceeded on these particular moments, i.e., on bus arrival times. To this end, for each stop $s \in \mathcal{S}^{\text{in}}$ and each bus-arrival time t_p^s, with $p \in \mathcal{P}_s$, we can calculate the set of parcels stored at s at time t by considering those delivered at s by all routes arriving before time t_p^s minus those already picked up from s by previous buses. These constraints have almost no impact on the pricing subproblem because they only add a dual cost for delivering at a stop at a given time. Storage capacities at out-stops can be modelled analogously.

References

Christofides, Nicos, Aristide Mingozzi and Paolo Toth (1981). 'State-space relaxation procedures for the computation of bounds to routing problems'. In: *Networks* 11.2, pp. 145–164. DOI: `10.1002/net.3230110207`.

Delle Donne, Diego et al. (2023). 'Freight-on-Transit for urban last-mile deliveries: A strategic planning approach'. In: *Transportation Research Part B: Methodological* 169, pp. 53–81. DOI: `10.1016/j.trb.2023.01.004`.

Mandal, Minakshi Punam and Claudia Archetti (2023). *A Decomposition Approach to Last Mile Delivery Using Public Transportation Systems.* arXiv: `2306.04219 [math.OC]`.

Pessoa, Artur et al. (2020). 'A generic exact solver for vehicle routing and related problems'. In: *Mathematical Programming* 183.1, pp. 483–523. DOI: `10.1007/s10107-020-01523-z`.

Pessoa, Artur et al. (2023). 'A unified exact approach for Clustered and Generalized Vehicle Routing Problems'. In: *Computers & Operations Research* 149. DOI: `10.1016/j.cor.2022.106040`.

Tagmouti, Mariam, Michel Gendreau and Jean-Yves Potvin (2007). 'Arc routing problems with time-dependent service costs'. In: *European Journal of Operational Research* 181.1, pp. 30–39.