Insights into the computational complexity of the single-source capacitated facility location problem with customer preferences

Christina Büsing Timo Gersing Sophia Wrede

Teaching and Research Area Combinatorial Optimization, RWTH Aachen University, Ahornstraße 55, 52074 Aachen, Germany.
{buesing,gersing,wrede}@combi.rwth-aachen.de

December 31, 2024

Abstract

Single-source capacitated facility location problems (SSCFLPs) are well known in the operations research literature. A set of facilities is opened and each customer is assigned to exactly one open facility so that the capacity at each facility is respected. This customer assignment, however, deprives customers from choosing facilities according to their individual preferences. If customers are allowed to deviate to their most preferred open facility within a solution for the SSCFLP, these re-assignments might lead to facilities serving more demand than they are capable of and, therefore, can render feasible solutions infeasible. Preference constraints take individual customer preferences into account before computing a solution and ensure that each customer is served at their most preferred open facility while respecting the capacity at each facility. This problem is called *single-source capacitated facility location problem with customer preferences* (*SSCFLPCP*) and is known to be strongly NP-hard as an extension of the SSCFLP, in which each customer is indifferent regarding all facilities. In this paper, we contribute new complexity results for the SSCFLPCP. We show that the type of customer preferences, e.g., geographically closest assignments or strict preferences, has a vital impact on the theoretical complexity of the considered instances. Afterwards, we focus on the special case in which preferences are defined by closest assignments and give several results for the SSCFLPCP which differ from complexity results known for the SSCFLP. Our findings provide a deeper understanding of the transition from polynomially solvable cases to strongly NP-hard cases of the computational complexity of the SSCFLPCP.

Keywords: facility location, capacities, preferences, computational complexity

1 Introduction

Facility location problems (FLPs) play an essential role in operations research literature (Laporte et al., 2019, Celik Turkoglu and Erol Genevois, 2020). In their basic version, facilities providing some sort of service for customers need to be opened at potential sites and customers are assigned to these facilities. The aim is to minimise the total cost consisting of opening costs for the facilities and costs for assigning the customers. We refer to this problem as the *uncapacitated facility location problem* (*UFLP*). Naturally, each customer is assigned to the open facility with smallest assignment cost. The class of FLPs is widely applicable to many real-world problems such as locating health care institutions (Ahmadi-Javid et al., 2017), charging stations for electrical cars (Ahmad et al., 2022), and many more (Celik Turkoglu and Erol Genevois, 2020).

There are no limitations on the service a facility can provide in UFLPs. However, such limitations are essential for many real-world problems, for example, when locating hospitals (Mestre et al., 2015). Often, customers incur a certain demand and facilities have a certain capacity for serving customers' demands. Such problems belong to the class of *capacitated facility location problems* (*CFLP*s). It is no longer guaranteed that each customer will be assigned to the open facility with lowest assignment cost when considering capacities. The problem is called *single-source* CFLP (SSCFLP) if each customer has to be assigned to exactly one open facility, that is, the demand can not be split.

If customers have an individual agenda regarding the facilities they want to be served at, then their assignment in the solution to the UFLP or CFLP might be in conflict with their preferences. In reality, however, customers then deviate to one of their most preferred open facilities. While this yields a feasible, although possibly worse, solution in the UFLP, it might turn a feasible solution of the (SS)CFLP infeasible. A real-world application for this class of location problems lies in the location of health care emergency centers. Here, facilities offer medical services, which are limited due to time and space. Customers' demands have to be served at these facilities without violating the facilities' capacities. In an emergency, it can be assumed that customers will seek the service of their closest open facility. Note that in the case where preferences coincide with closest assignments, any solution to the UFLP with closest assignments coincides with the classic UFLP.

We refer to the problem in which each customer has to be served at exactly one of their most preferred open facilities as the *(single-source capacitated) facility location problem with customer preferences* ((SSC)FLPCP). If preferences are defined by assignment costs, we call the problem *(single-source capacitated) facility location problem with closest assignments* ((SSC)FLPCA). In this work, we study the combinatorial structures of both the SSCFLPCP and the SSCFLPCA. The main contributions of this paper are the following results.

1. Determining feasibility of the SSCFLPCP can be done efficiently if each customer has a strict preference ordering of the facilities.
2. The SSCFLPCP with arbitrary strict preferences is strongly NP-hard, even when neglecting capacities and the assignment costs correspond to costs in an underlying path graph.
3. The SSCFLPCA on star graphs is weakly NP-hard if customer preferences are strict or customer demands are uniform.
4. The assignment problem of the SSCFLPCA with uniform customer demands and lower and upper bounds on the demand to be served at each open facility can be solved efficiently.
5. Determining feasibility of the SSCFLPCA with uniform customer demands and lower and upper bounds so that exactly two customers have to be served at each open facility is strongly NP-complete.

The addition of preference constraints has both positive and negative impact on the computational complexity of the UFLP and the SSCFLP. Contributions 1. and 3. show that the introduction of strict customer preferences or uniform customer demands can ease the computational complexity of the SSCFLP. Contributions 2. and 5. show that the introduction of customer preferences can increase the computational complexity compared to the traditional UFLP and SSCFLP, respectively. We can therefore conclude that customer preferences have to be considered with care.

The remainder of this paper is organized as follows. Section 2 provides an overview of related literature. A formal problem definition is given in Section 3, in which we also summarize used notation. In Section 4, we derive complexity results for the SSCFLPCP with strict preferences. In Section 5, we study the preference ordering defined by assignment costs and assume that distances are distinct. Afterwards, in Section 6, we allow indifferences within the customer preferences and set customer demands to one; this allows us to study the combinatorial structures of the location problem independent of the strong NP-hardness originating from the assignment problem within the SSCFLP. In Section 7, we consider the SSCFLPCA with lower bounds, uniform demands and capacities. Last but not least, we discuss our findings and future work in Section 8.

2 Related work

Facility location problems are thoroughly studied in the literature, see, e.g., Laporte et al. (2019), Celik Turkoglu and Erol Genevois (2020) for recent surveys. The *uncapacitated facility location problem* (UFLP) poses as the fundamental problem, which occurs in all facility location problems as a subproblem. From the complexity-theoretical point of view, it is easy to find a feasible solution for the UFLP but it is strongly NP-hard to find an optimal solution (Mirchandani and Francis, 1990). If the customers and facility locations are defined on an underlying graph, with assignment costs equating distances in the graph, then an optimal solution can be computed efficiently for the UFLP on tree graphs (Mirchandani and Francis, 1990). Conversely to the UFLP, finding a feasible solution for the *single-source capacitated facility location problem* (SSCFLP) is already strongly NP-hard; this can be seen via a reduction from 3-*partition* (Garey and Johnson, 1979). The relevance and complexity of FLPs has triggered a large number of scientific articles related to different aspects, such as exact solution methods (see, e.g., Avella and Boccia (2009), Görtz and Klose (2012), Fischetti et al. (2016)), the investigation of their polyhedral structure (see, e.g., Leung and Magnanti (1989), Aardal et al. (1995), Avella and Boccia (2009), Avella et al. (2021)) and heuristic solution methods (see, e.g., Mirchandani and Francis (1990), Korte and Vygen (2018)).

Most articles considering facility location problems with customer preferences focus on the uncapacitated case. The work of Hanjoul and Peeters (1987) is considered to be the first occurrence of preference constraints in the context of facility location problems. The authors propose an exact algorithm, which utilises a branch-and-bound procedure, as well as two heuristics for solving the UFLP with customer preferences. Several articles focus on preprocessing strategies and valid inequalities, see, e.g., Cánovas et al. (2007), Vasilyev et al. (2010) and Vasilyev et al. (2013). A semi-Lagrangian relaxation heuristic approach is proposed by Cabezas and García (2022).

Rojeski and ReVelle (1970), Wagner and Falkson (1975) and Gerrard and Church (1996) focus on the special case in which preferences are defined by distances, i.e., each customer prefers to be served at their closest open facility. Rojeski and ReVelle (1970) additionally consider capacities, allow to split customer demands and to expand capacities. Wagner and Falkson (1975) present models for the location of public facilities which maximise social welfare and propose a set of closest assignment constraints. Gerrard and Church (1996) review constraints for integer linear programming formulations for enforcing closest assignments and identify applications for facility location problems with closest assignment constraints. Espejo et al. (2012) theoretically compare all closest assignment constraints in the literature of discrete location theory up until 2011 and contribute a new set of constraints.

When studying CFLPs with customer preferences, there are two main approaches to dealing with overloaded facilities. In the first approach, the overloaded facility may be opened and some customers will be served at a facility they like less. In the second approach, each customer has to be served at their most preferred open facility and facilities that would be overloaded can never be opened. Note that whether a facility will be overloaded depends on which other facilities are opened. Most research revolves around the first approach (Casas-Ramírez et al., 2018, Calvete et al., 2020, Polino et al., 2023). Here, the authors consider variations of a bilevel setting where the leader opens facilities with the aim to minimise the total sum of opening and assignment costs; the follower assigns each customer, for which a ranking of all potential facilities is known, to the open facilities with the objective to optimise the sum of achieved preference rankings of the customers. Calvete et al. (2020) compare computational results of the first and second approach to dealing with overloaded facilities. In the second approach, instances are solvable in polynomial time if preferences are defined by assignment costs, assignment costs correspond to distances in an underlying graph, and the graph is a path or a cycle (Büsing et al., 2022). In contrast, the single-source CFLP is strongly NP-hard independently of the structure of assignment costs, as discussed at the beginning of this section.

To the best of our knowledge, no further research has been conducted on the second approach for dealing with overloaded facilities besides the work by Büsing et al. (2022). Polynomial time algorithms for special cases suggest further potential arising from combining capacities and closest assignments. With this work, we aim to contribute to a better understanding of the combinatorial structures of the single-source capacitated facility location problem with customer preferences when customers have to be served at their most preferred open facility.

3 Problem definition and notation

In the single-source capacitated facility location problem with customer preferences (SSCFLPCP), we are given a set of customers I with demands $d_i \in \mathbb{N}_0$ for each $i \in I$ and a set of potential facilities J. A capacity $Q_j \in \mathbb{N}_0$ and opening costs $f_j \geq 0$ are associated with each potential facility $j \in J$. Assigning a customer $i \in I$ to a facility $j \in J$ yields assignment costs $c_{ij} \geq 0$. Each customer has ranked all potential facilities according to customer specific preferences, inducing a complete, weak ordering of all facilities for every customer. To that end, $j <_i k$ indicates that customer $i \in I$ prefers potential facility $j \in J$ over potential facility $k \in J$ and $j =_i k$ indicates that i is indifferent about facilities $j, k \in J$. With this notation at hand, consider the formal definition of the single-source capacitated facility location problem with customer preferences next.

Definition 1. *Let $(I, J, (d_i)_{i\in I}, (Q_j)_{j\in J}, (f_j)_{j\in J}, (c_{ij})_{i\in I, j\in J})$ be an instance of the capacitated facility location problem. The* single-source capacitated facility location problem with customer preferences *(SSCFLPCP) consists in finding a subset $F \subseteq J$ of facilities that are opened and an assignment $\Lambda : I \to F$ of customers to these facilities such that*

1. *the capacity limit of each open facility is met, i.e., $\sum_{i\in\Lambda^{-1}(j)} d_i \leq Q_j$ holds for each $j \in J$,*
2. *each customer is assigned to a facility they prefer most among all open facilities, i.e., there is no $i \in I$ and $j \in F$ such that $j <_i \Lambda(i)$.*
3. *the cost of opening facilities and assigning customers is minimised, i.e., $\sum_{j\in F} f_j + \sum_{i\in I} c_{i\Lambda(i)} \to \min$.*

The third condition in the definition above aims for a solution which balances the opening cost of facilities and the total serving cost of customers. Note that this function is a standard objective function in the facility location literature (Laporte et al., 2019) and, more specifically, it is also considered in literature on the (SSC)FLPCP (Cánovas et al., 2007, Calvete et al., 2020).

The SSCFLPCP is strongly NP-hard as it contains the uncapacitated facility location problem as a special case when each facility has enough capacity to serve the total demand of all customers and preferences are defined by assignment costs. Therefore, it immediately follows that the SSCFLPCP is strongly NP-hard even if each customer has uniform demand.

Often, facility location problems are introduced by some underlying graph structure. We then identify facilities and customers with the nodes of the graph. Consequently, we define for each node a demand for the corresponding customer as well as a capacity and opening cost for the corresponding facility. Facilities and customers are connected via edges in the graph, which may represent streets and have non-negative travel cost. The assignment cost of a customer i to a facility j then corresponds to the cost of a shortest path between the corresponding nodes i and j in the graph. We can then classify facility location instances by the underlying network structure, like complete graphs, trees or paths.

4 SSCFLPCP with strict preferences

In this section, we discuss the computational complexity of the SSCFLPCP if each customer has a strict preference ordering of the facilities.

Recall that finding a feasible solution for the SSCFLPCP is strongly NP-hard. The difficulty arises, for example, from instances where solving the SSCFLPCP is equivalent to solving the strongly NP-hard problem *3-partition*. This is particularly the case when customers rate all facilities equally: finding a feasible solution is only a matter of partitioning customers such that capacities are met. By demanding strict preference orderings, we exclude such settings. In the following, we suppose that each customer has a *strict preference* ordering of the given potential facility locations, i.e., it is either $j <_i k$ or $k <_i j$ for all $i \in I$, $j, k \in J$. This restriction allows us to compute a feasible solution for the SSCFLPCP efficiently.

Theorem 1. *A feasible solution for an instance of the SSCFLPCP, where each customer has a strict preference ordering of the potential facilities, can be found in $\mathcal{O}(|I| \cdot |J| \cdot \log(|J|) + |J| \cdot |I|)$ steps. If we are given a list of facilities sorted by preference for each customer, then we only need $\mathcal{O}(|J| \cdot |I|)$ steps.*

Proof. If not already given, we construct for each customer a sorted list of facilities in $\mathcal{O}(|I| \cdot |J| \cdot \log(|J|))$ steps.

Afterwards, we start by considering a solution in which all facilities are opened and assign each customer to their most preferred open facility. Since each customer has a strict preference ordering of the facilities, the assignment of the customers is easy: assign each customer to their most-preferred open facility. If this corresponds to a feasible solution, we are done.

Suppose this solution is infeasible. Then, there must be an open facility which has to serve more demand than its capacity offers. Denote this facility with j. Since all facilities are open and each customer is assigned to their most-preferred facility, we can not open another facility to ease j's burden. Moreover, closing other facilities will either keep j's burden constant or increase it. Hence, there is no feasible solution in which facility j is open. We close facility j and reassign all customers that were assigned to j. If the solution is still infeasible, then we repeat the previous step until either the set of open facilities is empty, i.e., there is no feasible solution, or until we obtain a solution in which the capacity of each open facility is respected.

Note that reassigning a customer can be done in constant time, as we can assign them to the next facility in the preference list. While doing so, we update the aggregated demand that is assigned to the new facility and check whether this demand violates the capacity. Since there are at most $|I|$ customers to be reassigned, the process of reassignment and checking for violated capacities can be done in $\mathcal{O}(|I|)$ steps. In the worst case, $|J|$ facilities are closed until infeasibility is proven. This amounts to $\mathcal{O}(|J| \cdot |I|)$ steps. If the instance is feasible, we need less iterations. □

When applying the procedure described in the proof of Theorem 1, we close exactly those facilities that can not be open in any feasible solution. Hence, every feasible set of facilities is contained in the set of facilities computed by the above procedure.

Theorem 1 yields hope for further special cases of the SSCFLPCP in which a cost-minimising solution can be computed efficiently. Computing an optimal solution for the SSCFLPCP, however, is strongly NP-hard - even if the structure of the instance corresponds to a path - as we show next. Note that this result is independent of the capacities.

Theorem 2. *The uncapacitated facility location problem with strict customer preferences (UFLPSCP) is strongly NP-hard even if the instance is defined on a path graph.*

Proof. We prove the theorem via a reduction from the strongly NP-hard minimum vertex cover problem (Garey and Johnson, 1979). In the minimum vertex cover problem, we are given a graph $G = (V, E)$ and a number $k \in \mathbb{N}$. The goal is to determine a subset of the nodes $S \subseteq V$ with size of at most k so that at least one end node of each edge in graph G is element of set S.

We transform any vertex cover instance $\mathcal{I} = (G = (V, E), k)$ into an instance of the UFLPSCP on a path as follows. Introduce one node v'_i for each $i \in \{1, \dots, |V|\}$ in the vertex cover instance, and denote this set of nodes with V'_V. Furthermore, introduce one node e'_j for each $j \in \{1, \dots, |E|\}$ in graph G, and denote this set of nodes with V'_E. Finally, introduce dummy nodes $V'_\xi = \{\xi_1, \dots, \xi_k\}$ with $k \in \mathbb{N}$ being the number from the vertex cover instance. Set $V' = V'_V \cup V'_E \cup V'_\xi$ denotes the nodes in the underlying graph of the transformed UFLPSCP instance. Set the opening costs of facilities in set V'_ξ to 1 and the opening costs of facilities in set $V'_E \cup V'_V$ to $|V'|^2 + 1$. Connect the nodes in V' such that they form a path in the order $(e'_1, \dots, e'_{|E|}, v'_1, \dots, v'_{|V|}, \xi_1, \dots, \xi_k)$. Assign to each edge in the path costs of 1, except for the edge $\{v'_{|V|}, \xi_1\}$, to which we assign cost of $3 \cdot |V'|^2$. The edge costs are defined so that it is favorable to assign customers in $V'_E \cup V'_V$ to facilities in $V'_E \cup V'_V$ as well as customers in V'_ξ to facilities in V'_ξ. Otherwise, customers need to cross the expensive edge $\{v'_{|V|}, \xi_1\}$.

It remains to define the preferences of each customer. Consider customers located in $V'_V \cup V'_\xi$. Each customer in $V'_V \cup V'_\xi$ prefers facilities at nodes of the same type over all other facilities. That is, customers in V'_V prefer facilities in V'_V. Likewise, customers in V'_ξ prefer facilities in V'_ξ. Apart from that, the order of preferences is arbitrary. Consider a customer $e' \in V'_E$ who corresponds to an edge $e = \{u, w\} \in E$ in the vertex cover instance. Denote the nodes corresponding to u, w with u', w' in the SSCFLPCP instance. We define the most preferred facilities of e' to be those in $\{u', w'\}$, followed by the facilities in V'_ξ in arbitrary order. The preference order of the remaining facilities is also arbitrary. Assuming that at least one facility in V'_ξ is open, the preferences of e' ensure that we have to pay high assignment costs by crossing the expensive edge $\{v'_{|V|}, \xi_1\}$ if we do not open a facility corresponding to one of the end nodes of e. We will show that this yields a correspondence between open facilities in V'_V and vertex covers in V.

We prove that there exists a vertex cover $S \subseteq V$ of size $|S| \leq k$ *if and only if* there exists a solution to the SSCFLPCP instance with cost of at most $(k+1) \cdot |V'|^2 + k$.

Assume there exists a vertex cover $S \subseteq V$ with $|S| \leq k$. Let $V'_S \subseteq V'_V$ be the nodes in the SSCFLPCP instance corresponding to those in the vertex cover. Then we construct a solution by opening all facilities in $V'_S \cup V'_\xi$. This yields opening costs of k for the facilities in V'_ξ and at most $k \cdot (|V'|^2 + 1)$ for those in V'_S. We can assign all customers in V'_ξ to their own facility, which yields no assignment costs. Furthermore, we assign all customers in $V'_V \cup V'_E$ to facilities in V'_S, which yields assignment costs less than $(|V| + |E|)^2$, since $|V| + |E|$ customers need to be assigned and every assignment costs less than $|V| + |E|$. The assignment is feasible, as customers in V'_V and V'_ξ prefer facilities at nodes of their own type and the most preferred facilities of customers in V'_E are located in V'_S, since S is a vertex cover. Overall, we obtain a feasible solution of costs less than $k + k \cdot (|V'|^2 + 1) + (|V| + |E|)^2$. Since $k + (|V| + |E|)^2 \leq (k + |V| + |E|)^2 = |V'|^2$ holds, we obtain total costs of less than $k \cdot (|V'|^2 + 1) + |V'|^2 = (k+1) \cdot |V'|^2 + k$.

Now, assume we are given a solution to the SSCFLPCP instance of cost at most $(k+1) \cdot |V'|^2 + k$. Then there exists an open facility in V'_ξ. Otherwise, all customers in V'_ξ would need to cross edge $\{v'_{|V|}, \xi_1\}$ to their assigned facility, which amounts to assignment costs of more than $k \cdot 3 \cdot |V'|^2 \geq (k+2) \cdot |V'|^2 > (k+1) \cdot |V'|^2 + k$.

We can assume without loss of generality that customers in V'_E are not assigned to facilities in V'_ξ. Otherwise, if a customer $e' \in V'_E$ corresponding to $e = \{u, w\} \in E$ is assigned to a facility in V'_ξ, then we open one of the facilities at u' or w' and assign the customer accordingly. This is feasible, since the customer prefers the new facility over the previous facility in V'_ξ. The necessary reassignments implied by opening the facility only affect customers in $V'_V \cup V'_E$ and thus imply costs of less than $(|V| + |E|)^2$, since less than $|V| + |E|$ customers have to be reassigned and the new assignment costs are each less than $|V| + |E|$. Note that this includes the reassignment of e' itself. Together with the opening cost of $|V'|^2$, the total costs are less than the previous assignment cost of more than $3 \cdot |V'|^2$ for assigning the customer e' to a facility in V'_ξ. Hence, the cost of the new solution is less than that of the initial solution.

Then, there exists no customer $e' \in V'_E$ who is assigned to a facility in V'_ξ, although there is an open facility in V'_ξ. This implies that for all $e' \in V'_E$ corresponding to $e = \{u, w\} \in E$, there exists an open facility in $\{u', w'\} \subseteq V'_V$. Hence, the set of open facilities $S' \subseteq V'_V$ corresponds to a vertex cover $S \subseteq V$.

Assume that $|S'| > k$ holds. Then the opening cost of our solution is at least $(k+1)(|V'|^2 + 1) > (k+1) \cdot |V'|^2 + k$, which contradicts our assumption. Hence, S is a vertex cover of size $|S| = |S'| \leq k$, which completes the proof. $\square$

It is easy to see that this result also extends to the SSCFLPCP. In conclusion, solving the SSCFLPCP with strict preferences is already strongly NP-hard on basic graph classes, independent of the actual demands and capacities. This result emphasizes the importance of an underlying preference structure of all customers.

5 SSCFLPCA with distinct assignment costs on star graphs

In this section, we study the complexity of a special case of the SSCFLPCP in which customer preferences are based on the assignment costs of the customers. That is, we assume that a customer prefers facility j over facility k if their assignment cost to j are lower than their assignment cost to k. We refer to this special case as the *single-source capacitated facility location problem with customer preferences* (SSCFLPCA). In this section, we assume that each customer has distinct assignment costs to all potential facilities, i.e., there are no two potential facilities any customer is indifferent between.

Finding an optimal solution to the SSCFLPCA with distinct assignment costs continues to be strongly NP-hard. This can be easily seen via a reduction from the classical uncapacitated facility location problem (UFLP) to an instance of SSCFLPCA, where the capacity of each facility is large enough to serve all customers. From Theorem 1, we know that finding a feasible solution to SSCFLPCA is easy, just like for the UFLP. Moreover, the SSCFLPCA is at least polynomially solvable on paths (Büsing et al., 2022), while the UFLP is polynomially solvable on trees (Mirchandani and Francis, 1990). These observations suggest a complexity-theoretical similarity between both problems. In the next result, however, we show that the SSCFLPCA is harder on trees than the UFLP. In fact, it is already weakly NP-hard on stars.

Theorem 3. *The SSCFLPCA with distinct assignment costs, which are with respect to an underlying star graph, is at least weakly NP-hard.*

Proof. We prove our claim via a reduction from the weakly NP-hard *knapsack problem* (Garey and Johnson, 1979). Consider a knapsack instance consisting of a capacity $B \in \mathbb{N}$ and a set of items N, where each item $i \in N$ is assigned a profit $p_i \in \mathbb{N}$ and a weight $w_i \in \mathbb{N}$. The goal is to find a subset $S \subseteq N$ of items so that the total weight $\sum_{i \in S} w_i$ is at most B and the total profit $\sum_{i \in S} p_i$ is maximal. We transform a given knapsack instance into an instance of our problem by identifying the knapsack with the center of a star and the items with the leaves. The transformation is performed so that the items that are packed into the knapsack correspond to the leaves that are assigned to the center of the star.

Introduce node ξ as the center node in the star graph. The demand, capacity, and opening costs of ξ are defined as $d_\xi = 1$, $Q_\xi = B + 1$, and $f_\xi = 0$. Introduce for each item in set N a leaf node on the star. Denote the set of leaves with $L = \{\ell_1, \ell_2, \ldots, \ell_{|N|}\}$. Each node $\ell_i \in L$ has both a demand and a capacity equal to the weight of the corresponding item in the knapsack instance, i.e., $d_{\ell_i} = Q_{\ell_i} = w_i$. We define the costs $c_{\ell_i \xi} = i$ on the edges $\{\ell_i, \xi\}$ for all $\ell_i \in L$, which ensures that each customer has distinct distances, and thus distinct assignment costs, to all facilities. We define the opening costs f_{ℓ_i} so that we save costs of p_i when assigning ℓ_i to ξ instead of opening a facility at ℓ_i. That is, we define $f_{\ell_i} = p_i + c_{\ell_i \xi} = p_i + i$.

We now prove that there exists a knapsack solution of value at least P *if and only if* there exists a facility location solution of cost at most $\sum_{i \in N}(p_i + i) - P$.

Suppose that there exists a feasible solution $S \subseteq N$ for a given knapsack instance with $\sum_{i \in S} p_i \geq P$. We construct the corresponding facility location solution by opening facility ξ as well as all facilities in $\{\ell_i \in L | i \notin S\}$. Due to the closest assignment constraints, we need to assign all customers to their own facility, if it is open. This assignment meets the capacity constraints on the leaves, as we have $d_{\ell_i} = w_i = Q_{\ell_i}$ for all $\ell_i \in L$. The customers in $\{\ell_i \in L | i \in S\}$ need to be assigned to facility ξ. This meets the capacity constraint, as we have $d_\xi + \sum_{i \in S} d_{\ell_i} = 1 + \sum_{i \in S} w_i \leq 1 + B = Q_\xi$, since S is a feasible knapsack solution. The total opening costs of our solution are $f_\xi + \sum_{i \notin S} f_{\ell_i} = 0 + \sum_{i \notin S}(p_i + i)$. The total assignment costs are $\sum_{i \in S} c_{\ell_i \xi} = \sum_{i \in S} i$. The overall costs are thus

$$\sum_{i \notin S}(p_i + i) + \sum_{i \in S} i = \sum_{i \in N}(p_i + i) - \sum_{i \in S} p_i \leq \sum_{i \in N}(p_i + i) - P$$

due to the value $\sum_{i \in S} p_i \geq P$ of the knapsack solution.

Conversely, consider a solution to the considered SSCFLPCA instance of cost at most $\sum_{i \in N}(p_i + i) - P$, where $F \subseteq L \cup \{\xi\}$ corresponds to the set of nodes on which facilities are opened. First, note that $\xi \in F$ holds, since no other facility is capable of serving ξ in addition to the customer at the respective node. Hence, the customers in $L \setminus F$ are assigned to facility ξ, while all other customers are assigned to their own facility. We define $S = \{i \in N | \ell_i \in L \setminus F\}$ as the corresponding solution to the knapsack instance. Due to the capacity constraint of facility ξ, we have $\sum_{i \in S} w_i = d_\xi - 1 + \sum_{\ell_i \in L \setminus F} d_{\ell_i} \leq Q_\xi - 1 = B$. Hence, S is a feasible knapsack solution. The costs $\sum_{\ell_i \in L\setminus F} c_{\ell_i \xi} + \sum_{i\in F\cap L} f_{\ell_i} \leq \sum_{i\in N}(p_i + i) - P$ of our facility location solution imply

$$\begin{aligned}\sum_{i\in S} p_i &\geq \sum_{i\in S} p_i + \left(\sum_{\ell_i\in L\setminus F} c_{\ell_i\xi} + \sum_{i\in F\cap L} f_{\ell_i}\right) - \left(\sum_{i\in N}(p_i+i) - P\right)\\ &= \sum_{\ell_i\in L\setminus F}(f_{\ell_i} - c_{\ell_i\xi}) + \sum_{\ell_i\in L\setminus F} c_{\ell_i\xi} + \sum_{i\in F\cap L} f_{\ell_i} - \sum_{\ell_i\in L} f_{\ell_i} + P\\ &= P,\end{aligned}$$

which completes the proof □

Hence, solving the SSCFLPCA with distict assignment costs on a star graph is at least weakly NP-hard.

A pseudo-polynomial algorithm

In the following, we describe a pseudo-polynomial algorithm for solving instances of the SSCFLPCA with distinct assignment costs defined on a star graph. Consider a star graph $T = (V, E)$ and denote the center node with z. We start by describing the structure of a solution for our problem. Denote the set of nodes on which facilities are opened with $F \subseteq V$.

Unique closest facility to the center node and always-closed facilities In any feasible solution there is a unique open facility which has lowest assignment costs to the center node. Suppose that this closest open facility to the center node is $\xi \in V$, i.e., $c_{\xi z} < c_{iz}$ for any open facility $i \in F \setminus \{\xi\}$. If facility ξ is opened, it has to serve customer ξ as well as all customers $p \in V \setminus F$ that meet at least one of the following two conditions.

1. The assignment cost of p to center node z is lower than the assignment cost of ξ to z, i.e., $c_{pz} < c_{\xi z}$.
2. The capacity of facility p can not serve the demand of customer p, i.e., $Q_p < d_p$.

The set of locations with at least one of these properties is defined as

$$C(\xi) = \{\xi\} \cup \{p \in V : c_{pz} < c_{\xi z}\} \cup \{p \in V : c_{pz} > c_{\xi z},\, d_p > Q_p\}$$

and it has the following property.

All facilities in set $C(\xi) \setminus \{\xi\}$ are kept closed and facility ξ serves a demand of at least

$$D(\xi) = \sum_{p\in C(\xi)} d_p = d_\xi + \sum_{p\in V: c_{pz} < c_{\xi z}} d_p + \sum_{p\in V: c_{pz} > c_{\xi z},\, d_p > Q_p} d_p.$$

Proof. Let us begin with proving the first part of the claim. Consider a customer $p \in C(\xi) \setminus \{\xi\}$. If the assignment cost of p to the center node z is lower than the assignment cost of ξ to z, i.e., $c_{pz} < c_{\xi z}$, opening a facility at p is a contradiction to ξ being the open facility with lowest assignment costs to the center node among all open facilities. Therefore, the facility at p has to stay closed in any feasible solution in which the facility at ξ is the closest open facility to the center node. Suppose next that the capacity of the facility at location p can not serve the demand of customer p, i.e., $d_p > Q_p$ holds. Then, the facility at p has to stay closed as otherwise customer p would wish to be served there. Thus, all facilities in set $C(\xi) \setminus \{\xi\}$ are kept closed.

With the facilities at locations in set $C(\xi) \setminus \{\xi\}$ kept closed, each customer $p \in C(\xi) \setminus \{\xi\}$ has to be served at their closest open facility. Their closest open facility is the closest open facility to the center node, that is, facility ξ. Since facility ξ is open, the customer at ξ is also served at ξ. Therefore, facility ξ has to serve a demand of at least the demands of all customers in $C(\xi)$ and the second part of the claim holds. □

Next, we analyse the set of open facilities. Note that the set of open facilities F depends on the location of the closest open facility to the center node. We acknowledge this by renaming the set of open facilities from F to $F(\xi)$ with ξ the closest open facility to the center node.

Always-open facilities The set of open facilities $F(\xi)$ includes, besides ξ, any $i \in V$ that simultaneously fulfills the following three conditions, as we show in Corollary 1.

1. The assignment cost of customer i to center node z is strictly greater than the assignment cost of ξ to z, i.e., $c_{\xi z} < c_{iz}$.
2. The capacity of i is great enough to serve at least the demand of customer i, i.e., $d_i \leq Q_i$.
3. The assignment cost of customer i to facility ξ is more expensive than opening facility i, i.e., $f_i \leq c_{i\xi}$.

Let $\bar{F}(\xi) = \{i \in V : c_{\xi z} < c_{iz},\, d_i \leq Q_i,\, f_i \leq c_{i\xi}\}$ be the set of locations satisfying the three listed properties above.

Corollary 1. *Each facility located at an $i \in \bar{F}(\xi)$ is open.*

Proof. We prove our claim by showing that opening a facility at a location in $\bar{F}(\xi)$ does not increase the overall cost. First, note that we may open a facility at any $i \in \bar{F}(\xi)$: due to the first and second property of any element in $\bar{F}(\xi)$, the intersection of sets $\bar{F}(\xi)$ and $C(\xi)$ is empty. Second, suppose it is strictly more expensive to assign customer $i \in \bar{F}(\xi)$ to facility ξ than to open a facility at i, and suppose we do not open a facility at i. Then, customer i is assigned to their closest open facility, which is located at ξ. Due to the third property of any customer in set $\bar{F}(\xi)$, we can reduce the objective value of a solution in which customer i is served at ξ by opening a facility at i. Hence, opening a facility at a location in $\bar{F}(\xi)$ reduces the overall cost if it is strictly more expensive to assign customer $i \in \bar{F}(\xi)$ to facility ξ than to open a facility at i. Third, suppose assigning customer i to ξ is as expensive as to open a facility at i, and suppose we don't open a facility at i. In this case, the decision of whether or not we assign the customer at i to facility ξ contributes the same cost to the objective value of a solution as opening a facility at i. However, opening a facility at i makes room at facility ξ for customers at nodes $i \in V \setminus \bar{F}(\xi)$ for which it is cheaper to assign i to ξ than to open a facility at i. In conclusion, opening all facilities in set $\bar{F}(\xi)$ does not increase the objective value and is advantageous. Hence, we can assume that in any solution all facilities in $\bar{F}(\xi)$ are opened. □

Lastly, we consider the set of facility locations for which it is yet undecided whether a facility is opened at the location. Denote this set with $W(\xi)$.

Undecided locations The set of undecided locations $W(\xi)$ contains all customers $i \in V$ that simultaneously meet the following three properties.

1. The assignment cost of customer i to center node z is strictly greater than the assignment cost of ξ to z, i.e., $c_{\xi z} < c_{iz}$.
2. The capacity of i is great enough to serve at least the demand of customer i, i.e., $d_i \leq Q_i$.
3. The assignment cost of customer i to facility ξ is strictly less expensive than opening facility i, i.e., $c_{i\xi} < f_i$.

Hence, the set of undecided locations is the set of all locations that are neither element of $C(\xi)$ nor element of $\bar{F}(\xi)$, i.e., $W(\xi) = V \setminus (C(\xi) \cup \bar{F}(\xi))$.

Since the facility at ξ is the closest open facility to the center node and since we consider an SSCFLPCA instance defined on a star graph, none of the remaining facilities besides ξ and potentially $i \in W(\xi)$ is relevant for the assignment of the customer located at i. Furthermore, each location $i \in W(\xi)$ has the property that it is cheaper to assign the customer at i to facility ξ than to open a new facility at i. Thus, assigning all customers in $W(\xi)$ to facility ξ yields a cost-minimising solution given that ξ is the closest open facility to the center node. However, assigning all customers in $W(\xi)$ to ξ might violate the capacity at ξ, thus, yielding an infeasible solution. Therefore, our goal is to determine a subset of the customers in $W(\xi)$ who we assign to ξ so that the saved costs by assigning them to ξ and not opening a facility at their location is maximised.

Solution approach In order to decide which of the customers in set $W(\xi)$ are served at facility ξ so that the saved costs by assigning them to ξ and not opening a facility at their location are maximised, we transform this decision problem into an instance of the knapsack problem. For that, we introduce for each location $i \in W(\xi)$ one knapsack item i' with weight $w_{i'} = d_i$ and value $p_{i'} = f_i - c_{i\xi}$ and set the bound on the weight of packed items to $B = Q_\xi - D(\xi)$. Then, the following property holds.

Proposition 1. *There is a cost-preserving 1-1-correspondence between the saved costs by assigning customers from set $W(\xi)$ to facility ξ in a solution for the considered SSCFLPCA instance and a solution for the constructed knapsack instance.*

Proof. A feasible solution for the considered SSCFLPCA instance implies a feasible solution for the corresponding knapsack instance, as we discuss now. Suppose customers in $S \subseteq W(\xi)$ are served at ξ in a feasible solution for the SSCFLPCA. Then, the demand served at ξ corresponds to $D(\xi) + \sum_{i \in S} d_i \leq Q_\xi$ which is equivalent to $\sum_{i \in S} d_i \leq Q_\xi - D(\xi)$. Denote with S' the set of items in the considered knapsack instance corresponding to customers in S. Then, per construction, it is $\sum_{i' \in S'} w_{i'} = \sum_{i \in S} d_i \leq Q_\xi - D(\xi) = B$ and S' is a feasible solution for the knapsack instance. Conversely, let S' be a feasible solution for the considered knapsack instance. Then, denote the customers in the SSCFLPCA instance who correspond to items in S' with S, and assign the customers in S to facility ξ together with all customers in $C(\xi)$. The capacity at ξ is met since inequality

$$\sum_{i \in C(\xi)} d_i + \sum_{i \in S} d_i = D(\xi) + \sum_{i' \in S'} w_{i'} \leq D(\xi) + Q_\xi - D(\xi) = Q_\xi$$

holds. Assign the remaining customers to the facility at their own location. These remaining customers can be served at their location: otherwise, they were element of $C(\xi)$. In conclusion, we can derive a feasible solution for the SSCFLPCA from a feasible knapsack solution. The 1-1-correspondence follows immediately.

It remains to prove that this correspondence preserves costs. Denote with $S \subseteq W(\xi)$ the set of customers that are assigned to facility ξ and with S' the knapsack items corresponding to customers in S. Denote with $\Delta(S, \xi)$ the saved cost by assigning customers in set S to facility ξ instead of opening facilities at all nodes in S. Denote with $val(S', \xi)$ the value of the knapsack instance when S' corresponds to a feasible knapsack solution. Then, equality

$$val(S', \xi) = \sum_{i \in S'} p_i = \sum_{i \in S} (f_i - c_{i\xi}) = \Delta(S, \xi)$$

holds. Thus, there is indeed a cost-preserving 1-1-correspondence between the saved costs by assigning customers from set $W(\xi)$ to facility ξ in a solution for the considered SSCFLPCA instance and a solution for the constructed knapsack instance. □

We derive the following theorem from our previous results.

Theorem 4. *The SSCFLPCA with distinct assignment costs can be solved in pseudo-polynomial time if the underlying graph corresponds to a star. The computation takes $\mathcal{O}(|V|^2 \cdot \max_{v \in V}\{Q_v\})$ steps.*

Proof. The pseudo-polynomial algorithm starts by considering each location $\xi \in V$ as the closest open facility to the center node. Consider the sub-problem in which the closest open facility to the center node is located at ξ.

We first compute sets $C(\xi)$, $\bar{F}(\xi)$ and $W(\xi)$ as well as value $D(\xi)$. We can compute the three sets in $\mathcal{O}(|V|)$ steps: we have to check for each customer $p \in V$ their assignment cost to the center node as well as the difference between the capacity of the facility at their location and their demand, and the difference between the opening cost of a facility at their location and their assignment cost to ξ. This can be done in a constant number of steps for each customer since the relevant values needed for the comparisons are already given in the instance. While we compute sets $C(\xi)$, $\bar{F}(\xi)$ and $W(\xi)$, we can also compute value $D(\xi)$ by adding the demand of any customer who is added to $C(\xi)$ to the sum of demands of customers already added to set $C(\xi)$. Furthermore, while determining these three sets and $D(\xi)$, we can also assign each customer in $C(\xi)$ to ξ, open all facilities in $\bar{F}(\xi)$ and assign each customer in $\bar{F}(\xi)$ to the facility at their location. Then, we have already computed a partial solution for the sub-problem in which the closest open facility to the center node is located at ξ.

Note that the costs incurred by all customers in $C(\xi) \cup \bar{F}(\xi)$ are fixed for a given ξ; c.f. Section 5 and Corollary 1. It is the costs incurred by assigning customers in $W(\xi)$ to ξ which determine the costs for the sub-problem in which the facility at ξ is the closest open facility to the center node. In order to detect a cost-minimising assignment of customers in $W(\xi)$ to ξ, we transform this problem of assigning customers from $W(\xi)$ to ξ into a knapsack instance as discussed right in front of Proposition 1. This transformation takes $\mathcal{O}(|V|)$ steps as we have to introduce one knapsack item for each customer in set $W(\xi)$. As shown in Proposition 1, assigning customers in $W(\xi)$ according to an optimal solution for the considered knapsack instance maximises the saved cost regarding the assignment of customers in $W(\xi)$ to ξ. Therefore, this solution minimises the overall costs of the considered sub-problem. Solving the considered knapsack instance takes $\mathcal{O}(|V| \cdot \max_{v \in V}\{Q_v\})$ steps.

To sum up, computing a cost-minimising solution for each possible location of the closest open facility to the center node takes $\mathcal{O}\big(|V| \cdot (2 \cdot |V| + |V| \cdot \max_{v \in V}\{Q_v\})\big) = \mathcal{O}(|V|^2 \cdot \max_{v \in V}\{Q_v\})$ steps. Finally, we have to determine the solution with lowest overall costs among the $\mathcal{O}(|V|)$ solutions arising from computing a cost-minimising solution for each possible location of the closest open facility to the center node. This takes another $\mathcal{O}(|V|)$ steps. We conclude that computing an optimal solution for the SSCFLPCA with distinct assignment costs defined on a star graph takes $\mathcal{O}(|V|^2 \cdot \max_{v \in V}\{Q_v\} + |V|) = \mathcal{O}(|V|^2 \cdot \max_{v \in V}\{Q_v\})$ steps. □

To the best of our knowledge, an extension to general trees is not straight-forward - if there is one.

Recall that the knapsack problem can be solved efficiently if all items have a weight of one: choose the items with greatest value until no capacity is left. Thus, we immediately get the following result for SSCFLPCA instances with distinct assignment costs and an underlying graph corresponding to a star.

Corollary 2. *The SSCFLPCA with distinct assignment costs and uniform customer demands can be solved in polynomial time if the underlying graph corresponds to a star*

The previous result and Theorem 4 raise two questions. First, whether the computational complexity on star graphs extends to general trees. Second, how much impact uniform customer demands have on the computational complexity of the SSCFLPCA. Answering the first question goes beyond the scope of this work, which aims for an initial understanding of customer preferences in the context of capacitated facility location problems. We focus on the second question.

6 SSCFLPCA with uniform customer demands on star graphs

In the following, we consider instances of the SSCFLPCA in which customers may be indifferent between multiple facilities and each customer has uniform demand. Without loss of generality, we assume that all demands are equal to one. Otherwise, we apply scaling and rounding to demands and capacities.

Recall that a feasible solution for the traditional SSCFLP can be found in polynomial time if the demand of each customer is equal to one: open all facilities. This solution approach is no longer guaranteed to work when adding closest assignment constraints. Contrary to the case with strict customer preferences, the assignment of a customer to a facility is no longer unique due to indifferences. We first prove that a feasible solution to the SSCFLPCA with uniform demands can be computed in polynomial time. Afterwards, we show that the SSCFLPCA with uniform demands is weakly NP-hard on star graphs.

The next result states that the assignment problem of the SSCFLPCA can be solved in polynomial time.

Corollary 3. *Given an instance of the SSCFLPCA with uniform demands and a set of open facilities, a feasible assignment of customers to these facilities can be computed in polynomial time.*

Proof. We prove a generalisation of this result in Theorem 6, where each open facility has both a lower and an upper bound on the demand assigned to it. The claim in this corollary corresponds to the setting in which we set the lower bound on the demand to be served at any open facility to zero. Therefore, we omit the proof for this corollary here. □

Note that the result in Theorem 6 is independent of the preference ordering and therefore also holds for the assignment problem of the SSCFLPCP with uniform customer demands.

The idea used for proving Theorem 1 combined with the result from Corollary 3 allows us to prove that a feasible solution for the SSCFLPCA with uniform demands can be computed in polynomial time - even if customers are allowed to be indifferent between multiple facilities.

A feasible solution for the SSCFLPCA with uniform demands can be computed in polynomial time.

Proof. We prove this claim in a similar manner to the claim in Theorem 1 and assume that we are already given a list of facilities sorted by preference for each customer. Then, we start by considering a solution in which all facilities are opened and assign each customer to their most preferred open facility. Due to Corollary 3, we can decide in polynomial time whether there is a feasible assignment. If such an assignment exists, we have found a feasible solution and we are done.

Suppose this solution is infeasible. Then, there must be an open facility which has to serve more demand than its capacity offers. Denote this facility with j, and denote a customer who prefers no facility more than j and whose demand violates the capacity at j with i. We say that all customers violate the capacity at a facility if the sum of their demands exceed the facility's capacity. Denote with $\mathcal{P}(i,j) \subseteq J$ the set of facilities customer i is indifferent to regarding j. Since all facilities are open and each customer is assigned to their most-preferred facility, we can not open another facility to ease j's burden. Moreover, closing other facilities will either keep j's burden constant or increase it. Hence, there is no feasible solution in which facility j is open. It is important to also close all other facilities customer i is indifferent to regarding j, i.e., all facilities in $\mathcal{P}(i,j)$: if a facility $p \in \mathcal{P}(i,j)$ had been able to serve i, then facility j being at full capacity would not have been a problem. However, all facilities in set $\mathcal{P}(i,j)$ are incapable of serving i's demand and this will not change by closing facility j. Hence, we close facility j and all other facilities i is indifferent to. Then, we reassign all customers that were either assigned to j or to another facility in $\mathcal{P}(i,j)$. If the solution is still infeasible, then we repeat the previous step until either the set of open facilities is empty, i.e., there is no feasible solution, or until we obtain a solution in which the capacity of each open facility is respected.

Since Corollary 3 allows for a customer assignment in polynomial time and the remaining operations also take a polynomial number of steps, cf. Theorem 1, the overall running time of computing a feasible solution is polynomial. □

This result also holds true for general preference orderings as long as each customer has uniform demand. Furthermore, this result shows that adding preference constraints and uniform demands to the SSCFLP does not increase the computational complexity of determining a feasible solution compared to the computational complexity of determining a feasible solution in the uncapacitated facility location problem.

Like in the previous section, the result in Section 6 raises further questions regarding the complexity of our problem if the underlying graphs are restricted to simple graph structures. We can prove weak NP-hardness on star graphs for the SSCFLPCA with uniform demands, where customers may be indifferent between several facilities. We receive the following result.

Solving the SSCFLPCA with uniform demands on a star graph is at least weakly NP-hard.

Proof. We prove this claim via a reduction from the *min-knapsack problem.* Consider a min-knapsack instance consisting of a bound $B \in \mathbb{N}$ and a set of items N, where each item $i \in N$ is assigned a cost $p_i \in N$ and a weight $w_i \in N$. The goal is to find a subset $S \subseteq N$ of items so that the total weight $\sum_{i \in S} w_i$ is at least B and the total cost $\sum_{i \in S} p_i$ is minimal. This problem is known to be weakly NP-hard and is, for example, considered when solving the separation problem for cover inequalities. Without loss of generality, we assume that $B > |N|$ holds; note that we can tweak any given min-knapsack instance so that the bound is at least as big as the number of items by scaling the parameters. We transform a given min-knapsack instance into an instance of our problem by identifying the knapsack with the set of open facilities and the items with potential facility locations. The transformation is performed so that the items that are packed into the knapsack correspond to the locations at which facilities are opened.

Introduce a node z as the center node in the star graph. The demand, capacity and opening costs of z are defined as $d_z = 1$ and $Q_z = f_z = 0$. Introduce for each item in set N a leaf node on the star. Denote the set of leaves with $L = \{\ell_1, \ell_2, \ldots, \ell_{|N|}\}$. Each node $\ell_i \in L$ has a demand $d_{\ell_i} = 1$, a capacity equal to the weight of the corresponding item in the knapsack instance, i.e., $Q_{\ell_i} = w_i$, and opening cost equal to the cost of the corresponding item in the knapsack problem, i.e., $f_{\ell_i} = p_i$. Introduce $B - (|N| + 1)$ further leaf nodes and denote this set of additional leaves with $L' = \{\ell_{|N|+1}, \ldots, \ell_{B-1}\}$. We define the demand, capacity and opening costs of any $\ell \in L'$ as $d_z = 1$ and $Q_z = f_z = 0$. We define the costs $c_{\ell_i z} = 0$ on all edges $\{\ell_i, \xi\}$ for all $\ell_i \in L$. This implies that customers are indifferent between all open facilities. Note that due to this choice of assignment costs, this proof can also be used for showing that computing an optimal solution for the SSCFLP is weakly NP-hard if customers have uniform demands.

We now prove that there exists a min-knapsack solution of value at most P *if and only if* there exists a facility location solution of cost at most P.

Suppose that there exists a feasible solution $S \subseteq N$ for a given knapsack instance with $\sum_{i \in S} p_i \leq P$. We construct the corresponding facility location solution by opening all facilities in $\{\ell_i \in L | i \in S\}$. The capacity provided by all open facilities is $\sum_{i \in S} Q_{\ell_i} = \sum_{i \in S} w_i \geq B$ and the opening costs are $\sum_{i \in S} f_{\ell_i} = \sum_{i \in S} p_i \leq P$. Since the assignment costs of any customer to any open facility is equal to zero, the overall cost of the considered solution corresponds to the opening costs of the facilities in $\{\ell_i \in L | i \in S\}$ and is therefore at most P. It remains to prove that the constructed solution is indeed feasible. Since all customers are indifferent between all open facilities, we may assign the customers to open facilities according to our liking and any assignment meets the preference constraints. Since the provided capacity exceeds B and the total demand of all customers is equal to B, the set of open facilities allows for a feasible assignment.

Conversely, consider a solution to the considered SSCFLPCA instance of cost at most P, where $F \subseteq L$ corresponds to the set of nodes on which facilities are opened. First, note that facilities at nodes $\{z\} \cup L'$ might be opened but do not have any impact on the solution since they can not serve any customer due to the choice of demands and capacities and do not incur any opening costs; since all customers are indifferent between all facilities, customers at $\{z\} \cup L'$ may be served at a facility not located on their node while sharing their location with an open facility. For the sake of simplicity, however, we assume that all facilities at $\{z\} \cup L'$ stay closed. Hence, all customers in the considered instance have to be served at facilities in F. We define $S = \{i \in N | \ell_i \in F\}$ as the corresponding solution to the knapsack instance. Due to the total demand in the SSCFLPCA instance, we have $B \leq \sum_{\ell_i \in F} Q_{\ell_i} = \sum_{i \in S} w_i$. Hence, S is a feasible knapsack solution. The costs of the facility solution correspond to the costs for opening facilities. Therefore, $P \geq \sum_{\ell_i \in F} f_{\ell_i} = \sum_{i \in S} p_i$ holds which completes the proof. □

Note that this result is independent of the customer preferences and, therefore, also applicable to the traditional SSCFLP. While this result is already known for the SSCFLP, it's proof gives us important hints for solving the SSCFLPCA with uniform customer demands on a star graph.

A pseudo-polynomial algorithm

In the following, we present a pseudo-polynomial algorithm for the SSCFLPCA with uniform customer demands on a star graph, showing that the problem is indeed only weakly NP-hard. While some of the steps of our algorithm are similar to the algorithm presented in the previous section, new challenges occur if customers have uniform demand and are allowed to be indifferent between several facilities. Hence, we need a modified algorithm for our problem.

Consider an SSCFLPCA instance defined on a star graph $T = (V, E)$ and denote the center node with z. Order the assignment costs of all customers to the center node increasingly and ignore duplicates. The first entry is zero, which corresponds the assignment costs of the center node to itself, and the last entry corresponds to the assignment costs of the customer with greatest costs to the center node. Denote this ordering with R and note that it consists of at most $|V|$ entries. Due to the closest assignment constraints, each customer has to be served at their closest open facility. Hence, as in the proof for Theorem 4, a customer is either served at the facility at their node or at an open facility with lowest assignment costs to the center node. In the next algorithm, we utilise this knowledge.

Suppose we open a facility at a node with assignment costs of $r \in R$ to the center node and, among all other open facilities, there is no facility with strictly lower assignment costs to the center node. Perform the following steps.

Step 1: Categorisation of customers Consider all customers whose assignment costs to the center node are strictly lower than r, i.e., all customers $i \in V$ with $c_{iz} < r$, as well as all customers who can not be served at a potential facility at their node due to lack of capacity, i.e., all customers $i \in V$ with $d_i > Q_i$. Denote this set of customers with $C(r) = \{p \in V : c_{pz} < r\} \cup \{p \in V : d_p > Q_p\}$ and note that the facilities at nodes in $C(r)$ will always stay closed. Denote the set of customers whose assignment costs to the center node are exactly r with $\Xi(r) = \{\xi \in V : c_{\xi z} = r\}$. Then, set $C(r) \cup \Xi(r)$ is the set of customers $i \in V$ that *have* to be served at a facility with assignment costs of r to the center node.

Next, we define the set of facilities that will always be open. Open all potential facilities at nodes $V \setminus C(r)$ for which it is at most as expensive to serve the customer at the same location as to have the customer travel to a facility in set $\Xi(r)$, i.e., open facilities at nodes $i \in V \setminus C(r)$ with $c_{iz} + r \geq f_i$. Denote this set of locations with $\bar{F}(r) = \{p \in V : r \leq c_{pz},\, d_p \leq Q_p,\, f_p \leq c_{pz} + r\}$. For the sake of simplicity, we assume that set $\bar{F}(r)$ is empty and therefore omitted. We may make this assumption since facilities in $\bar{F}(r)$ do not have any impact on the assignment decisions of the remaining customers and are fixed for fixed $r \in R$. Finally, we denote the set of undecided customers with $W(r) = V \setminus (C(r) \cup \Xi(r))$.

Each customer in set $W(r)$ is either served at a facility located at their location or served at a facility in set $\Xi(r)$. Then, the choice of open facilities has to be made from set $L = \Xi(r) \cup W(r)$ and this choice determines the objective value of any solution in which the lowest assignment costs to the center equals r. We prove this claim in Steps 3 and 4.

Step 2: Choosing locations in L to be opened Let us describe a procedure that determines a subset of open facilities in set L that yields cost-minimising solutions next. We have to open enough facilities at nodes in set $\Xi(r)$ such that the demand of all customers in set $C(r) \cup \Xi(r)$ can be served at these facilities. Our goal is to open those facilities which minimise the loss resulting from opening them, i.e., the difference between the opening and the assignment costs of customers in set $C(r) \cup \Xi(r)$. We detect such facilities by solving a min-knapsack problem, in which we aim to minimise total costs while exceeding a given lower bound on the total weight. We model each facility $i \in \Xi(r)$ by an item i' of weight Q_i and cost f_i minus the assignment costs of i to another facility in $\Xi(r)$, i.e., cost of $\Delta(i', r) = f_i - 2 \cdot r$. In order to ensure that the detected solution serves the demand of all customers in set $C(r) \cup \Xi(r)$, we set the lower bound on the weight of the items in any min-knapsack instance to $B = |C(r) \cup \Xi(r)|$. We open facilities at all locations corresponding to items in the min-knapsack solution and denote this set of locations with F_B.

Step 3: Choosing locations in $W(r)$ In this step, we show that employing the min-knapsack solution from the previous step in order to find open facilities in $W(r)$ returns a cost-minimising sub-solution, given that a set of open facilities in set $\Xi(r)$ is provided. Denote the set of facilities chosen in the previous step with F_B.

Note that facilities in set F_B provide leftover capacity of at most $Q_B := \sum_{i \in F_B} Q_i - (|C(r) + \Xi(r)|)$ for customers in set $W(r)$. Suppose that $Q_B > 0$ holds. We aim to minimise the costs arising from opening facilities in set $W(r)$ and assigning customers in $W(r)$ without a facility at their location to a facility in F_B, and determine such a solution as follows. We compute for all nodes in $W(r)$ the difference between their opening costs and their assignment costs to a facility in F_B, which is $\Delta(i, r) = f_i - (c_{iz} + r)$. Thus, value $\Delta(i, r)$ coincides with the cost we save by assigning customer i to a facility in set F_B instead of opening a facility at location i. Assigning Q_B customers with highest saved costs to the facilities in F_B yields a cost-minimising assignment. We see this by transforming the considered instance into a knapsack instance so that each customer $i \in W(r)$ corresponds to one knapsack item i' and item i' has a weight of one and a profit of $\Delta(i,r)$; the bound on the weights is Q_B. Hence, the following result holds.

Proposition 2. *Consider the sub-problem of the SSCFLPCA with uniform customer demands on a star graph in which a part of facilities is already open and located at $F_B \subseteq \Xi(r)$; facilities in set $\Xi(r) \setminus F_B$ have to stay closed. Let F_B be capable of serving a demand of $B = |C(r) \cup \Xi(r)|$, and denote the unused capacity at facilities in F_B with $Q_B = \sum_{i \in F_B} Q_i - |C(r) \cup \Xi(r)|$. Then, we achieve an optimal solution for this sub-problem by assigning Q_B customers with highest values $\Delta(i,r) = f_i - (c_{iz} + r)$ in set $W(r)$ to facilities in F_B and opening facilities at the remaining locations in $W(r)$.*

Proof. Consider a solution for the SSCFLPCA with uniform customer demands on a star graph with open facilities in set $\Xi(r)$ located at nodes $F_B \subseteq \Xi(r)$, which is capable of serving a demand of $B = |C(r) \cup \Xi(r)|$. Denote the unused capacity at facilities in F_B after serving all customers in $C(r) \cup \Xi(r)$ with $Q_B = \sum_{i \in F_B} Q_i - |C(r) \cup \Xi(r)|$ and suppose that Q_B customers with highest values $\Delta(i,r) = f_i - (c_{iz} + r)$ from set $W(r)$ are served at this unused capacity provided by facilities in F_B while the remaining customers in $W(r)$ are served at a facility at their location. Suppose now that there is another assignment of customers in $i \in W(r)$ to facilities in F_B which further decreases the total costs of the considered sub-problem.

Due to capacity restrictions, we are not allowed to assign more customers in $W(r)$ to facilities in F_B. Assigning less customers in $W(r)$ to facilities in F_B than allowed raises the total costs as we have to open a new facility in $W(r)$, which incurs higher costs than assigning any customer in $W(r)$ to a facility in $\Xi(r)$. Hence, assigning as many customers in $W(r)$ to facilities in F_B as possible is economically better than having unused capacity.

Suppose we achieve a cost-minimising solution for the considered sub-problem by assigning at least one customer from $W(r)$ to a facility in F_B who is not found by assigning customers with highest $\Delta(\cdot, r)$-values to facilities in set F_B. Denote this customer with k. Then, there is a customer $p \in W(r)$ with $\Delta(p,r) > \Delta(k,r)$, who is not assigned to a facility in F_B. Denote with $F = \bar{F}(r) \cup F_B \cup W_B$ the set of open facilities if customer p is assigned to a facility in F_B, i.e., the solution we receive from assigning customers with highest $\Delta(\cdot, r)$-value to facilities in F_B. Denote with $F_k = (\bar{F}(r) \cup F_B \cup W_B \setminus \{k\}) \cup \{p\}$ the set of open facilities if customer k instead of customer p is assigned to a facility in F_B. Let $\Lambda : V \to F_k, \ell \mapsto \Lambda(\ell)$ be the assignment function of customers in V to facilities in F_k and $\Lambda' : V \to F, \ell \mapsto \Lambda'(\ell)$ be the assignment function of customers in V to facilities in F. Since the facilities in F_B are given and the number of customers from set $W(r)$ served at facilities in F_B is fixed, the assignment costs of customers $\ell \in C(r) \cup \Xi(r)$ is the same for both assignment functions Λ, Λ', i.e., $c_{\ell\Lambda(\ell)} = c_{\ell\Lambda'(\ell)}$. The same holds for customers in set $W(r)$ who are assigned to facilities in F_B under both Λ and Λ'. Moreover, the assignment cost of any customer $\ell \in W(r) \setminus F$ to a facility in F_B under assignment function Λ' is $c_{\ell\Lambda'(\ell)} = c_{\ell z} + r$. The same holds for assignment function Λ and facilities in F_k. Then, the following inequality holds for the costs.

$$\begin{aligned}
c(F_k) &= \sum_{i \in F_k} f_i + \sum_{i \in V \setminus F_k} c_{i\Lambda(i)} \\
&= \sum_{\substack{i \in (\bar{F}(r) \cup F_B \\ \cup W_B \setminus \{k\}) \cup \{p\}}} f_i + \sum_{\substack{i \in V \setminus ((\bar{F}(r) \cup F_B \\ \cup W_B \setminus \{k\}) \cup \{p\})}} c_{i\Lambda(i)} + (c_{p\Lambda'(p)} - c_{p\Lambda'(p)}) + (f_k - f_k) \\
&= \sum_{\substack{i \in \bar{F}(r) \cup F_B \\ \cup W_B}} f_i + \sum_{\substack{i \in V \setminus (\bar{F}(r) \cup F_B \\ \cup W_B)}} c_{i\Lambda'(i)} + (f_p - c_{p\Lambda'(p)}) + (c_{k\Lambda(k)} - f_k) \\
&= \sum_{i \in F} f_i + \sum_{i \in V \setminus F} c_{i\Lambda'(i)} + \Delta(p,r) - \Delta(k,r) \\
&> \sum_{i \in F} f_i + \sum_{i \in V \setminus F} c_{i\Lambda'(i)} \quad = \quad c(\bar{F}(r) \cup F_B \cup W_B) \quad = \quad c(F).
\end{aligned}$$

This is a contradiction of the assumption that set F_k yields an optimal, i.e., cost-minimising, solution for the considered sub-problem. In conclusion, we can lower the costs by assigning customers with greatest differences between the price of opening a facility and assigning the customer to a facility in set F_B and our claim follows. $\square$

In conclusion, the cost of opening facilities in set $W(r)$ is determined by the set of open facilities in set $\Xi(r)$, which dictates the leftover capacity for serving customers in $W(r)$.

Step 4: Optimality of facilities in L From the previous step we know that we can efficiently compute an optimal assignment of customers in $W(r)$ if we are given a set of open facilities in set $\Xi(r)$. Of course, we are not forced to stop providing opening further facilities in set $\Xi(r)$ once we provide enough capacity to serve all customers in set $C(r) \cup \Xi(r)$. With extra capacity, we can serve further customers in $W(r)$.

In order to determine the actual solution for the sub-problem of the SSCFLPCA in which the closest open facility to the center node has assignment costs of r to the center, we solve the min-knapsack problem for bounds $B = |C(r) \cup \Xi(r)|, |C(r) \cup \Xi(r)| + 1, \dots, \sum_{i \in \Xi(r)} Q_i$. We denote the set of open facilities which can serve a demand of exactly B with F_B; if we do not find a subset M of open facilities which can serve a demand of exactly B but only a subset which serves more than B, remove value B from the consideration: we achieve the same result by setting the lower bound on the provided capacity to the sum of capacities of facilities in set M. For each B, we solve the assignment problem of customers in $W(r)$ to facilities in F_B in polynomial time according to the procedure in Step 3: we assign $B - |C(r) \cup \Xi(r)|$ customers from $W(r)$ with highest $\Delta(\cdot, r)$-values to facilities in set F_B and open facilities at the remaining locations in $W(r)$. This solution corresponds to the cost-minimising solution for the sub-problem in which all open facilities are located at nodes in set $F_B \subseteq \Xi(r)$. Optimality can be shown analogously to the proof in Proposition 2. For any set of open facilities F_B with provided capacity of B which we derive by solving a min-knapsack instance, the following result holds.

Proposition 3. *In an instance of the SSCFLPCA, let r be the assignment cost of the closest open facility to the center node. Suppose facilities in Ξ provide a capacity of B, where $B = |C(r) \cup \Xi(r)|, \dots, \sum_{i \in \Xi(r)} Q_i$. We obtain a solution for the sub-problem of the SSCFLPCA in which the capacity provided by the open facilities in $\Xi(r)$ is B if facilities in set $\Xi(r)$ are opened according to the min-knapsack solution described in Step 2 with lower bound on the weight of min-knapsack items of B.*

Proof. Denote with $F_B \subseteq \Xi(r)$ the set of facilities according to the min-knapsack solution described in Step 2 with lower bound B. Set F_B is supposed to minimise the total costs while providing a capacity of exactly B. Denote with $W_B \subseteq W(r)$ the set of open facilities in set $W(r)$ which minimises the total costs if the facilities in set F_B are opened. We saw how to compute set W_B efficiently in Step 3. Denote with $\Lambda : V \to F_B \cup W_B$ the assignment function of customers to the open facilities.

Suppose $F_B \cup W_B$ does not yield a cost-minimising solution in our sub-problem, in which the closest open facility has assignment costs of r to the center node and the provided capacity by open facilities in $\Xi(r)$ is equal to B. From Proposition 2, we know that the error must have occurred in set F_B. Suppose facilities $F'_B \cup W'_B$ provide the actual cost-minimising solution and denote with $\Lambda' : V \to F'_B \cup W'_B$ the corresponding assignment function. Note that $c_{\ell\Lambda(\ell)} = c_{\ell\Lambda'(\ell)}$ holds for all $\ell \in (C(r) \cup \Xi(r)) \setminus (F_B \cup F'_B)$ and all $\ell \in F_B \cap F'_B$. Since F_B is the solution of the min-knapsack problem, inequality $\sum_{i \in F_B} (f_i - (c_{iz} + r)) \leq \sum_{i \in F'_B} (f_i - (c_{iz} + r))$ holds. Moreover, sets F_B and W_B are disjoint and so are sets F'_B and W'_B. Per assumption, we assign $Q_B = \sum_{i \in F_B} Q_i - |C(r) \cup \Xi(r)| = B - |C(r) \cup \Xi(r)|$ customers from $W(r)$ to facilities in set F_B as well as to facilities in set F'_B. We solve the assignment problem of customers in $W(r)$ according to the procedure described in the previous step. This yields $\sum_{k \in W'_B} \Delta(k, r) = \sum_{p \in W_B} \Delta(p, r)$, since customers in $W(r)$ are indifferent between open facilities in $\Xi(r)$ and the provided capacity is the same in both F_B and F'_B. Therefore, we may assume without loss of generality that $W_B = W'_B$ holds. Combining the listed observations, we show that the costs of solution $(F'_B \cup W'_B, \Lambda')$ are not minimal.

$$c(F'_B \cup W'_B) \quad = \quad \sum_{i \in F'_B \cup W'_B} f_i + \sum_{i \in V \setminus (F'_B \cup W'_B)} c_{i\Lambda'(i)}$$

$$
\begin{aligned}
&= \sum_{i \in F'_B \cup W'_B} f_i + \sum_{i \in V \setminus (F'_B \cup W'_B)} c_{i\Lambda'(i)} + \sum_{p \in F_B} (f_p - f_p) + \sum_{k \in F'_B} (c_{k\Lambda(k)} - c_{k\Lambda(k)}) \\
&= \left[\sum_{i \in F'_B} f_i + \sum_{i \in F_B} f_i + \sum_{i \in W'_B} f_i \right] + \left[\sum_{i \in V \setminus (F'_B \cup F_B \cup W'_B)} c_{i\Lambda(i)} + \sum_{i \in F_B} c_{i\Lambda(i)} + \sum_{k \in F'_B} c_{k\Lambda(k)} \right] \\
&\quad - \sum_{p \in F_B} f_p - \sum_{k \in F'_B} c_{k\Lambda(k)} \\
&= \sum_{i \in F_B \cup W_B} f_i + \sum_{i \in V \setminus (F_B \cup W_B)} c_{i\Lambda(i)} + \sum_{p \in F_B} (c_{p\Lambda'(p)} - f_p) + \sum_{k \in F'_B} (f_k - c_{k\Lambda(k)}) \\
&= \sum_{i \in F_B \cup W_B} f_i + \sum_{i \in V \setminus (F_B \cup W_B)} c_{i\Lambda(i)} + \sum_{k \in F'_B} (f_k - (c_{kz} + r)) - \sum_{p \in F_B} (f_p - (c_{pz} + r)) \\
&\geq c(F_B \cup W_B).
\end{aligned}
$$

This is a contradiction to our assumption that solution $(F'_B \cup W'_B, \Lambda')$, which does not correspond to a min-knapsack solution, implies a cost-minimising solution. Hence, our claim holds. □

Then, by performing Steps 3, 4 for each $B = |C(r) \cup \Xi(r)|, \ldots, \sum_{i \in \Xi(r)} Q_i$, where B is the provided capacity by open facilities in $\Xi(r)$, we are able to determine the cost-minimising solution in the sub-problem in which the closest open facility to the center node has assignment costs of r to the center node.

Step 5: Conclusion Repeat steps 1 to 4 for each assignment cost $r \in R$ with the assumption that there is no open facility with lower assignment costs to the center node than r. The runtime of the described algorithm is proportional to the maximum location capacity.

Theorem 5. *Solving the SSCFLPCA with uniform customer demands on a star graph takes* $\mathcal{O}((|V| \cdot \sum_{i \in \Xi(r)} Q_i)^2)$ *steps.*

Proof. First, note that we have to consider each value $r \in R$ as the assignment costs from the closest open facility to the center node to the center. We have to consider at most $|R| \leq |V|$ assignment costs. Let us consider the number of steps it takes to solve a sub-problem in which the closest open facility to the center node has assignment costs of r to the center. We have to solve one min-knapsack problem for each $B = |C(r) \cup \Xi(r)|, \ldots, \sum_{i \in \Xi(r)} Q_i$, which corresponds to solving at most $\mathcal{O}(\sum_{i \in \Xi(r)} Q_i)$ sub-problems. Solving one of these sub-problems takes $\mathcal{O}\big(|V| \cdot \sum_{i \in \Xi(r)} Q_i\big)$ steps since we have to solve a min-knapsack instance with B at most $\sum_{i \in \Xi(r)} Q_i$. In conclusion, determining the cost-minimising solution for each choice of $r \in R$ takes $\mathcal{O}((|V| \cdot \sum_{i \in \Xi(r)} Q_i)^2)$ steps. It remains to compute the globally optimal solution, which can be done in linear number of steps and has no impact on the asymptotical number of steps. The same holds for determining sets $C(r)$, $\Xi(r)$ and $W(r)$ for each $r \in R$. □

Note that computing a solution for the SSCFLPCA with uniform customer demands on a star graph takes more steps than computing a solution for the SSCFLPCA with strict preferences on a star graph; cf. Theorem 4.

Corollary 4. *The SSCFLPCA with uniform customer demands on a star graph is weakly NP-hard.*

We conclude this section with a comparison of the complexity between the SSCFLP, the SSCFLPCA with strict preferences and the SSCFLPCA with uniform demands. First, recall that solutions for the SSCFLP are not generally feasible for the SSCFLPCA as customers in the SSCFLP are allowed to be served at facilities which are not their closest open facility. Therefore, solution algorithms for the SSCFLP do not generally yield solutions for the SSCFLPCA. Second, note that preference constraints neither ease the computational complexity of the considered problem - compared to the computational complexity of the SSCFLP- nor do these constraints increase the difficulty.

7 SSCFLPCA with lower bounds

In this section, we consider a generalisation of the SSCFLPCA, where lower bounds on the customer demand to be served at each open facility have to be respected in addition to the upper bounds. That is, each facility has a certain demand for customer demand as well as a capacity on it and these bounds on the demand have to be met at any open facility. In the following, we denote with $\ell_j \in \mathbb{N}_0$ the lower bound on the customer demand to be served at facility $j \in J$ if it is open.

We first show that a feasible assignment can be computed efficiently if each customer has uniform demand and the set of open facilities is known. Afterwards, we prove that computing a feasible solution is strongly NP-complete if the decision on the open facilities is yet to be made.

The assignment problem of the SSCFLPCA is polynomially solvable if a set of open facilities is given and each customer has uniform demand.

Theorem 6. *Given a set of open facilities, a feasible assignment of the customers to these facilities can be computed in polynomial time for the SSCFLPCA with uniform demands as well as lower bounds on the number of served customers at open facilities.*

Proof. We can assume without loss of generality that all demands are equal to one. Otherwise, we apply scaling and rounding to demands and capacities.

Let F be the set of open facilities. We introduce $\tilde{F} = \bigcup_{j \in F} \{j_1, \ldots, j_{Qj}\}$ to be the set containing a number of copies of each facility according to its capacity. We construct a bipartite graph $G = (I \cup \tilde{F}, E)$ with edges $E = \{\{i, j_k\} | i \in I, j \in F, k \leq Q_j : \nexists j' \in F : j' <_i j\}$, i.e., $i \in I$ is connected to the copies of $j \in F$ if j is one of the most preferred facilities of i. It is evident that a feasible assignment of customers exists if and only if there is a matching $M \subseteq E$ that matches every customer $i \in I$ and at least ℓ_j many copies of j for all $j \in F$. Given such a matching, we simply assign customer $i \in I$ to the facility $j \in F$ with whose copy the customer was matched. Hence, it is sufficient to compute such a matching to find a feasible assignment. To this end, we first compute a matching that meets the lower bounds on the number of served customers and then extend it to a matching with the desired properties.

Assume there exists a matching $M \subseteq E$ with the desired properties. Then M is a maximum matching with respect to cardinality in G, since $|M| = |I|$. Moreover, every matching of maximum cardinality in G matches all customers I. There exists a subset $M' \subseteq M$ that matches exactly ℓ_j many copies of j for all $j \in F$. We can assume without loss of generality that M' matches the first ℓ_j copies of j. Then M' is also a matching in the reduced graph $G' = (I \cup \tilde{F}', E')$ with $\tilde{F}' = \bigcup_{j \in F} \{j_1, \ldots, j_{\ell j}\}$ and $E' = E(I \cup \tilde{F}')$. Moreover, M' is a maximum matching in G', since $|M| = |\tilde{F}'|$, and thus every maximum matching in G' matches all nodes in $\tilde{F}'$. We can compute a maximum matching in G' in polynomial time Korte and Vygen (2018). Let M'' be such a matching. If M'' already matches all customers in I, then M'' corresponds to a feasible solution of the assignment problem. Otherwise, M'' is not a maximum matching in G. According to Berge's Theorem Korte and Vygen (2018), there exists an augmenting M''-path P in G, i.e., a path that alternates between edges in $E \setminus M''$ and M'' and that starts and ends in different exposed nodes. Finding an augmenting path P can be done in polynomial time via breadth-first search when starting in exposed nodes and traversing edges in an alternating fashion. Note that the symmetric difference $M'' \triangle P$ is a matching of size $|M''| + 1$. Moreover, every node that is matched by M'' is also matched by $M'' \triangle P$. In particular, $M'' \triangle P$ matches at least ℓ_j copies of j for every facility $j \in F$. We can iteratively augment M'' until it is a maximum matching in G. Then every customer in I is matched and, again, at least ℓ_j copies of j are matched for every facility $j \in F$. Hence, the matching corresponds to a feasible assignment. □

Note that Theorem 6 also holds for the more general case in which customer preferences may be arbitrary.

Contrary to the results in the previous sections, which show that feasible solutions for the SSCFLPCA with either distinct assignment costs or uniform demands can be computed in polynomial time, it is no longer possible to compute feasible solutions in polynomial time.

Theorem 7. *Computing a feasible solution for the SSCFLPCA with uniform demands and upper and lower capacities so that exactly two customers have to be served at any open facility is strongly NP-complete even if the triangle inequality holds for the assignment costs and the number of open facilities is already known.*

Proof. We prove our claim via a reduction from the strongly NP-complete problem *3-SAT* (Garey and Johnson, 1979). We start by transforming an instance of 3-SAT into an instance of the considered SSCFLPCA version. Consider a 3-SAT instance X consisting of n truth variables $x_1, \dots, x_n$ and k clauses $C_1, \dots, C_k$ with three literals each. Without loss of generality, we assume that each literal occurs in at least one clause. Denote the instance of the SSCFLPCA we are transforming X into with T. Denote the set of customers with I and introduce $|I| = (8 + 12k)n + 10k$ customers. Set the demand of each customer to one. We denote the set of potential facility locations with J and define each customer as a potential facility location. We set both the lower and upper bound of each facility to two and set the opening costs to zero. Due to the choice of capacities and facility demands, we know that any feasible solution consists of $|I|/2$ facilities. Last but not least, we define assignment costs of customers to the facilities. For that, we define the problem on an underlying undirected graph G where each customer and each facility corresponds to a node and the edges correspond to possibilities to reach various nodes amongst each other. We denote the set of nodes with V and the set of edges with E. Our graph consists of $|E| = (10 + 6k)n + 15k$ edges. Each edge is assigned a non-negative value and the assignment cost of a customer i to a facility j corresponds to the shortest path from the node corresponding to i to the node corresponding to j in G. Figure 1 depicts the construction from 3-SAT instance $(x_1 \vee x_2 \vee x_3) \wedge (x_1 \vee x_2 \vee \bar{x}_3) \wedge (\bar{x}_1 \vee \bar{x}_1 \vee x_3) \wedge (\bar{x}_2 \vee \bar{x}_2 \vee \bar{x}_3)$ to the class of graphs we consider in this proof.

It remains to prove that there is a feasible solution for 3-SAT instance X *if and only if* there is a feasible solution for the constructed SSCFLPCA instance T.

In order to prove this claim, we first have to understand the structure and dependencies within the underlying graph of instance T. The graph consists of two layers: an upper layer, which we refer to as *Layer 1*, and a lower layer, denoted with *Layer 2*. We start with the description of Layer 1. Layer 1 consists of one connected component consisting of $8 + 12k$ nodes and $10 + 6k$ edges for each variable in 3-SAT instance X. Denote the subgraph corresponding to variable x_i with L_1^i; in Figure 1, we put a blue box around each of these components. Subgraph L_1^i consists of three parts. That is: a *center part*, consisting of nodes $a_i, b_i, y_i^+, y_i^-, e_i, g_i, h_i, l_i$, a *left hand*, consisting of nodes $u_i^1, \dots, u_i^{3k}, v_i^1, \dots, v_i^{3k}$, as well as a *right hand*, consisting of nodes $w_i^1, \dots, w_i^{3k}, z_i^1, \dots, z_i^{3k}$. For the definition of the set of edges as well as the edge weights, we refer to Figure 1. In the reduction, the decisions regarding where to open facilities in the center parts correspond to decisions in 3-SAT; the left and right hand are auxiliary subgraphs each, as we see later.

Layer 2 consists of one subgraph for each clause in instance X. We denote the subgraph corresponding to clause C_j with L_2^j; we put an orange box around each of these subgraphs in Figure 1. Each such subgraph consists of ten nodes and twelve edges. Subgraph L_2^j consists of a complete graph with four vertices $s_j^0, s_j^1, s_j^2, s_j^3$. In addition to the complete graph, the subgraph consists of six further nodes, $m_j, n_j, o_j, p_j, q_j, r_j$ with edges $\{m_j, p_j\}, \{n_j, q_j\}, \{o_j, r_j\}$ and edges $\{p_j, s_j^1\}, \{q_j, s_j^2\}, \{r_j, s_j^3\}$. For the actual edge weights, we refer again to Figure 1.

Layers 1 and 2 are connected as follows. There are edges connecting both layers which are only incident to node y_i^+ or y_i^- in each subgraph L_1^i. Furthermore, in each subgraph L_2^j nodes m_j, n_j, o_j are incident to exactly one edge connecting both layers. The decision regarding which two nodes to connect through an edge follows from the 3-SAT instance X. We say that node y_i^+ corresponds to the positive literal of variable x_i; node y_i^- corresponds to the negative literal of x_i; nodes m_j, n_j, o_j correspond to the three literals in clause C_j, $j \in \{1, 2, \dots, k\}$. We introduce an edge from y_i^+ (y_i^-) to a node $v_j \in \{m_j, n_j, o_j\}$ which is not yet connected to a node in Layer 1 for each literal in clause C_j if the positive (negative) literal corresponding to node y_i^+ (y_i^-) occurs in clause C_j; cf. Figure 1. We then say that the literal in C_j corresponding to node v_j coincides with the positive (negative) literal corresponding to y_i^+ (y_i^-). Again, we refer to Figure 1 for the actual choice of edge weights.

Note that the triangle inequality holds for the assignment costs in the constructed graph. Any solution for instance T has the following properties, as we show after the main proof.

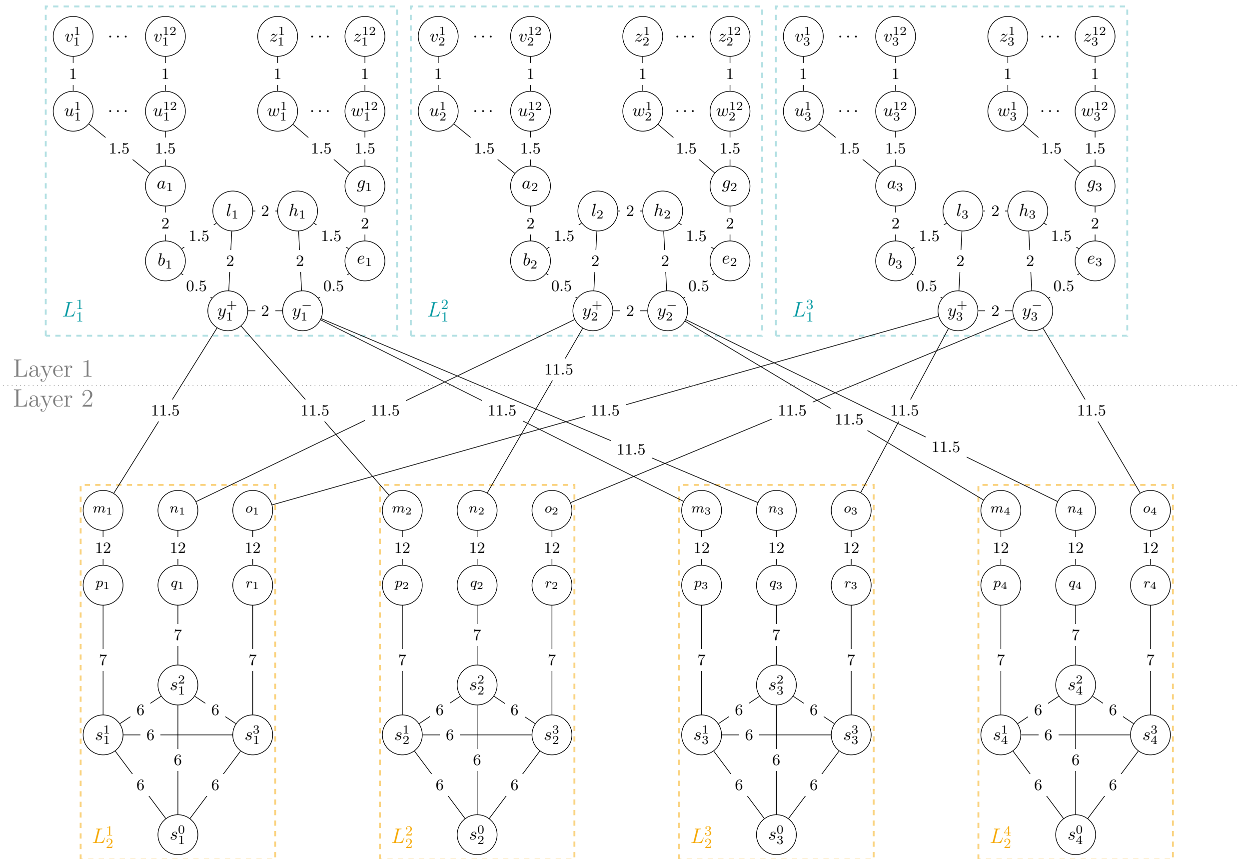

Figure 1: Transformation of 3-SAT instance $(x_1 \vee x_2 \vee x_3) \wedge (x_1 \vee x_2 \vee \bar{x}_3) \wedge (\bar{x}_1 \vee \bar{x}_1 \vee x_3) \wedge (\bar{x}_2 \vee \bar{x}_2 \vee \bar{x}_3)$ to considered SSCFLPCA instance.

Claim 1 There are exactly four open facilities in the center part of each subgraph L_1^i in Layer 1: either facilities $\{a_i, y_i^-, e_i, h_i\}$ or facilities $\{b_i, y_i^+, g_i, l_i\}$ are opened in subgraph L_1^i, for $i = 1, 2, \ldots, n$, in any feasible solution for instance T. Moreover, each customer is served within their subgraph.

Claim 2 At least one facilitiy has to be opened among nodes p_j, q_j, r_j in each subgraph L_2^j in any feasible solution for instance T.

Equipped with these properties, we prove the theorem. Consider a feasible solution of 3-SAT instance X. We transform this solution into a solution of T as follows. If variable x_i is set to TRUE, open facilities at nodes a_i, y_i^-, e_i, h_i in subgraph L_1^i as well as at all auxiliary nodes $v_i^1, \ldots, v_i^{3k}, z_i^1, \ldots, z_i^{3k}$. Since there is an open facility at node y_i^-, there has to be an open facility at all its neighbouring nodes in Layer 2 - otherwise, these neighbouring customers in Layer 2 would deviate to y_i^-, which is a contradiction to Claim 1's property that each customer is served within their own subgraph. If variable x_i is set to FALSE, open facilities at nodes b_i, y_i^+, g_i, l_i in subgraph L_1^i as well as at all auxiliary nodes $v_i^1, \ldots, v_i^{3k}, z_i^1, \ldots, z_i^{3k}$. Due to Claim 1, we have to open facilities at all nodes in Layer 2 which are neighbouring node y_i^+ in order to prevent them from deviating to a different subgraph. Given this setting of open facilities in Layer 1, assign each customer in Layer 1 with no open facility at their location to their closest open facility. We can compute this assignment in polynomial time. Note that we open $4 + 6k$ facilities in each subgraph in Layer 1 and due to the choice of open facilities, each facility serves exactly two customers. Since we are given a feasible truth assignment, there is at least one TRUE literal in each clause. Hence, amongst nodes m_j, n_j, o_j, at most two nodes are neighbouring an open facility in Layer 1 and at least one is not. Therefore, at most two facilities *have* to be opened amongst nodes m_j, n_j, o_j in each subgraph L_2^j; for the last node, no implications are made. Without loss of generality, suppose no facility has to be opened at node m_j, i.e., the literal in clause C_j corresponding to m_j is satisfied. Then, we open a facility at node p_j as well as at nodes s_j^0, s_j^1, and assign customer m_j to p_j, customer s_j^2 to s_j^1 and customer s_j^3 to s_j^0. Only the decisions of where to open facilities among nodes n_j, o_j, q_j, r_j are missing. If n_j, o_j are neighbouring an open facility in Layer 1, we know that we have to open facilities on these nodes too. If these nodes are not neighbouring an open facility in Layer 1, open facilities at nodes q_j, r_j. Assign customer n_j to q_j and customer o_j to r_j. Opening facilities as described yields a feasible solution for instance T as each open facility serves exactly two customers and each customer is served at their most preferred open facility in the constructed solution.

Conversely, consider a feasible solution to instance T. Due to Claim 1, there are either open facilities at nodes a_i, y_i^-, e_i, h_i or at nodes b_i, y_i^+, g_i, l_i in each subgraph on Layer 1. In our 3-SAT instance X, we assign a TRUE value to variable x_i if facilities a_i, y_i^-, e_i, h_i are opened in the solution; otherwise, we assign FALSE. It remains to prove that this assignment is feasible. Suppose there is at least one clause in which all literals are FALSE; let C_j be this clause. In this case, all of nodes m_j, n_j, o_j are neighbouring an open facility in Layer 1. Due to Claim 2, at least one facility has to be opened amongst nodes p_j, q_j, r_j; then, their neighbouring facilities at m_j, n_j, o_j would not meet their lower bounds on their demands - a contradiction. Hence, there is at least one TRUE literal in each clause. Furthermore, each variable is assigned to exactly one truth value since we must never open facilities simultaneously at locations y_j^+ and y_j^- in any subgraph L_1^i, as dictated by Claim 1. In conclusion, the constructed assignment is a feasible instance for 3-SAT instance X and the claim holds. Note that if a facility is opened at all nodes p_j, q_j, r_j in subgraph L_2^j, this corresponds to all literals in clause j being satisfied.

The problem lies in NP as we can check in polynomial time whether a set of open facilities implies a feasible solution.

It remains to prove that Claims 1 and 2 are indeed true. We start by proving Claim 1, i.e., that (1) either facilities $\{a_i, y_i^-, e_i, h_i\}$ or facilities $\{b_i, y_i^+, g_i, l_i\}$ are opened in the center part of subgraph L_1^i, for $i = 1, 2, \ldots, n$, in any feasible solution for instance T, and (2) each customer is served within their subgraph.

Claim 1 First, we study the case in which at least one facility is opened in a considered subgraph L_1^i. Afterwards, we study the case in which no facilities are opened.

Suppose at least one facility is opened in subgraph L_1^i. Then, none of the customers in this subgraph will deviate to another subgraph since any node in L_1^i can be reached from node y_i^+ (y_i^-) at cost strictly less than 11.5. Therefore, any customer in L_1^i prefers any facility in this subgraph to any potential facility in a different subgraph.

If at least one facility is open, we have to open exactly $4 + 6k$ facilities in considered subgraph, with four open facilities in the center part and $3k$ facilities in the left and right hand, each, as we show next. First, due to the capacities, the demands and due to the observation that no customer in L_1^i wants to deviate from their subgraph as soon as one facility is opened in it, we have to open $\frac{1}{2}(8 + 12k) = 4 + 6k$ facilities in said subgraph in order to serve all customers. Let us discuss the locations of these facilities next.

Consider the left hand of subgraph L_1^i first. Each pair u_i^t, v_i^t with $t = 1, 2, \ldots, 3k$ has to be served at the same facility, as we see next. First, suppose that we open a facility at location v_i^t. This is u_i^t's most-preferred open facility since its assignment costs to any other open facility are at least 1.5. If we open a facility at location u_i^t, it has to serve customer $v_i^{;otherwise,customer}$ v_i^t is either served at a facility they like less or, if we open another facility at their location, the facility at v_i^t does not serve two customers. Hence, in these two cases the claim holds. Second, suppose that customers u_i^t, v_i^t are served at other facilities. They must not be served at a facility located in $S := \{u_i^1, \ldots, u_i^{3k}, v_i^1, \ldots, v_i^{3k}, w_i^1, \ldots, w_i^{3k}, z_i^1, \ldots, z_i^{3k}\} \setminus \{u_i^t, v_i^t\}$: assume that either u_i^t or v_i^t is served at a facility located at $u \in S \setminus \{u_i^t, v_i^t\}$, the node incident to u in S is also served at u, as we can show in a similar manner to the behaviour of u_i^t, v_i^t if a facility is opened at one of them.

This violates the capacity of the facility at u - a contradiction. Then, customers u_i^t, v_i^t must be served at the center part. Suppose at least one facility is opened at a location in $S \setminus \{u_i^t, v_i^t\}$. Then, customers u_i^t, v_i^t must be served either at facility a_i, b_i or y_i^+: otherwise, both customers u_i^t, v_i^t deviate to open facilities in S. Now, both customers u_i^t, v_i^t want to be served at the same of these three facilities due to the assignment costs and since they all have distinct assignment costs to a_i. This incurs a demand of at least three at either a_i, b_i or y_i^+ - a violation of the capacities. Suppose no facility is opened in set S. Then, at least six customers have to be served at the center part (if $k = 1$) - we already know that the customers in the left hand must not be served at the right hand. However, there are less than six nodes in the center part with the same distance to node a_i. Therefore, there is no feasible solution in which facilities at the center node can serve all customers in the left hand without violating capacities or preference constraints. We therefore derive that each customer in the left hand is served at a facility located in the left hand if at least one facility is opened in subgraph L_1^i.

Analogously, we can show that the same holds for the right hand of subgraph L_1^i. The results for the left and right hand imply that customers in the center part have to be served at facilities located in the center part, as there is no capacity left in both hands. Suppose we open a facility at node a_i. Then, no customer in the left hand wishes to deviate to a_i as we showed just before. To satisfy a_i's lower bound, we must ensure that b_i is served at a_i. In order to achieve this, we have to keep facilities at locations b_i, l_i, y_i^+ closed. If we open a facility at node b_i, it will serve the customer at b_i and facility a_i will never meet their demand - a contradiction. If we open a facility at node l_i or y_i^+, customer b_i will deviate from a_i to one of them; then, neither b_i nor any other customer will ever be served at a_i - again, a contradiction to the demand constraint on the facility at a_i. Claim 1, which we want to prove, states that we are to open facilities at nodes y_i^- and h_i as well as at node e_i if we open a facility at a_i. There is no other alternative to this, as we see next. Since there are eight nodes in the center part and we already know that each customer located in the center part has to be served at a facility at the center part, we have to open three more facilities besides a_i. Moreover, facilities at b_i, y_i^+ and l_i have to stay closed. Thus, we have to open three facilities amongst locations y_i^-, e_i, g_i, h_i. Suppose we don't follow the proposed facilities from Claim 1. Then, we have to open a facility at h_i. Due to symmetry, we can argue that e_i has to be served at h_i analogously to b_i being served at a_i. Thus, if we open a facility at h_i, the facility at e_i has to stay closed. Then, however, we must keep the facilities at y_i^- and h_i closed too in order to stop e_i from deviating to one of them. We derive that we are only allowed to open two facilities, namely a_i, g_i, if these two facilities are opened simultaneously. This yields a violation of their capacities as these two facilities have to serve all customers in the center part. In conclusion, if we open a facility at a_i, we must not open a facility at g_i (and vice versa). Then, we open facilities at nodes y_i^-, e_i, g_i and it is easy to see that this indeed yields a feasible solution for the center part. Thus, if we open a facility at a_i, we must open facilities at y_i^-, e_i, h_i, as stated in Claim 1. If no facility is opened at node a_i, due to symmetry, we can show analogously to a_i being open that facilities at $\{b_i, y_i^+, g_i, l_i\}$ have to be opened. In summary, if we open a facility in subgraph L_1^i, the first part of Claim 1 holds and no customer wishes to deviate from Layer 1 to another subgraph.

Suppose there are no open facilities in subgraph L_1^i. Then, all $8+12k$ customers in the considerd subgraph have to be served at facilities in other subgraphs. Subgraph L_1^i has at most $3k$ outgoing edges to Layer 2: at most three edges to each clause. Then, there is at least one edge which is used by $12k/3k + 1 = 5$ customers to enter a subgraph in Layer 2.

If a facility in a subgraph in Layer 2 which is connected with L_1^i serves at least one of the customers from subgraph L_1^i, all customers from L_1^i are served by a subgraph in Layer 2 which is connected with L_1^i. Note that there might be multiple subgraphs neighbouring L_1^i. Let subgraph L_2^j be a neighbour of L_1^i, and suppose these two subgraphs are connected by one edge. Suppose that five customers use this edge to be served at a facility in L_2^j. Without loss of generality, let these customers enter L_2^j via node m_j. Suppose all of these five customers are served in L_2^j. This is impossible since there aren't five nodes with equal distance from node m_j in L_2^j, which we need in order to meet the capacity of each open facility. Thus, we must not open any facility in subgraph L_2^j at all.

If there is no open facility in subgraph L_1^i, there must not be an open facility in any subgraph on

Layer 2 which is neighbouring L_1^i. Hence, at least 15 customers, consisting of ten customers from L_2^j and at least five customers from L_1^i, have to be served in a subgraph on Layer 1 which is not L_1^i. Suppose there is such a serving graph on Layer 1, which we denote with $L_1^{i'}$. Suppose $L_1^{i'}$ is neighbouring subgraph L_2^j. Then, the facilities in $L_1^{i'}$ will serve at least $\lceil (5+10)/2 \rceil = 8$ customers from L_1^i and L_2^j: the minimum number of customers from L_1^i that take an edge to L_2^j plus all customers in L_2^j divided by two, as subgraph L_2^j is connected to at most two subgraphs which are not L_1^i - which is important since we assume that there is no open facility in L_1^i. If these eight customers enter subgraph $L_1^{i'}$ via node $y_{i'}^+$ $(y_{i'}^-)$, we need at least eight nodes with equivalent distance to them. Nodes with this property occur in set $\{u_{i'}^1, \ldots, u_{i'}^{3k}, v_{i'}^1, \ldots, v_{i'}^{3k}, w_{i'}^1, \ldots, w_{i'}^{3k}, z_{i'}^1, \ldots, z_{i'}^{3k}\}$. However, if we open facilities at nodes in set $\{u_{i'}^1, \ldots, u_{i'}, w_{i'}^1, \ldots, w_{i'}^{3k}\}$, these facilities have to serve their neighbouring nodes in set $\{v_{i'}^1, \ldots, v_{i'}^{3k}, z_{i'}^1, \ldots, z_{i'}^{3k}\}$ as well, and vice versa. This is a contradiction and the eight additional customers from L_1^i and L_2^j can not be served in $L_1^{i'}$ but have to be served in another subgraph. Then, no facility must be open in subgraph $L_1^{i'}$ or any other subgraph in Layer 1 neighbouring L_2^j in order to prevent the eight unassigned customers from deviating to one of their open facilities. Hence, the customers from subgraph $L_1^{i'}$ have to be served at a different subgraph as well. From before we already know that these customers can neither be served in a subgraph on Layer 2 nor in a subgraph on Layer 1. This is a contradiction to our assumption that there is a feasible solution in which all facilities in L_1^i are kept closed. Therefore, we have to open at least one facility in each sub-graph in Layer 1 and no customer in Layer 1 deviates to a facility in Layer 2. It remains to prove that no customer in Layer 2 deviates to a facility outside of their subgraph.

Finally, suppose some customers deviate from their subgraph in Layer 2 to another subgraph. Let L_2^j be the subgraph from which customers are deviating from. Suppose the customers deviate to a subgraph L_1^i in Layer 1. Then, at least one facility is opened in L_1^i, all customers located in L_1^i are served by facilities in L_1^i, and, in addition, at least one customer from L_2^j is served at a facility in L_1^i. In order to meet the lower and upper bounds on each open facility in L_1^i, at least two external customers have to be served at a facility in L_1^i.

First, suppose that these external customers enter through the same node and let this node be y_i^+. Then, these two customers and y_i^+ have to be served at a facility with the same distance. The only locations fulfilling this property are locations in set $\{u_{i'}^1, \ldots, u_{i'}^{3k}, v_{i'}^1, \ldots, v_{i'}^{3k}, w_{i'}^1, \ldots, w_{i'}^{3k}, z_{i'}^1, \ldots, z_{i'}^{3k}\}$; from above, however, we know that it is impossible to serve customers from outside the right or left hand at facilities located in the right or left hand. Due to symmetry of L_1^i, we also derive a contradiction if both external customers enter L_1^i through the node at y_i^-. Hence, the two external nodes must not enter through the same node.

Suppose at least one external customer enters via y_i^+ and at least one external customer enters via y_i^-. We must not open facilities at locations y_i^+ and y_i^-: if we do so, we must open facilities at b_i, e_i as well in order to stop the customers at these nodes from deviating to y_i^+, y_i^-, respectively. This setting does not yield a feasible solution, as we see next. Customer a_i has to be served at b_i and customer g_i has to be served at e_i since otherwise facilities at a_i, g_i would not meet their minimum demand. Then, customers h_i, l_i have to be served at the same facility; if we open a facility at h_i (l_i), customer l_i (h_i) deviates to b_i (e_i), which yields an infeasible solution.

Suppose we open a facility at y_i^+ and no facility at y_i^-. Then, the external customer entering via y_i^+ is served at y_i^+ and the external customer entering via y_i^- as well as y_i^- must be served at a facility with costs of at most two from y_i^- which is not y_i^+ - otherwise, they deviate to y_i^+ which will then have its capacity violated. This yields an immediate contradiction as there are no two nodes with the same distance to y_i^- in L_1^i with a distance of at most two, which are not y_i^+. Due to symmetry of L_1^i, we derive a contradiction if we keep the facility at y_i^+ closed and open the facility at y_i^-.

Thus, a customer from L_2^j can not be served at a facility in a neighbouring subgraph in Layer 1. The same holds for such a customer being served at a facility in a different subgraph in Layer 2: in this case, there is at least one subgraph in Layer 1, in which we don't open any facilities; from before, we know that this is impossible for any feasible solution. In conclusion, Claim 1 holds true.

Claim 2 Our goal is to show that at least one facility has to be opened among nodes p_j, q_j, r_j in each subgraph L_2^j in any feasible solution for instance T. Suppose there is no open facility at nodes p_j, q_j, r_j in subgraph L_2^j. From Claim 1 we know that each customer has to be served within their own subgraph and we open a facility either at y_i^+ or at y_i^- in any subgraph in Layer 1. To stop customers m_j, n_j, o_j neighbouring an open facility in Layer 1 from deviating to such a facility, we then have to open facilities at each m_j, n_j, o_j neighbouring an open facility in Layer 1. Suppose the customer at node m_j is not neighbouring an open facility in Layer 1. The assignment costs to their closest open facility in Layer 1 is then equal to 13.5. Thus, the customer in m_j has to be served at a facility in L_2^j with assignment costs of at most 13.5 in order to stop them from deviating to Layer 1. Since we assume that we keep the facilities at locations p_j, q_j, r_j closed, we, therefore, have to open a facility at m_j. We can show this property for all remaining customers in m_j, n_j, o_j, who are not neighbouring an open facility in Layer 1. Hence, we have to open facilities at nodes m_j, n_j, o_j in order to stop any customer in L_2^j from deviating to another subgraph if we keep facilities at p_j, q_j, r_j closed. In order to meet the capacity and demand constraints of the facilities at m_j, n_j, o_j, they have to serve customers p_j, q_j, r_j, respectively.

Furthermore, we have to open two facilities among nodes $s_j^0, s_j^1, s_j^2, s_j^3$ besides the facilities in m_j, n_j, o_j in order to serve the whole demand in subgraph L_2^j. Then, we have to open at least one facility among nodes s_j^1, s_j^2, s_j^3. Opening a facility at one of these three locations makes their neighbouring customer in p_j, q_j, r_j deviate from their serving facility in m_j, n_j, o_j to their neighbouring facility in s_j^1, s_j^2, s_j^3. This leaves at least one facility in m_j, n_j, o_j with too little served demand - a contradiction. In conclusion, we have to open at least one facility among locations p_j, q_j, r_j in any subgraph in Layer 2 in any feasible solution for instance T. □

Theorem 7 shows that adding both a lower and an upper bound to the UFLPSCP changes the combinatorial structure compared to adding either a lower or an upper bound.

Finally, note that a feasible solution for the SSCFLP with the same restrictions on the lower and upper bound on the capacities of facilities as in the previous theorem can be computed efficiently.

Corollary 5. *A feasible solution for the SSCFLP with uniform demands as well as upper and lower bounds so that exactly two customers can be served at any open facility can be computed in polynomial time.*

Proof. Given that each customer has a demand of one and each open facility an upper and lower bound of two, $|I|/2$ facilities have to be opened in any feasible solution. Hence, a feasible solution can be computed by opening $|I|/2$ facilities and assigning two customers to each open facility. □

The results in this section reveal that adding closest assignment constraints might increase the computational complexity of the traditional SSCFLP and has to be considered with care.

8 Conclusion

In this paper, we investigate the computational complexity of the single-source capacitated facility location problem with customer preferences (SSCFLPCP). While the complexity of (single-source capacitated) facility location problems is already well studied, less research has been conducted on the SSCFLPCP. Our results show that customer preferences can ease the strong NP-hardness, which holds for the SSCFLP. The main contributions in this work are that

- a feasible solution for the SSCFLPCP can be found in polynomial time if each customer has a strict preference ordering of the facilities;
- finding an optimal solution for the SSCFLPCP with arbitrary strict preferences is strongly NP-hard, even when neglecting capacities and the assignment costs are within respect to an underlying path graph;

- finding an optimal solution for the SSCFLPCA on star graphs is weakly NP-hard if customer preferences are strict or customer demands are uniform;
- a feasible solution for the *assignment problem* of the SSCFLPCA with uniform customer demands and lower and upper bounds on the demand to be served at each open facility can be found in polynomial time;
- finding a feasible solution for the SSCFLPCA with uniform customer demands and lower and upper bounds so that two customers have to be served at each open facility is strongly NP-complete.

Our work reveals that there are yet a lot of open research questions waiting. First, the polynomial algorithm that finds feasible solutions for the SSCFLPCA with strict preferenecs can be used in heuristics starting from feasible solutions. This raises the question whether heuristics for the traditional UFLP or the UFLP with customer preferences can be adapted to our problem. Second, the result that arbitrary preference constraints have a strong impact on the computational complexity of the SSCFLPCP raises the question whether there are further types of preference constraints besides closest assignments for which there are polynomially solvable or weakly NP-hard special cases. Third, the weak NP-hardness on star graphs with either strict preferences or uniform demands raises the question whether it is possible to extend the result to general trees. The consideration of special trees, such as spiders or caterpillars, might help in investigating this question. Finally, it is yet to be studied where the change from polynomial solvable cases to strongly NP-hard cases of the computational complexity of the SSCFLPCA with upper and lower bounds takes place: Büsing et al. (2022) show that this problem can be solved efficiently on paths while Theorem 7 shows that determining feasibility is already strongly NP-hard if customers have uniform demands and any open facility has to serve exactly two customers. Further mapping of the changes in the complexity is needed in order to answer these questions.

Acknowledgments

The authors would like to warmly thank Mariia Anapolska for her careful reading of this paper, her valuable comments, and her patience while enduring a lot of misplaced commas.

References

K. Aardal, Y. Pochet, and L. A. Wolsey. Capacitated facility location: Valid inequalities and facets. *Mathematics of Operations Research 20*, page 562–82, 1995.

F. Ahmad, A. Iqbal, I. Ashraf, M. Marzband, and I. Khan. Optimal location of electric vehicle charging station and its impact on distribution network: A review. *Energy Reports*, 8:2314–2333, 2022.

A. Ahmadi-Javid, P. Seyedi, and S. S. Syam. A survey of healthcare facility location. *Computers & Operations Research*, 79:223–263, 2017.

P. Avella and M. Boccia. A cutting plane algorithm for the capacitated facility location problem. *Computational Optimization and Applications*, 43:39–65, 2009.

P. Avella, M. Boccia, S. Mattia, and F. Rossi. Weak flow cover inequalities for the capacitated facility location problem. *European Journal of Operational Research*, 289(2):485–494, 2021. ISSN 0377-2217.

C. Büsing, T. Gersing, and S. Wrede. Analysing the complexity of facility location problems with capacities, revenues, and closest assignments. *Proceedings of the 10th International Network Optimization Conference (INOC), Aachen, Germany, June 7–10, 2022*, pages 81–86, 2022.

X. Cabezas and S. García. A semi-lagrangian relaxation heuristic algorithm for the simple plant location problem with order. *Journal of the Operational Research Society*, 0(0):1–12, 2022.

H. Calvete, C. Galé, J. Iranzo, J. Camacho-Vallejo, and M. Casas-Ramírez. A matheuristic for solving the bilevel approach of the facility location problem with cardinality constraints and preferences. *Computers & Operations Research*, 124:105066, 2020.

L. Cánovas, S. García, M. Labbé, and A. Marín. A strengthened formulation for the simple plant location problem with order. *Operations Research Letters*, 35(2):141–150, 2007.

M. Casas-Ramírez, J. Camacho-Vallejo, and I. Martínez-Salazar. Approximating solutions to a bilevel capacitated facility location problem with customer's patronization toward a list of preferences. *Applied Mathematics and Computation*, 319:369–386, 2018. Recent Advances in Computing.

D. Celik Turkoglu and M. Erol Genevois. A comparative survey of service facility location problems. *Annals of Operations Research*, 292:399–468, 2020.

I. Espejo, A. Marín, and A. M. Rodríguez-Chía. Closest assignment constraints in discrete location problems. *European Journal of Operational Research*, 219(1):49–58, 2012.

M. Fischetti, I. Ljubić, and M. Sinnl. Benders decomposition without separability: A computational study for capacitated facility location problems. *European Journal of Operational Research*, 253(3):557–569, 2016.

M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of $\mathcal{NP}$-completeness.* W. H. Freeman & Co., New York, USA, 1979. ISBN 978-0716710455.

R. A. Gerrard and R. L. Church. Closest assignment constraints and location models: Properties and structure. *Location Science*, 4(4):251–270, 1996.

S. Görtz and A. Klose. A simple but usually fast branch-and-bound algorithm for the capacitated facility location problem. *INFORMS Journal on Computing*, 24(4):597–610, 2012.

P. Hanjoul and D. Peeters. A facility location problem with clients' preference orderings. *Regional Science and Urban Economics*, 17(3):451–473, 1987.

B. H. Korte and J. Vygen. *Combinatorial Optimization: Theory and Algorithms.* Springer-Verlag, New York, NY, 2018. ISBN 978-3-662-56039-6.

G. Laporte, S. Nickel, and F. Saldanha da Gama, editors. *Location Science.* Springer, 2019. ISBN 978-3-030-32177-2.

J. M. Y. Leung and T. L. Magnanti. Valid inequalities and facets of the capacitated plant location problem. *Mathematical Programming*, 44:271–291, 1989.

A. M. Mestre, M. D. Oliveira, and A. P. Barbosa-Póvoa. Location–allocation approaches for hospital network planning under uncertainty. *European Journal of Operational Research*, 240(3):791–806, 2015.

P. Mirchandani and R. Francis, editors. *Discrete Location Theory.* Wiley Interscience, 1990. ISBN 0-471-89233-5.

S. Polino, J.-F. Camacho-Vallejo, and J. G. Villegas. A facility location problem for extracurricular workshop planning: bi-level model and metaheuristics. *International Transactions in Operational Research*, 0:1–43, 2023.

P. Rojeski and C. ReVelle. Central facilities location under an investment constraint. *Geographical Analysis*, 2(4):343–360, 1970.

I. Vasilyev, X. Klimentova, and Y. A. Kochetov. New lower bounds for the facility location problem with clients' preferences. *Computational Mathematics and Mathematical Physics*, 49:2010–2020, 2010.

I. Vasilyev, X. Klimentova, and M. Boccia. Polyhedral study of simple plant location problem with order. *Oper. Res. Lett.*, 41:153–158, 2013.

J. L. Wagner and L. M. Falkson. The optimal nodal location of public facilities with price-sensitive demand. *Geographical Analysis*, 7(1):69–83, 1975.