



# Complexity, Algorithmic, and Computational Aspects of a Dial-a-Ride Type Problem

Mourad Baïou, Rafael Colares, Hervé Kerivin

## ► To cite this version:

Mourad Baïou, Rafael Colares, Hervé Kerivin. Complexity, Algorithmic, and Computational Aspects of a Dial-a-Ride Type Problem. 2022. hal-03578358

**HAL Id: hal-03578358**

**<https://hal.science/hal-03578358>**

Preprint submitted on 17 Feb 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Complexity, Algorithmic, and Computational Aspects of a Dial-a-Ride Type Problem

Mourad Baïou<sup>a</sup>, Rafael Colares<sup>a,\*</sup>, Hervé Kerivin<sup>a,b</sup>

<sup>a</sup>*LIMOS, CNRS UMR 6158, Université Clermont-Auvergne, Complexe Scientifique des Cézeaux, Aubière, 63177, France*

<sup>b</sup>*School of Mathematical and Statistical Sciences, Clemson University, Clemson, 29634, SC, USA*

---

## Abstract

In dial-a-ride systems involving autonomous vehicles circulating along a circuit, a fleet of vehicles must serve clients who request rides between stations of the circuit such that the total number of pick-up and drop-off operations is minimized. In this paper, we focus on a unitary variant where each client requests a single place in the vehicles and all the clients must be served within a single tour of the circuit. Such unitary variant induces a combinatorial optimization problem for which we introduce a nontrivial special case that is polynomially solvable when the capacity of each vehicle is at most 2 but it is NP-Hard otherwise. We also study the polytope associated with the solutions to this problem. We introduce new families of valid inequalities and give necessary and sufficient conditions under which they are facet-defining. Based on these inequalities, we devise an efficient branch-and-cut algorithm that outperforms the state-of-the-art commercial solvers.

*Keywords:* Combinatorial optimization, autonomous vehicles, dial-a-ride, computational complexity, polyhedral study

---

## 1. Introduction

It is undeniable that the demand for faster, cheaper, and more convenient forms of mobility has unceasingly evolved throughout the history of human civilization. Transport systems have also developed along. Today

---

\*Corresponding author

*Email addresses:* `baïou@isima.fr` (Mourad Baïou), `rafael.colares@gmail.com` (Rafael Colares), `kerivin@isima.fr` (Hervé Kerivin)

more than ever, mobility represents an important issue when dealing with sustainability challenges. Autonomous vehicles and electric vehicles have been key actors in transforming the transportation sector and reshaping mobility, through new types of business and service models for the former and more environment-friendly alternatives for the latter. Several studies can be found in the literature that link electric autonomous vehicles to improvements of air quality, traffic mitigation, and safety. (See, e.g., Fagnant and Kockelman (2015) and Pelletier et al. (2016).)

One of the many issues and challenges of deploying electric autonomous vehicles is the management of fleets of such vehicles when operating in semi-closed or closed sites such as medical complexes, commercial or industrial areas, airports, train stations, and campuses. VIPA, a French acronym standing for Automomous Individual Public Vehicle, is an example of an electric driverless vehicle designed to transport goods or people in such environments. (See, e.g., EasyMile (2015) whose EZ10 shuttles have been operating in a dozen US cities, on a trial basis, for over a year.) VIPA shuttles may perform in three different modes of operation (Bsaybes et al., 2018). In this paper we only focus on the so-called tram mode where same-capacity VIPA shuttles continuously go around a circuit, always in the same direction, and upon requests stop at predefined stations to pick up and drop off users. Each of these dial-a-ride requests, later on called *demands*, is specified by a pick-up station, a drop-off station, and a load that corresponds to the number of places the user will be needing in a VIPA shuttle.

Due to infrastructure restrictions, the stations are not located on the circuit but are attached to it as illustrated in Figure 1. This peculiarity significantly impacts the management of the fleet of vehicles. Indeed, a VIPA shuttle responding to a request must slow down and deviate from its original course. This necessary detour increases the travel time of on-board users as well as the shuttle’s battery consumption, a key resource for electrical vehicles. If the detour lengths are supposed to be approximately the same, the quality of service can then be improved by minimizing the total number of stops the VIPA shuttles have to perform. Pimenta et al. (2017) pointed out that a reduction of the total number of stops also ensures a steady flow of vehicles which helps improve the reliability of such VIPA-based systems. The *Stop Number Problem (SNP)* we consider in this paper hence consists of assigning the demands to the shuttles of a fleet such that no shuttle ever is overloaded and the total number of stops is minimized.

In practice, a VIPA-based system obviously needs to dynamically manage an ongoing flow of demands. This is usually handled through online algorithms. However an analysis of the solutions and difficulties one would

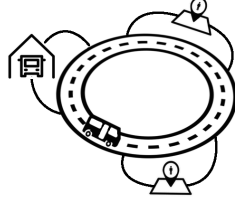


Figure 1: An illustration of the VIPA circuit scheme.

encounter in an offline management (i.e., when all the demands are known beforehand) may help gain important insights into the development of better-suited online algorithms. A reasonable option therefore is to deeply investigate the offline version of SNP and thus determine where its complexity lies. In this paper, we explore the combinatorial core of SNP through the simplest version of the problem where (i) as many same-capacity vehicles as necessary are available to serve all the demands, (ii) each user only requests a single place (i.e., unit load), and (iii) a vehicle may only serve a demand during its first tour of the circuit (i.e., no waiting tours). This version of SNP, hereafter denoted *Unit Stop Number Problem (USNP)*, is formally defined as follows.

### 1.1. The Unit Stop Number Problem

We first go over some notation we use throughout this paper. Let the edge set (resp. node set) of an undirected graph  $G$  be denoted by  $E(G)$  (resp.  $V(G)$ ) or simply  $E$  (resp.  $V$ ) when  $G$  is clear from the context. For  $F \subseteq E$ , let  $V(F)$  denote the set of nodes spanned by the edges in  $F$ . For  $v \in V$ , the degree of node  $v$  in  $G$  is denoted  $\deg_G(v)$ . Given a subset of edges  $F \subseteq E$ , the undirected graph  $(V(F), F)$  induced by  $F$  is denoted  $G(F)$ .

Let  $V = \{1, \dots, n\}$  be the set of predefined stations as they appear along the circuit network. Even though all the vehicles leave from a same depot, we assume without loss of generality that the depot does not belong to  $V$ . Let  $E$  be the set of  $m$  unit-load demands, where each demand  $e$  is specified by a pick-up station  $o_e \in V$  and a drop-off station  $d_e \in V$ . Without loss of generality, we assume that each station of  $V$  appears as the endpoint of at least one demand of  $E$ . We also assume that  $o_e < d_e$  for any demand  $e = (o_e, d_e)$  in  $E$ . Notice that if this is not the case, one can provide an equivalent instance by setting  $V = \{1, \dots, 2n\}$  and replacing demands  $(o_e, d_e)$  by  $(o_e, d_e + n)$  for any  $e$  in  $E$  where  $o_e > d_e$ . To serve these  $m$  demands, we are given a fleet  $K$  of  $p$  identical vehicles, each of them having the same capacity  $C \in \mathbb{Z}_+$ . Since the decision maker can use

as many vehicles as needed to reduce the total number of stops, the number  $p$  of available vehicles is set to the trivial upper bound  $m$  unless otherwise specified. With any USNP instance  $\mathcal{I} = (V, E, C)$  a graph  $G_{\mathcal{I}} = (V, E)$  is associated where stations and demands may be referred to as nodes and edges, respectively. When instance  $\mathcal{I}$  is clear from the context, we may omit the subscript  $\mathcal{I}$  and only write  $G$ . Figure 2 provides an example of an USNP instance and its associated graph.

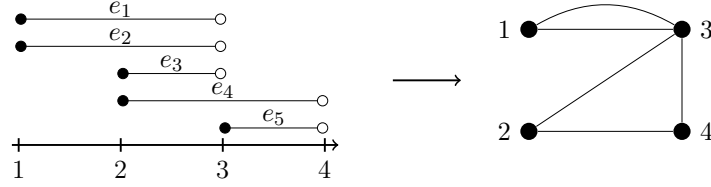


Figure 2: Example of an USNP instance. On the left, the set of five demands is represented as intervals over four stations. On the right, the instance is defined by its associated graph.

Demands having the same pick-up and drop-off stations (e.g.,  $e_1$  and  $e_2$  in Figure 2) are referred to as *parallel demands* and correspond to multiple edges in  $G_{\mathcal{I}}$ . Without loss of generality, we suppose that  $G_{\mathcal{I}}$  is a (loopless) connected multigraph for otherwise solving USNP on  $G_{\mathcal{I}}$  would reduce to solving as many independent USNPs as  $G_{\mathcal{I}}$  has connected components.

For any  $F \subseteq E$  and any  $v \in V$ , let  $\Delta_F(v) = \{e \in F : o_e \leq v < d_e\}$  denote the set of demands of  $F$  that need to go through station  $v$ . Notice that demands having  $v$  as their drop-off station do not belong to  $\Delta_F(v)$ . The demands belonging to  $\Delta_E(v)$  are said to *intersect* station  $v$ . Every station  $v$  whose set  $\Delta_E(v)$  is inclusionwise maximal (i.e., there is no  $v'$  in  $V$  such that  $\Delta_E(v) \subset \Delta_E(v')$ ) is referred to as a *maximal-intersection station*. For the example depicted in Figure 2,  $\Delta_E(1) = \{e_1, e_2\}$ ,  $\Delta_E(2) = \{e_1, e_2, e_3, e_4\}$ ,  $\Delta_E(3) = \{e_4, e_5\}$ , and  $\Delta_E(4) = \emptyset$ , and both stations 2 and 3 are maximal-intersection.

Any feasible solution to USNP can hence be represented as a partition of  $E$  into  $p$  (possibly empty) subsets  $E_1, \dots, E_p$  that satisfies  $|\Delta_{E_i}(v)| \leq C$  for all  $i = 1, \dots, p$  and all  $v \in V$ . USNP thus consists of finding a partition  $\{E_1, \dots, E_p\}$  that minimizes the cost function

$$c(E_1, \dots, E_p) = \sum_{i=1}^p |V(E_i)|,$$

where  $V(E_i)$  is composed of all the stations where vehicle  $i$  stops.

## 1.2. Literature review

SNP was first introduced in Pimenta et al. (2017) where it was proven to be (weakly) NP-Hard through a simple reduction from the classic partition problem (Garey and Johnson (2002)). The authors presented a GRASP-based heuristic for SNP along with a branch-and-price method that turned out to be inefficient due to a lack of strong bounds. USNP was also mentioned in Pimenta et al. (2017) where it was conjectured to be NP-hard. This conjecture was proven by Baïou et al. (2021) where USNP is showed to be NP-hard for any fixed capacity  $C \geq 2$  even when the graph associated with the instance is planar bipartite.

USNP clearly corresponds to a specific case of the well-studied Dial-A-Ride Problem (DARP). Indeed DARP arises from door-to-door transportation services, particularly for elderly and disabled people, and its objective is twofold. It not only consists of assigning demands to vehicles but also of designing minimum-cost vehicle routes to determine a picking-up and delivery order that fulfills client-driven requirements (e.g., time windows, maximum riding time, and vehicle occupancy).

From an exact-method standpoint, a usual strategy for dealing with DARP is to use branch-and-cut algorithms. The first branch-and-cut algorithm for DARP was proposed by Cordeau (2006) who introduced a three-index formulation and derived families of valid inequalities from well-known inequalities valid for the Vehicle Routing Problem and the Traveling Salesman Problem. Ropke et al. (2007) proposed a tighter two-index formulation and also introduced new families of valid inequalities based on the idea of incompatible time windows. Parragh (2011) and Braekers et al. (2014) adapted known branch-and-cut algorithms to deal with heterogeneous vehicles. Liu et al. (2015) proposed new valid inequalities for taking into account problem-specific properties such as driver lunch breaks. For more extensive and detailed surveys on DARP, the reader is referred to Cordeau and Laporte (2007), Parragh et al. (2008) and Ho et al. (2018).

A common point in the DARP literature is the use of valid inequalities to tackle scheduling issues such as precedence requirements, time-windows compatibility, or subtour prevention. However these specific aspects of DARP appear irrelevant to USNP where they are implicitly taken into account since the vehicles must serve all the demands within a single tour of a predefined circuit. These inequalities, necessary for DARP, therefore become of no interest when dealing with USNP.

It is also worth noticing that the combinatorics behind USNP is closely related to other well-studied telecommunication problems such as the traffic

grooming problem as pointed out by Baïou et al. (2021) and the  $k$ -edge partitioning problem. For these relations require a deeper understanding of the solutions to USNP, they are postponed to Section 2.

### 1.3. Our contribution

In the next section, the complexity of USNP is investigated and connections with other well-studied problems are drawn. We present special cases where the problem can be solved in polynomial time as well as cases where USNP is proven to be NP-Hard. In Section 3 a polyhedral analysis is conducted and yields new facet-defining inequalities that strengthen an integer-programming formulation known in the literature. Section 4 is devoted to a branch-and-cut framework with a focus on symmetry and variable management, and separation problems. Computational results demonstrating the efficiency of our branch-and-cut approach are provided at the end of that section. In Section 5, we conclude the paper pointing out interesting directions of research. Notice that many results presented along this study can be extended to other related problems. These extensions are exposed when necessary. The complete proofs of theorems are provided in the Appendix.

## 2. Complexity analysis

USNP is trivially solvable in polynomial time if  $C = 1$ . Baïou et al. (2021) have proven that USNP becomes NP-hard for any fixed capacity  $C \geq 2$  even when the associated graph  $G$  is planar bipartite. In this section we focus on a variant of USNP where all the demands intersect some station  $v' \in V$ , that is,  $\Delta_E(v') = E$ . Let *Intersection-USNP* (*I-USNP*) denote this variant. Notice that in I-USNP, each vehicle overall serves at most  $C$  demands. This last fact will be deeply explored throughout this paper and is the main argument behind both the development of our complexity results and our polyhedral study. The study of I-USNP is therefore crucial since it helps to better understand and solve the more general USNP.

**Remark 1.** *For every I-USNP instance  $\mathcal{I} = (V, E, C)$ , the set of stations  $V$  can be partitioned into  $U = \{1, \dots, v'\}$  and  $V \setminus U = \{v' + 1, \dots, n\}$ , where  $v' \in V$  is the unique station with  $\Delta_E(v') = E$ . In this case every demand in  $E$  has its pick-up station in  $U$  and its drop-off station in  $V \setminus U$ , and the associated graph  $G_{\mathcal{I}} = (V, E)$  hence is bipartite.*

We now show that I-USNP can be solved in  $O(m)$  time for  $C \leq 2$ . Surprisingly, for  $C \geq 3$ , even this very restrictive variant remains NP-Hard.

### 2.1. Polynomial cases

For every  $v \in V$ , let  $\delta^+(v) = \{e \in E : o_e = v\}$  (resp.  $\delta^-(v) = \{e \in E : d_e = v\}$ ) denote the set of demands having  $v$  as their pick-up (resp. drop-off) station. It can therefore be easily seen that at least  $\left\lceil \frac{\max\{|\delta^+(v)|, |\delta^-(v)|\}}{C} \right\rceil$  vehicles need to stop at station  $v$ . Consequently a lower bound on the number of stops is

$$\sum_{v \in V} \left\lceil \frac{\max\{|\delta^+(v)|, |\delta^-(v)|\}}{C} \right\rceil. \quad (1)$$

Notice that for I-USNP, either  $\delta^+(v) = \emptyset$  or  $\delta^-(v) = \emptyset$  for every station  $v \in V$ . Moreover, each vehicle serves no more than  $C$  demands in I-USNP. It then follows that when  $C = 1$ , sequentially assigning a demand to an available vehicle solves I-USNP. The number of stops of this trivial solution equals  $2m$  and by (1), it is optimal. The next theorem states that for  $C = 2$ , I-USNP can be solved in polynomial time as opposed to general USNP.

**Theorem 1.** *Intersection-USNP can be solved in  $O(m)$  time when  $C = 2$ .*

*Proof.* Consider an instance  $\mathcal{I} = (V, E, 2)$  of I-USNP. Suppose  $E$  has two parallel demands  $e$  and  $e'$ . We claim that there exists an optimal solution wherein  $e$  and  $e'$  are served by the same vehicle. Assume on the contrary that  $e$  and  $e'$  are served by different vehicles in every optimal solution. Let  $\{E_1^*, \dots, E_p^*\}$  be an optimal solution with, without loss of generality,  $e \in E_1^*$  and  $e' \in E_2^*$ . We trivially have  $|E_1^*| = 2$  for otherwise  $\{E_1^* \cup \{e'\}, E_2^* \setminus \{e'\}, E_3^*, \dots, E_p^*\}$  would be a feasible solution with no more stops than  $\{E_1^*, \dots, E_p^*\}$ . Similarly  $|E_2^*| = 2$ . Let  $\{f\} = E_1^* \setminus \{e\}$  and  $\{f'\} = E_2^* \setminus \{e'\}$ . Consider now the feasible solution  $\{\{e, e'\}, \{f, f'\}, E_3^*, \dots, E_p^*\}$  and let  $\alpha = |V(\{f, f'\})|$ . We have

$$c(\{e, e'\}, \{f, f'\}, E_3^*, \dots, E_p^*) = c(E_1^*, \dots, E_p^*) - (|V(E_1^*)| + |V(E_2^*)|) + 2 + \alpha. \quad (2)$$

If  $e$  and  $f$  or  $e'$  and  $f'$  are parallel demands, we then have  $|V(E_1^*)| + |V(E_2^*)| = 2 + \alpha$  and  $\{\{e, e'\}, \{f, f'\}, E_3^*, \dots, E_p^*\}$  is optimal by (2). If none of the sets  $E_1^*$  and  $E_2^*$  is composed of parallel demands, we have  $|V(E_1^*)| + |V(E_2^*)| \geq 6 \geq 2 + \alpha$ , the last inequality coming from  $\alpha \leq 4$ . Solution  $\{\{e, e'\}, \{f, f'\}, E_3^*, \dots, E_p^*\}$  hence is optimal. Consequently we may assume that in the optimal solution  $\{E_1^*, \dots, E_p^*\}$ :

- i. the first  $p' \leq p$  vehicles serve pairs of parallel demands, that is, subsets  $E_i^*$  are composed of parallel demands for  $1 \leq i \leq p'$ , and
- ii.  $\bigcup_{i=p'+1}^p E_i^*$  contains no parallel demands.



Packing together pairs of parallel demands (i.e., computing  $E_i^*$  for  $1 \leq i \leq p'$ ) can be done in linear time. So we now consider the I-USNP instance  $\mathcal{I}' = (V, E', 2)$ , where  $E' = E \setminus \bigcup_{i=1}^{p'} E_i^*$  contains no parallel demands. We now show that an optimal solution  $\{E_1'^*, \dots, E_p'^*\}$  to  $\mathcal{I}'$  can be constructed in  $O(m)$  time which results in  $\mathcal{I}$  also being solved in  $O(m)$  time.

We assume, w.l.o.g., that for every vehicle  $i$ ,  $G(E_i'^*)$  is a connected graph, for otherwise the two demands assigned to vehicle  $i$  could be split between two vehicles and the obtained solution would be as good as  $\{E_1'^*, \dots, E_p'^*\}$ . Therefore, solving  $\mathcal{I}'$  amounts to maximizing the number of subsets  $E_i'^*$  containing two connected demands. To do so, consider the line-graph of  $G_{\mathcal{I}'}$ , where every node represents a demand and two nodes are adjacent if their associated demands share an endpoint. A matching  $M$  in such line-graph provides a feasible solution to  $\mathcal{I}'$ . Indeed, every edge in  $M$  corresponds to the pair of demands served by a vehicle. The remaining demands (i.e., the ones associated with the nodes incident to no edges in  $M$ ) can be each assigned to different vehicles. Consequently solving the maximum-cardinality matching problem in the line-graph of  $G_{\mathcal{I}'}$  clearly solves  $\mathcal{I}'$ .

Notice that every edge in the line-graph of  $G_{\mathcal{I}'}$  corresponds to a 2-path (i.e., a simple path of length 2) in  $G_{\mathcal{I}'}$ . The maximum-cardinality matching problem in the line-graph of  $G_{\mathcal{I}'}$  thus is equivalent to the maximum edge-disjoint 2-path packing problem in  $G_{\mathcal{I}'}$ . Masuyama and Ibaraki (1991) gave an algorithm, that runs in  $O(m)$  time, to solve the latter problem.  $\square$

## 2.2. NP-hardness

After having proven that I-USNP can be solved in polynomial time for  $C \leq 2$ , the question of whether or not I-USNP is NP-Hard for larger capacities arises naturally. We first point out that I-USNP can be solved in polynomial time for  $C \leq 2$  mainly because of two properties its optimal solutions satisfy. The first one is related to the number of nonempty vehicles used in an optimal solution.

Let  $p_{\min} \leq p$  denote the minimum number of vehicles necessary to serve the  $m$  demands of a USNP instance and let  $p_{\text{opt}} \geq p_{\min}$  denote the minimum number of nonempty vehicles used by an optimal solution. For the demands can be viewed as (half closed - half open) intervals of the real line,  $p_{\min}$  clearly can be computed in polynomial time by a first-fit algorithm where the demands are picked out according to the order of their increasing pick-up station. More precisely, we have  $p_{\min} = \max_{v \in V} \left\lceil \frac{|\Delta_E(v)|}{C} \right\rceil$ .

**Property 1.** *For  $C \leq 2$ , there exists an optimal solution of Intersection-USNP whose number of nonempty vehicles equals the minimum number of necessary vehicles (i.e.,  $p_{opt} = p_{min}$ ).*

Indeed, for  $C = 2$ , whenever two vehicles only serve one demand each, those two demands can be merged into a single vehicle and the obtained solution is as good as the initial one. Consequently the optimal solution obtained through the procedure described in the proof of Theorem 1 can easily be tweaked to only use  $p_{min}$  vehicles. Therefore  $p_{opt} = p_{min}$  for I-USNP when  $C = 2$ .

Based on empirical observation, Pimenta et al. (2017) claimed that Property 1 holds for USNP whatever the value of  $C$  is. This claim was disproven by Baïou et al. (2021) who gave a counter-example for which  $p_{opt} > p_{min}$ .

For I-USNP with  $C \geq 3$ , we now show that this property may not always be satisfied. To do so, we consider the instance of I-USNP with  $C = 3$  depicted in Figure 3. This instance is composed of 3 sets of 3 parallel demands, all being picked up at station 1, and 3 pairs of demands  $\delta^+(2)$ ,  $\delta^+(3)$ , and  $\delta^+(4)$ . Notice that  $v' = 4$ . By assigning each of the 6 foregoing sets to different vehicles we obtain a feasible solution using 6 vehicles and yielding 15 stops. Since the lower bound (1) on the number of stops equals 15, this solution is optimal.

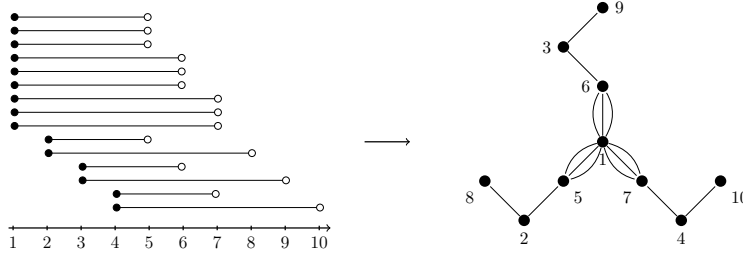


Figure 3: Counter-example I-USNP instance for proving  $p_{min} \neq p_{opt}$ .

The minimum number of necessary vehicles clearly is  $p_{min} = \left\lceil \frac{|\Delta(v')|}{C} \right\rceil = \left\lceil \frac{15}{3} \right\rceil = 5$ . We claim that no optimal solution using  $p_{min}$  vehicles exists. In any solution with 15 stops, exactly  $\left\lceil \frac{\max\{|\delta^-(v)|, |\delta^+(v)|\}}{C} \right\rceil$  vehicles must stop at every station  $v \in V$ , that is, one vehicle stops at each of the stations 2, 3, and 4, two vehicles at each of the stations 5, 6, and 7, and three vehicles at station 1. Moreover if only five vehicles are used then every vehicle needs to serve three demands. For  $i \in \{2, 3, 4\}$ , both demands in  $\delta^+(i)$  together with one of the parallel demands hence need to be assigned to the same

vehicle. We thus already have three vehicles that stop at station 1 but six of the parallel demands have not been assigned yet. Consequently at least two more vehicles need to stop at station 1. The number of vehicles stopping at station 1 therefore exceeds the three vehicles required in any optimal solution. So, no optimal solution with 5 vehicles exists.

A second property of optimal solutions to I-USNP, that seems quite intuitive, is about parallel demands being assigned to the same vehicle providing no overloading happens.

**Property 2.** *For any pair of parallel demands, there exists an optimal solution to Intersection-USNP,  $C = 2$ , where these parallel demands are served by the same vehicle.*

The optimal solution to I-USNP with  $C = 2$  described in the proof of Theorem 1 relies on Property 2. However for larger values of  $C$ , this second property does not have to be satisfied by any optimal solution to I-USNP as illustrated by the following counter-example. Consider the instance of I-USNP with  $C = 4$  described in Figure 4. This instance contains two parallel demands, namely  $d$  and  $e$ . The solution  $\{E_1, E_2\}$  with  $E_1 = \{a, b, c, d\}$  and  $E_2 = \{e, f, g, h\}$  is feasible to this instance and yields 8 stops. Notice that parallel demands  $d$  and  $e$  are not assigned to the same vehicle. If these parallel demands were assigned to the same vehicle then no feasible solution with less than 9 stops would exist. In fact, every subset of 3 or 4 demands of  $\{a, b, c, f, g, h\}$  assigned to the same vehicle would result in at least 4 stops for that vehicle, and every subset of 2 demands in at least 3 stops. Overall, the total number of stops would be at least 9.

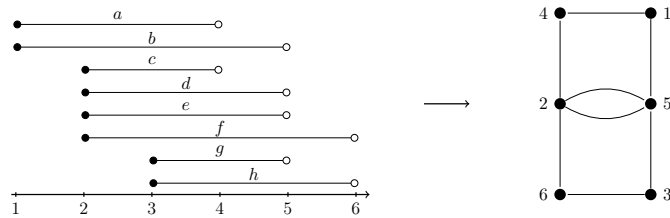


Figure 4: Counter-example I-USNP instance for proving that parallel demands might need to be served separately.

The fact that Properties 1 and 2 do not hold for higher capacities starts to indicate that solving I-USNP might not be that easy. Before proving that I-USNP with  $C \geq 3$  is indeed NP-Hard, we draw a tight connection between I-USNP and the well-known *k-Edge-Partitioning Problem (k-EPP)*.

In the literature,  $k$ -EPP appears as a uniform version of the intra-ring synchronous optical network design problem (Goldschmidt et al., 2003). This latter problem is a network design problem arising from the deployment of synchronous optical networks (i.e., a North-American standardized communication protocol, known as SONET, used to transmit a large amount of data over relatively large distances using optical fibers). In SONET networks, a set of telecommunication centers are linked in a circular way by a cable composed of optic fibers, called rings, each having a fixed capacity. Given a set of demands, each specified by two distinct centers and a requested bandwidth, the intra-ring synchronous optical network design problem consists of assigning the demands to the rings such that the total bandwidth assigned to a ring does not exceed its capacity. For each ring, an electronic termination called Add-Drop Multiplexer (ADM) must be placed at each center being the endpoint of at least one demand assigned to this ring. Notice that several demands assigned to the same ring and incident to the same center may use the same ADM. Due to the high cost of an ADM and the insignificant cost of using or opening a ring, the intra-ring synchronous optical network design problem aims at minimizing the number of installed ADMs.

When all the requested bandwidths are equal and each ring can serve up to  $k$  demands, the intra-ring synchronous optical network design problem is known as  $k$ -EPP. This problem is formally defined as follows. Given a simple graph  $G = (V, E)$ , where  $V$  represents the set of  $n$  centers and  $E$  the set of  $m$  demands (no parallel edges are allowed) and an integer  $k \geq 1$ , find a partition of  $E$  into  $R$  subsets  $E_1, \dots, E_R$  such that  $|E_i| \leq k$  for  $i \in \{1, \dots, R\}$  and  $\sum_{i=1}^R |V(E_i)|$  is minimized. Notice that if  $G$  is bipartite then  $k$ -EPP is equivalent to I-USNP with no parallel demands.

Goldschmidt et al. (2003) proved that  $k$ -EPP is NP-Hard for general graphs and any fixed  $k \geq 3$ . Unfortunately, their proof cannot be directly applied to bipartite graphs. A characterization of the complexity of I-USNP cannot thus be deduced from the one of  $k$ -EPP. We next prove that I-USNP is indeed NP-Hard even when  $G$  is planar bipartite. Therefore, the complexity result of  $k$ -EPP may be extended to the more restricted class of planar bipartite graphs.

To this end, we use a reduction from the *3-Dimensional Matching Problem* (3DMP) which can be stated as follows. Given three disjoint sets  $X, Y, Z$  with  $|X| = |Y| = |Z| = q$  and a set  $T \subseteq X \times Y \times Z$  of triples, 3DMP problem consists of deciding whether or not there exists a 3-dimensional matching, that is, a subset  $M \subseteq T$  such that  $|M| = q$  and every element of  $S = X \cup Y \cup Z$  appears in exactly one triple of  $M$ . 3DMP is a well-known NP-hard problem (Karp, 1972). In Dyer and Frieze (1986), 3DMP

was showed to be NP-hard even when the bipartite graph  $G = (T, S, E)$ , where  $E = \bigcup_{t=(x,y,z) \in T} \{(t, x), (t, y), (t, z)\}$ , is restricted to be planar. An illustration of graph  $G$  is given in Figure 5.

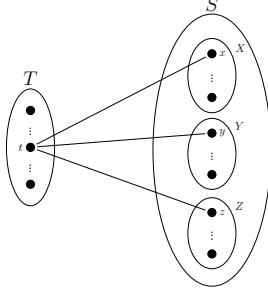


Figure 5: The bipartite graph  $G = (T, S, E)$  associated with an instance of 3DMP.

**Theorem 2.** *Intersection-USNP is NP-Hard for any fixed odd capacity  $C \geq 3$  even when  $G$  is restricted to be planar bipartite.*

*Proof.* Consider a 3DMP instance as described above such that bipartite graph  $G$  is planar. Let  $k \geq 1$ . We first transform  $G$  into a graph  $G' = (V', E')$  associated with an instance  $\mathcal{I} = (\mathcal{V}', \mathcal{E}', \in || + \infty)$  of I-USNP as follows. For every node  $v \in Y \cup Z$ , we replace edge  $(t, v)$  in  $E$  with a path  $P_v^t$  of length  $k$  and add  $\deg_G(v) - 1$  node-disjoint paths  $P_v^1, \dots, P_v^{\deg_G(v)-1}$  of length  $k + 1$ , all having  $v$  as an endnode. For every node  $x \in X$ , we add  $\deg_G(x) - 1$  node-disjoint paths  $P_x^1, \dots, P_x^{\deg_G(x)-1}$  of length  $2k$ , all having  $x$  as an endnode. The number of nodes in  $G'$  clearly equals  $(6k + 1)|T| - (4k - 1)q$  and its number of edges  $(2k + 1)(3|T| - 2q)$ . Notice that since  $G$  is planar bipartite, so is  $G'$ . See Figure 6 for an illustration of graph  $G'$ .

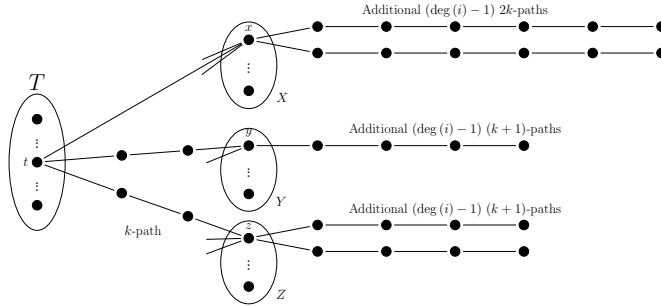


Figure 6: Final construction of graph  $G' = (V', E')$ .

To obtain instance  $\mathcal{I}$  of I-USNP, the stations in  $V'$  now need to be labeled such that there exists a station  $v'$  where all demands intersect. From  $G'$  being bipartite,  $V'$  can be partitioned into independent sets  $W$  and  $\overline{W}$ . A feasible labeling of  $V'$  thus consists of first mapping the nodes in  $W$  onto  $\{1, \dots, |W|\}$  and then those in  $\overline{W}$  onto  $\{|W| + 1, \dots, |V'|\}$ . Notice that  $v'$  is the node in  $W$  with the largest label. The construction of  $\mathcal{I}$  can clearly be done in polynomial time. We claim that this 3DMP instance has a 3-dimensional matching of cardinality  $q$  if and only if I-USNP instance  $\mathcal{I}$  has a solution with  $|E'| + \frac{|E'|}{2k+1}$  stops.

Consider a solution  $\{E'_1, \dots, E'_p\}$  to  $\mathcal{I}$  with  $|E'| + \frac{|E'|}{2k+1}$  stops. By definition,  $|E'_i| \leq 2k + 1$  for all  $i \in \{1, \dots, p\}$ . Since  $G'$  contains no cycle of size less than  $2k + 2$ , every nonempty subgraph  $G'(E'_i)$  is a forest with  $f_i \geq 1$  components and every nonempty vehicle  $i \in \{1, \dots, p\}$  thus makes  $|E'_i| + f_i$  stops. The total number of stops of this given solution therefore is  $|E'| + \sum_{i=1}^p f_i$ . We then obtain  $\sum_{i=1}^p f_i = \frac{|E'|}{2k+1} = p_{min}$ , that is, only  $p_{min}$  vehicles are nonempty and each of them serves  $2k + 1$  demands and stops  $2k + 2$  times. Consequently for every station  $v$  in  $X \cup Y \cup Z$ , the  $\deg_G(v) - 1$  paths  $P_v^1, \dots, P_v^{\deg_G(v)-1}$  are assigned to different vehicles. Each of these  $3|T| - q$  vehicles also serves either  $P_v^t$  if  $v \in Y \cup Z$  or  $(t, v)$  if  $v \in X$  as  $G'(E'_1), \dots, G'(E'_p)$  are trees with  $2k + 1$  edges if nonempty. This last property of  $G'(E'_1), \dots, G'(E'_p)$ , along with the number of nonempty vehicles being  $\frac{|E'|}{2k+1}$ , then implies that each of the remaining  $q$  nonempty vehicles serves an edge  $(t, x)$  and two paths  $P_y^t$  and  $P_z^t$  for some  $t = (x, y, z) \in T$ . The loads of these last  $q$  vehicles clearly induce a 3-dimensional matching of cardinality  $q$  in  $G$ .

Conversely from any 3-dimensional matching of cardinality  $q$  in  $G$ , a solution to  $\mathcal{I}$  with  $|E'| + \frac{|E'|}{2k+1}$  stops can easily be obtained by following the foregoing process in a reverse order.  $\square$

**Corollary 1.**  *$k$ -EPP is strongly NP-Hard even when  $G$  is restricted to be planar bipartite.*

Using a similar idea as in the proof of Theorem 2, we prove the NP-Hardness of I-USNP for all even values of  $C$  greater than or equal to 6.

**Theorem 3.** *Intersection-USNP is NP-Hard for any fixed even capacity  $C \geq 6$  even when  $G$  is restricted to be planar bipartite.*

*Proof.* The proof follows a similar approach as in Theorem 2. Let  $k \geq 1$ . To obtain an I-USNP instance  $\mathcal{I}$  with  $C = 2k + 4$  from a 3DMP instance

with planar bipartite graph  $G$ , we construct graph  $G' = (V', E')$  from  $G$  by (i) replacing every edge  $(t, v)$ ,  $v \in Y \cup Z$ , with a path  $P_v^t$  of length  $k + 1$ , (ii) replacing every edge  $(t, x)$ ,  $x \in X$ , with a  $P_v^t$  of length 2, (iii) attaching  $\deg_G(v) - 1$  node-disjoint paths  $P_v^1, \dots, P_v^{\deg_G(v)-1}$  of length  $k + 3$  to every node  $v \in Y \cup Z$ , and (iv) attaching  $\deg(x) - 1$  disjoint paths  $P_x^1, \dots, P_x^{\deg_G(x)-1}$  of length  $2k + 2$  to every node  $x \in X$ . See Figure 7 for an illustration of bipartite planar graph  $G'$  when  $C = 10$  (i.e.,  $k = 3$ ).

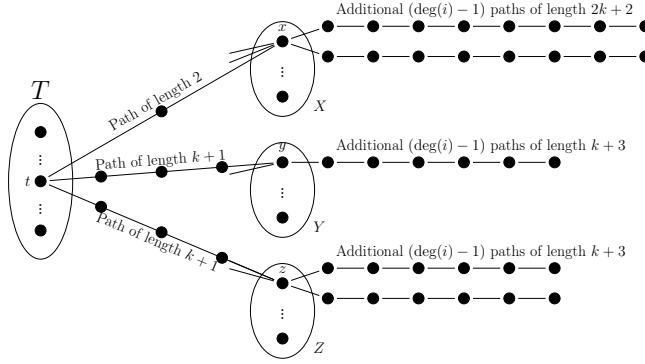


Figure 7: Construction of graph  $G'$  for Theorem 3 with  $k = 3$ .

With a partition of  $V'$  into two independent sets, we associate a feasible labeling of the nodes in  $V'$  and we thus obtain instance  $\mathcal{I} = (V', E', 2k + 4)$  in polynomial time. Using similar arguments as in the proof of Theorem 2 we prove that the 3DMP instance has a 3-dimensional matching of cardinality  $q$  if and only if I-USNP instance  $\mathcal{I}$  has a solution with  $|E'| + \frac{|E'|}{2k+4}$ .  $\square$

The only value of  $C$  for which the computational complexity of I-USNP remains open is  $C = 4$ . The key argument the proofs of Theorems 2 and 3 rely on is twofold:  $G'$  contains no cycle of size less than  $C + 1$  and  $G'$  contains for every node  $t \in T$ , a tree with  $C$  edges that covers  $t$  and its three associated nodes in  $X \cup Y \cup Z$ . This was achieved through successive subdivisions of the edges of  $G$ . If  $C$  was equal to 4, only one of the three edges in  $\{(t, x), (t, y), (t, z)\}$  could be subdivided for each  $t = (x, y, z) \in T$  and cycles of size 4 might thus appear. That therefore prevents us from using a similar argument as in the two foregoing proofs to prove that I-USNP with  $C = 4$  is NP-Hard on planar bipartite graphs. Since there is no reason to believe a polynomial-time algorithm could be devised for I-USNP with  $C = 4$ , we consequently conjecture.

**Conjecture 1.** *The Intersection-USNP is NP-Hard for  $C = 4$  even when*

restricted to the case where  $G$  is a planar bipartite graph.

### 3. Polyhedral Analysis

This section is devoted to a polyhedral study of USNP. A usual approach when tackling NP-Hard problems is to first model them as Mixed Integer Programs (MIPs) and then apply branch-and-cut based algorithms. A natural integer programming formulation for USNP presented in Pimenta et al. (2017) is described below

$$\min \sum_{v \in V} \sum_{i \in K} y_v^i \quad (3)$$

subject to

$$\sum_{i \in K} x_e^i = 1 \quad \forall e \in E, \quad (4)$$

$$\sum_{e \in \Delta_E(v)} x_e^i \leq C \quad \forall v \in V, i \in K, \quad (5)$$

$$x_e^i - y_v^i \leq 0 \quad \forall i \in K, e \in E, v \in \{o_e, d_e\}, \quad (6)$$

$$x_e^i \in \{0, 1\} \quad \forall e \in E, i \in K, \quad (7)$$

$$y_v^i \in \{0, 1\} \quad \forall v \in V, i \in K. \quad (8)$$

For every vehicle  $i \in K$  and every demand  $e \in E$ , the variable  $x_e^i$  indicates whether or not demand  $e$  is served by vehicle  $i$  (i.e.,  $x_e^i = 1$  iff  $e \in E_i$ ). The variable  $y_v^i$  represents the fact that vehicle  $i$  stops or not at station  $v$  (i.e.,  $y_v^i = 1$  iff  $v \in V(E_i)$ ). Objective function (3) minimizes the total number of stops. The assignment constraints (4) ensure that every demand is served by exactly one vehicle. The capacity constraints (5) guarantee that the vehicle's capacity never is violated all along the circuit. The stop constraints (6) impose that a vehicle must stop at the pick-up and drop-off stations of every demand assigned to it. Finally, constraints (7) and (8) make the  $x$ - and  $y$ -variables be binary. The formulation (3)-(8) is hereinafter referred to as USNP-XY. Its linear relaxation is obtained by replacing constraints (7) and (8) by the inequalities

$$0 \leq x_e^i \leq 1 \quad \forall e \in E, i \in K, \quad (9)$$

$$0 \leq y_v^i \leq 1 \quad \forall v \in V, i \in K. \quad (10)$$

Given an instance  $\mathcal{I} = (V, E, C)$  of USNP, let  $\text{USNP}(V, E, C)$  denote the convex hull of the solutions to (4)-(8) and  $P(V, E, C)$  denote the feasible set



of the linear relaxation of USNP-XY, that is,

$$P(V, E, C) = \{(x, y) \in \mathbb{R}^{mp} \times \mathbb{R}^{np} : (x, y) \text{ satisfies (4)-(6), (9) and (10)}\}.$$

The next theorem states the dimension of polytope  $USNP(V, E, C)$  and gives the necessary and sufficient conditions for inequalities (5), (6), (9), and (10) to be facet-defining for  $USNP(V, E, C)$ ; the proofs are omitted, since they use standard techniques. (See Colares (2019).)

**Theorem 4.** *Let  $G = (V, E)$  be a graph and  $K$  be a set of  $p$  vehicles of same capacity  $C \in \mathbb{Z}_+$  such that  $p > p_{\min}$ .*

- (i)  $\dim(USNP(V, E, C)) = (n + m)p - m$ .
- (ii) *Inequality (5) associated with station  $v \in V$  and vehicle  $i \in K$  is facet-defining for  $USNP(V, E, C)$  if and only if*
  - a.  $\Delta_E(v) \setminus \Delta_E(u) \neq \emptyset$  for every station  $u \in V \setminus v$  and
  - b.  $|\Delta_E(v) \setminus \delta(u)| \geq C$  for every station  $u \in V$ .
- (iii) *Inequalities (6) are facet-defining for  $USNP(V, E, C)$ .*
- (iv) *Inequality  $x_e^i \geq 0$  for demand  $e \in E$  and vehicle  $i \in K$  is facet-defining for  $USNP(V, E, C)$  if and only if  $p \geq 3$ .*
- (v) *Inequality  $y_v^i \leq 1$  for station  $v \in V$  and vehicle  $i \in K$  is facet-defining for  $USNP(V, E, C)$ .*

### 3.1. The weakness of formulation USNP-XY

The strength of formulations for USNP can be measured by the dual (lower) bounds its linear relaxations provide. The larger the bound is, the stronger the formulation is. We next investigate how strong formulation USNP-XY is.

**Theorem 5.** *Let  $z_{P(V, E, C)} = \min\{\sum_{v \in V} \sum_{i \in K} y_v^i : (x, y) \in P(V, E, C)\}$ . We have  $z_{P(V, E, C)} = |V| = n$ .*

*Proof.* We first show that  $z_{P(V, E, C)} \geq n$ . For every station  $v \in V$ , consider a demand  $e_v \in \delta^+(v) \cup \delta^-(v)$ . Summing up the  $p$  inequalities (6) associated with  $v$  and  $e_v$  yields  $\sum_{i \in K} y_v^i \geq \sum_{i \in K} x_{e_v}^i$ . The  $n$  inequalities thus obtained, combined with (4), gives

$$\sum_{i \in K} y_v^i \geq 1 \quad \forall v \in V$$

and therefore  $z_{P(V,E,C)} \geq n$ .

To complete the proof we now provide a feasible solution  $(\bar{x}, \bar{y})$  whose cost equals  $n$ . Let  $\bar{x}_e^i = \frac{1}{p}$  for all  $e \in E, i \in K$ , and  $\bar{y}_v^i = \frac{1}{p}$  for all  $v \in V, i \in K$ . Solution  $(\bar{x}, \bar{y})$  clearly satisfies constraints (4)-(6), (9), and (10). Therefore,  $(\bar{x}, \bar{y})$  is a feasible solution whose cost is  $\sum_{v \in V} \sum_{i \in K} \bar{y}_v^i = |V|$ .  $\square$

Notice that  $n = |V|$  is a trivial lower bound on the number of stops. Since each station is assumed to be the pick-up or drop-off station for at least one demand, it indeed must be visited by at least one vehicle. The following example illustrates how the lower bound given in Theorem 5 may be very weak. Consider an instance where all demands go from station 1 to station 2. The optimal value of the linear relaxation of USNP-XY is equal to 2, while the optimal value of USNP-XY is  $2 \lceil \frac{m}{C} \rceil$ . In other words one may find a relative integrality gap as large as  $1 - \frac{1}{\lceil \frac{m}{C} \rceil}$ . This negative result actually indicates that there exists quite some room for improvements on formulation USNP-XY. We therefore investigate new valid inequalities to strengthen it.

### 3.2. Strong capacity inequalities

The first family of inequalities we introduce relies on using  $C$  as an upper bound on the number of demands a vehicle may pick up (or drop off) at a station providing it stops there.

**Theorem 6.** *The strong capacity inequalities*

$$\sum_{e \in \delta^-(v)} x_e^i - C y_v^i \leq 0 \quad \forall v \in V, i \in K, \quad (11)$$

$$\sum_{e \in \delta^+(v)} x_e^i - C y_v^i \leq 0 \quad \forall v \in V, i \in K. \quad (12)$$

are valid for  $USNP(V, E, C)$ . Moreover, inequalities (11) (resp. (12)) associated with  $v \in V$  are facet-defining if and only if  $|\delta^-(v)| \geq C + 1$  (resp.  $|\delta^+(v)| \geq C + 1$ ).

*Proof.* Let  $i \in K$  be a vehicle and  $v \in V$  be a station. We first prove the validity of (11) and (12). If vehicle  $i$  does not stop at station  $v$  (i.e.,  $y_v^i = 0$ ) then it cannot serve any demand whose pick-up (or drop-off) station is  $v$ . If vehicle  $i$  stops at station  $v$  (i.e.,  $y_v^i = 1$ ), then its capacity only allows vehicle  $i$  to serve at most  $C$  demands whose pick-up (or drop-off) station is  $v$ . Inequalities (11) and (12) hence are valid for  $USNP(V, E, C)$ .

The proof of necessary and sufficient conditions under which inequalities (11) and (12) are facet-defining for  $USNP(V, E, C)$  is given in the Appendix.  $\square$

Using Chvátal-Gomory procedure (Chvatal, 1973; Gomory et al., 1958), we derive new valid inequalities from (4), (11), and (12).

**Theorem 7.** *The following min-stop inequalities*

$$\sum_{i \in K} y_v^i \geq \left\lceil \frac{\max\{|\delta^-(v)|, |\delta^+(v)|\}}{C} \right\rceil \quad \forall v \in V \quad (13)$$

are valid for  $USNP(V, E, C)$ .

*Proof.* Let  $v \in V$  be a station. Summing up the  $p$  inequalities (11) associated with  $v$  yields

$$\sum_{e \in \delta^-(v)} \sum_{i \in K} x_e^i - C \sum_{i \in K} y_v^i \leq 0.$$

By (4), we have  $\sum_{e \in \delta^-(v)} \sum_{i \in K} x_e^i = |\delta^-(v)|$ . After having divided the previous

inequality by  $C$  and then rounded-up the constant term  $\frac{|\delta^-(v)|}{C}$ , we obtain  $\sum_{i \in K} y_v^i \geq \left\lceil \frac{|\delta^-(v)|}{C} \right\rceil$ . A similar argument gives  $\sum_{i \in K} y_v^i \geq \left\lceil \frac{|\delta^+(v)|}{C} \right\rceil$  from (12).  $\square$

Adding inequalities (13) to formulation  $USNP\text{-}XY$  guarantees that its linear relaxation's optimal value is at least the lower bound provided by (1). It is also worth noticing that inequalities (11)-(13) cut off the fractional extreme point considered in the proof of Theorem 5. To assess the effectiveness of considering such inequalities in our formulation, we compare the root node's relative MIP gaps provided by CPLEX 12.8 when inequalities (11)-(13) are added or not to formulation  $USNP\text{-}XY$ .

To the best of our knowledge, there are no standard test sets for  $USNP$  in the literature. We therefore created our own set of 54  $USNP$  instances that is used throughout our study. Each instance is randomly generated with respect to scenarios based on a number of demands  $m \in \{30, 35, 40, 45, 50, 55\}$ , a capacity  $C \in \{2, 5, 8\}$ , and a density  $\rho = \frac{m}{n} \in \{1.5, 3.0, 4.5\}$ . Each generated instance is then referred to as  $m\_C\_ \rho$ .

Table 1 summarizes the results we obtained for instances with  $C = 5$  and either low density (i.e.,  $\rho = 1.5$ ) or high density (i.e.,  $\rho = 4.5$ ). For each formulation (i.e.,  $USNP\text{-}XY$  and  $USNP\text{-}XY$  with (11)-(13)), Table 1 presents the lower bounds  $LB_r$  and upper bound  $UB_r$  obtained after having processed the root node of our branch-and-cut framework. The relative MIP gaps  $\frac{UB_r - LB_r}{UB_r}$  then are given in columns labeled **gap<sub>r</sub>(%)**, and for formulation  $USNP\text{-}XY$  with (11)-(13) we also provide the numbers of added Strong Capacity inequalities (11)-(12) in column labeled **cut**.

Table 1: How Strong Capacity inequalities reinforce formulation USNP-XY.

Instances			USNP-XY			USNP-XY with (11)-(13)			
$m$	$C$	$\rho$	LB <sub>r</sub>	UB <sub>r</sub>	gap <sub>r</sub> (%)	LB <sub>r</sub>	UB <sub>r</sub>	gap <sub>r</sub> (%)	cut
30	5	1.5	19.5	27	27.78	20.4	29	29.66	71
35	5	1.5	23.4	34	31.18	23.8	39	38.97	83
40	5	1.5	28.3	49	42.24	28.3	39	27.44	50
45	5	1.5	30.6	42	27.14	30.6	42	27.14	0
50	5	1.5	34.5	52	33.65	35.7	69	48.26	146
55	5	1.5	37.6	76	50.53	37.7	73	48.36	81
30	5	4.5	9.4	16	41.25	14	14	0.00	210
35	5	4.5	10.3	21	50.95	16.6	19	12.63	337
40	5	4.5	10.8	24	55.00	15.2	20	24.00	436
45	5	4.5	12.6	34	62.94	19.6	28	30.00	926
50	5	4.5	13.9	45	69.11	21.6	33	34.55	979
55	5	4.5	14.5	54	73.15	22.2	36	38.33	1174

Inequalities (11)-(13) appear quite effective for high-density instances, with an average lower-bound increase of 52.48%, a reduction in the upper-bound values, and consequently smaller relative MIP gaps. For low-density instances, the benefits of adding (11)-(13) seem very slim if any. Even though the lower bound averagely increased by 1.68%, half the upper bounds did worsen which thus led to larger relative MIP gaps. Such deterioration in the upper bounds is due to the way CPLEX handles its calls for primal heuristics.

For relatively sparse graphs, inequalities (11)-(13) may fail to effectively strengthen formulation USNP-XY as illustrated in Table 1 when  $\rho = 1.5$ . The most symptomatic instances are those where the degree of each node is at most  $C$ . Inequalities (11)-(13) then become redundant and therefore yield no reinforcement of formulation USNP-XY. The instance given in Figure 8 where  $C = 2$  illustrates this fact. The lower bound provided by the linear relaxation of USNP-XY with (11)-(13) is  $n$ , as stated in Theorem 5, while the optimal value to (3)-(8) equals  $\frac{3n}{2}$ . To overcome this limitation, we introduce a new family of facet-defining inequalities that are based on the notion of  $k$ -cardinality tree.

### 3.3. $k$ -cardinality tree inequalities

Given a positive integer  $k$ , a  $k$ -cardinality tree is an undirected, connected, and acyclic graph having exactly  $k$  edges (Fischetti et al., 1994). Let  $\mathcal{CT}(G, k)$  denote the set of edge sets that induce  $k$ -cardinality trees of  $G$ ,

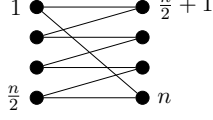


Figure 8: Example of instance where the inclusion of *strong capacity inequalities* is useless but *k-cardinality tree inequalities* are important.

that is,  $\mathcal{CT}(G, k) = \{S \subseteq E(G) : G(S) \text{ is a } k\text{-cardinality tree}\}$ .

**Theorem 8.** *The k-cardinality tree inequalities*

$$\sum_{e \in S} x_e^i - \sum_{u \in V(S)} (\deg_{G(S)}(u) - 1) y_u^i \leq 0 \quad \forall i \in K, v \in V, S \subseteq \Delta_E(v) \\ \text{with } S \in \mathcal{CT}(G, C + 1) \quad (14)$$

are valid for  $USNP(V, E, C)$ .

*Proof.* Let  $i \in K$  be a vehicle,  $v \in V$  be a station, and  $S \subseteq \Delta_E(v)$  with  $T = G(S)$  being a  $(C + 1)$ -cardinality tree. We prove that the associated inequality (14) can be obtained as a rank-1 Chvátal-Gomory cut. Select an arbitrary node  $r$  of  $T$  and think about  $T$  as a tree rooted at  $r$ . For every pair of distinct nodes  $t$  and  $u$  of  $T$ , we write  $t \prec u$  if  $t$  lies on the path of  $T$  from  $r$  to  $u$ . For every node  $u \in V(T)$ , let  $T_u$  denote the subtree of  $T$  rooted at  $u$ . (Note that  $T_r = T$ .) Consider the following linear combination of inequalities (5), (6), and (9)

$$\begin{aligned} (C - |E(T_u)|)(x_e^i - y_t^i) &\leq 0 & \forall e = (t, u) \in S, t \prec u, \\ |E(T_u)|(x_e^i - y_u^i) &\leq 0 & \forall e = (t, u) \in S, t \prec u, \\ \sum_{e \in \Delta_E(v)} x_e^i &\leq C, \\ -x_e^i &\leq 0 & \forall e \in \Delta_E(v) \setminus S. \end{aligned}$$

The sum of the above inequalities results in

$$(C + 1) \sum_{e \in S} x_e^i - (C + 1) \sum_{v \in V(S)} (\deg_{G(S)}(v) - 1) y_v^i \leq C.$$

Dividing it by  $(C + 1)$  and rounding down the right-hand side yields (14).  $\square$

For the instance given in Figure 8 where  $C = 2$ , the optimal value of the linear relaxation of formulation  $USNP\text{-}XY$  reinforced with *k-cardinality tree inequalities* (14) equals  $\frac{3m}{2}$  (i.e., the relative MIP gap has vanished). Figure 9 plots the root node's relative MIP gaps we obtained by considering either  $USNP\text{-}XY$  with (11)-(13) or  $USNP\text{-}XY$  with (11)-(14) for the foregoing

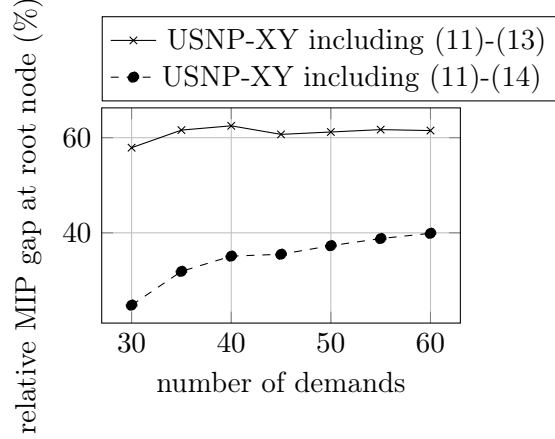


Figure 9: Improvement in the relative MIP gap at the root node after adding inequalities (14) for instances where  $C = 2$  and  $\rho = 1.5$ .

instances where  $C = 2$  and  $\rho = 1.5$ . The addition of (14) makes these gaps drop below 40% when they were around 60% before.

The next theorem provides necessary and sufficient conditions under which inequalities (14) are facet-defining for  $\text{USNP}(V, E, C)$ .

**Theorem 9.** *The  $k$ -cardinality tree inequalities (14) are facet-defining if and only if*

- (i.) *there does not exist a node  $u \in V(S)$  such that  $S \subset \delta_G^-(u)$  or  $S \subset \delta_G^+(u)$ , and*
- (ii.) *there does not exist an edge  $(u_1, u_2) \in E \setminus S$  such that  $u_1$  and  $u_2$  are the only internal nodes of  $G(S)$ .*

*Proof.* The complete proof can be found in the Appendix.  $\square$

Although not used in the branch-and-cut framework described in Section 4, it is worth mentioning that a generalization of  $k$ -cardinality tree inequalities to  $S \subseteq E$  such that  $G(S)$  is a forest with  $|S \cap \Delta_E(v)| = C + 1$  have been introduced by Colares (2019).

### 3.4. Girth inequalities

So far, we only have made use of sparsest structures, such as stars and trees, to derive valid inequalities for  $\text{USNP}(V, E, C)$ . We now introduced new families of valid inequalities by looking at some denser structures. The *girth*

of a graph  $G$  corresponds to the length of its smallest cycle. If  $G$  is a graph of girth  $k$ , then  $G(S)$  is a forest for any  $S \subseteq E(G)$  such that  $|S| \leq k - 1$ . Our next family of valid inequalities is based on this remark. Let  $\mathcal{G}(G, k)$  denote the set of edge sets that induce subgraphs of girth at least  $k$ , that is,  $\mathcal{G}(G, k) = \{S \subseteq E(G) : G(S) \text{ has girth at least } k\}$ .

**Theorem 10.** *The girth inequalities*

$$\sum_{e \in S} (C + 1)x_e^i - \sum_{u \in V(S)} Cy_u^i \leq 0 \quad \forall i \in K, v \in V, S \subseteq \Delta_E(v) \quad (15)$$

*with  $S \in \mathcal{G}(G, C + 1)$*

are valid for  $USNP(V, E, C)$ .

*Proof.* Let  $i \in K$  be a vehicle,  $v \in V$  be a station, and  $S \subseteq \Delta_E(v)$  with  $G(S)$  of girth greater than or equal to  $C + 1$ . Consider an arbitrary feasible solution  $(\bar{x}, \bar{y}) \in USNP(V, E, C) \cap \mathbb{Z}^{(n+m)p}$  and let  $S_i \subseteq S$  denote the set of demands of  $S$  served by vehicle  $i$  in the solution  $(\bar{x}, \bar{y})$ , that is,  $S_i = \{e \in S : \bar{x}_e^i = 1\}$ . If  $S_i = \emptyset$ , then (15), associated with  $i, v$ , and  $S$ , is obviously satisfied by  $(\bar{x}, \bar{y})$ . Hence assume  $|S_i| \geq 1$ . By definition,

$$\sum_{e \in S} (C + 1)\bar{x}_e^i = \sum_{e \in S_i} (C + 1) = C|S_i| + |S_i|.$$

Since  $S \subseteq \Delta_E(v)$ , we must have  $|S_i| \leq C$ .  $G[S]$  having girth greater than or equal to  $C + 1$  thus implies that  $G[S_i]$  must be a forest and vehicle  $i$  hence stops at at least  $|S_i| + 1$  stations in  $V(S)$ . Therefore,

$$\sum_{u \in V(S)} C\bar{y}_u^i \geq \sum_{u \in V(S_i)} C \geq (|S_i| + 1)C = C|S_i| + C.$$

Consequently,

$$\sum_{e \in S} (C + 1)\bar{x}_e^i = C|S_i| + |S_i| \leq C|S_i| + C \leq \sum_{u \in V(S)} C\bar{y}_u^i,$$

and our proof is complete.  $\square$

Using Chvátal-Gomory procedure, we derive new valid inequalities from girth inequalities (15).

**Theorem 11.** *The following inequalities*

$$\sum_{u \in V(S)} \sum_{i \in K} y_u^i \geq |S| + \left\lceil \frac{|S|}{C} \right\rceil \quad \forall v \in V, S \subseteq \Delta_E(v) \text{ with } S \in \mathcal{G}(G, C + 1) \quad (16)$$

are valid for  $USNP(V, E, C)$ .

*Proof.* Let  $v \in V$  be a station and  $S \subseteq \Delta_E(v)$  with  $G(S)$  of girth greater than or equal to  $C + 1$ . Summing up the  $p$  inequalities (15) yields

$$\sum_{u \in V(S)} \sum_{i \in K} C y_u^i \geq \sum_{e \in S} \sum_{i \in K} (C + 1) x_e^i.$$

Using (4), we then obtain

$$\sum_{u \in V(S)} \sum_{i \in K} y_u^i \geq \frac{C + 1}{C} \sum_{e \in S} \sum_{i \in K} x_e^i = \frac{C + 1}{C} |S| = |S| + \frac{|S|}{C}.$$

Rounding up  $|S| + \frac{|S|}{C}$  thus give inequalities (16). Note that  $S \in \mathcal{G}(G, C + 1)$  implies  $|S| > C$ .  $\square$

If  $(V, E, C)$  is an instance of I-USNP with  $G = (V, E)$  of girth greater than or equal to  $C + 1$ , then Theorem 11 immediately provides

$$m + \left\lceil \frac{m}{C} \right\rceil$$

as a lower bound on the total number of stops.

#### 4. Branch-and-cut and computational results

The branch-and-cut framework we have developed from the results obtained in Section 3 is hereafter presented. To boost its performance, several features, such as (i) breaking the symmetry inherent in USNP, (ii) eliminating variables, (iii) relaxing variables, and (iv) integrating the strengthening inequalities of Section 3, have been carefully handled and are next reviewed.

##### 4.1. Symmetry

In Margot (2010), a MIP is defined as being symmetric if its variables can be permuted without changing the problem's structure. In our case, since vehicles are considered to be identical, USNP hides a complete symmetry with respect to vehicles. In other words, a feasible solution  $\{E_1, \dots, E_p\}$  ensures the existence of  $p!$  distinct equivalent solutions since every permutation of  $\{E_1, \dots, E_p\}$  is feasible and has the same number of stops.

Numerous authors have highlighted the importance of eliminating, or at least reducing, symmetry that may appear in MIPs (e.g., Sherali and Smith (2001); Kaibel and Pfetsch (2008); Denton et al. (2010); Ostrowski et al. (2010)). When solving a symmetric MIP through a branch-and-bound-based procedure, isomorphic subproblems are encountered in the enumeration tree



yielding a useless duplication of efforts and slowing down the resolution process.

Kaibel et al. (2011) proposed a linear-time algorithm called *Orbitopal Fixing* that completely removes symmetry for partition problems. At each node of the enumeration tree, the method fixes variables based on both a predefined lexicographic order and the already-fixed variables at this current node. With respect to USNP, such lexicographic order would amount to impose that (i.) the  $e$ -th demand,  $e \in \{1, \dots, m\}$ , must be assigned to one of the first  $e$  vehicles, and (ii.) a demand  $e \in E$  cannot be assigned to an empty vehicle  $i$  if there exists another empty vehicle  $i'$  such that  $i' < i$ . In our branch-and-cut framework to solve USNP, we have applied this Orbitopal-Fixing approach at every branching step. Classic symmetry-breaking constraints were also investigated but did not perform as well as the Orbitopal-Fixing approach. Detailed results on this feature may be found in Colares (2019).

#### 4.2. Variable elimination

In Section 2, we have showed that, unlike stated by Pimenta et al. (2017), the minimum number of vehicles needed to obtain an optimal solution to USNP may differ from the minimum number of vehicles needed to obtain a feasible solution, that is,  $p_{opt} \neq p_{min}$ . For this reason, the number of available vehicles  $p$  has been so far assumed to be  $m$ . Notice however, that the number of variables in our formulation is directly related to the number of available vehicles. More precisely, formulation USNP-XY uses  $(m + n)p$  variables. Providing a tighter upper bound on  $p_{opt}$  would therefore allow to considerably reduce the number of variables in the formulation. Moreover, as mentioned in the previous section, the symmetry inherent in USNP is closely related to the number of available vehicles, for each feasible solution corresponds to  $p!$  equivalent solutions. The benefits of decreasing the value of  $p$  would hence be twofold: reducing the number of variables in formulation USNP-XY and lessening the MIP symmetry. The next theorem provides a better upper bound on  $p_{opt}$ .

**Theorem 12.** *For any USNP instance  $\mathcal{I} = (V, E, C)$ , we have*

$$p_{opt} \leq \left\lceil \frac{m}{\left\lfloor \frac{C}{2} \right\rfloor + 1} \right\rceil.$$

*Proof.* Consider a USNP instance  $\mathcal{I}$  for which  $p_{opt} = \left\lceil \frac{m}{\left\lfloor \frac{C}{2} \right\rfloor + 1} \right\rceil + 1$ .

Let  $\{E_1, \dots, E_{p_{opt}}, \dots, E_p\}$  be an optimal solution for  $\mathcal{I}$  where without loss of generality,  $E_i \neq \emptyset$  for  $1 \leq i \leq p_{opt}$  and  $E_i = \emptyset$  for  $p_{opt} + 1 \leq i \leq p$ . We clearly have  $|E_{i_1}| + |E_{i_2}| > C$  for any distinct vehicles  $1 \leq i_1, i_2 \leq p_{opt}$  since otherwise, merging both vehicles  $i_1$  and  $i_2$  would yield an optimal solution using one less vehicle. Therefore, at most one vehicle  $i \in \{1, \dots, p_{opt}\}$  serves less than  $\lfloor \frac{C}{2} \rfloor + 1$  demands, that is, this vehicle serves  $1 \leq d \leq \lfloor \frac{C}{2} \rfloor$  demands. All the other  $(p_{opt} - 1)$  vehicles serve at least  $\lfloor \frac{C}{2} \rfloor + 1$  demands each, implying

$$m \geq (p_{opt} - 1) \left( \left\lfloor \frac{C}{2} \right\rfloor + 1 \right) + d = \left( \left\lceil \frac{m}{\lfloor \frac{C}{2} \rfloor + 1} \right\rceil \right) \left( \left\lfloor \frac{C}{2} \right\rfloor + 1 \right) + d. \quad (17)$$

Since

$$m = (p_{opt} - 1) \left( \left\lfloor \frac{C}{2} \right\rfloor + 1 \right) + r$$

with  $0 \leq r < \lfloor \frac{C}{2} \rfloor + 1$ , we obtain from (17)  $d \leq 0$ , a contradiction to vehicle  $i$  being nonempty. Therefore,  $p_{opt} \leq \left\lceil \frac{m}{\lfloor \frac{C}{2} \rfloor + 1} \right\rceil$ .  $\square$

It is worth noticing that the upper bound on  $p_{opt}$  given in Theorem 12 is tight for some USNP instances. Indeed, if we consider an instance of *Intersection-USNP* composed of  $C$  disjoint sets of  $\lfloor \frac{C}{2} \rfloor + 1$  parallel demands, then the optimal solution consists of assigning each set to a different vehicle and hence,  $p_{opt} = \left\lceil \frac{m}{\lfloor \frac{C}{2} \rfloor + 1} \right\rceil = C$ .

#### 4.3. Variable relaxation

It can be easily seen that once the  $x$ -variables are fixed to 0-1 values (i.e., once the set of demands each vehicle serves is known), the problem of deciding where the vehicles stop becomes trivial. For every vehicle  $i \in K$  and every station  $v \in V$ , we simply set variable  $y_v^i$  to  $\max\{x_e^i : e \text{ is incident to } v\}$ . This means that the integrality constraints (8) for the  $y$ -variables can be dropped. Let USNP-X denote the formulation defined by (3)-(7),(10).

The opposite question of deciding whether or not an optimal assignment of demands to vehicles can be deduced in polynomial time once the stops of each vehicle are known is not straightforward. Such question may be answered by showing that the polytope associated with  $P(V, E, C)$  when the  $y$ -variables are fixed to 0-1 values is integral. We next positively answer this question for I-USNP but we provide a counterexample for general USNP.

For any  $\bar{y} \in \{0, 1\}^{n \times p}$  let

$$P_x(V, E, C, \bar{y}) = \{(x, y) \in P(V, E, C) : y = \bar{y}\}$$

denote the face of the polytope  $P(V, E, C)$  induced by  $y$  being fixed to  $\bar{y}$ .

**Theorem 13.** *Given an instance  $(V, E, C)$  of Intersection-USNP and a vector  $\bar{y} \in \{0, 1\}^{n \times p}$ , the polyhedron  $P_x(V, E, C, \bar{y})$  is integral.*

*Proof.* Since  $(V, E, C)$  is an instance of I-USNP, the capacity constraints (5) reduce to

$$\sum_{e \in E} x_e^i \leq C \quad \forall i \in K. \quad (18)$$

The coefficient matrix of the system (4),(18) corresponds to the node-edge incidence matrix of the complete bipartite graph whose node set is  $E \cup K$ , and hence is well-known for being totally unimodular. Since appending the identity matrix to a totally-unimodular matrix preserves the totally-unimodular property and so does duplicating a row or multiplying a row by -1, the coefficient matrix of the system defining  $P_x(V, E, C, \bar{y})$  also is totally unimodular. Consequently,  $P_x(V, E, C, \bar{y})$  is an integral polytope.  $\square$

The complete bipartite graph the proof of Theorem 13 relies on cannot be achieved for USNP instances. The total unimodularity of the system defining  $P_x(V, E, C, \bar{y})$  then is not preserved for those instances. The following example shows that if there is no station  $v \in V$  for which  $\Delta_E(v) = E$ , then  $P_x(V, E, C, \bar{y})$  may have a fractional extreme point. Consider the instance depicted in Figure 10 with  $C = 1$ . If  $\bar{y} = \mathbf{1}$ , then the following fractional solution  $\bar{x}$  is an extreme point of  $P_x(V, E, 1, \mathbf{1})$  as it is the unique solution to the system of inequalities defining  $P_x(V, E, 1, \mathbf{1})$  that are binding at  $\bar{x}$ :

$$\begin{aligned} \bar{x}_e^1 &= \begin{cases} \frac{1}{2} & \text{if } e \in \{a, c, d\}, \\ 0 & \text{if } e \in \{b\}, \end{cases} & \bar{x}_e^2 &= \begin{cases} \frac{1}{2} & \text{if } e \in \{a, b\}, \\ 0 & \text{if } e \in \{c, d\}, \end{cases} \\ \bar{x}_e^3 &= \begin{cases} \frac{1}{2} & \text{if } e \in \{b, d\}, \\ 0 & \text{if } e \in \{a, c\}, \end{cases} & \bar{x}_e^4 &= \begin{cases} \frac{1}{2} & \text{if } e \in \{c\}, \\ 0 & \text{if } e \in \{a, b, d\}. \end{cases} \end{aligned}$$

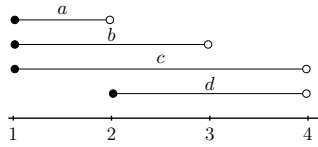


Figure 10: Counter-example USNP instance to prove the nonintegrality of polyhedron  $P_x(V, E, C, \bar{y})$  when  $C = 1$ .

For USNP, keeping the integrality of the  $y$ -variables while relaxing that of the  $x$ -variables therefore does not guarantee the integrality of the associate solution. So let USNP-Y denote the formulation USNP-XY with the additional rules that prioritizes branching on a  $y$ -variable over an  $x$ -variable.

In our computational experiments we therefore consider three formulations, USNP-XY, USNP-X, and USNP-Y that, although equally strong as their linear relaxations are identical, differ in how the integrality constraints (7),(8) and the branching rules are managed. In USNP-XY, no particular branching rule is applied and the optimization solver uses its default rules to choose which variable it branches on. In USNP-X, where (8) are relaxed, only  $x$ -variables can be branched on. In USNP-Y, branching on an  $x$ -variable is only possible once all the  $y$ -variables have values in  $\{0, 1\}$ . Notice that the information derived from fixing either an  $x$ -variable or a  $y$ -variable is much different. On the one hand, for a station  $v \in V$  and a vehicle  $i \in K$ , fixing  $y_v^i$  to 0 implies  $x_e^i = 0$  for every demand  $e \in \delta(v)$ . On the other hand, for a demand  $e \in E$  and a vehicle  $i \in K$ , fixing  $x_e^i$  to 1 only implies  $y_{\delta_e}^i = y_{d_e}^i = 1$ . Furthermore, the denser the associated graph  $G$  is the larger is the difference between the numbers of  $x$ - and  $y$ -variables. We may therefore expect USNP-Y to produce better results when  $G$  is dense.

Table 2 provides a sample of the results obtained with each formulation for the instances described in Section 3.2 where  $C = 5$ . In this experiment, Orbitopal Fixing symmetry-breaking method is applied and the number of available vehicles is set to the new upper bound on  $p_{opt}$  given by Theorem 12. For each formulation, the total number of binary variables is given in column **bin**. The total amount of time (in seconds) required by the branch-and-cut framework is displayed in column **CPU**. A time limit of 2 hours is imposed and once exceeded, the relative MIP gap is displayed in column **gap**. The number of nodes (in thousands) explored in the enumeration tree is given in column **node**.

Table 2: Comparison between formulations USNP-XY, USNP-X and USNP-Y.

Inst.		USNP-XY				USNP-X				USNP-Y			
$m$	$\rho$	bin	CPU	gap	node	bin	CPU	gap	node	bin	CPU	gap	node
30	1.5	500	14.1	-	9.8	300	16.5	-	8.3	500	5	-	3.1
35	1.5	696	522	-	227	420	111	-	39.6	696	755	-	296
40	1.5	924	1896	-	546	560	2183	-	501	924	7200	4.8	1637
45	1.5	1125	7200	5.7	1144	675	7200	7.1	1205	1125	7200	8.4	1063
50	1.5	1411	7200	9.5	858	850	7200	11.0	781	1411	7200	11.8	827
55	1.5	1729	7200	15.3	533	1045	7200	12.7	507	1729	7200	12.9	416
30	4.5	360	1.8	-	0.6	300	68.9	-	48.9	360	1.8	-	0.9
35	4.5	504	7200	8.2	3070	420	4330	-	2012	504	186	-	67.5
40	4.5	672	7200	16.0	1310	560	7200	18.4	1388	672	2130	-	499
45	4.5	825	7200	26.2	1056	675	7200	25.0	1082	825	7200	19.5	809
50	4.5	1037	7200	27.9	755	850	7200	30.2	778	1037	7200	24.9	604
55	4.5	1273	7200	34.1	494	1045	7200	34.1	473	1273	7200	30.3	431

As expected, formulation USNP-Y clearly outperforms the other two formulations for instances with dense graphs (i.e.,  $\rho = 4.5$ ). Considering only instances with  $\rho = 4.5$ , formulation USNP-Y was generally able to either solve faster the problem to optimality or provide smaller relative MIP gaps when compared to USNP-XY and USNP-X. For sparse instances (i.e.,  $\rho = 1.5$ ), formulation USNP-Y did not outperform USNP-X or USNP-XY. Even if neither USNP-XY nor USNP-X seems to perform consistently better than the other for sparse instances in Table 2, a deeper analysis over the whole set of sparse instances shows that formulation USNP-X has, in average, slightly better performance. For this reason in the follow-up experiments, we have decided to only apply formulation USNP-X when the instance is sparse (i.e.,  $\rho = 1.5$ ). For every other instance, we have chosen to only apply formulation USNP-Y.

#### 4.4. Integrating cuts

We now look at how the strengthening valid inequalities (11), (12), (13), (14), (15), and (16) are managed throughout our branch-and-cut framework. Since the strong-capacity inequalities (11), (12), and (13) appear in polynomial numbers, storing them in a pool and checking whether they all are satisfied remains an efficient way of handling them. For the  $k$ -cardinality tree inequalities (14) and the girth inequalities (15), and (16), we next focus on their separation problems. Recall that for a family  $\mathcal{F}$  of valid inequalities, the *separation problem* for  $\mathcal{F}$  consists of either finding an inequality in  $\mathcal{F}$  violated by a given a vector  $(\bar{x}, \bar{y}) \in \mathbb{R}^{(m+n) \times p}$  or proving that  $(\bar{x}, \bar{y})$  satisfies all the inequalities in  $\mathcal{F}$ .

**Theorem 14.** *The separation problem for the  $k$ -cardinality tree inequalities (14) is NP-Hard.*

*Proof.* We give a reduction from the  $k$ -Minimum Spanning Tree decision problem ( $k$ -MST) which can be defined as follows. Given a graph  $G = (V, E)$  with edge-weight vector  $w \in \mathbb{R}^E$  and a scalar  $B$ , does there exist a tree  $T = (V', E')$  spanning exactly  $k$  edges (i.e., so that  $|E'| = k$ ) with total weight at most  $B$ ? This problem can be solved in polynomial time when  $G$  is itself a tree or when  $k$  is a fixed constant, but it is NP-Hard in the general case. (See Fischetti et al. (1994); Ravi et al. (1996).) Notice that for proving the NP-Hardness of  $k$ -MST, Fischetti et al. (1994) used a reduction from the Steiner Tree problem that preserves bipartiteness. Since the Steiner Tree problem is NP-Hard even on bipartite graphs (see Garey and Johnson (2002, p. 208)),  $k$ -MST also is NP-Hard on bipartite graphs.

Consider thus a  $k$ -MST instance specified by a bipartite graph  $G = (V, E)$ , negative edge-weights  $w \in \mathbb{R}^E$ , and a scalar  $B \in \mathbb{R}$ , and the I-USNP instance  $(V, E, k - 1)$ . Let  $(\bar{x}, \bar{y}) \in \mathbb{R}^{(m+n) \times p}$  be defined by  $\bar{x}_e^i = -w_e$  for every vehicle  $i \in K$  and every demand  $e \in E$ , and  $\bar{y}_v^i = -\frac{B}{k-1}$  for every vehicle  $i \in K$  and every station  $v \in V$ . We now show that any tree  $T'$  of  $G$  spanning exactly  $k$  edges such that  $w(E(T')) \leq B - \epsilon$  for  $\epsilon > 0$  corresponds to a  $k$ -cardinality tree inequality violated by  $(\bar{x}, \bar{y})$ . Let  $T$  be any  $k$ -cardinality tree in  $G$ . Notice that by definition,  $|V(T)| = k + 1$ . Since all the  $y$ -variables have the same value, we then obtain

$$\begin{aligned} \sum_{u \in V(T)} (\deg_T(u) - 1) \bar{y}_u^i &= \sum_{u \in V(T)} \deg_T(u) \bar{y}_u^i - \sum_{u \in V(T)} \bar{y}_u^i \\ &= -2k \frac{B}{k-1} + (k+1) \frac{B}{k-1} = -B. \end{aligned}$$

Therefore, the  $k$ -cardinality inequality (14) associated with vehicle  $i \in K$  and set  $E(T) \in \mathcal{CT}(G, k)$  of demands is equivalent to

$$\sum_{e \in E(T)} \bar{x}_e^i \leq -B.$$

If this inequality is violated by  $(\bar{x}, \bar{y})$ , then the  $k$ -cardinality tree  $T$  solves  $k$ -MST since  $\sum_{e \in E(T)} \bar{x}_e^i = -w(E(T))$ . Conversely, any  $k$ -cardinality tree  $T$  of  $G$  such that  $w(E(T)) \leq B - \epsilon$  induces a  $k$ -cardinality tree inequality violated by  $(\bar{x}, \bar{y})$ .  $\square$

Since the separation of inequalities (14) is NP-Hard, we have chosen to heuristically solve it through a simple greedy procedure. Given a solution  $(\bar{x}, \bar{y}) \in \mathbb{R}^{(m+n) \times p}$  and a vehicle  $i \in K$ , this heuristic constructs a maximum-weight  $(C+1)$ -cardinality tree of graphs  $G[\Delta_E(v)]$  with the restriction of  $\bar{x}^i$  to  $\Delta_E(v)$  as edge weights and the restriction of  $-\bar{y}^i$  to  $V(\Delta_E(v))$  as node costs, for  $v \in V$ , by iteratively adding the highest-valued edge to the tree. Such algorithm runs in  $\mathcal{O}(Cm)$  time.

**Theorem 15.** *The separation problem for girth inequalities (15) is NP-Hard.*

*Proof.* We show that the separation problem for the girth inequalities (15) is equivalent to the Hamiltonian Cycle (HC) problem. Given a graph  $G = (V, E)$ , HC problem consists of finding a cycle  $H = (V, S)$  of  $G$ . This problem was shown to be NP-Hard even if  $G$  is bipartite (see Krishnamoorthy (1975)).

From a bipartite graph  $G = (V, E)$ , construct an I-USNP instance  $(V, E, |V| - 1)$  and define  $(\bar{x}, \bar{y}) \in \mathbb{R}^{(m+n) \times p}$  as follows. For every vehicle  $i \in K$

and every demand  $e \in E$ , let  $\bar{x}_e^i = \frac{1}{|V|}$  and for every vehicle  $i \in K$  and every station  $v \in V$ , let  $\bar{y}_v^i = \frac{1}{|V|}$ . We show that any girth inequality violated by  $(\bar{x}, \bar{y})$  induces an Hamiltonian Cycle in  $G$  and conversely.

Let  $S \subseteq E$  be the edge set of an Hamiltonian cycle  $H$  of  $G$ . The left-hand side of inequality (15) associated with vehicle  $i \in K$  and set  $S \in \mathcal{G}(G, |V|)$  of demands then satisfies

$$\sum_{e \in S} |V| \bar{x}_e^i - \sum_{u \in V(S)} (|V| - 1) \bar{y}_u^i = |V| - (|V| - 1) > 0$$

and this girth inequality is violated by  $(\bar{x}, \bar{y})$ .

To complete the proof, let  $S \subseteq E$  be such that the girth of  $G[S]$  is greater than or equal to  $|V|$  and  $G[S]$  is not an Hamiltonian cycle of  $G$ . Then,  $G[S]$  is acyclic and hence  $|V| \geq |V(S)| \geq |S| + 1$ . The left-hand side of inequality (15) associated with vehicle  $i \in K$  and set  $S \in \mathcal{G}(G, |V|)$  of demands then satisfies

$$\sum_{e \in S} |V| \bar{x}_e^i - \sum_{u \in V(S)} (|V| - 1) \bar{y}_u^i = |S| - |V(S)| \frac{|V| - 1}{|V|} \leq 0.$$

Consequently if a girth inequality induced by a set  $S \subseteq E$  and violated by  $(\bar{x}, \bar{y})$  is found, then the cycle  $G[S]$  solves HC problem.  $\square$

Since the separation problem for inequalities (15) is NP-Hard, we propose a greedy procedure to solve it heuristically. Given a solution  $(\bar{x}, \bar{y}) \in \mathbb{R}^{(m+n) \times p}$  and a vehicle  $i \in K$ , this heuristic is based on the idea of building a maximum-cost spanning tree  $T$  of graphs  $G[\Delta_E(v)]$  with the restriction of  $(C+1)\bar{x}^i$  to  $\Delta_E(v)$  as edge weights, for  $v \in V$ , and then including additional edges to  $T$  such that the resulting graph does not contain any cycle of size less than or equal to  $C$ . The proposed algorithm runs in  $\mathcal{O}(m \log n)$  time and is also used to heuristically separate inequalities (16).

#### 4.5. Numerical experiments

To implement the features described in this section, we have made use of the callback functionality of CPLEX 12.8. When doing so, CPLEX deactivates some of its default settings (e.g., dynamic search and dual reductions). For a fair comparison, a dummy callback is used to deactivate such features in our baseline CPLEX solver. This is a common practice in the literature for proving the efficiency of branching, cutting planes, and node-selection methods. (See, e.g., Carvajal et al. (2014), Sabharwal et al. (2012), Fischetti and Monaci (2014).) Table 3 provides a performance comparison of

the branch-and-cut framework described in Section 4 (displayed as scenario **branch-and-cut**) and the formulation USNP-XY solved using CPLEX’s traditional branch-and-cut algorithm (displayed as scenario **CPLEX**).

Columns **time<sub>g</sub>**, **LB<sub>g</sub>**, **UB<sub>g</sub>**, **gap<sub>g</sub>**, and **Nodes** provide information on the global optimization process by showing, respectively, the total amount of time (in seconds), the lower and upper bounds obtained at the end of the optimization process, the relative MIP gap and the total number of nodes investigated in the enumeration tree. Column **CP** refers to the total number of cuts CPLEX automatically generated during the whole optimization process. Columns **cut<sub>SC</sub>**, **cut<sub>KT</sub>**, and **cut<sub>G</sub>** indicate the total numbers of violated strong capacity inequalities (11)-(13),  $k$ -cardinality tree inequalities (14), and girth inequalities (15)-(16), respectively, added during the optimization process. Finally, columns **cut<sub>tot</sub>** and **time<sub>cut</sub>** provide the total number of added user-defined cuts (11)-(16) and the total amount of time spent solving the separation problems for (14)-(16), respectively. All numerical experiments were performed on a 3.1 GHz Intel Xeon E5 computer with 32 Gb of memory.

The results achieved by **branch-and-cut** clearly outperform the ones obtained with **CPLEX**. Indeed, while only one out of the 54 instances described in Section 3.2 could be solved to optimality within the time limit of 2 hours by **CPLEX**, **branch-and-cut** could optimally solve 24 of them. The only instance solved by **CPLEX** required almost 40 minutes while **branch-and-cut** solved it in only 2 seconds. Moreover, the number of nodes explored to prove optimality for such instance went down from 538,000 to 100.

Even if the total number of explored nodes required to prove optimality is greatly reduced with **branch-and-cut**, the time spent at each node of the enumeration tree might considerably increase. Despite the fact that the linear-relaxation size (i.e., the numbers of variables and constraints) is significantly reduced by the elimination of variables and the integration of symmetry-breaking methods described in Section 4, longer node evaluation does not come as a surprise since many violated cuts are added to the formulation, yielding several re-optimization of the node’s linear program. Notice that for some unsolved instances (e.g., 55\_2\_3.0), **branch-and-cut** spent the whole optimization process at the root node, searching for violated cuts and re-optimizing the root’s linear program with such new cuts. Such behavior reveals a trade-off between quickly solving the nodes’ linear programs and hence further exploring the enumeration tree, or instead, strengthening the nodes’ formulations with the continuous search for violated cuts. Using our separation heuristics, we believe we have found a good balance between these two aspects. Indeed, for the 30 instances that could not be solved



Table 3: Final computational results.

Instances			CPLEX						Branch-and-Cut										
$m$	$C$	$\rho$	time <sub>g</sub>	LB <sub>g</sub>	UB <sub>g</sub>	gap <sub>g</sub>	Nodes	CP	time <sub>g</sub>	LB <sub>g</sub>	UB <sub>g</sub>	gap <sub>g</sub>	Nodes	CP	cut <sub>SC</sub>	cut <sub>KT</sub>	cut <sub>G</sub>	cut <sub>tot</sub>	time <sub>cut</sub>
30	2	1.5	7200	22.6	35	35.3	808.0	143	36.5	35.0	35	-	1.9	0	143	491	81	715	2.5
30	5	1.5	7200	19.5	25	22.1	1318	40	27.3	25.0	25	-	2.9	0	16	739	101	856	2.1
30	8	1.5	2351	24.0	24	-	538.0	4	2	24.0	24	-	0.1	1	0	190	9	199	0.2
35	2	1.5	7200	25.1	41	38.9	591.1	109	6007	41.0	41	-	433.1	0	207	835	75	1117	7.0
35	5	1.5	7200	23.4	31	24.6	627.6	26	571	31.0	31	-	24.8	0	24	1387	440	1851	11.5
35	8	1.5	7200	25.0	29	14.0	701.3	4	5.1	29.0	29	-	1.1	0	0	129	7	136	0.4
40	2	1.5	7200	26.8	51	47.4	356.9	15	7200	45.5	50	9.0	308.4	0	252	1662	115	2029	19.4
40	5	1.5	7200	27.7	37	25.0	572.7	6	6112	37.0	37	-	281.5	2	15	1283	109	1407	10.1
40	8	1.5	7200	26.6	32	17.0	394.5	3	9.7	32.0	32	-	1.9	0	0	147	9	156	0.3
45	2	1.5	7200	29.8	50	40.4	141.1	236	7200	45.7	49	6.7	206.4	0	287	1318	168	1773	14.5
45	5	1.5	7200	30.5	40	23.8	337.9	4	7200	36.9	40	7.8	390.4	3	0	1400	187	1587	10.2
45	8	1.5	7200	31.3	38	17.6	259.2	4	227	38.0	38	-	13.4	0	0	1197	103	1300	5.2
50	2	1.5	7200	35.5	63	43.6	59.5	231	7200	57.0	62	8.1	87.5	0	390	2109	335	2834	52.9
50	5	1.5	7200	33.0	46	28.3	192.6	5	7200	41.0	46	10.9	132.0	0	34	3406	727	4167	75.4
50	8	1.5	7200	33.1	41	19.3	193.7	4	1625	41.0	41	-	59.6	0	0	1216	129	1345	13.1
55	2	1.5	7200	37.2	65	42.8	47.4	290	7200	58.6	64	8.4	36.2	0	435	1913	171	2519	24.9
55	5	1.5	7200	35.8	52	31.1	69.2	19	7200	45.0	51	11.8	195.6	1	19	1701	152	1872	33.1
55	8	1.5	7200	34.7	44	21.1	99.9	3	7200	40.8	44	7.3	162.1	2	0	2983	260	3243	32.7
30	2	3.0	7200	18.0	31	42.1	775.0	170	7200	28.7	31	7.4	1155.4	9	160	509	82	751	6.1
30	5	3.0	7200	12.8	19	32.4	1040	59	64.3	19.0	19	-	14.5	2	43	476	53	572	1.2
30	8	3.0	7200	12.4	16	22.5	982.9	32	0.4	16.0	16	-	0.2	8	7	2	0	9	0.0
35	2	3.0	7200	14.1	35	59.6	616.4	121	7200	32.6	35	6.9	451.5	3	228	1389	156	1773	9.6
35	5	3.0	7200	13.2	22	39.8	588.6	14	762	22.0	22	-	19.5	6	52	3491	214	3757	9.5
35	8	3.0	7200	13.0	18	27.8	665.1	39	2	18.0	18	-	0.5	2	13	128	9	150	0.3
40	2	3.0	7200	16.7	38	56.0	244.0	161	7200	34.8	38	8.4	168.3	4	294	1861	36	2191	12.2
40	5	3.0	7200	16.0	26	38.4	361.8	91	7200	23.9	26	8.1	133.1	1	75	3675	975	4725	37.3
40	8	3.0	7200	14.5	22	34.1	402.4	8	423	22.0	22	-	92.5	2	0	628	30	658	2.5
45	2	3.0	7200	17.4	44	60.5	69.8	289	7200	39.7	44	9.8	125.2	1	360	1848	132	2340	24.0
45	5	3.0	7200	16.5	31	46.8	198.1	87	7200	24.8	31	20.0	23.3	2	64	6931	3147	10142	92.3
45	8	3.0	7200	16.5	25	34.1	276.9	21	416	25.0	25	-	16.1	3	30	1927	352	2309	16.4
50	2	3.0	7200	19.9	50	60.2	75.1	251	7200	45.5	50	9.0	34.9	1	490	2134	315	2939	45.4
50	5	3.0	7200	17.3	31	44.2	146.6	54	7200	26.6	33	19.4	37.6	1	108	6303	1650	8061	83.5
50	8	3.0	7200	17.0	28	39.3	155.6	6	7200	23.5	28	16.1	50.3	4	11	7749	1387	9147	63.1
55	2	3.0	7200	20.4	56	63.7	29.2	297	7200	48.0	57	15.8	0.0	0	609	1827	0	2436	11.5
55	5	3.0	7200	18.9	39	51.7	61.2	115	7200	29.8	41	27.3	7.9	0	140	9647	5309	15096	260.3
55	8	3.0	7200	18.8	32	41.4	99.7	10	7200	25.2	32	21.3	62.7	5	24	6806	1554	8384	114.9
30	2	4.5	7200	15.8	24	34.2	1251	124	10.7	24.0	24	-	0.9	1	141	333	18	492	0.8
30	5	4.5	7200	10.2	14	27.3	1282	73	1.3	14.0	14	-	0.6	1	49	0	0	49	0.0
30	8	4.5	7200	8.6	11	21.6	970.5	94	0.3	11.0	11	-	0.1	1	18	0	0	18	0.0
35	2	4.5	7200	19.3	32	39.7	253.8	745	7200	31.0	32	3.1	433.1	6	165	1052	160	1377	9.1
35	5	4.5	7200	10.9	18	39.6	717.0	99	13.3	18.0	18	-	2.6	4	69	6	1	76	0.1
35	8	4.5	7200	9.4	12	21.7	732.0	99	0.5	12.0	12	-	0.2	4	15	0	0	15	0.0
40	2	4.5	7200	13.9	35	60.4	284.2	173	7200	33.5	35	4.3	416.8	6	227	1031	139	1397	10.3
40	5	4.5	7200	10.7	19	43.6	348.9	108	1983	19.0	19	-	206.9	2	103	743	78	924	3.6
40	8	4.5	7200	10.2	14	27.4	432.8	96	0.9	14.0	14	-	0.2	1	16	0	0	16	0.0
45	2	4.5	7200	17.2	42	59.1	162.2	235	7200	39.0	42	7.1	220.7	3	286	1171	182	1639	15.9
45	5	4.5	7200	12.2	25	51.3	201.9	96	7200	20.7	25	17.2	215.9	3	121	2418	482	3021	23.8
45	8	4.5	7200	11.6	19	38.9	290.6	60	42.9	19.0	19	-	11.4	2	40	21	0	61	0.2
50	2	4.5	7200	12.9	47	72.6	54.5	203	7200	44.0	47	6.4	56.7	2	389	3198	585	4172	67.4
50	5	4.5	7200	12.6	29	56.7	117.0	127	7200	23.3	30	22.3	31.8	3	153	5319	3881	9353	128.1
50	8	4.5	7200	12.4	22	43.8	179.5	68	7200	18.7	22	15.0	510.3	3	33	1051	23	1107	8.4
55	2	4.5	7200	13.1	49	73.3	27.7	201	7200	46.0	49	6.1	50.3	1	548	1754	223	2525	41.9
55	5	4.5	7200	12.8	31	58.6	79.8	98	7200	23.0	32	28.1	10.7	0	199	12641	2448	15288	146.1
55	8	4.5	7200	13.0	24	45.8	130.5	20	7200	21.1	24	12.1	22.2	1	36	14472	517	15025	91.3

to optimality within the time limit by any of the two approaches, the average lower bound obtained by **Cplex** was 20.45 compared with 35.13 for **branch-and-cut**. Moreover, the average relative MIP gap went down from 48.12% with **Cplex** to 12.04% with **branch-and-cut**.

It is important to highlight that even if the **Cplex** automatic cut generation was not disabled in **branch-and-cut**, very few cuts were introduced through this feature when compared with **Cplex**. In average, 105 such cuts were added using **Cplex** while only 2 such cuts were added in **branch-and-cut**. This reveals that the vast majority of the cuts automatically added by **Cplex** might be actually dominated by the ones added through our separation routines. This presumption is enhanced by the fact that the cuts we add are often facet-defining inequalities. Moreover, our heuristics presented in Section 4.4 have proved to be quite effective in the search for violated cuts. Indeed, we added in average 2445.3  $k$ -cardinality tree inequalities (14) and 505.8 girth inequalities (15)-(16) during the whole optimization process. The relatively small number of generated girth inequalities generated is explained by the fact their separation routine is run only when no violated strong capacity inequality nor  $k$ -cardinality tree inequality is identified. This explains why no girth inequality was included for some instances (e.g., 55\_2\_3.0). The total number of cuts added was, in average, of 2983.4 cuts.

## 5. Concluding remarks

In this paper we have focused on USNP for which we have obtained some complexity results and new facet-defining inequalities that, combined with variable elimination and symmetry-breaking rules, have significantly improved the behavior of our branch-and-cut framework. These valid inequalities can be easily generalized to SNP, that is, when demands may request as many places as needed and the vehicles may perform several tours of the circuit. Intensive computational experiments would then definitely assess the impact that such inequalities would have on the performance of our branch-and-cut framework to solve SNP (and for other related problems such as the  $k$ -Edge Partitioning as well). Future research lines may also involve the investigation of reformulating our formulation to provide tighter linear relaxations. This might be a way to bypass the need of so many additional cuts to reinforce the formulation.

For the complexity aspect, we have showed that I-USNP for  $C = 2$  can be solved in polynomial time by reducing it to a maximum-cardinality matching problem. It follows that a complete description of the I-USNP

polytope for  $C = 2$  can be obtained from the matching polytope. The study of the projection of that polytope onto  $\mathbb{R}^{(m+n)p}$  (i.e., the  $(x, y)$ -space of natural variables) may reveal new families of facet-defining inequalities for SNP. Finally, the answer to our conjecture stating that the I-USNP is NP-Hard for  $C = 4$  even when  $G$  is restricted to be planar bipartite remains an open question that deserves attention.

## Acknowledgments

This work was funded by a public grant overseen by the French National Research Agency as part of the “Investissements d’Avenir” through the IMobS3 Laboratory of Excellence (ANR-10-LABX-0016) and the IDEX-ISITE initiative CAP 20-25 (ANR-16-IDEX-0001). Financial support was also received from the European Union through the European Regional Development Fund (ERDF) program and by the Auvergne-Rhône-Alpes region.

## Appendix A. Additional notations

Next we present the complete proofs of Theorems 6 and 9. Before doing so we introduce some additional notation and present two auxiliary lemmas.

**Definition 1** (Idleness). *Given an instance  $\mathcal{I} = (V, E, C)$  of USNP and a solution  $(\bar{x}, \bar{y})$  in  $\text{USNP}(V, E, C)$ , vehicle  $i \in K$  is said to be perfectly idle if  $\bar{x}_e^i = 0$  for every  $e \in E$  and  $\bar{y}_v^i = 0$  for every  $v \in V$ .*

**Definition 2** (Sequential permutation). *Given an instance  $\mathcal{I} = (V, E, C)$  of USNP, a solution  $(\bar{x}, \bar{y})$  in  $\text{USNP}(V, E, C)$ , and a subset of vehicles  $S = \{s_1, \dots, s_{|S|}\} \subseteq K$ , the solution  $(\hat{x}, \hat{y})$  in  $\text{USNP}(V, E, C)$  is said to be a sequential permutation of  $(\bar{x}, \bar{y})$  over  $S$  if for every  $i \in K$  and every  $e \in E$*

$$\hat{x}_e^i = \begin{cases} \bar{x}_e^{s_{\ell+1}} & \text{if } i = s_\ell \text{ with } \ell \in \{1, \dots, |S| - 1\}, \\ \bar{x}_e^{s_1} & \text{if } i = s_{|S|}, \\ \bar{x}_e^i & \text{if } i \notin S, \end{cases}$$

and for every  $i \in K$  and every  $v \in V$

$$\hat{y}_v^i = \begin{cases} \bar{y}_v^{s_{\ell+1}} & \text{if } i = s_\ell \text{ with } \ell \in \{1, \dots, |S| - 1\}, \\ \bar{y}_v^{s_1} & \text{if } i = s_{|S|}, \\ \bar{y}_v^i & \text{if } i \notin S. \end{cases}$$

Notice that if  $S = K$ , a solution  $\{\hat{E}_1, \dots, \hat{E}_p\}$  is said to be a sequential permutation of solution  $\{\bar{E}_1, \dots, \bar{E}_p\}$  if for every  $i \in \{1, \dots, p\}$ , we have  $\hat{E}_i = \bar{E}_{i+1}$ . (All the indices are interpreted modulo  $p$ .)

## Appendix B. Auxiliary lemmas

**Lemma 1.** *Given an instance  $(V, E, C)$  of USNP, a station  $v' \in V$ , and a vehicle  $i' \in K$ , let  $(\bar{x}, \bar{y})$  in  $\text{USNP}(V, E, C)$  with  $\bar{y}_{v'}^{i'} = 0$ , and let  $(\hat{x}, \hat{y})$  be the solution in  $\text{USNP}(V, E, C)$  such that  $\hat{x} = \bar{x}$  and for every  $i \in K$  and every  $v \in V$ ,*

$$\hat{y}_v^i = \begin{cases} \bar{y}_v^i & \text{if } i \in K \setminus i', \\ \hat{y}_v^{i'} & \text{if } i = i' \text{ and } v \in V \setminus v', \\ 1 & \text{if } i = i' \text{ and } v = v'. \end{cases}$$

*If both  $(\bar{x}, \bar{y})$  and  $(\hat{x}, \hat{y})$  belong to the same face*

$$\mathcal{F}' = \{(x, y) \in \text{USNP}(V, E, C) : \alpha^T x + \beta^T y = \gamma\},$$

*then  $\beta_{v'}^{i'} = 0$ .*

*Proof.* If  $(\bar{x}, \bar{y}) \in \mathcal{F}'$ , the following equality holds

$$\sum_{e \in E} \sum_{i \in K} \alpha_e^i \bar{x}_e^i + \sum_{v \in V} \sum_{i \in K} \beta_v^i \bar{y}_v^i = \gamma. \quad (\text{B.1})$$

If  $(\hat{x}, \hat{y}) \in \mathcal{F}'$ , the following equality holds

$$\sum_{e \in E} \sum_{i \in K} \alpha_e^i \hat{x}_e^i + \sum_{v \in V} \sum_{i \in K} \beta_v^i \hat{y}_v^i + \beta_{v'}^{i'} = \gamma. \quad (\text{B.2})$$

By subtracting (B.1) from (B.2), one obtains  $\beta_{v'}^{i'} = 0$ .  $\square$

**Lemma 2.** *Given an instance  $(V, E, C)$  of USNP, a demand  $e' \in E$  and vehicle  $i' \in K$ , let  $(\bar{x}, \bar{y}) \in \text{USNP}(V, E, C)$  be a solution where vehicle  $i'$  is perfectly idle, and let  $(\hat{x}, \hat{y}) \in \text{USNP}(V, E, C)$  constructed by transferring demand  $e'$  that was served by some vehicle  $k \in K \setminus \{i'\}$  to vehicle  $i'$ . If both  $(\bar{x}, \bar{y})$  and  $(\hat{x}, \hat{y})$  belong to the same face*

$$\mathcal{F}' = \{(x, y) \in \text{USNP}(V, E, C) : \alpha^T x + \beta^T y = \gamma\},$$

*then  $\alpha_{e'}^{i'} + \beta_{o_{e'}}^{i'} + \beta_{d_{e'}}^{i'} = \alpha_{e'}^k$ .*

*Proof.* If  $(\bar{x}, \bar{y}) \in \mathcal{F}'$ , the following equality holds

$$\sum_{e \in E} \sum_{i \in K} \alpha_e^i \bar{x}_e^i + \sum_{v \in V} \sum_{i \in K} \beta_v^i \bar{y}_v^i = \gamma. \quad (\text{B.3})$$

If  $(\hat{x}, \hat{y}) \in \mathcal{F}'$ , the following equality holds

$$\sum_{e \in E} \sum_{i \in K} \alpha_e^i \hat{x}_e^i - \alpha_{e'}^k + \alpha_{e'}^{i'} + \sum_{v \in V} \sum_{i \in K} \beta_v^i \hat{y}_v^i + \beta_{o_{e'}}^{i'} + \beta_{d_{e'}}^{i'} = \gamma. \quad (\text{B.4})$$

By subtracting (B.3) from (B.4), one obtains  $\alpha_{e'}^{i'} + \beta_{o_{e'}}^{i'} + \beta_{d_{e'}}^{i'} = \alpha_{e'}^k$ .  $\square$

### Appendix C. Proof of Theorem 6.

**Theorem 6.** *The following holds:*

- (i) *Strong Capacity inequalities (11) (resp. (12)) are valid for USNP( $V, E, C$ ), and*
- (ii) *are facet-defining if and only if  $|\delta^-(v)|$  (resp.  $|\delta^+(v)|$ )  $\geq C + 1$ .*

*Proof.* Let  $i \in K$  be a vehicle and  $v \in V$  be a station. We first prove the validity of (11) and (12). If vehicle  $i$  does not stop at station  $v$  (i.e.,  $y_v^i = 0$ ) then it cannot serve any demand whose pick-up (or drop-off) station is  $v$ . If vehicle  $i$  stops at station  $v$  (i.e.,  $y_v^i = 1$ ), then its capacity only allows vehicle  $i$  to serve at most  $C$  demands whose pick-up (or drop-off) station is  $v$ . Inequalities (11) and (12) hence are valid for USNP( $V, E, C$ ).

We now prove the necessary and sufficient conditions for the face  $\mathcal{F}$  induced by inequality (11) associated with  $i$  and  $v$  to be a facet of USNP( $V, E, C$ ). The proof for the face induced by inequality (12) is analogous. Assume  $|\delta^-(v)| \leq C$ . Summing up the  $|\delta^-(v)|$  inequalities (6) associated with  $i$  and  $v$  then yields the inequality

$$\sum_{e \in \delta^-(v)} x_e^i - |\delta^-(v)| y_v^i \leq 0$$

which dominates (11) if  $|\delta^-(v)| < C$  since

$$\sum_{e \in \delta^-(v)} x_e^i - C y_v^i \leq \sum_{e \in \delta^-(v)} x_e^i - |\delta^-(v)| y_v^i \leq 0.$$

Notice that if  $C = |\delta^-(v)|$  then  $\mathcal{F} \subsetneq \{(x, y) \in \text{USNP}(V, E, C) : x_e^i - y_v^i = 0\} \subsetneq \text{USNP}(V, E, C)$  for all  $e \in \delta^-(v)$ . The necessity of  $|\delta^-(v)| \geq C + 1$  for inequality (11) to be facet-defining of USNP( $V, E, C$ ) thus is proven.

For the sufficiency of this condition, we prove that if  $|\delta^-(v)| \geq C + 1$ , then  $\mathcal{F}$  is an inclusionwise maximal proper face of USNP( $V, E, C$ ). More specifically, we show that if  $\mathcal{F} \subseteq \mathcal{F}'$ , where  $\mathcal{F}'$  is a proper face of USNP( $V, E, C$ ) defined as

$$\mathcal{F}' = \left\{ (x, y) \in \text{USNP}(V, E, C) : \sum_{e \in E} \sum_{k \in K} \alpha_e^k x_e^k + \sum_{u \in V} \sum_{k \in K} \beta_u^k y_u^k = \gamma \right\},$$

for  $\alpha \in \mathbb{R}^{mp}$ ,  $\beta \in \mathbb{R}^{np}$ , and  $\gamma \in \mathbb{R}$ , then  $\mathcal{F}'$  is induced by a linear combination of inequality (11) associated with  $i$  and  $v$ , and equations (4), that is,

$$\mathcal{F}' = \left\{ (x, y) \in \text{USNP}(V, E, C) : \sum_{e \in E} \sum_{k \in K} \lambda_e x_e^k + \sum_{e \in \delta^-(v)} \omega x_e^i - \omega C y_v^i = \sum_{e \in E} \lambda_e \right\},$$

for  $\omega \in \mathbb{R}$  and  $\lambda \in \mathbb{R}^m$ . This implies  $\mathcal{F} = \mathcal{F}'$ .

To do so, we follow this six-step process:

- (i) show all the components of vector  $\beta$ , but  $\beta_v^i$ , equal 0;
- (ii) for every demand  $e \in E$ , show all the components  $\alpha_e^k$ , but  $\alpha_e^i$ , of vector  $\alpha$ , equal  $\lambda_e$ ;
- (iii) for every demand  $e \in E$  not having  $v$  as its drop-off station and every vehicle  $k \in K$ , show all the components  $\alpha_e^k$  of vector  $\alpha$  equal  $\lambda_e$ ;
- (iv) for every two demands  $e_1$  and  $e_2$  in  $E$  having  $v$  as their drop-off station and every vehicle  $k \in K$  different from  $i$ , show  $\alpha_{e_1}^i - \lambda_{e_1} = \alpha_{e_2}^i - \lambda_{e_2}$ ;
- (v) for every demand  $e$  having  $v$  as its drop-off station and every vehicle  $k \in K$  different from  $i$ , show  $\alpha_e^i - \lambda_e = \omega = -\frac{\beta_v^i}{C}$ ;
- (vi) for every vehicle  $k \in K$  different from  $i$ , show  $\gamma = \sum_{e \in E} \lambda_e$ .

*Step (i)* Let  $(\bar{x}, \bar{y}) \in \text{USNP}(V, E, C)$  be such that vehicle  $i$  is perfectly idle. We immediately have  $(\bar{x}, \bar{y}) \in \mathcal{F}$ . Lemma 1 then yields

$$\beta_u^i = 0 \quad \forall u \in V \setminus \{v\}.$$

For every vehicle  $k \in K \setminus \{i\}$ , consider  $(\tilde{x}, \tilde{y})$  in  $\text{USNP}(V, E, C)$  such that  $k$  is perfectly idle and  $C$  demands of  $\delta^-(v)$  are served by vehicle  $i$ . Clearly  $(\tilde{x}, \tilde{y}) \in \mathcal{F}$ . Lemma 1 then implies  $\beta_u^k = 0$  for every  $u \in V$ . We therefore have

$$\beta_u^k = 0 \quad \forall u \in V, k \in K \setminus \{i\}.$$

Defining  $\omega = -\frac{\beta_v^i}{C}$ , we thus obtain

$$\mathcal{F}' = \left\{ (x, y) \in \text{USNP}(V, E, C) : \sum_{e \in E} \sum_{k \in K} \alpha_e^k x_e^k - \omega C y_v^i = \gamma \right\}.$$

*Step (ii)* For every vehicle  $k \in K \setminus \{i\}$ , consider  $(\tilde{x}, \tilde{y}) \in \mathcal{F}$  defined in Step (i). Since  $|\delta^-(v)| \geq C + 1$ , we may additionally suppose that demand  $e \in E$  is served by vehicle  $k' \in K \setminus \{i, k\}$ , that is,  $\tilde{x}_e^{k'} = 1$ . Lemma 2 then yields  $\alpha_e^k = \alpha_e^{k'}$ . Since vehicles  $k$  and  $k'$  were arbitrarily chosen in  $K \setminus \{i\}$  and demand  $e$  in  $E$ , we have

$$\alpha_e^k = \lambda_e \quad \forall e \in E, k \in K \setminus \{i\},$$

with  $\lambda_e \in \mathbb{R}$ . Face  $\mathcal{F}'$  can therefore be rewritten as

$$\mathcal{F}' = \left\{ (x, y) \in \text{USNP}(V, E, C) : \sum_{e \in E} \sum_{k \in K \setminus \{i\}} \lambda_e x_e^k + \sum_{e \in E} \alpha_e^i x_e^i - \omega C y_v^i = \gamma \right\}.$$

*Step (iii)* Consider  $(\bar{x}, \bar{y}) \in \mathcal{F}$  defined in Step (i). Suppose demand  $e \in E \setminus \delta^-(v)$  is served by vehicle  $k \in K \setminus \{i\}$ , that is,  $\bar{x}_e^k = 1$ . From Lemma 2, we obtain  $\alpha_e^i = \alpha_e^k$ . Consequently we have

$$\alpha_e^i = \lambda_e \quad \forall e \in E \setminus \delta^-(v)$$

and

$$\begin{aligned} \mathcal{F}' = \left\{ (x, y) \in \text{USNP}(V, E, C) : \sum_{e \in E} \sum_{k \in K \setminus \{i\}} \lambda_e x_e^k + \sum_{e \in E \setminus \delta^-(v)} \lambda_e x_e^i \right. \\ \left. + \sum_{e \in \delta^-(v)} \alpha_e^i x_e^i - \omega C y_v^i = \gamma \right\}. \end{aligned} \quad (\text{C.1})$$

*Step (iv)* Consider  $(\tilde{x}, \tilde{y}) \in \mathcal{F}$  defined in Step (i). Let  $S$  denote the set of  $C$  demands of  $\delta^-(v)$  that are served by vehicle  $i$ . Notice that  $\delta^-(v) \setminus S \neq \emptyset$ . Suppose that demand  $e_1 \in \delta^-(v) \setminus S$  is the only demand served by vehicle  $k_1 \in K \setminus \{i\}$ . Since  $\mathcal{F} \subseteq \mathcal{F}'$ , we have

$$\sum_{e \in E} \sum_{k \in K \setminus \{i\}} \lambda_e \tilde{x}_e^k + \sum_{e \in E \setminus \delta^-(v)} \lambda_e \tilde{x}_e^i + \sum_{e \in \delta^-(v)} \alpha_e^i \tilde{x}_e^i - \omega C \tilde{y}_v^i = \gamma.$$

Let  $(\hat{x}, \hat{y})$  be the solution obtained from  $(\tilde{x}, \tilde{y})$  by swapping demand  $e_1$  and any demand  $e_2 \in S$ . We clearly have  $(\hat{x}, \hat{y})$  also belong to  $\mathcal{F}$  and

$$\sum_{e \in E} \sum_{k \in K \setminus \{i\}} \lambda_e \hat{x}_e^k + \sum_{e \in E \setminus \delta^-(v)} \lambda_e \hat{x}_e^i + \sum_{e \in \delta^-(v)} \alpha_e^i \hat{x}_e^i - \omega C \hat{y}_v^i - \lambda_{e_1} + \alpha_{e_1}^i - \alpha_{e_2}^i + \lambda_{e_2} = \gamma.$$

From (C.1), we then deduce  $\alpha_{e_1}^i - \lambda_{e_1} = \alpha_{e_2}^i - \lambda_{e_2}$ . Consequently, we obtain

$$\alpha_{e_1}^i - \lambda_{e_1} = \alpha_{e_2}^i - \lambda_{e_2} \quad \forall e_1 \in \delta^-(v), e_2 \in \delta^-(v).$$

*Step (v)* Consider  $(\bar{x}, \bar{y}) \in \mathcal{F}$  defined in Step (i). We then have

$$\sum_{e \in E} \sum_{k \in K \setminus \{i\}} \lambda_e \bar{x}_e^k = \gamma. \quad (\text{C.2})$$

Suppose now that  $C$  demands of  $\delta^-(v)$  are served by vehicle  $k \in K \setminus \{i\}$ . Let  $S$  denote the set of these  $C$  demands. The solution obtained from  $(\bar{x}, \bar{y})$  by making all the demands in  $S$  be served by vehicle  $i$  clearly belongs to  $\mathcal{F}$ . We therefore have

$$\sum_{e \in E} \sum_{k \in K \setminus \{i\}} \lambda_e \bar{x}_e^k + \sum_{e \in S} \alpha_e^i - \sum_{e \in S} \lambda_e - \omega C = \gamma. \quad (\text{C.3})$$

Subtracting (C.2) from (C.3) yields

$$\sum_{e \in S} (\alpha_e^i - \lambda_e) - \omega C = 0.$$

Using Step (iv), this equation implies

$$\alpha_e^i - \lambda_e = \omega = -\frac{\beta_v^i}{C} \quad \forall e \in \delta^-(v)$$

and consequently

$$\mathcal{F}' = \left\{ (x, y) \in \text{USNP}(V, E, C) : \sum_{e \in E} \sum_{k \in K} \lambda_e x_e^k + \sum_{e \in \delta^-(v)} \omega x_e^i - \omega C y_v^i = \gamma \right\}.$$

*Step (vi)* Consider  $(\bar{x}, \bar{y}) \in \mathcal{F}$  defined in Step (i). Construct  $(p-2)$  solutions, starting from  $(\bar{x}, \bar{y}) \in \mathcal{F}$ , by sequentially permuting the vehicles in  $K \setminus \{i\}$ . Since vehicle  $i$  is perfectly idle in these  $(p-1)$  solutions, they all belong to  $\mathcal{F}$  and satisfy the following equation

$$\sum_{e \in E} \sum_{k \in K \setminus \{i\}} \lambda_e x_e^k = \gamma. \quad (\text{C.4})$$

Notice that this construction guarantees that no demand in  $E$  is assigned twice to the same vehicle of  $K \setminus \{i\}$  in these  $(p-1)$  solutions. Summing up the equations associated with these  $(p-1)$  solutions then yields

$$(p-1) \sum_{e \in E} \lambda_e = (p-1)\gamma,$$

that is,  $\gamma = \sum_{e \in E} \lambda_e$ . We therefore deduce

$$\mathcal{F}' = \left\{ (x, y) \in \text{USNP}(V, E, C) : \sum_{e \in E} \sum_{k \in K} \lambda_e x_e^k + \sum_{e \in \delta^-(v)} \omega x_e^i - \omega C y_v^i = \sum_{e \in E} \lambda_e \right\}$$

and thereby  $\mathcal{F}' = \mathcal{F}$ .  $\square$



## Appendix D. Proof of Theorem 9

**Theorem 9.** *The  $k$ -cardinality tree inequalities (14) are facet-defining if and only if*

- (i.) *there does not exist a node  $u \in V(S)$  such that  $S \subset \delta_G^-(u)$  or  $S \subset \delta_G^+(u)$ , and*
- (ii.) *there does not exist an edge  $(u_1, u_2) \in E \setminus S$  such that  $u_1$  and  $u_2$  are the only internal nodes of  $G(S)$ .*

*Proof.* Let  $i \in K$  be a vehicle,  $v \in V$  be a station, and  $S \subseteq \Delta_E(v)$  with  $S \in \mathcal{CT}(G, C+1)$ .

We first prove the necessity. Suppose there exists  $u \in V(S)$  such that  $S \subset \delta_G^-(u)$  or  $S \subset \delta_G^+(u)$ , that is,  $G(S)$  is a star with internal node  $u$ . Since  $\deg_{G(S)}(u) = C+1$ , inequality (14) reads

$$\sum_{e \in S} x_e^i - C y_u^i \leq 0$$

which is dominated by either (11) or (12).

Suppose there exists an edge  $e' = (u_1, u_2) \in E \setminus S$  with  $u_1$  and  $u_2$  being the only internal nodes of  $G(S)$ . (See Figure D.11.) By definition,  $\deg_{G(S)}(u_1) - 1 + \deg_{G(S)}(u_2) - 1 = C$ . Since  $S \cup e' \subseteq \Delta_E(v)$  and hence  $x_{e'}^i + \sum_{e \in S} x_e^i \leq C$ , the inequality

$$x_{e'}^i + \sum_{e \in S} x_e^i - \sum_{u \in V(S)} (\deg_{G(S)}(u) - 1) y_u^i \leq 0$$

is valid and clearly dominates (14). (Notice that such idea can be used to lift inequalities (14) whenever Condition (ii.) does not hold.)

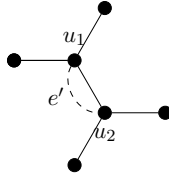


Figure D.11: Nonfacet-inducing  $k$ -cardinality tree where  $C = 4$ ; solid edges belong to  $S$ , dashed edge does not.

Let  $\mathcal{F}$  be the face of  $\text{USNP}(V, E, C)$  induced by inequality (14) associated with  $i, v$ , and  $S$ . For the sufficiency of Conditions (i.) and (ii.), we show that

if  $\mathcal{F} \subseteq \mathcal{F}'$ , where  $\mathcal{F}'$  is a proper face of  $\text{USNP}(V, E, C)$  defined as

$$\mathcal{F}' = \left\{ (x, y) \in \text{USNP}(V, E, C) : \sum_{e \in E} \sum_{k \in K} \alpha_e^k x_e^k + \sum_{u \in V} \sum_{k \in K} \beta_u^k y_u^k = \gamma \right\},$$

for  $\alpha \in \mathbb{R}^{mp}$ ,  $\beta \in \mathbb{R}^{np}$ , and  $\gamma \in \mathbb{R}$ , then  $\mathcal{F}'$  is induced by a linear combination of inequality (14) associated with  $i, v$ , and  $S$  and equations (4), that is,

$$\mathcal{F}' = \left\{ (x, y) \in \text{USNP}(V, E, C) : \sum_{e \in E} \sum_{k \in K} \lambda_e x_e^k + \sum_{e \in S} \omega x_e^i - \sum_{u \in V(S)} \omega (\deg_{G(S)}(u) - 1) y_u^i = \sum_{e \in E} \lambda_e \right\},$$

for  $\omega \in \mathbb{R}$  and  $\lambda \in \mathbb{R}^m$ . This implies  $\mathcal{F} = \mathcal{F}'$ .

We first notice that if there exists  $u \in V(S)$  such that  $S = \delta_G^-(u)$  or  $S = \delta_G^+(u)$ , then inequality (14) is nothing but inequality (11) or (12) associated with station  $u$  and vehicle  $i$ . By Theorem Appendix C,  $\mathcal{F}$  thus is a facet of  $\text{USNP}(V, E, C)$ . So, in the remainder of the proof, Condition (i.) corresponds to  $G(S)$  not being a star. We now follow this seven-step process:

- (i) for every station  $u \in V$  and every vehicle  $k \in K$  different from  $i$ , show  $\beta_u^k = 0$ ;
- (ii) for every demand  $e \in E$  and every two vehicles  $k$  and  $k'$  in  $K$  different from  $i$ , show  $\alpha_e^k = \alpha_e^{k'}$ ;
- (iii) for every station  $u \in V$  not being an internal node of  $G(S)$ , show  $\beta_u^i = 0$ ;
- (iv) for every vehicle  $k \in K$  different from  $i$ , show  $\gamma = \sum_{e \in E} \alpha_e^k$ ;
- (v) for every two demands  $e_1$  and  $e_2$  in  $S$  and every vehicle  $k \in K$  different from  $i$ , show  $\alpha_{e_1}^i - \alpha_{e_1}^k = \alpha_{e_2}^i - \alpha_{e_2}^k$ ;
- (vi) for every demand  $e \in E \setminus S$  and every vehicle  $k \in K$ , show  $\alpha_e^i = \alpha_e^k$ ;
- (vii) for every station  $u \in V(S)$ , every demand  $e \in S$ , and every vehicle  $k \in K$  different from  $i$ , show  $\beta_u^i = -(\deg_{G(S)}(u) - 1)(\alpha_e^i - \alpha_e^k)$ .

*Step (i)* For every vehicle  $k \in K \setminus \{i\}$  and every demand  $e' \in S$ , consider  $(\bar{x}, \bar{y}) \in \text{USNP}(V, E, C)$  such that vehicle  $k$  is perfectly idle and all the demands in  $S \setminus \{e'\}$  are served by vehicle  $i$ . We immediately have  $(\bar{x}, \bar{y}) \in \mathcal{F}$ . Lemma 1 then yields  $\beta_u^k = 0$  for every  $u \in V$ . We therefore deduce

$$\beta_u^k = 0 \quad \forall u \in V, k \in K \setminus \{i\}.$$

Face  $\mathcal{F}'$  can now be rewritten as

$$\mathcal{F}' = \left\{ (x, y) \in \text{USNP}(V, E, C) : \sum_{e \in E} \sum_{k \in K} \alpha_e^k x_e^k + \sum_{u \in V} \beta_u^i y_u^i = \gamma \right\}.$$

*Step (ii)* For every vehicle  $k \in K \setminus \{i\}$  and every demand  $e' \in S$ , consider  $(\bar{x}, \bar{y}) \in \text{USNP}(V, E, C)$  defined in Step (i). We may additionally suppose that demand  $e \in E$  is served by vehicle  $k' \in K \setminus \{i, k\}$ , that is,  $\tilde{x}_e^{k'} = 1$ . Lemma 2 then yields  $\alpha_e^k = \alpha_e^{k'}$ . Since vehicles  $k$  and  $k'$  were arbitrarily chosen in  $K \setminus \{i\}$  and demand  $e$  in  $E$ , we have

$$\alpha_e^k = \lambda_e \quad \forall e \in E, k \in K \setminus \{i\},$$

with  $\lambda_e \in \mathbb{R}$ . Face  $\mathcal{F}'$  can therefore be rewritten as

$$\mathcal{F}' = \left\{ (x, y) \in \text{USNP}(V, E, C) : \sum_{e \in E} \sum_{k \in K \setminus \{i\}} \lambda_e x_e^k + \sum_{e \in E} \alpha_e^i x_e^i + \sum_{u \in V} \beta_u^i y_u^i = \gamma \right\}.$$

*Step (iii)* Let  $I(S)$  denote the set of internal nodes of  $G(S)$ , that is,  $I(S) = \{u \in V(S) : \deg_{G(S)} u \geq 2\}$ . Let  $(\tilde{x}, \tilde{y}) \in \text{USNP}(V, E, C)$  be such that vehicle  $i$  is perfectly idle. We immediately have  $(\tilde{x}, \tilde{y}) \in \mathcal{F}$ . For every station  $u \in V(S) \setminus I(S)$ , Lemma 1 yields  $\beta_u^i = 0$ . Therefore, face  $\mathcal{F}'$  can now be rewritten as

$$\mathcal{F}' = \left\{ (x, y) \in \text{USNP}(V, E, C) : \sum_{e \in E} \sum_{k \in K \setminus \{i\}} \lambda_e x_e^k + \sum_{e \in E} \alpha_e^i x_e^i + \sum_{u \in I(S)} \beta_u^i y_u^i = \gamma \right\}.$$

*Step (iv)* Consider  $(\tilde{x}, \tilde{y}) \in \mathcal{F}$  defined in Step (iii). Construct  $(p-2)$  solutions, starting from  $(\tilde{x}, \tilde{y}) \in \mathcal{F}$ , by sequentially permuting the vehicles in  $K \setminus \{i\}$ . Since vehicle  $i$  is perfectly idle in these  $(p-1)$  solutions, they all belong to  $\mathcal{F}$  and satisfy the following equation

$$\sum_{e \in E} \sum_{k \in K \setminus \{i\}} \lambda_e x_e^k = \gamma.$$

Notice that this construction guarantees that no demand in  $E$  is assigned twice to the same vehicle of  $K \setminus \{i\}$  in these  $(p-1)$  solutions. Summing up the equations associated with these  $(p-1)$  solutions then yields

$$(p-1) \sum_{e \in E} \lambda_e = (p-1)\gamma,$$

that is,  $\gamma = \sum_{e \in E} \lambda_e$ . We therefore deduce

$$\mathcal{F}' = \left\{ (x, y) \in \text{USNP}(V, E, C) : \sum_{e \in E} \sum_{k \in K \setminus \{i\}} \lambda_e x_e^k + \sum_{e \in E} \alpha_e^i x_e^i + \sum_{u \in I(S)} \beta_u^i y_u^i = \sum_{e \in E} \lambda_e \right\}.$$

*Step (v)* For every demand  $e_1 \in S$ , let  $(\hat{x}, \hat{y}) \in \text{USNP}(V, E, C)$  be such that all the demands in  $S \setminus \{e_1\}$  are served by vehicle  $i$ . Suppose that demand  $e_1$  is the only demand served by vehicle  $k \in K \setminus \{i\}$ . Since  $(\hat{x}, \hat{y}) \in \mathcal{F}$ , the equation

$$\sum_{e \in E} \sum_{k \in K \setminus \{i\}} \lambda_e \hat{x}_e^k + \sum_{e \in E} \alpha_e^i \hat{x}_e^i + \sum_{u \in I(S)} \beta_u^i \hat{y}_u^i = \sum_{e \in E} \lambda_e$$

holds. Considering the solution in  $\mathcal{F}$ , obtained from  $(\hat{x}, \hat{y})$  by swapping  $e_1$  and any demand  $e_2 \in S \setminus \{e_1\}$  yields

$$\sum_{e \in E} \sum_{k \in K \setminus \{i\}} \lambda_e \hat{x}_e^k + \sum_{e \in E} \alpha_e^i \hat{x}_e^i + \sum_{u \in I(S)} \beta_u^i \hat{y}_u^i - \lambda_{e_1} + \alpha_{e_1}^i - \alpha_{e_2}^i + \lambda_{e_2} = \sum_{e \in E} \lambda_e.$$

holds. Therefore  $\alpha_{e_1}^i - \lambda_{e_1} = \alpha_{e_2}^i - \lambda_{e_2}$ . Consequently, we obtain

$$\alpha_e^i = \lambda_e + \omega \quad \forall e \in S$$

with  $\omega \in \mathbb{R}$ , and face  $\mathcal{F}'$  can then be rewritten as

$$\mathcal{F}' = \left\{ (x, y) \in \text{USNP}(V, E, C) : \sum_{e \in E} \sum_{k \in K \setminus \{i\}} \lambda_e x_e^k + \sum_{e \in E \setminus S} \alpha_e^i x_e^i + \sum_{e \in S} (\lambda_e + \omega) x_e^i + \sum_{u \in I(S)} \beta_u^i y_u^i = \sum_{e \in E} \lambda_e \right\}.$$

*Step (vi)* Let  $e'$  be any demand in  $E \setminus S$ . We consider three cases depending on whether  $e'$  has zero, one, or two stations of  $I(S)$  appearing as its endpoints (i.e.,  $|I(S) \cap \{o_{e'}, d_{e'}\}|$  equals 0, 1, or 2).

*Case 1:* Demand  $e'$  is incident with no internal nodes of  $G(S)$ . Consider  $(\tilde{x}, \tilde{y}) \in \mathcal{F}$  defined in Step (iii). The equation

$$\sum_{e \in E} \sum_{k \in K \setminus i} \lambda_e \hat{x}_e^k = \sum_{e \in E} \lambda_e$$

clearly holds. Suppose that demand  $e'$  is served by vehicle  $k \in K \setminus \{i\}$ . The solution, obtained from  $(\tilde{x}, \tilde{y})$  by making demand  $e'$  be served by vehicle  $i$ , still belongs to  $\mathcal{F}$ , and hence

$$\sum_{e \in E} \sum_{k \in K \setminus i} \lambda_e \tilde{x}_e^k + \alpha_{e'}^i - \lambda_{e'} = \sum_{e \in E} \lambda_e.$$

Therefore,  $\alpha_{e'}^i = \lambda_{e'}$  for every demand  $e' \in E \setminus S$  such that  $I(S) \cap \{o_{e'}, d_{e'}\} = \emptyset$ .

*Case 2:* Demand  $e'$  is incident with exactly one internal node  $u$  of  $G(S)$ . Consider any leaf node  $r \in V(S)$  such that  $u$  and  $r$  are not adjacent in  $G(S)$ . Since  $G(S)$  is not a star, node  $r$  always exists. Thinking about  $G(S)$  as a tree rooted at  $r$ , let  $T_u$  denote the subtree of  $G(S)$  rooted at  $u$ . We clearly have  $|E(T_u)| < C$ . Let  $(\check{x}, \check{y}) \in \text{USNP}(V, E, C)$  be such that the only demands served by vehicle  $i$  are those in  $E(T_u)$ . We then have

$$\sum_{e \in S} \check{x}_e^i = |E(T_u)|$$

and since  $u \in I(S)$ ,

$$\sum_{t \in V(S)} (\deg_{G(S)}(t) - 1) \check{y}_t^i = \sum_{t \in V(T_u)} (\deg_{G(S)}(t) - 1) = |E(T_u)|.$$

Consequently,  $(\check{x}, \check{y}) \in \mathcal{F}$  and the equation

$$\sum_{e \in E} \sum_{k \in K \setminus \{i\}} \lambda_e \check{x}_e^k + \sum_{e \in E \setminus S} \alpha_e^i \check{x}_e^i + \sum_{e \in S} (\lambda_e + \omega) \check{x}_e^i + \sum_{t \in I(S)} \beta_t^i \check{y}_t^i = \sum_{e \in E} \lambda_e$$

holds. Suppose that demand  $e'$  is served by vehicle  $k \in K \setminus \{i\}$ . Since  $|E(T_u)| < C$ , the solution, obtained from  $(\check{x}, \check{y})$  by making demand  $e'$  be served by vehicle  $i$ , still belongs to  $\mathcal{F}$ , and hence

$$\sum_{e \in E} \sum_{k \in K \setminus \{i\}} \lambda_e \check{x}_e^k + \sum_{e \in E \setminus S} \alpha_e^i \check{x}_e^i + \sum_{e \in S} (\lambda_e + \omega) \check{x}_e^i + \sum_{t \in I(S)} \beta_t^i \check{y}_t^i + \alpha_{e'}^i - \lambda_{e'} = \sum_{e \in E} \lambda_e.$$

Thus  $\alpha_{e'}^i = \lambda_{e'}$  for every demand  $e' \in E \setminus S$  such that  $|I(S) \cap \{o_{e'}, d_{e'}\}| = 1$ .  
*Case 3:* Demand  $e'$  is incident with two internal nodes  $u$  and  $u'$  of  $G(S)$ . Suppose first that  $e'$  is parallel to no demands in  $S$ . Let  $P$  denote the path between  $u$  and  $u'$  in  $G(S)$ . Path  $P$  clearly has at least one internal node. Thinking about  $G(S)$  as a tree rooted at any internal node of  $P$ , let  $T_u$  and  $T_{u'}$  be the subtrees of  $G(S)$  rooted at  $u$  and  $u'$ , respectively. Since  $|E(T_u)| + |E(T_{u'})| < C$ , following a similar approach as in Case 2 with respect to the solution in  $\text{USNP}(V, E, C)$  where the only demands served by vehicle  $i$  are those in  $E(T_u) \cup E(T_{u'})$  yields  $\alpha_{e'}^i = \lambda_{e'}$  for every demand  $e' \in E \setminus S$  such that  $\{o_{e'}, d_{e'}\} \subseteq I(S)$  and  $e'$  is parallel to no demands in  $S$ .

Suppose now that there exists a demand in  $S$  whose endpoints are  $u$  and  $u'$ . Consider any node  $r \in I(S) \setminus \{u, u'\}$ . By Condition (ii.), node  $r$  exists. W.l.o.g., suppose that  $u$  is closer to  $r$  in  $G(S)$  than  $u'$  is. Since  $|E(T_u)| < C$ , following a similar approach as in Case 2 with respect to the solution in  $\text{USNP}(V, E, C)$  where the only demands served by vehicle  $i$  are those in  $E(T_u)$  yields  $\alpha_{e'}^i = \lambda_{e'}$  for every demand  $e' \in E \setminus S$  such that  $\{o_{e'}, d_{e'}\} \subseteq I(S)$  and  $e'$  is parallel a demand in  $S$ . This completes Step (vi) and face  $\mathcal{F}'$  can now be rewritten as

$$\mathcal{F}' = \left\{ (x, y) \in \text{USNP}(V, E, C) : \sum_{e \in E} \sum_{k \in K} \lambda_e x_e^k + \sum_{e \in S} \omega x_e^i + \sum_{u \in I(S)} \beta_u^i y_u^i = \sum_{e \in E} \lambda_e \right\}.$$

*Step (vii)* Consider any leaf node  $r$  of  $G(S)$  and think about  $G(S)$  as a tree rooted at  $r$ . Let  $u'$  be the parent in  $G(S)$  of a deepest leaf node of  $G(S)$  and  $T_{u'}$  be the subtree of  $G(S)$  rooted at  $u'$ . Notice that  $T_{u'}$  is star and since  $G(S)$  is not a star,  $|E(T_{u'})| = \deg_{G(S)}(u') - 1 < C$ . Let  $(\dot{x}, \dot{y}) \in \text{USNP}(V, E, C)$  be such that the only demands served by vehicle  $i$  are those in  $E(T_{u'})$ . We then have  $(\dot{x}, \dot{y}) \in \mathcal{F}$  and the equation

$$\sum_{e \in E \setminus E(T_{u'})} \sum_{k \in K \setminus \{i\}} \lambda_e \dot{x}_e^k + \sum_{e \in E(T_{u'})} (\lambda_e + \omega) + \beta_{u'}^i = \sum_{e \in E} \lambda_e$$

holds. Constructing  $(p-2)$  solutions from  $(\dot{x}, \dot{y})$  by sequentially permuting the vehicles in  $K \setminus \{i\}$  and applying the same reasoning as in Step (iv), we obtain

$$(p-1) \sum_{e \in E \setminus E(T_{u'})} \lambda_e + (p-1) \sum_{e \in E(T_{u'})} (\lambda_e + \omega) + (p-1) \beta_{u'}^i = (p-1) \sum_{e \in E} \lambda_e,$$

that is,

$$\beta_{u'}^i = - \sum_{e \in E(T_{u'})} \omega = -|E(T_{u'})| \omega = -(\deg_{G(S)}(u') - 1) \omega.$$

For any internal node  $u$  of  $G(S)$ , let  $T_u$  denote the subtree of  $G(S)$  rooted at  $u$  and  $C(u)$  denote the set of nodes adjacent to  $u$  in  $T_u$ . We clearly have

$$E(T_u) = \bigcup_{c \in C(u)} (E(T_c) \cup \{(u, c)\}).$$

We now use this decomposition of the edge sets of the subtrees of  $G(S)$  rooted at the internal nodes and a similar approach as above for determining  $\beta_{u'}^i$  to recursively obtain  $\beta_u^i$  for  $u \in I(S)$ . Starting from the internal nodes of  $G(S)$  only having leaf nodes in their rooted subtree and moving towards the only internal node adjacent to  $r$  in  $G(S)$ , we deduce for any internal node  $u \in I(S)$

$$\beta_u^i + \sum_{c \in C(u)} \sum_{u' \in V(T_c)} \beta_{u'}^i = -(\deg_{G(S)}(u) - 1)\omega - \sum_{c \in C(u)} \sum_{u' \in V(T_c)} (\deg_{G(S)}(u') - 1)\omega.$$

that is,  $\beta_u^i = -(\deg_{G(S)}(u) - 1)\omega$  for every  $u \in V(S)$ . Therefore, face  $\mathcal{F}'$  can finally be rewritten as

$$\mathcal{F}' = \left\{ (x, y) \in \text{USNP}(V, E, C) : \sum_{e \in E} \sum_{k \in K} \lambda_e x_e^k + \sum_{e \in S} \omega x_e^i - \sum_{u \in V(S)} \omega (\deg_{G(S)}(u) - 1) y_u^i = \sum_{e \in E} \lambda_e \right\}$$

which concludes the proof.  $\square$

## References

- Baïou, M., Colares, R., Kerivin, H., 2021. The complexity of the Unit Stop Number Problem and its implications to other related problems. URL: <https://hal.archives-ouvertes.fr/hal-03120087>. preprint.
- Braekers, K., Caris, A., Janssens, G.K., 2014. Exact and meta-heuristic approach for a general heterogeneous dial-a-ride problem with multiple depots. *Transportation Research Part B: Methodological* 67, 166–186.
- Bsaybes, S., Quilliot, A., Wagler, A.K., 2018. Fleet management for autonomous vehicles using multicommodity coupled flows in time-expanded networks, in: 17th International Symposium on Experimental Algorithms (SEA 2018), Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.

- Carvajal, R., Ahmed, S., Nemhauser, G., Furman, K., Goel, V., Shao, Y., 2014. Using diversification, communication and parallelism to solve mixed-integer linear programs. *Operations Research Letters* 42, 186–189.
- Chvatal, V., 1973. Edmonds polytopes and a hierarchy of combinatorial problems. *Discrete mathematics* 4, 305–337.
- Colares, R., 2019. Exploring combinatorial aspects of the stop number problem. Ph.D. thesis. University Clermont Auvergne, France.
- Cordeau, J.F., 2006. A branch-and-cut algorithm for the dial-a-ride problem. *Operations Research* 54, 573–586.
- Cordeau, J.F., Laporte, G., 2007. The dial-a-ride problem: models and algorithms. *Annals of operations research* 153, 29–46.
- Denton, B.T., Miller, A.J., Balasubramanian, H.J., Huschka, T.R., 2010. Optimal allocation of surgery blocks to operating rooms under uncertainty. *Operations research* 58, 802–816.
- Dyer, M.E., Frieze, A.M., 1986. Planar 3dm is np-complete. *Journal of Algorithms* 7, 174–184.
- EasyMile, 2015. Ez10 passenger shuttle. [www.easymile.com/vehicle-solutions/ez10-passenger-shuttle](http://www.easymile.com/vehicle-solutions/ez10-passenger-shuttle).
- Fagnant, D.J., Kockelman, K., 2015. Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. *Transportation Research Part A: Policy and Practice* 77, 167–181.
- Fischetti, M., Hamacher, H.W., Jørnsten, K., Maffioli, F., 1994. Weighted k-cardinality trees: Complexity and polyhedral structure. *Networks* 24, 11–21.
- Fischetti, M., Monaci, M., 2014. Exploiting erraticism in search. *Operations Research* 62, 114–122.
- Garey, M.R., Johnson, D.S., 2002. Computers and intractability. volume 29. W.H. Freeman and Company, New York.
- Goldschmidt, O., Hochbaum, D.S., Levin, A., Olinick, E.V., 2003. The sonet edge-partition problem. *Networks: An International Journal* 41, 13–23.



- Gomory, R.E., et al., 1958. Outline of an algorithm for integer solutions to linear programs. *Bulletin of the American Mathematical society* 64, 275–278.
- Ho, S.C., Szeto, W., Kuo, Y.H., Leung, J.M., Petering, M., Tou, T.W., 2018. A survey of dial-a-ride problems: Literature review and recent developments. *Transportation Research Part B: Methodological* 111, 395–421.
- Kaibel, V., Peinhardt, M., Pfetsch, M.E., 2011. Orbitopal fixing. *Discrete Optimization* 8, 595–610.
- Kaibel, V., Pfetsch, M.E., 2008. Packing and partitioning orbitopes. *Mathematical Programming* 114, 1–36.
- Karp, R.M., 1972. Reducibility among combinatorial problems, in: *Complexity of computer computations*. Springer, pp. 85–103.
- Krishnamoorthy, M.S., 1975. An np-hard problem in bipartite graphs. *ACM SIGACT News* 7, 26–26.
- Liu, M., Luo, Z., Lim, A., 2015. A branch-and-cut algorithm for a realistic dial-a-ride problem. *Transportation Research Part B: Methodological* 81, 267–288.
- Margot, F., 2010. Symmetry in integer linear programming, in: *50 Years of Integer Programming 1958-2008*. Springer, pp. 647–686.
- Masuyama, S., Ibaraki, T., 1991. Chain packing in graphs. *Algorithmica* 6, 826–839.
- Ostrowski, J., Anjos, M.F., Vannelli, A., 2010. Symmetry in scheduling problems. *Citeseer*.
- Parragh, S.N., 2011. Introducing heterogeneous users and vehicles into models and algorithms for the dial-a-ride problem. *Transportation Research Part C: Emerging Technologies* 19, 912–930.
- Parragh, S.N., Doerner, K.F., Hartl, R.F., 2008. A survey on pickup and delivery problems. *Journal für Betriebswirtschaft* 58, 21–51.
- Pelletier, S., Jabali, O., Laporte, G., 2016. 50th anniversary invited article—goods distribution with electric vehicles: review and research perspectives. *Transportation Science* 50, 3–22.

- Pimenta, V., Quilliot, A., Toussaint, H., Vigo, D., 2017. Models and algorithms for reliability-oriented dial-a-ride with autonomous electric vehicles. *European Journal of Operational Research* 257, 601–613.
- Ravi, R., Sundaram, R., Marathe, M.V., Rosenkrantz, D.J., Ravi, S.S., 1996. Spanning trees—short or small. *SIAM Journal on Discrete Mathematics* 9, 178–200.
- Ropke, S., Cordeau, J.F., Laporte, G., 2007. Models and branch-and-cut algorithms for pickup and delivery problems with time windows. *Networks: An International Journal* 49, 258–272.
- Sabharwal, A., Samulowitz, H., Reddy, C., 2012. Guiding combinatorial optimization with uct, in: *International Conference on Integration of Artificial Intelligence (AI) and Operations Research (OR) Techniques in Constraint Programming*, Springer. pp. 356–361.
- Sherali, H.D., Smith, J.C., 2001. Improving discrete model representations via symmetry considerations. *Management Science* 47, 1396–1407.