Integrality gaps for strengthened linear relaxations of Capacitated Facility Location*

Stavros G. Kolliopoulos† Yannis Moysoglou‡

Abstract

Metric uncapacitated facility location is a well-studied problem for which linear programming methods have been used with great success in deriving approximation algorithms. Capacitated facility location (Cfl) is a generalization for which there are local-search-based constant-factor approximations, while there is no known compact relaxation with constant integrality gap.

This paper produces, through a host of impossibility results, the first comprehensive investigation of the effectiveness of mathematical programming for metric capacitated facility location, with emphasis on lift-and-project methods. We show that the relaxations obtained from the natural LP at $\Omega(n)$ levels of the semidefinite Lovász-Schrijver hierarchy for mixed programs, and at $\Omega(n)$ levels of the Sherali-Adams hierarchy, have an integrality gap of $\Omega(n)$, where n is the number of facilities, partially answering an open question of [41, 5]. For the families of instances we consider, both hierarchies yield at the nth level an exact formulation for Cfl. Thus our bounds are asymptotically tight. Building on our methodology for the Sherali-Adams result, we prove that the standard Cfl relaxation enriched with the submodular inequalities of [1], a generalization of the flow-cover valid inequalities, has also an $\Omega(n)$ gap and thus not bounded by any constant. This disproves a long-standing conjecture of [39]. We finally introduce the family of proper relaxations which generalizes to its logical extreme the classic star relaxation and captures general configuration-style LPs. We characterize the behavior of proper relaxations for Cfl through a sharp threshold phenomenon.

Mathematics Subject Classification (2010) 90B80 Operations Research, Mathematical Programming; discrete location and assignment

1 Introduction

Facility location is one of the most well-studied families of models in combinatorial optimization. In the *uncapacitated facility location* problem *(*Ufl*)* we are given a set F of

*This work was co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: "Thalis: Investing in knowledge society through the European SocialFund'.

†Department of Informatics and Telecommunications, National and Kapodistrian University of Athens, Panepistimiopolis Ilissia, Athens 157 84, Greece; (www.di.uoa.gr/~sgk). Part of this work conducted while visiting the IEOR Department, Columbia University, New York, NY 10027.

‡Department of Informatics and Telecommunications, National and Kapodistrian University of Athens, Panepistimiopolis Ilissia, Athens 157 84, Greece; (gmoys@di.uoa.gr)

facilities and a set C of clients. We may open facility i by paying its opening cost f_i and we may assign client j to facility i by paying the connection cost c_{ij}. We are asked to open a subset $F' \subseteq F$ of the facilities and assign each client to an open facility. The goal is to minimize the total opening and connection cost. A *ρ-approximation algorithm,* $\rho \geq 1$, outputs in polynomial time a feasible solution with cost at most ρ times the optimum. The approximability of UFL was settled by an $O(\log |C|)$-approximation [27], which via a reduction from Set Cover is asymptotically best possible, unless $\mathsf{P} = \mathsf{NP}$ [47]. In *metric* UFL the service costs satisfy the following variant of the triangle inequality: $c_{ij} \leq c_{ij'} + c_{i'j'} + c_{i'j}$ for any $i, i' \in F$ and $j, j' \in C$. This natural special case of UFL is approximable within a constant-factor, and many improved results have been published over the years. In those, LP-based methods, such as filtering, randomized rounding and the primal-dual method have been particularly prominent (see, e.g., [55]). After a long series of papers the currently best approximation ratio for metric UFL is 1.488 [40], while the best known lower bound is 1.463, unless $\mathsf{P} = \mathsf{NP}$. The hardness result is due to an unpublished theorem by Sviridenko, stated in [54], which generalizes the complexity assumption of [26]. In this paper we focus on a generalization of metric UFL, the *capacitated facility location (*CFL*)*.

CFL is the generalization of metric UFL where every facility i has a capacity u_i that specifies the maximum number of clients that may be assigned to i. In *uniform* CFL all facilities have the same capacity U. Finding an approximation algorithm for CFL that uses a linear programming lower bound was until recently a notorious open problem. The natural LP relaxations have an unbounded integrality gap and up to the recent work of [6], the only known $O(1)$-approximation algorithms were based on local search, with the currently best ratios being 5 [10] for the non-uniform and 3 [4] for the uniform case respectively. In the special case where all facility costs are equal, CFL admits an LP-based 5-approximation [39]. Comparing the LP optimum against the solution output by an LP-based algorithm establishes a guarantee that is at least as strong as the one established a priori by worst-case analysis. In contrast, when a local search algorithm terminates, it is not at all clear what the lower bound is. Williamson and Shmoys [55], stated the design of a relaxation-based algorithm for CFL as one of the top 10 open problems in approximation algorithms. Very recently, An et al. [6] gave a polynomial-time LP-based 288-approximation algorithm, thus answering the open question of [55]. The LP in [6] has exponential size and is not known to be separable in polynomial time. Therefore the question on the existence of an efficient, compact, linear relaxation for CFL remains open.

A lot of effort has been devoted to understanding the quality of relaxations of 0-1 polytopes obtained by an iterative lift-and-project procedure. Such procedures define hierarchies of successively stronger relaxations, where valid inequalities are added at each level. At level at most d, where d is the number of variables, all valid inequalities have been added and thus the integer polytope is expressed. Relevant methods include those developed by Balas et al. [9], Lovász and Schrijver [42] (for linear and semidefinite programs), Sherali and Adams [50], Lasserre [36] (for semidefinite programs). See [37] for a comparative discussion.

The seminal work of Arora et al. [8], studied integrality gaps of families of relaxations for Vertex Cover, including relaxations in the Lovász-Schrijver (LS) hierarchy. This paper introduced the use of hierarchies as a restricted model of computation for obtaining LP-based hardness of approximation results. Proving that the integrality gap for a problem remains large for many levels of a hierarchy is an unconditional guarantee against the class of relaxation-based algorithms obtainable through the specific method. At the same time, if an LP relaxation maintains a gap of g for a linear number of levels, one can take this as evidence that compact, polynomial-size, relaxations are unlikely to yield approximations better than g (see also [49]). In fact, the former belief is now a theorem for maximum constraint satisfaction problems: in terms of approximation, LPs of size d^k, are exactly as powerful as $O(k)$-level Sherali-Adams (SA) relaxations [15]. From the algorithmic side, relaxations at the first $O(1)$ levels of LP hierarchies can be optimized in polynomial time as long as a weak separation oracle exists [42, 50].

We note that there are known examples of polytopes, such as the matching polytope, whose convex hull is obtained at high levels of the LS and the SA hierarchies [46], while at the same time there are known linear descriptions of the integer hull that can be optimized in polynomial time. The separation algorithm for the matching LP however has algorithmic steps that were not known to be translatable into a polynomial-size linear program. After the recent result of Rothvoß [48] we positively know this separation algorithm cannot be captured by a compact linear program and thus a polynomial-size formulation cannot be obtained. So, at least for matching, the weakness of hierarchies is due to an intrinsic difficulty of any compact LP formulation for the problem.

The result of Rothvoß [48] concerns the more general setting of *extended formulations* of a polytope P, where one may use additional auxiliary variables in the formulation as long as after projection to the original variable space, one obtains P. There are known examples, such as the permutahedron and the spanning tree polytope, where introducing extra variables decreases dramatically the size of the formulation. See [20] for further background. Techniques for lower bounding the size of extended formulations were introduced in the seminal paper of Yannakakis [56]. We note that the LP of [6] uses the standard variable space for the problem. It remains open whether a compact extended formulation exists for CFL.

This paper presents the first systematic theoretical study of the power of linear programming for approximating CFL, focusing especially on lift-and-project methods. In follow-up work [31] we show that linear relaxations in the space of the classic variables require at least an exponential number of constraints to achieve a bounded integrality gap. Note that it is well-known that hierarchies may produce an exponential number of inequalities already at the first level. For related problems there are some recent interesting results. Improved approximations were given for k-median [41] and capacitated k-center [22, 5], problems closely related to facility location. For both, the improvements are obtained by LP-based techniques that include preprocessing of the instance in order to defeat the known integrality gap. For k-median, the authors of [41] state that their $(1 + \sqrt{3} + \epsilon)$-approximation algorithm can be converted to a rounding algorithm on an $O(\frac{1}{\epsilon^2})$-level LP in the Sherali-Adams (SA) lift-and-project hierarchy. They propose exploring the direction of using SA for approximating CFL. An et al. [5] raise as an important question to understand the power of lift-and-project methods for capacitated location problems, including whether those methods capture "automatically" relevant preprocessing steps.

1.1 Our results

We give impossibility results on arguably the most promising directions for linear strengthened relaxations for CFL and in doing so we answer open problems from the literature.

Our first result (cf. Theorem 3.1) is that there is an instance with $\Theta(n)$ facilities and $\Theta(n^4)$ clients on which the relaxations produced at $\Omega(n)$ levels when the LS procedure is applied on the natural CFL LP have an integrality gap of $\Omega(n)$. The natural LP, defined in Section 2, has a facility opening variable y_i, for every $i \in F$, and an assignment variable x_{ij}, for every $i \in F$, and client $j \in C$.

The definition of the LS hierarchy is inductive. In order to obtain the relaxation at level $t + 1$, one applies the LS procedure to the relaxation at level t. Therefore, instead of levels, one may equivalently refer to the number of *rounds* for which the procedure has been applied. It is well-known that the LS procedure extends to mixed 0-1 programs [42, 9] such as CFL with general client demands. By lifting only the binary variables, the convex hull of the mixed integer feasible set is known to be obtained the latest at the pth level of the LS hierarchy, where p is the number of those binary variables ([42], [9, Theorem 2.6]). For CFL, p equals the number $|F|$ of facilities. In the instances we consider clients have unit demands and the capacities are integer; it is well-known that in this case the integer polytope and the mixed integer one (where fractional client assignments are allowed) are the same. In the proof of Theorem 3.1, we treat all variables as binary. In every round we obtain a polytope which is at least as tight as the one obtained when only the facility-opening variables are binary. Therefore our lower bound of $\Omega(n)$ applies also to the mixed integer LS procedure and is linear in the parameter p. In our proof we use a simple reformulation of the feasibility conditions of [42] for the strengthened relaxations. In a nutshell, the LS procedure reduces the survival at level t of the original bad solution s to the existence of a set of vectors with appropriate structure. These vectors act as "witnesses" of the survival of s and after suitable scaling they can be collected in the so-called *protection matrix* [8] at that level. Our proof gives an explicit, inductive definition of these witnesses for $t = \Omega(n)$. We believe that our constructive approach could be of independent interest in the context of the LS literature.

The LS_+ procedure is the stronger version of LS where one additionally requires that every protection matrix is positive semidefinite. The *mixed* LS_+ *procedure* for a mixed integer program is the version of LS_+ where one lifts only the 0-1 variables and requires that the resulting protection matrix is positive semidefinite (see, [9], [21]). Theorem 3.4, which follows almost immediately from the proof of our LS result, shows that the $\Omega(n)$ gap applies for $\Omega(n)$ rounds of mixed LS_+ as well. In fact, Theorem 3.4 shows this result for the stronger procedure where one lifts also the client assignment variables, but one requires that only the minor of the protection matrix that corresponds to the facility opening variables is positive semidefinite.

We then show that the LPs obtained from the natural relaxation for CFL at $\Omega(n)$ levels of the SA hierarchy have a gap of $\Omega(n)$ on the same family of instances used for the LS result, with $|F| = \Theta(n)$ and $|C| = \Theta(n^4)$. This result answers the questions of [41] and [5] stated above as far as the natural LP is concerned. Our bound is asymptotically tight since the relaxation obtained at every level of the SA hierarchy is at least as strong as the one obtained at the same level of LS .

We use the *local-to-global* method which was implicit in [8] for local-constraint relaxations and was then extended to the SA hierarchy in [23]. See also [24] for an explicit description and [17] for applications to Max Cut and other problems. In this approach, the feasibility of a solution for the t-level SA relaxation is established through the design of a set of appropriate distributions over feasible integer solutions for each constraint such that these global distributions agree with each other locally on relevant events. To prove the feasibility of a bad solution for CFL we devise first an intuitive method to construct an initial set of distributions for a constraint. These initial distributions are inadequate for constraints where all facilities appear as indices. An alteration procedure fixes that problem and produces the final set of distributions.

We note that, while the lower bound on the SA hierarchy implies the bound on the weaker LS hierarchy, it is in principle incomparable to the bound on the mixed LS_+ hierarchy. On the other hand, there are problems for which LS_+ is known to perform better than the LP hierarchies. For Stable Set one round of LS_+ is strictly stronger that one round of LS [42] and for Max Cut one round of LS_+ is strictly stronger than $\Omega(n^\delta)$ levels of SA [17]. From a qualitative aspect, we give the first, to our knowledge, hierarchy bounds for a relaxation where variables have more than one type of semantics, namely the facility opening and the client assignment type. Compare this, for example, with the Knapsack and Max Cut LPs that contain each one type of variable.

Our third contribution (cf. Theorem 5.1) is that the *submodular* inequalities introduced in [1, 2] for CFL fail to reduce the gap of the classic relaxation to constant. These constraints generalize the flow-cover inequalities for CFL. Thus we disprove the long-standing conjecture of [39] that the addition of the latter to the classic LP suffices for a constant integrality gap. Our proof deviates from standard integrality gap constructions by applying the local-global method. The bad solution fools every inequality π because its part that is *visible* to π, i.e., the variables in the support of π, can be extended to a solution that is a convex combination of feasible integer solutions.

We finally introduce the family of proper relaxations which are configuration-like linear programs. The so-called *Configuration LP* was used by Bansal and Sviridenko [11] for the Santa Claus problem and has yielded valuable insights, mostly for resource allocation and scheduling problems (e.g., [52]). The analogue of the Configuration LP for facility location already exists, it is the *star relaxation* (see, e.g., [28]). We take the idea of a star to its logical extreme by introducing classes. A *class* consists of a set with an arbitrary number of facilities and clients together with an assignment of each client to a facility in the set. A *proper relaxation* for an instance is defined by a collection $\mathcal{C}$ of classes and a decision variable for every class. We allow great freedom in defining $\mathcal{C}$: the only requirement is that the resulting formulation is symmetric and valid. The *complexity* α of a proper relaxation is the maximum fraction of the available facilities that are contained in a class of $\mathcal{C}$. In Theorem 6.1 we characterize the behavior of proper relaxations for CFL through a threshold result: anything less than maximum complexity results in a gap that is not bounded by any constant, while there are proper relaxations of maximum complexity with a gap of 1. The proof of Theorem 6.1 is combinatorial in nature and relies on Lemma 6.2. The latter lemma establishes that any proper relaxation which is valid for a family of instances we define must include a variable for a certain class with specific properties. Using that class and its symmetric ones, one can then define a bad fractional solution.

1.2 Other related work

Koropulu et al. [35] gave the first constant-factor approximation algorithm for uniform Cfl. Chudak and Williamson [19] obtained a ratio of 6, subsequently improved to 5.83 [16]. Pál et al. [45] gave the first constant-factor approximation for non-uniform Cfl. This was improved by Mahdian and Pál [43] and Zhang et al. [57] to a 5.83-approximation algorithm. As mentioned, the currently best guarantee is 5, due to Bansal et al. [10], while for the uniform case a 3-approximation exists [4]. All these results use local search.

In the *soft-capacitated* facility location problem one is allowed to open multiple copies of the same facility. Work on this problem includes [51, 18, 19, 29]. As observed in [28] a ρ-approximation for Ufl yields a 2ρ-approximation for the case with soft capacities. Mahdian, Ye and Zhang [44] noticed a sharper tradeoff and obtained a 2-approximation. A tradeoff between the blowup of capacities and the cost approximation for Cfl was studied in [3].

For hard capacities and general demands the feasibility of the *unsplittable* case, where the demand of each client has to be assigned to a single facility, is NP-complete, as Partition reduces to it. Bateni and Hajiaghayi [12] considered the unsplittable problem with an $(1+\epsilon)$ violation of the capacities and obtained an $O(\log n)$-approximation.

1.3 Outline of the paper

The outline of this paper is as follows. Section 2 contains elementary definitions. Section 3 includes the definition of the LS hierarchy in Subsection 3.1, the reformulation of the latter that we use in our proofs in Subsection 3.2, and the proof of our results for the LS and LS_+ hierarchies for Cfl in Subsections 3.3 and 3.4 respectively. In Section 4 we present our result on the SA hierarchy for Cfl. In particular, Subsection 4.1 contains necessary background, Subsection 4.2 the proof of the result, and in Subsection 4.3 we comment on the robustness of our proof. In Section 5 we give the definitions of the flow-cover, effective capacity and submodular inequalities and then the proof of our theorem regarding their strength as far as the quality of approximation is concerned. Finally, in Section 6 we introduce the proper relaxations and in Subsections 6.1 and 6.2 we prove a theorem that characterizes their strength through a sharp threshold phenomenon. We conclude in Section 7 with a brief discussion.

A preliminary version of the results in Sections 4, 5 and 6 appears in the conference paper [34]. Versions that are by now obsolete appeared in [32, 33].

2 Preliminaries

Let $I(F, C)$ be an instance of Cfl. We will show our negative results for uniform, integer capacities. Each client can be thought of as representing one unit of demand. It is well-known that in such a setting the splittable and unsplittable versions of the problem are equivalent. The following 0-1 IP is the standard valid formulation of uncapacitated facility location with unsplittable unit demands. Variable y_i has the meaning of opening facility $i \in F$ and variable x_{ij} stands for assigning client $j \in C$ to facility $i \in F$.

$$\min \sum_{i \in F} f_i y_i + \sum_{i \in F} \sum_{j \in C} x_{ij} c_{ij} \tag{1}$$

$$x_{ij} \leq y_i \qquad \forall i \in F, \forall j \in C \tag{2}$$

$$\sum_{i \in F} x_{ij} = 1 \qquad \forall j \in C \tag{3}$$

$$x_{ij} \in \{0,1\} \qquad \forall i \in F, \forall j \in C \tag{4}$$

$$y_i \in \{0,1\} \qquad \forall i \in F \tag{5}$$

Relaxing the integrality constraints from the above IP yields:

$$0 \leq x_{ij} \leq 1 \qquad \forall i \in F, \forall j \in C \tag{6}$$

$$0 \leq y_i \leq 1 \qquad \forall i \in F \tag{7}$$

To obtain the standard LP relaxation for uniform CFL, where each facility has capacity U, the following constraints are added:

$$\sum_{j \in C} x_{ij} \leq U y_i \qquad \forall i \in F \tag{8}$$

The standard LP relaxation for CFL consists of (1), (2), (3), (6), (7), (8). In the rest of the paper we refer to this LP by the term *(LP-classic)*. It is easy to see that (LP-classic) has a super-constant integrality gap: consider the following family of instances $\mathcal{H}$. For every $n \geq 1$, $\mathcal{H}$ contains the following instance. There are 2 facilities with $f_1 = 0$ and $f_2 = 1$. The number of clients is $n+1$, all facilities and clients are co-located at the same point and both capacities are equal to n. By opening the first facility integrally and the second at a fraction of $\frac{1}{n}$ and assigning each client with a fraction of $\frac{n}{n+1}$ to the first and with a fraction of $\frac{1}{n+1}$ to the second, we get a solution that is feasible for (LP-classic) with a cost of $1/n$ (see Figure 1). On the other hand, each integral solution has a cost of 1 as both facilities must be open in order to satisfy the client demand.

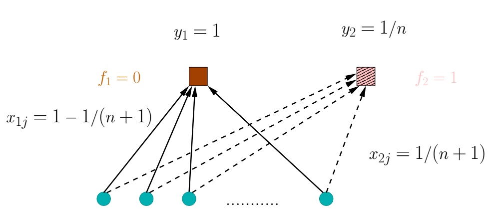

Figure 1: Representative instance from family $\mathcal{H}$ and the associated fractional solution to (LP-classic).

The instances and solutions that we will use in our proofs in Sections 3, 4 and 5 generalize the former example. The parameterized family $\mathcal{I}$ contains the following bad instance for every $n \geq 1$. Consider a set of n facilities which have 0 opening cost. We call that set *Cheap*. Moreover, consider a set of n facilities that have an opening cost of 1 each. Call that set *Costly*. The set of facilities F is $Cheap \cup Costly$. Let all the facilities have the same capacity $U = n^3$, and let there be a total of $nU + 1$ clients in the set C. All clients and facilities are at a distance of 0 from each other. Clearly all integer solutions to the instance have a cost of at least 1. The bad fractional solution s to (LP-classic) is the following: for each facility $i \in Cheap$, $y_i = 1$, and for each client j, set $x_{ij} = \frac{(1-a)}{n}$, $a = n^{-2}$. For each facility $i \in Costly$, $y_i = b$ where $b = 10/n^2$, and for each client j, set $x_{ij} = a/n$ (see Figure 2). The constructed solution incurs a cost of $\Theta(n^{-1})$.

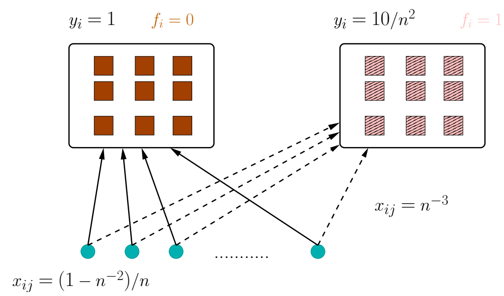

Figure 2: Representative instance from family $\mathcal{I}$ and the associated fractional solution s that will be used in the upcoming sections.

Given two vectors $y = (y_1, \ldots, y_n)^T$ and $x = (x_1, \ldots, x_m)^T$ by (y, x) we denote the vector $(y_1, \ldots, y_n, x_1, \ldots, x_m)^T$. We slightly abuse notation and use $(1, x)$ to denote the vector $(1, x_1, \ldots, x_m)^T$.

3 LS gaps for Cfl

3.1 The LS Hierarchy

The Lovász-Schrijver hierarchy of relaxations was defined in [42]. For a comprehensive presentation and various reformulations see [53]. In this section we give the definitions necessary for our proof.

In [42] an operator N was defined through the use of which we can refine a convex set $P \subseteq [0,1]^d$. Starting with the polytope of interest $P \subseteq [0,1]^d$ we define $\text{cone}(P) = \{y = (\lambda, \lambda z_1, ..., \lambda z_d)^T \mid \lambda \geq 0, (z_1, ..., z_d)^T \in P\}$. For such a conified polytope in $\mathbb{R}^{d+1}_{\geq 0}$ we follow the convention that the indices of the coordinates range from 0 to d. Observe that the projection of $\text{cone}(P)$ onto $x_0 = 1$ equals P. One can define the N operator through lift-and-project operations over the linear system defining P (see, e.g., [42, 9]). For our purposes it is sufficient to use the following formalism. For a cone $K \subseteq [0,1]^{d+1}$, and $t \geq 1$, we define the notation $N^0(K) = K$ and $N^t(K) = N(N^{t-1}(K))$.

Definition 3.1 ([42]) *Let K be a cone in $\mathbb{R}^{d+1}$. For $m \geq 1$, $N^m(K)$ is the set of vectors $z \in \mathbb{R}^{d+1}$ for which there is a $(d+1) \times (d+1)$ symmetric matrix Y satisfying*

1. $Ye_0 = diag(Y) = z$.

2. *For $1 \leq i \leq d$, both Ye_i and $Y(e_0 - e_i)$ are in $N^{m-1}(K)$.*

Y is called the protection matrix *of z.*

Lovász and Schrijver [42] showed that for any polytope $P \subseteq [0,1]^d$, the projection of $N^d(\text{cone}(P))$ onto $x_0 = 1$ equals the integer hull of P.

The operator N_+ is defined as above with the additional restriction that Y has to be positive semidefinite. The LS *procedure* consists of repeated applications of the operator N starting from $K = \text{cone}(P)$, where P is the polytope of the original relaxation. Every such application is called a *round.* If there is a protection matrix Y for z, we say that z *survives* one round of LS. Given a cost vector c, the LS procedure *on P produces after m rounds,* $m \geq 1$, the linear relaxation

$$\min\{c^T x \mid (1, x) \in N^m(K)\}.$$

Similarly, the LS_+ procedure on P *produces after m rounds,* $m \geq 1$, the linear relaxation

$$\min\{c^T x \mid (1, x) \in N_+^m(K)\}.$$

The discussion so far has treated P as the relaxation of an integer 0-1 program with d variables. For a mixed integer program with $d + p$ variables where one optimizes over $\text{conv}(P \cap (\{0,1\}^d \times \mathbb{R}^p))$, to derive an exact formulation it is sufficient to lift in each round only the d integer variables [9, 21]. To reflect this we define the operators M and M_+ as follows.

Definition 3.2 ([42, 9]) *Let K be a cone in $\mathbb{R}^{d+p+1}$ and let $M^0(K) = K$. For $m \geq 1$, $M^m(K)$ is the set of vectors $z = (y, x) \in \mathbb{R}^{d+p+1}$, $y \in \mathbb{R}^{d+1}$, $x \in \mathbb{R}^p$, for which there is a $(d+1) \times (d+1)$ symmetric matrix Y satisfying*

1. *$Ye_0 = diag(Y) = y$.*
2. *For $1 \leq i \leq d$, there exists $x^{(i)} \in \mathbb{R}^p$ such that both $(Ye_i, x^{(i)})$ and $(Y(e_0 - e_i), x - x^{(i)})$ are in $M^{m-1}(K)$.*

When in the definition above Y is constrained to be positive semidefinite, we obtain the operator M_+. Clearly, for any cone K, $N(K) \subseteq M(K)$ and $N_+(K) \subseteq M_+(K)$. The *mixed* LS_+ *procedure* applies repeatedly the M_+ operator.

3.2 Reformulation of the LS procedure

In this section we highlight some simple properties of protection matrices and explain how they will be used in our proofs.

Since we are interested in the projection of the cones resulting from the successive application of the operator N on the hyperplane $z_0 = 1$ which contains our original polytope P, we restate the conditions of survival of z as the following lemma which is immediate from Definition 3.1.

Lemma 3.1 ([7]) *Let $P \subseteq [0,1]^d$ be a polytope and denote by K the set* $\text{cone}(P)$ *in* $\mathbb{R}^{d+1}_{\geq 0}$. *A vector $z \in \mathbb{R}^{d+1}$ with $z_0 = 1$ is in $N^m(K)$, $m \geq 1$, iff there is a $(d+1) \times (d+1)$ symmetric matrix Y satisfying*

1. $Ye_0 = diag(Y) = z$.
2. *For $1 \leq i \leq d$: If $z_i = 0$ then $Ye_i = \mathbf{0}$; If $z_i = 1$ then $Ye_i = z$; Otherwise, Ye_i/z_i, $Y(e_0 - e_i)/(1 - z_i)$ both lie in the projection of $N^{m-1}(K)$ onto the hyperplane $z_0 = 1$.*

Let Y_i denote the vector $Y^T e_i$, i.e., the ith row of Y. Lemma 3.1 makes it convenient to work with individual vector solutions that can be combined as rows to build the protection matrix. Given a protection matrix Y of z, we define a set of at most $2d$ *witnesses* of vector z. For each variable z_i, $1 \leq i \leq d$, there are at most 2 such witnesses: the one that equals Y_i/z_i (if $z_i \neq 0$), which we call *type 1 witness of z corresponding to variable z_i*, and the vector $\frac{Y_0 - Y_i}{1 - z_i}$ (if $z_i \neq 1$), which we call *type 2 witness of z corresponding to variable z_i*.

Before proceeding, we give a brief preview of the proof strategy for our main LS result, Theorem 3.1. To prove the existence of a protection matrix Y for a vector z, we will use the following steps. We will define a set $S(z)$ of vectors, which consists of a "candidate" type 1 and a "candidate" type 2 witness for every non-integer variable and one of the appropriate type for each integer variable, with respect to some as of yet unknown protection matrix. We will ensure that the vectors in $S(z)$ meet a set of conditions (described in Lemma 3.2 below) which are sufficient so that $S(z)$ corresponds to a set of *actual* witnesses from a correct protection matrix Y of z. To prove the survival of a vector for many rounds we just embed the strategy above in an inductive argument. The following fact is immediate from Lemma 3.1: if, for all i the vectors Y_i/z_i and $\frac{Y_0 - Y_i}{1 - z_i}$ (that are defined) witnessing the survival of our initial vector z, survive themselves k rounds of LS, then z survives actually $k + 1$ rounds of LS.

We are now ready to start presenting a reformulation of Lemma 3.1. For the validity of the following observation recall that if $z_i = 0$, and hence the type 1 witness corresponding to i is undefined, $Y_i = \mathbf{0}$.

Observation 3.1 *The fact that Y's main diagonal is equal to the vector Y_0 is equivalent to the following* Condition (i): the variable z'_i of the type 1 witness z' corresponding to variable $z_i \neq 0$ is equal to 1, and equal to 0 if $z_i = 0$.

The rows of Y that correspond to zero variables in z are filled with zeros. Moreover if $z_i = 1$, $Y_i = z$. To account for these requirements it is not enough that the integer values in Y_0 appear on the main diagonal. The following observation states that they are replicated across all witnesses.

Observation 3.2 *Let z' be a witness of z.* [Condition (ii):] if for some i, $z_i \in \{0, 1\}$, then $z'_i = z_i$.

To enforce symmetry for a row $Y_i = \mathbf{0}$ that corresponds to a variable $z_i = 0$, it must be the case that the ith column is set to zero as well. This is ensured by Observation 3.2 for all entries Y_{ki} of the column for which $z_k \neq 0$. (The remaining entries of the column belong to zero rows and are equal to zero anyway). For the remaining rows, it will be convenient to express the type 1 witness z' of z corresponding to some variable z_i, by defining the factors by which the variables of z' differ from the corresponding variables of z. Then the symmetry condition of Y is satisfied by ensuring that the condition of the following proposition on those factors holds.

Proposition 3.1 *Let indices q, t take values in $\{1, \ldots, d\}$. The symmetry condition of the protection matrix of z holds iff Condition (ii) from Observation 3.2 holds together with the following* Condition (iii): for each type 1 witness z' of z corresponding to variable z_q, for which $z'_t = z_t f$, $z_t \neq 0$, then, for the type 1 witness z'' of z corresponding to variable z_t, we have $z''_q = z_q f$.

The above proposition is obtained by just observing that $Y_{qt} = z_q z'_t = z_q z_t f = z_t z''_q = Y_{tq}$ for non-integer z_q, z_t and by the fact that $Y_i = z$ for integer z_i. Note that when we construct a type 1 witness z' corresponding to z_t, the type 2 witness z'' corresponding to z_t is automatically defined. We say that z'' is the *twin* of z' .

Proposition 3.2 *Let indices q, t take values in $\{1, \ldots, d\}$. If Y is a protection matrix of z, the following* Condition (iv) must hold: if $z'_q = z_q(1+\epsilon)$ in the type 1 witness z' corresponding to z_t, $z_t \neq 1$, then $z''_q = z_q(1 - \frac{z_t \epsilon}{1-z_t})$ where z'' is the type 2 twin of z'.

Based on Observation 3.1 and Propositions 3.1, 3.2, the following lemma is now a straightforward reformulation of Lemma 3.1.

Lemma 3.2 *Let $z \in \mathbb{R}^d$ and let $S(z)$ be a collection of vectors that lie in the projection of $N^{m-1}(K)$ onto the hyperplane $z_o = 1$. The vectors in $S(z)$ satisfy the Conditions (i)-(iv) stated in Observation 3.1 and Propositions 3.1 and 3.2 iff there is a protection matrix Y for z such that for each $z_i \neq 0$ there is exactly one $z' \in S(z)$ such that $z' = Y_i/z_i$ and for each $z_i \neq 1$ there is exactly one $z'' \in S(z)$ such that $z'' = \frac{Y_0 - Y_i}{1-z_i}$.*

Lemma 3.2 shows how to produce a protection matrix that ensures the survival of a vector z for a single round of LS . To show survival for several rounds, we implement the inductive argument mentioned above by defining inductively an appropriate tree structure. In particular, we define the *evolution tree T_z of z*. Every node in T_z is *associated* with a vector. The tree is defined recursively: vector z is associated with the root of the tree, and if v is a node of T_z, associated with vector $z(v)$, then a set of vectors witnessing the survival of $z(v)$ is associated in one-to-one manner with the children of v. If there is no such set of witnesses, the fractional solution $z(v)$ does not survive one round of LS. The length of the shortest path from the root of an evolution tree T_z to a childless node is a lower bound on the number of rounds that our initial vector z survives.

Given a root vector, we will show that as long as we have walked down the evolution tree at depth k, where k is the target number of rounds, then the protection matrices of the root and all its descendants are well-defined. The inductive step shows how to define all children of a node v and therefore increase the depth of the tree by one. From now on we refer interchangeably to a node and its associated solution vector. Accordingly, if v' is a child of node v, z' (z) is associated with v' (resp. v) and z' is a type 1 (2) witness of z corresponding to variable z_i, we will refer to node v' as a *type* 1 *(resp.* 2*) child of node-solution z corresponding to variable z_i.*

Finally, the following fact will be useful for the feasibility proof.

Lemma 3.3 *Given a solution z in the evolution tree that satisfies an equality constraint $\sum_i a_i z_i = b$, and given a child of z that is a type* 1 *solution z' corresponding to some z_t that satisfies $\sum_i a_i z'_i = b$, then the twin type* 2 *solution z'' of z' also satisfies $\sum_i a_i z''_i = b$.*

Proof. Let $z'_i = z_i(1 + \epsilon_i)$. From $\sum_i a_i z_i = b$ and $\sum_i a_i z'_i = b$ we get $\sum_i a_i z_i \epsilon_i = 0$. Then by Proposition 3.2,

$$\sum_i a_i z''_i = \sum_i a_i z_i (1 - \frac{z_t \epsilon_i}{1 - z_t}) = \sum_i a_i z_i - \frac{z_t}{1 - z_t} \sum_i a_i z_i \epsilon_i = b.$$

∎

3.3 LS gaps for approximate Cfl

In this section we show that the integrality gap on a suitable instance of Cfl remains poor even after a large number of iterations of the LS procedure. More precisely we prove the following result.

Theorem 3.1 *There is a family of instances $\mathcal{I}$ of uniform* Cfl*, such that for every sufficiently large n, there is an instance in the family with $2n$ facilities and $n^4 + 1$ clients for which the relaxation produced by the* LS *procedure in $\Omega(n)$ rounds has integrality gap $\Omega(n)$.*

For the families of instances we consider (unit demands and integer capacities) our bound is the best possible asymptotically. The following is well-known and is included for the sake of completeness.

Theorem 3.2 *[9] For every instance of* Cfl *with unit demands and integer capacities, the* LS *and mixed* LS$_+$ *procedures obtain each the integer hull after a number of rounds that is equal to the number of the facilities.*

Proof. In the case of unit demands and integer capacities the fully integer and the mixed Cfl polytopes are the same. The LS procedure applies to mixed integer programs as well where only the integer variables are lifted and the mixed integer polytope is obtained after a number of rounds that is equal to the number of the integer variables [9]. For Cfl, the lifting of just the facility variables produces a weaker relaxation than the lifting of both the facility and the assignment variables that our LS procedure implements. Thus our LS procedure obtains the integral polytope after at most a number of rounds that is equal to the number of the facilities. ∎

It is well-known that the LS procedure may take a large number of rounds to produce some simple valid inequalities. In the case of CFL, Theorem 3.1 implies that $\Theta(n)$ rounds are required to obtain, e.g., the inequality

$$\sum_{i \in F} y_i \geq \lceil |C|/U \rceil \tag{9}$$

which is facet-inducing [38, p. 283] for our instance. However, this inequality is not critical for our proof. It is easy to modify the input by adding one facility f' with zero opening cost and one client c' which are co-located at a large distance from the rest of the instance. There is a bad fractional solution in which f' is fully opened and client c' is integrally assigned to f'; this solution satisfies the inequality above. Theorem 3.1 continues to hold for the augmented instance: variables with integer values are handled trivially in our construction, cf. Observation 3.2 and the proof of Theorem 3.3 below.

To prove Theorem 3.1, we will consider an instance with $2n$ facilities from the family $\mathcal{I}$ and the associated bad solution $s = (y, x)$ defined in Section 2. Before proving a lower bound on the number of rounds s survives, we establish an easy upper bound.

Observation 3.3 *Solution s survives less than n rounds of the LS procedure.*

For the proof of Observation 3.3 the following more general lemma suffices:

Lemma 3.4 *Let $K \subseteq [0,1]^d$ be a polytope. Let S be a set of indices of variables of vector $g \in [0,1]^d$, s.t. $\sum_{i \in S} g_i < 1$. If $(1, g)$ belongs to $N^{|S|}(\text{cone}(K))$, then there is a vector g^* such that $(1, g^*)$ belongs to $\text{cone}(K)$ and for all $i \in S$ $g_i^* = 0$.*

Proof. Every vector in the feasible set obtained after r rounds of LS has the property that for every tuple of r coordinates it can be expressed as a convex combination of vectors that are feasible for the starting relaxation K and have integer values on those coordinates [42]. Thus, at least one such vector g^* in the corresponding convex combination that yields g must contain 0 values at all positions indexed by $i \in S$. ∎

To prove Observation 3.3, assume that s survives n rounds. Apply Lemma 3.4 with K being the feasible set of (LP-classic). There must be a vector in K in which all the y_i variables, for $i \in Costly$, are set to 0. This cannot be a feasible solution to (LP-classic) since at least one costly facility has to be opened by a nonzero fraction, a contradiction.

We are ready to state the main theorem of this section which implies that the solution s survives $n/10$ rounds of LS. We do not make any attempt to optimize the constant. The intuition is that at every level of the induction the new witness solutions cannot differ drastically from their parent node. We identify a set of invariants that express this controlled evolution of the values. Recall that $b = 10/n^2$ is the value of the costly facility variables in s.

Theorem 3.3 *Let n be sufficiently large. We can construct an evolution tree T_s with root s such that any node u of T_s at depth $k \leq \frac{n}{10}$ is associated with a feasible solution (y, x) that satisfies the following invariants:*

1 For variable $y_i \notin \{0,1\}$, $i \in Costly$, $b - 2k\frac{a}{n} \le y_i \le b + 2k\frac{a}{n}$.

2 (a) For variable $x_{ij} \notin \{0,1\}$, $i \in Cheap$, $\frac{1-a}{n} - 2k\frac{a}{n^2}b^{-1} \le x_{ij} \le \frac{1-a}{n} + 2k\frac{1-a}{n^2}$.

(b) For variable $x_{ij} \neq 0$, $i \in Costly$, and $y_i \notin \{0,1\}$, $\frac{a}{n} \le x_{ij} \le \frac{a}{n} + 2k\frac{a(1-a)}{n^2}$.

(c) For variable $x_{ij} \notin \{0,1\}$, $i \in Costly$, and $y_i = 1$, $\frac{a}{n} \le x_{ij} \le (\frac{a}{n} + 2k\frac{a(1-a)}{n^2})b^{-1}(1+\frac{1}{10})$.

3 For $i \in Cheap$, $\sum_j x_{ij} \le (nU+1)\frac{1-a}{n} + 2k(nU+1)\frac{a}{n^2}$.

4 For $i \in Costly$,

(a) if $y_i \neq 1$, $\sum_j x_{ij} \le (nU+1)\frac{a}{n} + k$.

(b) if $y_i = 1$, $\sum_j x_{ij} \le ((nU+1)\frac{a}{n} + k)(1 + \frac{1}{10})b^{-1}$.

Theorem 3.3 implies that solution s survives $\Omega(n)$ rounds. It remains to prove Theorem 3.3 and then the proof of Theorem 3.1 will be complete.

3.3.1 Proof of Theorem 3.3

The proof is by induction on the depth of node u. More specifically, by assuming that the invariants hold for an arbitrary node v at depth less than $n/10$, we show how to construct all the children nodes of v so that they are well-defined and the invariants are met.

In the proof, whenever we give the construction of a type 1 or type 2 child of v corresponding to some variable z_i, we refer to z_i as the *touched variable* – we also say that z_i is *touched* as type 1 or type 2 in the current step. We will consider cases according to which variable is touched and whether it is touched as type 1 or as type 2. When we touch a variable $z_i \notin \{0,1\}$ as type 1, the z'_i variable of the corresponding type 1 witness z' always takes the value 1 so by Observation 3.1 we satisfy the condition that the diagonal of the underlying protection matrix is equal to the 0th row. Note that we will not give the construction for the case in which y_i, $i \in Cheap$, is touched, since y_i is always 1 and the construction is trivial in those cases. The same applies to the cases of all variables that have integral values in the node-solution v of the inductive hypothesis, as we simply enforce Observation 3.2.

Another feature of our construction, which is actually a necessary property towards LS feasibility, is the following: when a fractional variable $x_{i'j}$ is touched as type 1, it is set to 1, and for all $i \neq i'$, x_{ij} becomes 0. If $x_{i'j}$ is touched as type 2, it is set to 0 and its previous value must be distributed among the other assignments of client j. Thus for every j, either there is some i' such that $x_{i'j} = 1$ and for all other $i \neq i'$, $x_{ij} = 0$ (e.g., when cases 1b, 1c below have happened for an ancestor of v), or there are at most k facilities to which the assignment of j is 0, where k is the depth of the tree (if there are type 2 nodes, through cases 2a, 2b, 2c, along the path of the tree that leads to v). In fact, as far as assignments to cheap facilities are concerned, the upper bound of k holds cumulatively across all clients, since no more than k assignment variables can be touched as Type 2 along a path of length k. Specifically, let C' be the set of clients j for which, for all $i \in F$, $x_{ij} < 1$. We will use the fact that $|\{x_{ij}, i \in Cheap, j \in C' \mid x_{ij} = 0\}| < k$.

Note that the invariants of Theorem 3.3 imply the satisfaction of constraints (2),(6),(7) and (8) for the number of rounds we consider.

Lemma 3.5 *Let (y, x) be a node-solution defined at depth $k \leq \frac{n}{10}$ of the evolution tree T_s. If (y, x) satisfies Invariants 1–4, then (y, x) meets constraints (2), (6), (7) and (8).*

Proof. Constraints (6) and (7) are obviously satisfied by Invariants 1 and 2. For $i \in \mathit{Costly}$, if $y_i \notin \{0, 1\}$ we have, by Invariant 1, that y_i is always "close" to b while, by Invariant 2.b, for every j a non-integer x_{ij} is "close" to $\frac{a}{n}$ and the total demand assigned to i is, by Invariant 4, $\sum_j x_{ij} \leq (nU + 1)\frac{a}{n} + k \leq U(b - 2k\frac{a}{n}) \leq Uy_i$. So (2) and (8) are satisfied.

Also note that if $y_i = 0$, by our construction $\sum_j x_{ij} = 0$, thus again (2) and (8) are satisfied.

For $i \in \mathit{Costly}$, if $y_i = 1$ we have by Invariant 2.c, that for every j a non-integer x_{ij} is at most $(\frac{a}{n} + 2k\frac{a(1-a)}{n^2})b^{-1}(1 + \frac{1}{10})$ and the total demand assigned to i is by Invariant 4 at most $\sum_j x_{ij} \leq ((nU + 1)\frac{a}{n} + k)(1 + \frac{1}{10})b^{-1} \leq U$. So (2) and (8) are satisfied.

Similarly for (2) and (8) involving some facility in *Cheap*. ∎

Thus, each time we prove the feasibility of the constructed solution, we only have to ensure that (3) holds.

We now explain the structure of the inductive step that produces the children of node v, where v is at depth $k < n/10$. We distinguish cases according to the variable that is touched. For every case, the proof has three parts. First, we give the construction of the child node (y', x') based on the current node-solution (y, x). Second, we prove that (y', x') is feasible for (LP-classic). Third, we show that (y', x') satisfies the four invariants in the statement of the theorem.

In order to understand intuitively the upcoming calculations for the invariants, the reader should focus on the dominant term in the rhs (or lhs) of each invariant. The loss in accuracy in the upper (lower) bounds obtained this way will be offset in the inductive step by increasing the coefficients of the minor terms. For example, in Invariant 2a the dominant term in the lhs is $(1 - a)/n$ and the minor term is $2kab^{-1}/n^2$. In the inductive step, $2k$ will become $2k + 2$.

Case 1: type 1 children

<u>subcase 1a:</u> **touched variable is y_{i_k}, $i_k \in \mathit{Costly}$**

CONSTRUCTION

Consider the type 1 child (y', x') of v corresponding to variable y_{i_k}. Variables $y_{i_k}, x_{i_k j}$ for all j are multiplied by a factor of $1/y_{i_k}$ and so $y'_{i_k} = 1$. Note that since we only consider cases where y_{i_k} is fractional, by the inductive hypothesis we have that for all variables $x_{i_k j}$, Invariant 2.b holds. For the variables involving facilities $i' \in \mathit{Costly} - \{i_k\}$ and for all j, we set $y'_{i'} = y_{i'}$, $x'_{i'j} = x_{i'j}$. For all j and for all $i' \in \mathit{Cheap}$ such that $x_{i'j} \neq 0$ we set $x'_{i'j} = x_{i'j}(1 - \frac{(1/y_{i_k} - 1)x_{i_k j}}{x_{i'j}t})$, where t is the number of facilities in *Cheap* for which j is assigned with a non-zero fraction (so $t \geq n - k$).

FEASIBILITY

Constraint (3) is satisfied by construction:

$\sum_i x'_{ij} = \sum_i x_{ij} + (1/y_{i_k} - 1)x_{i_k j} - \sum_{i \in Cheap | x_{ij} > 0} \frac{(1/y_{i_k}-1)x_{i_k j}}{t} =$
$\sum_i x_{ij} = 1$

Invariants

Invariant 1. For $i \in Costly - \{i_k\}$, y_i remains unchanged so Invariant 1 holds by the inductive hypothesis (from now on abbreviated as *i.h.*).

Invariant 2. For $i \in Cheap$ we have $2.a$:

$x'_{ij} = x_{ij} - \frac{(1/y_{i_k}-1)x_{i_k j}}{t} \geq$ (by Invariants 1, 2 of i.h. and being generous)
$\frac{1-a}{n} - 2k\frac{a}{n^2}b^{-1} - 2b^{-1}\frac{a}{n^2} \geq$
$\frac{1-a}{n} - 2(k+1)\frac{a}{n^2}b^{-1}$

For $i \in Costly - \{i_k\}$, $2.b$ holds since variables x_{ij} were not changed. For $x'_{i_k j}$ we need to show $2.c$:

$x'_{i_k j} = x_{i_k j}\frac{1}{y_{i_k}} \leq$ (by Invariants $2.b$, 1)
$(\frac{a}{n} + 2k\frac{a(1-a)}{n^2})(b + 2k\frac{a}{n})^{-1} \leq$
$(\frac{a}{n} + 2k\frac{a(1-a)}{n^2})b^{-1}(1 + \frac{1}{10})$

Invariant 3. Observe than the total demand assigned to each facility in $Cheap$ was decreased so Invariant 3 holds by the inductive hypothesis.

Invariant 4. For $i \in Costly - \{i_k\}$ Invariant 4 holds by inductive hypothesis. For i_k we have $4.b$:

$\sum_j x'_{i_k j} = 1/y_{i_k} \sum_j x_{i_k j} \leq$ (by the invariants of i.h.)
$b^{-1}(1 + \frac{1}{10})((nU+1)\frac{a}{n} + k)$

<u>subcase $1b$</u>: touched variable is $x_{i_k j*}$, $i_k \in Costly$

Construction

Consider the type 1 children (y', x') of v corresponding to variable $x_{i_k j*}$. We obtain y'_{i_k} by multiplying y_{i_k} by a factor of $1/y_{i_k}$ and so $y'_{i_k} = 1$ (and of course $x'_{i_k j*} = 1$, and $x'_{ij*} = 0$ for $i \neq i_k$). Every other variable of (y', x') is the same as in (y, x).

Feasibility

The feasibility of this case is trivial.

Invariants

The Invariants 1, 2, 3 in this case are satisfied trivially. For 4 we have for facility i_k:

$\sum_j x'_{i_k j} \leq$ (variable $x_{i_k j*}$ becomes 1)
$\sum_j x_{i_k j} + 1 \leq$ (by 4 of i.h.)
$(nU+1)\frac{a}{n} + k + 1$ if $y_{i_k} \neq 1$ or
$((nU+1)\frac{a}{n} + k + 1)b^{-1}$ if $y_{i_k} = 1$

In either of the two cases Invariant $4.b$ holds for the new value y'_{i_k}.

<u>subcase $1c$:</u> touched variable is $x_{i_k j*}$, $i_k \in Cheap$

Construction

Consider the type 1 children (y', x') of v corresponding to variable $x_{i_k j*}$. We obtain $y'_i, i \in Costly$ with $y_i \notin \{0,1\}$ by multiplying y_i by a factor of $(1 - \frac{(1/y_i - 1)x_{ij*}}{x_{i_k j*} t})$, where t is again the number of facilities in $Cheap$ for which $j*$ is assigned with a non zero fraction (so $t \geq n-k$). Of course $x'_{i_k j*} = 1$, and $x'_{ij*} = 0$ for $i \neq i_k$ as usual. Every other variable of (y', x') is the same as in (y, x).

Feasibility

Obviously (3) is satisfied. All other constraints are satisfied by Lemma 3.5.

Invariants

Invariant 1. For each $i \in Costly$ such that $y_i \notin \{0,1\}$ we have:

$y'_i = y_i(1 - \frac{(1/y_i - 1)x_{ij*}}{x_{i_k j*} t}) \geq$ (by Invariant 1 of i.h.)

$b - 2k\frac{a}{n} - (\frac{(1-y_i)x_{ij*}}{x_{i_k j*} t}) \geq$ (by Invariants 1, $2.b$ of i.h.)

$b - 2k\frac{a}{n} - (\frac{(1-b+2k\frac{a}{n})(\frac{a}{n}+2k\frac{a(1-a)}{n^2})}{(\frac{1-a}{n}-2k\frac{a}{n^2}b^{-1})t})$

Observe that the numerator of the last fraction is almost $\frac{a}{n}$ and the denominator is almost $9/10$ since $t \geq \frac{9n}{10}$. The additive constant 2 in the last term of the following equation is enough to absorb the "noise" in our arguments. Therefore the last quantity is at least
$b - 2k\frac{a}{n} - 2\frac{a}{n} = b - (2k+2)\frac{a}{n}$

Invariant 2. Variables x'_{ij} remain unchanged for $j \neq j*$. For $j*$, $x'_{i_k j*} = 1$ while for $i \neq i_k$ we have $x'_{ij*} = 0$, so 2 is trivially satisfied.

Invariant 3. For $i \in Cheap - \{i_k\}$ the total demand is decreased (because of $j*$). For i_k:

$\sum_j x'_{i_k j} \leq \sum_j x_{i_k j} + 1 \leq$ (by 3 of i.h.)

$(nU+1)\frac{1-a}{n} + 2k(nU+1)\frac{a}{n^2} + 1 \leq (nU+1)\frac{1-a}{n} + 2(k+1)(nU+1)\frac{a}{n^2}$

Invariant 4. The demand assigned to facilities in $Costly - \{i_k\}$ is decreased (because of $j*$) so $4.a$, $4.b$ trivially hold.

Case 2: type 2 children

<u>subcase $2a$:</u> touched variable is y_{i_k}, $i_k \in Costly$

Construction

Consider the type 2 children (y', x') of v corresponding to variable $y_{i_k} \notin \{0,1\}$. Let $f = \frac{y_{i_k}}{1-y_{i_k}}$. Solution (y', x') is dictated by its twin type 1 solution (case 1a): variables $y_{i_k}, x_{i_k j}$ for all j, are multiplied by a factor of $(1 - f(1/y_{i_k} - 1))$ and so $y'_{i_k} = 0$ and $x'_{i_k j} = 0$, that is facility i_k closes. The variables involving facilities $i' \in Costly - \{i_k\}$, namely $y'_{i'}, x'_{i'j}$ for all j, have the same value as in (y, x). For all j and for $i' \in Cheap$ such that $x_{i'j} \neq 0$ we have

$x'_{i'j} = x_{i'j}(1 + \frac{f(1/y_{i_k}-1)x_{i_kj}}{x_{i'j}t})$, where t is again the number of facilities in *Cheap* for which j is assigned with a non zero fraction (so $t \geq n-k$).

Feasibility

Constraint (3) is satisfied by Lemma 3.3.

Invariants

Invariant 1. For $i \in Costly - \{i_k\}$, y_i remain unchanged so Invariant 1 holds by inductive hypothesis.

Invariant 2. For $i \in Cheap$ we have 2.a:

$x'_{ij} = x_{i'j} + \frac{f(1/y_{i_k}-1)x_{i_kj}}{t} \leq$ (by Invariants 1, 2 of i.h.)
$\frac{1-a}{n} + 2k\frac{1-a}{n^2} + 2\frac{a}{n^2} \leq$ (being very generous)
$\frac{1-a}{n} + (2k+2)\frac{1-a}{n^2}$

For $i \in Costly - \{i_k\}$, 2.b holds since variables x_{ij} were not changed.

Invariant 3. For $i \in Cheap$ we have:

$\sum_j x'_{ij} = \sum_j x_{ij} + \frac{1}{n}\sum_j x_{i_kj} + o(1) \leq$ (by Invariants 3, 4 of i.h.)
$(nU+1)\frac{1-a}{n} + 2k(nU+1)\frac{a}{n^2} + \frac{(nU+1)\frac{a}{n}+k}{n} + o(1) \leq$
$(nU+1)\frac{1-a}{n} + (2k+2)(nU+1)\frac{a}{n^2}$

The $o(1)$ above is due to the fact that at most k assignment variables for some cheap facilities may have been touched as type 2 and are 0. For those same clients the assignment to i_k is fractional, so the demand corresponding to them that was assigned to i_k, must be distributed among the, at least $n-k$, available cheap facilities. That additional demand is at most $\frac{k(\frac{a}{n}+2k\frac{a(1-a)}{n^2})}{n-k} = o(1)$.

Invariant 4. For $i \in Costly - i_k$ Invariant 4 holds by inductive hypothesis. For i_k we have $\sum_j x_{i_kj} = 0$.

<u>subcase 2b:</u> touched variable is x_{i_kj*}, $i_k \in Costly$

Construction

Consider the type 2 children (y', x') of v corresponding to variable x_{i_kj*}. Let $f = \frac{x_{i_kj*}}{1-x_{i_kj*}}$. Solution (y', x') is dictated by its twin type 1 node-solution (case 1b): variable y'_{i_k} is obtained by multiplying y_{i_k} by a factor of $(1 - f(1/y_{i_k} - 1))$ and for $i \neq i_k$, $x'_{ij*} = x_{ij*}(1+f)$ and $x'_{i_kj*} = 0$. Every other variable of (y', x') is the same as in (y, x).

Feasibility

The feasibility of this case is trivial by Lemma 3.3.

Invariants

Invariant 1. For facilities $i \in Costly - \{i_k\}$ the proof is trivial (no change). Same if $y_{i_k} = 1$. If $y_{i_k} \notin \{0,1\}$ we have:

$y'_{i_k} = y_{i_k}(1 - f(1/y_{i_k} - 1)) =$
$y_{i_k} - (1 - y_{i_k})\frac{x_{i_k j*}}{1 - x_{i_k j*}} \geq$ (by Invariants 1, 2 of i.h.)
$b - 2k\frac{a}{n} - 2\frac{a}{n} \geq$
$b - (2k+2)\frac{a}{n}$

Invariant 2. For client $j*$ and facility $i \in Cheap$ we have 2.a:

$x'_{ij*} = x_{ij*}(1+f) \leq$ (by Invariant 2 of i.h.)
$\frac{1-a}{n} + 2k\frac{1-a}{n^2} + 2\frac{(1-a)ab^{-1}}{n^2} \leq$
$\frac{1-a}{n} + (2k+2)\frac{(1-a)}{n^2}$ (recall $a, b = \Theta(n^{-2})$)

For client $j*$ and facility $i \in Costly$ we have 2.b:

$x'_{ij*} = x_{ij*}(1+f) \leq \frac{a}{n} + 2k\frac{a(1-a)}{n^2} + 2\frac{a^2b^{-1}}{n^2} \leq$
$\frac{a}{n} + 2(k+1)\frac{a(1-a)}{n^2}$

Similarly for Invariant 2.c.

Invariant 3. For $i \in Cheap$:

$\sum_j x'_{ij} \leq \sum_j x_{ij} + 1 \leq$ (by 3 of i.h.)
$(nU+1)\frac{1-a}{n} + 2(k+1)(nU+1)\frac{a}{n^2}$

Invariant 4. For i_k the total demand is decreased while for $i \in Costly - \{i_k\}$:

$\sum_j x'_{ij} \leq \sum_j x_{ij} + 1 \leq$ (by 3 of i.h.)
$(nU+1)\frac{a}{n} + k + 1$ if $y_i \neq 1$ or
$((nU+1)\frac{a}{n} + k + 1)b^{-1}$ if $y_i = 1$

subcase 2c: touched variable is $x_{i_k j*}$, $i_k \in Cheap$

Construction

Consider the type 2 children (y', x') of v corresponding to variable $x_{i_k j*}$. Let $f = \frac{x_{i_k j*}}{1 - x_{i_k j*}}$. Solution (y', x') is dictated by its twin type 1 node-solution (case 1c): variables $y'_i \notin \{0,1\}, i \in Costly$, are obtained by multiplying y_i by a factor of $(1 + f\frac{(1/y_i - 1)x_{ij*}}{x_{i_k j*} t})$, where t is again the number of facilities in $Cheap$ for which j is assigned with a non zero fraction (so $t \geq n - k$). For $i \neq i_k$, $x'_{ij*} = x_{ij*}(1+f)$ while $x'_{i_k j*} = 0$. Every other variable of (y', x') is the same as in (y, x).

Feasibility

The satisfaction of (3) is ensured by Lemma 3.3.

Invariants

Invariant 1. For facility $i \in Costly$ such that $y_i \notin \{0,1\}$ we have:

$y'_i = y_i(1 + f\frac{(1/y_i-1)x_{ij*}}{x_{i_k j*}n}) =$
$y_i + (1-y_i)\frac{x_{ij*}}{(1-x_{i_k j*})t} \leq$ (by Invariants 1, 2 of i.h.)
$b + 2k\frac{a}{n} + 2\frac{a}{n^2} \leq$
$b + (2k+2)\frac{a}{n}$

Invariant 2. For client $j*$ and facility $i \in Cheap$ we have $2.a$:

$x'_{ij*} = x_{ij*}(1+f) \leq$ (by Invariant 2 of i.h.)
$\frac{1-a}{n} + 2k\frac{1-a}{n^2} + 2\frac{(1-a)^2}{n^2} \leq$
$\frac{1-a}{n} + (2k+2)\frac{1-a}{n^2}$

For client $j*$ and facility $i \in Costly$ with $y_i \notin \{0,1\}$ we have $2.b$:

$x'_{ij*} = x_{ij*}(1+f) \leq$ (by Invariant 2 of i.h.)
$\frac{a}{n} + 2k\frac{a(1-a)}{n^2} + 2\frac{(1-a)a}{n^2} \leq$
$\frac{a}{n} + (2k+2)\frac{a(1-a)}{n^2}$

For client $j*$ and facility $i \in Costly$ with $y_i = 1$ we have $2.c$:

$x'_{ij*} = x_{ij*}(1+f) \leq$ (by Invariant 2 of i.h.)
$(\frac{a}{n} + 2k\frac{a(1-a)}{n^2})b^{-1}(1+\frac{1}{10}) + 2\frac{(1-a)ab^{-1}}{n^2} \leq$
$\frac{a}{n} + (2k+2)\frac{a(1-a)}{n^2}$

Invariant 3. The demand assigned to i_k is decreased. For $i \in Cheap - \{i_k\}$:

$\sum_j x'_{ij} \leq \sum_j x_{ij} + 1 \leq$ (by 3 of i.h.)
$(nU+1)\frac{1-a}{n} + 2(k+1)(nU+1)\frac{a}{n^2}$

Invariant 4. For $i \in Costly$:

$\sum_j x'_{ij} \leq \sum_j x_{ij} + 1 \leq$ (by 3 of i.h.)
$(nU+1)\frac{a}{n} + k + 1$ if $y_i \neq 1$ or
$((nU+1)\frac{a}{n} + k + 1)b^{-1}$ if $y_i = 1$

The case analysis is complete. It is easy to verify that the constructed vectors satisfy the conditions stated in Observation 3.1 and Proposition 3.2. It remains to show that the vectors we constructed for node v satisfy the symmetry requirements, i.e., the conditions in Proposition 3.1 and thus are indeed witnesses.

Lemma 3.6 *The symmetry Conditions (ii) and (iii) stated in Proposition 3.1 are satisfied for the children of node-solution v.*

Proof. By construction we never alter integer values of variables, as dictated by the LS procedure, therefore Condition (ii) of Observation 3.2 holds.

When a variable y_i, $i \in \mathit{Costly}$, is touched then for the symmetry between y_i and each other variable we have:

For all j, variables x_{ij} are multiplied by $1/y_i$ (case $1a$), and when some x_{ij} is touched, variable y_i is multiplied by $1/y_i$ (case $1b$).

For all j, variables $x_{i'j}$, $i' \in \mathit{Cheap}$, are multiplied by $(1 - (1/y_i - 1)\frac{x_{ij}}{x_{i'j}t})$ (case $1a$), and when some $x_{i'j}$ is touched, variable y_i is multiplied by $(1 - (1/y_i - 1)\frac{x_{ij}}{x_{i'j}t})$ (case $1c$).

For all j, variables $y_{i''}$, $x_{i''j}$, $i'' \in \mathit{Costly} - \{i\}$, are multiplied by 1 (case $1a$), and when $y_{i''}$ or some $x_{i''j}$ is touched, variable y_i is multiplied by 1 (cases $1a$, $1b$).

When a variable x_{ij}, $i \in \mathit{Costly}$, is touched then for the symmetry between x_{ij} and each other variable we have:

For all $j' \neq j$ and all i', variables $x_{i'j'}$ are multiplied by 1 (case $1b$), and when some $x_{i'j'}$ is touched, variable x_{ij} is multiplied by 1 (cases $1b$, $1c$).

For $i' \neq i$, variables $x_{i'j}$ are multiplied by 0 (case $1b$), and when some $x_{i'j}$ is touched, variable x_{ij} is multiplied by 0 (cases $1b$, $1c$).

Finally, when variable x_{ij}, $i \in \mathit{Cheap}$, is touched then for the symmetry between x_{ij} and each other variable, the remaining cases that have not been covered above are:

For all $j' \neq j$ and all $i' \in \mathit{Cheap}$, variables $x_{i'j'}$ are multiplied by 1 (case 1c), and when $x_{i'j'}$ is touched, variable x_{ij} is multiplied by 1 (case $1c$).

For all $i' \in \mathit{Cheap}$, variables $x_{i'j}$ are multiplied by 0 (case $1c$), and when $x_{i'j}$ is touched, variable x_{ij} is multiplied by 0 (case $1c$). ∎

By Lemma 3.2, the proof of Theorem 3.3 is now complete.

The proof can yield a tradeoff between the number of rounds as a function of the dimension of the instance and the integrality gap, which can be obtained by toying with the quantities U, a, and b that are left as parameters. One can obtain a higher gap that survives for a smaller number of rounds.

3.4 The mixed LS_+ case

It is not hard to see that the proof of Theorem 3.1 also yields the same lower bound for the mixed LS_+ procedure: simply restrict the constructed protection matrices to the facility opening y variables. The resulting matrices are of the form $(1, y)(1, y)^T + \mathrm{Diag}(0, y - y^2)$ which are well-known to be positive semidefinite (see Theorem 4.1 in [25]). The following theorem is immediate.

Theorem 3.4 *There is a family of instances $\mathcal{I}$ of uniform* CFL*, such that for every sufficiently large n, there is an instance in the family with $2n$ facilities and $n^4 + 1$ clients for* *which the relaxation produced by the mixed* LS_+ *procedure in* $\Omega(n)$ *rounds has integrality gap* $\Omega(n)$.

4 Sherali-Adams gap for CFL

4.1 The SA hierarchy

We proceed to define the Sherali-Adams hierarchy [50]. Consider a polytope $P \subseteq \mathbb{R}^d$ defined by the linear constraints $Ax - b \leq 0$, which include the subsystem $0 \leq x_i \leq 1$, $i = 1, \ldots, d$. We define the polytope $\text{SA}^k(P) \subseteq \mathbb{R}^d$ as follows. For every constraint $\pi(x) \leq 0$ of P, for every set of variables $X \subseteq \{x_i \mid i = 1, \ldots, d\}$ such that $|X| \leq k$, and for every $W \subseteq X$, consider the nonlinear inequality: $\pi(x) \prod_{x_i \in X-W} x_i \prod_{x_i \in W}(1 - x_i) \leq 0$. We call $\prod_{x_i \in X-W} x_i \prod_{x_i \in W}(1 - x_i)$ an (X, W)-*multiplier.* When the sets X, W are clear from the context, we will refer simply to a *multiplier.* After expanding the lhs of the inequalities, we linearize the resulting non-linear system as follows: (i) substitute x_i for x_i^2 for all i (ii) replace $\prod_{x_i \in I} x_i$ with x_I for each set $I \subseteq \{x_i \mid i = 1, \ldots, d\}$. Call the resulting lifted polyhedron Q. $\text{SA}^k(P)$ is the projection of Q onto the original variables x_i, where variables $x_{\{x_i\}}$ are treated as being equal to x_i. We call $\text{SA}^k(P)$ the polytope *obtained from* P *at level* k of the SA hierarchy. Given a cost vector $c \in \mathbb{R}^d$, the *relaxation obtained from* P *at level* k of SA is $\min\{c^T x \mid x \in \text{SA}^k(P)\}$.

Recall the family of instances $\mathcal{I}$ and the solution s. We will show in Section 4.2 that the solution s to this same instance [1] is feasible for a number of SA levels, which is linear in the number $2n$ of facilities, more specifically for $n/10$ levels. On the other hand, by Theorem 3.2, at level $2n$ the relaxation obtained expresses the integral polytope since SA refines LS. As an example of the robustness of our SA construction against the addition of new constraints, we show in Section 4.3 that a bad fractional solution remains feasible for $\Theta(n)$ levels of SA even if we add the valid inequality (9).

The following lemma, which is implicit in previous work [23, 17, 24] gives sufficient conditions for a solution to be feasible at level k of the SA hierarchy.

Lemma 4.1 *[23, 17, 24] Let* s *be a feasible solution to the relaxation and let* $v(\pi, z)$ *be the set of variables appearing in a nonlinear constraint obtained from* π *when multiplied by an* (X, W)*-multiplier* $z = \prod_{i \in X-W} x_i \prod_{i \in W}(1 - x_i)$*, for some* X *and* W*. Solution* s *is feasible for* k *levels of* SA *if for every constraint* π *and each such multiplier* z *with at most* k *distinct variables there is:*

1 A solution $s_{\pi,z}$ *which agrees with* s *on* $v(\pi, z)$ *such that* $s_{\pi,z}$ *is a convex combination* E_d *of feasible integer solutions (and thus* E_d *defines a distribution over integer solutions) and*

[1]The reader should notice that any similarity with Knapsack is superficial. Theorem 4.1 is about the CFL polytope. Moreover, it is easy to embed our instance in a slightly larger one, with a non-trivial metric, so that the projection of the bad CFL solution to the y-variables, is in the integral polytope of the "underlying" knapsack instance.

2 (Consistency Condition) *For any two sets $v(\pi_1, z_1)$ and $v(\pi_2, z_2)$, let $x_1 x_2 \cdot \ldots \cdot x_l$, $l \leq k+1$, be a product appearing in both constraints obtained from π_1 and π_2 after multiplying with z_1 and z_2 respectively. Then the probability $P[x_1 = 1 \wedge x_2 = 1 \wedge \ldots \wedge x_l = 1]$ is the same in both distributions E_{d_1} and E_{d_2} associated with $v(\pi_1, z_1)$ and $v(\pi_2, z_2)$ respectively.*

4.2 SA gaps for approximate Cfl

In this subsection, we state and prove our main SA result, namely Theorem 4.1. We show how to design appropriate distributions for every constraint of (LP-classic) and for every multiplier z so that they fulfill the conditions of Lemma 4.1. The proof groups the constraints of (LP-classic) into two cases.

4.2.1 Distributions for constraints (2), (6), (7), (8)

First consider a multiplier z and an inequality π from the family of constraints (8), i.e., $\pi\colon \sum_{j \in C} x_{i^\pi j} \leq U y_{i^\pi}$. After multiplying by z and expanding, we obtain a linear combination of monomials (products). Then, for the $k < n-1$ levels we consider there must be some costly facility $i_b \notin v(\pi, z)$. We construct a solution $s_{\pi,z} = (y', x')$ by setting $y'_{i_b} = 1 - \sum_{i \in Costly - \{i_b\}} y_i$ and leaving all other variables the same as in the original bad solution s. We say that facility i_b *takes the blame for* $v(\pi, z)$. When $v(\pi, z)$ is implied from the context, we simply say that i_b *takes the blame.* We will prove that $s_{\pi,z}$ can be obtained as a convex combination E_d of a set of integer solutions satisfying $\sum_{i \in Costly} y_i = 1$. While $s_{\pi,z}$ can be obtained as a convex combination E_d in a variety of ways, we require that the assignments of clients to the cheap facilities are indistinguishable in E_d and the same must be true for the assignments to costly facilities other than i_b. In the upcoming definition, we use the product $p = z_1 z_2 \ldots z_l$ as an abbreviation of the event $\mathcal{E}_p := \bigwedge_{i=1}^{l} (z_i = 1)$.

Definition 4.1 *Let i_b be the facility that takes the blame. We say that a distribution E_d over feasible integer solutions is* assignment-symmetric *if the following are true:*

1 *$P_{E_d}[x_{i_{a_1} j_{b_1}} \ldots x_{i_{a_t} j_{b_t}} y_{i_{a_{t+1}}} \ldots y_{i_{a_l}}]$, with $t + l \leq k+1$ is the same if we exchange all occurrences of cheap facility i_r with cheap facility $i_{r'}$ (i.e., if we relabel facilities). We allow repetitions of facilities and clients in the description of the event.*

2 *$P_{E_d}[x_{i_{a_1} j_{b_1}} \ldots x_{i_{a_t} j_{b_t}} y_{i_{a_{t+1}}} \ldots y_{i_{a_l}}]$ is the same if we exchange all occurrences of client j_q with client $j_{q'}$.*

3 *$P_{E_d}[x_{i_{a_1} j_{b_1}} \ldots x_{i_{a_t} j_{b_t}} y_{i_{a_{t+1}}} \ldots y_{i_{a_l}}]$ is the same if we exchange all occurrences of costly facility i_1 with costly facility i_2, such that $i_1, i_2 \neq i_b$.*

We can always obtain $s_{\pi,z}$ from such an assignment-symmetric distribution E_d as shown in the following lemma.

Lemma 4.2 *Solution $s_{\pi,z}$ is the expected (y, x) vector obtained from an assignment-symmetric distribution E_d over feasible integer solutions.*

Proof. We describe a probabilistic experiment which induces an assignment-symmetric distribution E_d over integer solutions satisfying $\sum_{i \in Costly} y_i = 1$. Let i_b be the facility that takes the blame for $v(\pi, z)$. We begin by defining some parameters for the integer solutions that will make up the support of the distribution E_d.

Let $w_{i_{co}} = \frac{\sum_{j \in C} x'_{i_{co}j}}{y'_{i_{co}}}$ be the number of clients assigned to facility i_{co} in the integer solutions in the support of E_d, when facility $i_{co} \in Costly$ is opened. To simplify the presentation let us assume for now that $w_{i_{co}}$ and the values we subsequently define are integers (we discuss at the end of the proof how to handle fractional values). Let $w^1_{ch} = \frac{|C| - w_{i_b}}{|Cheap|}$ be the number of clients assigned to any facility in $Cheap$ when i_b is the only opened costly facility. Likewise, for any costly facility $i_{co} \neq i_b$, let $w^2_{ch} = \frac{|C| - w_{i_{co}}}{|Cheap|}$ be the number of clients assigned to any facility in $Cheap$ in each integer solution in E_d where facility i_{co} is the only opened costly facility. Observe that all the defined values are less than U. The following procedure produces the assignment-symmetric distribution E_d.

Pick costly facility i_{co} with probability $y'_{i_{co}}$. If $i_{co} = i_b$ ($i_{co} \neq i_b$) then consider n bins corresponding to the n cheap facilities each one having w^1_{ch} (w^2_{ch}) slots and 1 bin corresponding to i_{co} having $w_{i_{co}}$ slots. Randomly distribute $|C|$ balls to the slots of the $n + 1$ bins, with exactly one ball in each slot. Note that the above experiment induces a distribution over feasible integer solutions satisfying $\sum_{i \in Costly} y_i = 1$ since all the defined bin capacities are less than U and every client is assigned to exactly one opened facility in each outcome and exactly 1 costly facility is opened. It is easy to see that the induced distribution E_d is assignment-symmetric because the clients and all the facilities other than i_b are handled symmetrically. It remains to show that the expected (y, x) vector with respect to E_d is solution $s_{\pi,z}$.

We argue now that $s_{\pi,z}$ is the convex combination induced by E_d. The cheap facilities are always open, and the costly are open a fraction of the time that is equal to the value of their corresponding y' variable. The expected demand assigned to each $i_{co} \in Costly$ is $y'_{i_{co}} w_{i_{co}}$ which is the total demand assigned to i_{co} by $s_{\pi,z}$. Since the clients have the same probability of being tossed in the bin corresponding to i_{co}, the expected assignment of each client j to i_{co} is the same as in $s_{\pi,z}$.

Similarly, we can prove that the expected assignments to the cheap facilities are as required. Observe that in every outcome of the experiment the demand not assigned to costly facilities is exactly the demand assigned to cheap facilities. Since we have proved that the expected assignments to the costly facilities are those of the bad solution $s_{\pi,z}$, by linearity of expectation we get that the total assignments to all cheap facilities are $\sum_{i \in Cheap} \sum_j x'_{ij}$ (the total assignment of each client adds up to 1). By the symmetric way the cheap facilities are handled in the experiment, we have that the total expected demand assigned to each $i \in Cheap$ is $\sum_{j \in C} x'_{ij}$. Moreover, by the symmetric way the clients are assigned to i in the experiment, we get that the expected assignment of each $j \in C$ to i is $\frac{\sum_j x'_{ij}}{|C|} = x'_{ij}$.

To handle the case where the $w_{i_{co}}, w^1_{ch}, w^2_{ch}$ are not integers (which is actually always the case), we do the following: each time costly facility i_b ($i_{co} \neq i_b$) is picked, we set the number of slots of the corresponding bin to $\lfloor w_{i_b} \rfloor$ ($\lfloor w_{i_{co}} \rfloor$) with probability $1 - (w_{i_b} - \lfloor w_{i_b} \rfloor)$ $(1 - (w_{i_{co}} - \lfloor w_{i_{co}} \rfloor))$, otherwise set the slots to $\lceil w_{i_b} \rceil (\lceil w_{i_{co}} \rceil)$; this ensures that the expected number of slots is w_{i_b} ($w_{i_{co}}$). The same rationale applies to the remaining cases of the construction. If the number of slots of i_b (i_{co}) is set to $\lfloor w_{i_b} \rfloor$ ($\lfloor w_{i_{co}} \rfloor$), then we pick some $n(\frac{|C|-\lfloor w_{i_b} \rfloor}{n} - \lfloor(\frac{|C|-\lfloor w_{i_b} \rfloor}{n})\rfloor)$ ($n(\frac{|C|-\lfloor w_{i_{co}} \rfloor}{n} - \lfloor(\frac{|C|-\lfloor w_{i_{co}} \rfloor}{n})\rfloor)$) cheap facilities at random and set their corresponding number of slots to $\lceil \frac{|C|-\lfloor w_{i_b} \rfloor}{n} \rceil$ ($\lceil \frac{|C|-\lfloor w_{i_{co}} \rfloor}{n} \rceil$) and the number of slots of the rest of the cheap facilities to $\lfloor \frac{|C|-\lfloor w_{i_b} \rfloor}{n} \rfloor$($\lfloor \frac{|C|-\lfloor w_{i_{co}} \rfloor}{n} \rfloor$). Otherwise, pick some $n(\frac{|C|-\lceil w_{i_b} \rceil}{n} - \lfloor(\frac{|C|-\lceil w_{i_b} \rceil}{n})\rfloor)$ ($n(\frac{|C|-\lceil w_{i_{co}} \rceil}{n} - \lfloor(\frac{|C|-\lceil w_{i_{co}} \rceil}{n})\rfloor)$) cheap facilities at random and set their corresponding number of slots to $\lceil \frac{|C|-\lceil w_{i_b} \rceil}{n} \rceil$ ($\lceil \frac{|C|-\lceil w_{i_{co}} \rceil}{n} \rceil$) and the number of slots of the rest to $\lfloor \frac{|C|-\lceil w_{i_b} \rceil}{n} \rfloor$($\lfloor \frac{|C|-\lceil w_{i_{co}} \rceil}{n} \rfloor$). Once again the capacities are respected since $w_{i_{co}} = \Theta(n)$ (if $i_{co} = i_b$), or $w_{i_{co}} = \Theta(n^3)$ (if $i_{co} \neq i_b$). In every case the expected number of slots per facility is the same as in the initial description of the experiment where we assumed that $w_{i_{co}}, w^1_{ch}, w^2_{ch}$ are integers. ∎

Constraints $x_{ij} \leq y_i$, $x_{ij} \leq 1$, $y_i \leq 1$, are handled in the exact same way; the set of variables appearing in them is a subset of those appearing in the more complex constraints.

4.2.2 Distributions for constraints (3)

The second and more challenging case is when constraint π is $\sum_{i \in F} x_{ij^\pi} = 1$ for some client j^π. Let again z be a multiplier with at most k terms, which is used at the lifting step of the SA procedure. Observe that all facilities in F appear in $v(\pi, z)$ as indexes of at least the variables x_{ij}. We select one costly facility i_b not appearing in multiplier z to *take the blame*. Let $s_{\pi,z} = (y', x')$ be the corresponding extended solution that is induced in expectation by the assignment–symmetric distribution E_d of Lemma 4.2. In this case there is a major obstacle to the consistency of the distributions: the facility that takes the blame depends on the multiplier z, but this time products containing that facility do appear in the corresponding nonlinear constraint and thus there is the danger of violating the consistency required by the 2nd condition of Lemma 4.1. For example, due to the different way the facility i_b is handled by the experiment of Lemma 4.2 compared to the rest of the costly facilities, conditioning on the event ($x_{i_b j} = 1$) the probability of an event ($x_{ij'} = 1$), $i \in Cheap$, for some $j' \neq j$, is higher than it would be if we were to condition on the event ($x_{i'j} = 1$), $i' \in Costly - \{i_b\}$. The same is true for more complex events involving assignments to cheap facilities conditioning on an assignment of facility i_b compared to the analogous event conditioning on some other costly facility. We will make alterations to E_d and construct a distribution E_f where the probabilities of the aforementioned events are equal.

An example of the required alterations to E_d

We first display the intuition behind the alteration with the following example: consider the event A: ($x_{i_b j} = 1 \wedge x_{i_{ch} j'} = 1$) and the event B: ($x_{i_{co} j} = 1 \wedge x_{i_{ch} j'} = 1$) with $i_{co} \in Costly - \{i_b\}$ and $i_{ch} \in Cheap$. The probability of A is $P[A] = P[x_{i_b j} = 1] P[x_{i_{ch} j'} = 1 \mid x_{i_b j} = 1] = x'_{i_b j} \frac{w^1_{ch}}{|C|-1}$ and the probability of B is $P[B] = P[x_{i_{co} j} = 1] P[x_{i_{ch} j'} = 1 \mid x_{i_{co} j} = 1] = x'_{i_{co} j} \frac{w^2_{ch}}{|C|-1}$. Note that $P[A] \approx P[B](1 + 1/n)$ so $P[A]$ is only slightly greater. We nullify the difference between those probabilities by performing an alteration step to distribution E_d that we shall call transfusion of probability. We pick some probability measure q of an integer solution s_1 (or of some set of solutions) for which $(x_{i_{ch}j'} = 1 \wedge x_{i_b j} = 1 \wedge x_{i_b j''} = 0)$ for some client j''. We then pick the same quantity q of measure of some integer solution (or of some set of solutions) s_2 for which $(x_{i_{ch}j'} = 0 \wedge x_{i_b j} = 0 \wedge x_{i_b j''} = 1)$ and we exchange the values of the assignments $x_{i_b j}, x_{i_b j''}$ of the solutions whose measure we picked. Let q be $P[A] - P[B]$, it is easy to see that each set of solutions has enough measure to perform the transfusion. The resulting distribution has now $P[A] = P[B]$. In general, when transfusing probabilistic measure for complex events, we must be careful not to change the probability of events involving only assignments to cheap facilities, as opposed to the simplified example above.

Characterizing the events whose probability must be corrected

We define now the events whose probabilities must be corrected. We then give some technical propositions concerning their probabilities in E_d. Let p be a product appearing after multiplying constraint π by multiplier z and expanding the resulting expression. Wlog, exactly one assignment variable of the form $x_{i_b j}$, for some $j \in C$, appears in p. Recall that we chose i_b so that it does not appear in z; thus we cannot have y_{i_b} or more than one assignments of i_b appearing in p. We may also assume that there is no y_i variable in p: if there is one for some $i \in Costly - \{i_b\}$, the probability of $\mathcal{E}_p$ is simply 0, and, if $i \in Cheap$, then we can ignore the effect of $y_i = 1$ since it is always true. Likewise, we assume there is no assignment variable of another costly facility. A product p with the above properties is called *critical* and the associated event $\mathcal{E}_p$ is a *critical event.* It is exactly the probabilities of critical events that need to be corrected.

The following proposition relates the probability of $\mathcal{E}_p$ with that of $\mathcal{E}_{p'} = \mathcal{E}_{px_{ij}}$, an event with the additional requirement that $x_{ij} = 1$.

Proposition 4.1 *Consider the critical products* $p = x_{i_b j} x_{i_{a_1} j_{b_1}} x_{i_{a_2} j_{b_2}} \dots x_{i_{a_l} j_{b_l}}$ *and* $p' = p x_{i_{a_{l+1}} j_{b_{l+1}}}$. *Then in* E_d $(1 - o(1))P[\mathcal{E}_p]/n \leq P[\mathcal{E}_{p'}] \leq (1 + o(1))P[\mathcal{E}_p]/n$.

Proof. Recall that we defined the assignment-symmetric distribution E_d via a balls-and-bins experiment. Event $\mathcal{E}_p$ is equivalent to the event where ball j falls in some slot of bin i_b and balls j_{b_t}, $t = 1, \dots, l$, fall each in some slot of bins i_{a_t}, $t = 1, \dots, l$, with at most one ball in each slot. Without loss of generality, we may assume that balls j and j_{b_t}, $t = 1, \dots, l$, are the first $l + 1$ balls to be thrown during the experiment. Since bin i_b has $\Theta(n)$ slots and bins i_{a_t}, $t = 1, \dots, l$, have each $\Theta(n^3)$ slots and the balls are at most n, it is easy to see that $(1 - o(1))\frac{P[\mathcal{E}_p]}{n} \leq P[\mathcal{E}_{p'}] \leq (1 + o(1))\frac{P[\mathcal{E}_p]}{n}$. ∎

Consider critical product $p = x_{i_b j} x_{i_{a_1} j_{b_1}} \dots x_{i_{a_{k-i+1}} j_{b_{k-i+1}}}$, for some $i \in \{1, \dots, k\}$, and the event $\mathcal{E}_p$: $(x_{i_b j} = 1 \wedge x_{i_{a_1} j_{b_1}} = 1 \wedge \dots \wedge x_{i_{a_{k-i+1}} j_{b_{k-i+1}}} = 1)$. We wish in E_f the probability $P[\mathcal{E}_p]$ to be equal to $P[\mathcal{E}_{p/fixed}] := P[x_{i^* j} = 1 \wedge x_{i_{a_1} j_{b_1}} = 1 \wedge \dots \wedge x_{i_{a_{k-i+1}} j_{b_{k-i+1}}} = 1]$ in E_d for $i^* \in Costly - \{i_b\}$. We bound the ratio $\frac{P[\mathcal{E}_p]}{P[\mathcal{E}_{p/fixed}]}$:

Proposition 4.2 *Let critical event* $\mathcal{E}_p$ *and* $\mathcal{E}_{p/fixed}$ *be defined as above. Then in* E_d $(1 + (1 - o(1))/n)^{k-i+1} \leq \frac{P[\mathcal{E}_p]}{P[\mathcal{E}_{p/fixed}]} \leq (1 + (1 + o(1))/n)^{k-i+1} < e^2$.

Proof. Consider again the random experiment of the proof of Proposition 4.1. Recall that, ignoring constant factors, $w^1_{ch} = n^3 - 1$ and $w^2_{ch} = n^3 - n^2$. $P[\mathcal{E}_p] = x_{i_b j} P[x_{i_{a_1} j_{b_1}} = 1 \wedge x_{i_{a_2} j_{b_2}} = 1 \wedge \ldots \wedge x_{i_{a_{k-i+1}} j_{b_{k-i+1}}} = 1 \mid x_{i_b j} = 1]$ and $P[\mathcal{E}_{p/fixed}] = x_{i^* j} P[x_{i_{a_1} j_{b_1}} = 1 \wedge x_{i_{a_2} j_{b_2}} = 1 \wedge \ldots \wedge x_{i_{a_{k-i+1}} j_{b_{k-i+1}}} = 1 \mid x_{i^* j} = 1]$ and since $x_{i_b j} = x_{i^* j}$ it is enough to compute the ratio of the probability of success of the tossing of the $k-i+1$ balls when $x_{i_b j} = 1$, and thus the capacity of the bins corresponding to cheap facilities is w^1_{ch}, to the probability of success of the tossing of the same $k-i+1$ balls when $x_{i^* j} = 1$ and thus the capacity of the bins corresponding to cheap facilities is w^2_{ch}. In order to do that we use that $P[x_{i_{a_1} j_{b_1}} = 1 \wedge \ldots \wedge x_{i_{a_{k-i+1}} j_{b_{k-i+1}}} = 1 \mid x_{i_b j} = 1] = \prod_{q=1}^{q=k-i+1} P[x_{i_{a_q} j_{b_q}} = 1 \mid (\bigwedge_{z<q} x_{i_{a_z} j_{b_z}} = 1) \wedge x_{i_b j} = 1]$ and we bound the ratio of the corresponding products of the two cases.

When tossing the ball j_{b_r} given the successful tossing of balls j_{b_q} with $q < r$, the probability of success is $\frac{w^1_{ch} - o}{|C| - r + 1}$ and $\frac{w^2_{ch} - o}{|C| - r + 1}$ respectively, where $0 \leq o \leq r$ is the number of balls already placed in some slot of the bin corresponding to cheap facility a_r. We have that $(1 + (1 - o(1))/n) \leq \frac{w^1_{ch} - o}{w^2_{ch} - o} \leq (1 + (1 + o(1))/n)$, since $w^1_{ch} = n^3 - 1$, $w^2_{ch} = n^3 - n^2$ and $o \leq k \leq n/10$. Therefore

$$(1 + (1 - o(1))/n)^{k-i+1} \leq \frac{P[\mathcal{E}_p]}{P[\mathcal{E}_{p/fixed}]} \leq (1 + (1 + o(1))/n)^{k-i+1} < e^2$$

using that $\lim_{x \to \infty} (1 + d/x)^x = e^d$. ∎

Construction of the altered distribution E_f

Now we define the *transfusion of probability process* (see Algorithm 1). It is an iterative alteration process. We shall make corrections of probability in a top-down manner: we fix the probabilities of critical events $\mathcal{E}_p$ in decreasing order of cardinality of the set of variables appearing in p. It is crucial to remember throughout the proof the correct interpretation of the notation: *in the probability space, event $\mathcal{E}_{px_{ij}}$ is a subset of the event $\mathcal{E}_p$*. For the rest of this subsection, $P[\mathcal{E}_p]$ stands for $P_{E_d}[\mathcal{E}_p]$.

Input: Probability distribution E_d
(defined by two arrays of probabilities: M with one entry $M[s]$ for each integer solution s and N with one entry $N[\mathcal{E}_p]$ for each event $\mathcal{E}_p$);
Set of variables $v(\pi, z)$; Multiplier z
Output: Probability distribution E_f (defined by the arrays M, N)

$i = 1$; $\mu, p_1, p_2 = 0$
while $i \leq k$ **do**
  forall the $p = x_{i_b j} q$ *s.t.* q *is a product of* $k - i + 1$ *assignments to cheap* **do**
    $\mu = N[\mathcal{E}_p] - P[\mathcal{E}_{p/fixed}]$;
    while $N[\mathcal{E}_p] \neq P[\mathcal{E}_{p/fixed}]$ **do**
      find solution s s.t. $M[s] > 0$, $x^s_{i_b j} = 0$, $x^s_{i_b j'} = 1$ for some $j' \notin v(\pi, z)$, and all the assignments in z are false;
      $p_1 = M[s]$;
      find solution s' s.t. $M[s'] > 0$, $x^{s'}_{i_b j} = 1$, $x^{s'}_{i_b j'} = 0$, $\mathcal{E}_p$ is true, and all the other assignments in z are false;
      $p_2 = M[s']$;
      s'' := solution s with the assignments of j, j' exchanged;
      s''' := solution s' with the assignments of j, j' exchanged;
      $M[s] := M[s] - \min\{\mu, p_1, p_2\}$;
      $M[s'] := M[s'] - \min\{\mu, p_1, p_2\}$;
      $M[s''] := M[s''] + \min\{\mu, p_1, p_2\}$;
      $M[s'''] := M[s'''] + \min\{\mu, p_1, p_2\}$;
      Update the values $N[\mathcal{E}_{\hat{p}}]$ for each event $\mathcal{E}_{\hat{p}}$ that is true for any of the s, s', s'', s''';
      $\mu := \mu - \min\{\mu, p_1, p_2\}$;
    end
  end
  $i := i + 1$
end

Algorithm 1: Transfusion of Probability Process

The above Algorithm 1 takes as input the probability space E_d, the variable set $v(\pi, z)$ and the multiplier z. The algorithm repeatedly performs alterations to distribution E_d, which are described by the inner while-loop, and produces intermediate distributions. The arrays M and N store the probabilities of the solutions and events with respect to the current distribution; they are initialized to the probabilities defined by E_d. At termination, the entries of N correspond to the probabilities of the desired distribution E_f.

A distribution E_g is a *valid alteration* of E_d if it has the following two properties: (i) for any event $\mathcal{E}_p$ such that p contains only assignments to cheap facilities, with the corresponding variables being members of $v(\pi, z)$, $P_{E_d}[\mathcal{E}_p] = P_{E_g}[\mathcal{E}_p]$ (ii) for $p = x_{i_b j}$, $P_{E_g}[\mathcal{E}_p] = P_{E_d}[\mathcal{E}_p] = x_{i_b j}$, i.e., the probability of the event $x_{i_b j} = 1$ must not be altered. Note that, by the consistency condition of Lemma 4.1, we do not care about altering the probability of events with variables not in $v(\pi, z)$. We shall prove that the distribution E_f output by Algorithm 1 is a valid alteration of E_d in which the probabilities of all critical events have been corrected.

We consider first the outer while-loop. At the ith iteration, $i = 1, \ldots, k$, of the outer while-loop we fix the probability of all the events $(x_{i_b j} = 1 \wedge x_{i_{a_1} j_{b_1}} = 1 \wedge \ldots \wedge x_{i_{a_{k-i+1}} j_{b_{k-i+1}}} = 1)$ where $x_{i_b j} x_{i_{a_1} j_{b_1}} \ldots x_{i_{a_{k-i+1}} j_{b_{k-i+1}}}$ is a product p appearing in constraint π multiplied by z. The corrections of the probabilities of events of previous iterations have affected the probabilities of the events of the current iteration and thus in general $N[\mathcal{E}_p] \neq P[\mathcal{E}_p]$. We bound this effect on the probability of an event $\mathcal{E}_p$ of the current iteration i by considering the total alteration of the probabilities of the events $\mathcal{E}_{p'} = \mathcal{E}_p \wedge x_{ij} = 1$, with x_{ij} in the set of variables appearing in z and $x_{ij} \notin \mathcal{E}_p$, of the previous iterations and using the union bound.[2] There are exactly i events needed to be taken into consideration for each such $\mathcal{E}_p$ of the current step i. Denote the amount of the effect of the correction of the previous iterations by δ_i. By Proposition 4.2, δ_i is at most $i((1 + (1 + o(1))/n)^{k-i+2} - 1)P[\mathcal{E}_{p'/fixed}]$ while the measure $\delta = P[\mathcal{E}_p] - P[\mathcal{E}_{p/fixed}]$ of the needed correction for $\mathcal{E}_p$ is at least $((1 + (1 - o(1))/n)^{k-i+1} - 1)P[\mathcal{E}_{p/fixed}]$, which by Proposition 4.1 and by the at most $\frac{n}{10}$ number of rounds we consider is higher, in particular

$$\begin{aligned} ((1 + (1 - o(1))/n)^{k-i+1} - 1)P[\mathcal{E}_{p/fixed}] \geq \\ n(1 - o(1))((1 + (1 - o(1))/n)^{k-i+1} - 1)P[\mathcal{E}_{p'/fixed}] > \\ i((1 + (1 + o(1))/n)^{k-i+2} - 1)P[\mathcal{E}_{p'/fixed}] \end{aligned}$$

Thus, at the current iteration i, we have to subtract from $N[\mathcal{E}_p]$ a nonnegative amount $\mu = \delta - \delta_i$. The following has been proved.

Proposition 4.3 *At Line 4 Algorithm 1 always assigns a nonnegative value to variable* μ.

To subtract from $N[\mathcal{E}_p]$ a measure of μ, the inner while-loop implements the following *transfusion step*: pick a measure μ of solutions from the current distribution such that $x_{i_b j} = 0$, $x_{i_b j'} = 1$ for any j' that does not appear as index of any variable in $v(\pi, z)$, all the other events of $\mathcal{E}_p$ are false, and so are all the remaining events corresponding to assignments in z. Then pick an equal measure μ of solutions such that $x_{i_b j} = 1$, $x_{i_b j'} = 0$, event $\mathcal{E}_p$ is true, and all the remaining events corresponding to assignments in z are false. Now exchange the values of the assignments of j and j' of the solutions of the two sets. In Algorithm 1 this is performed step-by-step by selecting a solution s with positive measure ($M(s) > 0$) from the first family, a solution s' with positive measure ($M(s') > 0$) from the second family and by decreasing equally their measure while at the same time increasing equally the measures of the solutions s'', s''' which are the s, s' with the assignments $x_{i_b j}, x_{i_b j'}$ flipped.

Proposition 4.4 *At the end of the inner while-loop of Algorithm 1 the distribution defined by the array N is a valid alteration of E_d. Moreover, when the for-loop ends at Line 20 for all the critical events $\mathcal{E}_p$ where p contains at least $k - i + 1$ assignments to cheap we have that $N[\mathcal{E}_p] = P[\mathcal{E}_{p/fixed}]$.*

Proof. The resulting distribution has the probability of $\mathcal{E}_p$ corrected to the desired value $P[\mathcal{E}_{p/fixed}]$ since we decreased by μ the measure of solutions for which $\mathcal{E}_p$ is true and at the same time $\mathcal{E}_p$ is false in those solutions whose measure was increased. Moreover, again by the

[2]Notice that any effect of iteration $j < i - 1$ on $P[\mathcal{E}_p]$, originates from events that are subsets of $\mathcal{E}_{p'}$ and has therefore been accounted for.

choice of the families of solutions on which we perform the transfusion step, the probabilities of the events corrected in the previous iterations were not altered since, in all the solutions involved, for any product p' of variables in $v(\pi, z)$ containing p the corresponding event $\mathcal{E}_{p'}$ is false. The probability of an event $\mathcal{E}_{p*}$ containing only assignments to cheap facilities in $v(\pi, z)$ was not changed since the event is always false for s, s'', and if $\mathcal{E}_{p*}$ is true for s' then and only then it is also true for s'''. Lastly, by the fact that $x_{i_b j}$ is true in s' and s'' and false in s and s''', we have that the probability of the event $x_{i_b j}$ remains invariant. ∎

It remains to show that the algorithm terminates. It suffices to show that the transfusion step can always be performed, i.e., that there is enough measure μ of integer solutions of the first family so that the algorithm always finds an appropriate s in the inner while-loop.

Proposition 4.5 *The transfusion step of Algorithm 1 can always be performed: as long as $N[\mathcal{E}_p] \neq P[\mathcal{E}_p/fixed]$, the inner while-loop always finds s with $M[s] > 0$.*

Proof. The intuition behind the proof is that the "donor" solutions that supply the required measure are much more likely to occur than the events that require the transfusion.

Consider the measure t in E_d of the set of integer solutions satisfying $y_{i_b} = 1$ and all events encountered at any iteration being false, namely $x_{i_b j} = 0 \wedge x_{i_1 j_1} = 0 \wedge x_{i_2 j_2} = 0 \wedge \ldots \wedge x_{i_k j_k} = 0$. Then, by the random experiment of the construction of E_d, this event is equivalent to the event that facility i_b is picked, $x_{i_b j} = 0$ and the k balls corresponding to the clients of the rest of the events are not tossed in their corresponding bins. Using again that both w^1_{ch}, w^2_{ch} are $\Theta(n^3)$ and $k < n$, we can bound the probability of the k balls by that of k Bernoulli trials with probability of success $2/n$ (we are once again very generous). Then the probability that all events fail is at least $(1 - 2/n)^k > \lim_{n\to\infty}(1 - 2/n)^n = 1/e^2$. Thus measure t is at least $(y_{i_b} - x_{i_b j})1/e^2$ which is constant. On the other hand, the measure required by the transfusion step for each event $\mathcal{E}_p$ of iteration i that needs to be fixed is, by Proposition 4.2, at most $(e^2-1)P[\mathcal{E}_{p/fixed}]$, which, by Proposition 4.1, is $\Theta(1/n^i)$. There are $\binom{k+1}{k-i+1}$ such events of iteration i, and summing over all the iterations of our construction we get $\sum_{i=1}^{k} \binom{k+1}{k-i+1}\Theta(1/n^i)$ which quantity is less than $(y_{i_b} - x_{i_b j})1/e^2$ for the $k = n/10$ levels of SA we consider. Therefore, we can always pick the required amount of measure. ∎

4.2.3 The main result

Theorem 4.1 *There is a family of instances $\mathcal{I}$ of uniform* CFL*, such that for every sufficiently large n, there is an instance in the family with $2n$ facilities and $n^4 + 1$ clients for which the relaxations obtained from (LP-classic) at $\Omega(n)$ levels of the Sherali-Adams hierarchy have an integrality gap of $\Omega(n)$.*

Proof. For each constraint π multiplied by multiplier z at level t, the corresponding distribution E_d or E_f is clearly a distribution over integer solutions, so the first condition of Lemma 4.1 is satisfied. For the second condition, we examine the probability of an event $\mathcal{E}_p$, for any monomial p appearing in the expanded form of π multiplied by z.

Observe that if an event $\mathcal{E}_p$ involves more than one costly facility, it has 0 probability in all distributions. If an event $\mathcal{E}_p$ involves only cheap facilities, it has the same probability in all distributions E_f and E_d, since in the construction of a distribution E_f we took care not to change the probability of such events. An event $\mathcal{E}_p$ that involves more than one assignment of a costly facility (but no other costly) has in every distribution E_f the same probability (which is the same as in every E_d) since in the construction of E_f we did not alter the probabilities of such events. Indeed, notice that the facility i_b in the arguments of Subsection 4.2.2, is chosen so that it does not appear in the multiplier z. And lastly, when an event $\mathcal{E}_p$ involves exactly one assignment of some costly facility i_x, note that in some cases i_x takes the blame but in other cases it does not, depending on $v(\pi, z)$. But due to the transfusion of probability process, the probability of event $\mathcal{E}_p$ in a distribution in which i_x is not the facility that takes the blame is equal to the probability of the same event in the distributions that i_x takes the blame. So Lemma 4.1 is applicable. It is then easy to see that the bad solution has cost $\Theta(n^{-1})$ while any feasible solution to the instance has cost $\Omega(1)$. ∎

We remark that even though in the proofs of Theorems 3.1 and 4.1 we show the survival of the same fractional solution s, the lifted solutions obtained in our proofs are in generally different. Consider the situation at level $t \geq 1$ of both hierarchies. The product $Y_{\{y_i, y_i'\}} = y_i y_{i'}$ for any two distinct costly facilities i and i', is zero as the random experiment we define in the SA proof opens exactly one costly facility. In contrast, the associated entry of the protection matrix we construct for the LS proof is in general nonzero, because as long as none of the two facilities have been touched as Type 2, their product is a small but nonzero quantity.

4.3 Robustness of the SA gap

In this section we explain how adding simple valid inequalities to (LP-classic) does not affect our arguments on the SA hierarchy.

As an example we address the valid inequality (9), which was discussed also in the context of LS in Subsection 3.3. Of course, this inequality is rendered useless by slight modifications to the instance and the bad solution. Identifying "areas" of a fractional solution where the demand exceeds the available capacity seems impossible for polynomially sized relaxations without some yet unknown form of preprocessing. In fact, part of the motivation behind Theorem 4.1 was to demonstrate that the SA hierarchy is inadequate for such preprocessing purposes.

We modify the family $\mathcal{I}$ as follows. To instance I with n cheap and n costly facilities and $Un+1$ clients we add a *dummy* facility a with 0 opening cost, at distance 1 from the rest. The underlying metric space is a line, and thus we have an instance of the so called *facility location on a line*. Call I_a the resulting instance and $\mathcal{I}_a$ the resulting family. We define bad solution s^a to agree with s on the old variables, additionally we set $y_a = 1$ and $x_{aj} = 0$ for all clients j. Inequality (9) is obviously satisfied by s^a. Moreover s^a induces an integrality gap of $\Omega(n)$.

We state first three generic properties of the SA hierarchy from which the claim will be obtained. This line of proof was suggested by an anonymous reviewer.

Lemma 4.3 *Consider a polytope $P \subseteq [0,1]^d$ of dimension at least one defined by the system $Ax \leq b$. Let $\pi_i(x) = a_i^T x - b_i \leq 0$ be a redundant inequality of the system. Let $P' \subseteq [0,1]^d$*

be the polytope defined by the system $A'x \leq b'$ which is obtained from $Ax \leq b$ by omitting π_i. Then $\mathrm{SA}^k(P) = \mathrm{SA}^k(P')$, i.e., the SA *system does not depend on the syntactic description of P.*

Proof. It suffices to show that the lifted polytopes are actually the same. Obviously $\mathrm{SA}^k(P) \supseteq \mathrm{SA}^k(P')$. Since $\pi_i(x) \leq 0$ is redundant, it can be obtained by a conic combination $\sum_j \lambda_j \pi_j(x)$ of inequalities from $Ax \leq b$. Now consider a constraint $z\pi_i(x) \leq 0$ of the lifted P for some k-level (X,W)-multiplier z. This constraint is redundant in the system of the lifted P' as it can obtained as $z\pi_i(x) = \sum_j \lambda_j z\pi_j(x) \leq 0$. Thus $\mathrm{SA}^k(P) = \mathrm{SA}^k(P')$. ∎

Lemma 4.4 *Let $\bar{x} \in \mathrm{SA}^k(P)$ and $\bar{x}_i = 1$ ($\bar{x}_i = 0$). Then for a feasible vector z for the lifted polytope that projects to $\bar{x}$ and for $I \not\ni x_i$ it is true that $z_I = z_{I\cup\{x_i\}}$ ($z_{I\cup\{x_i\}} = 0$). (For notational simplicity we assume $z_\emptyset = 1$.)*

Proof. All box constraints $0 \leq x_j \leq 1$ are valid by definition. First observe that by multiplying $x_i \leq 1$ by $\prod_{l\in I} x_l$ we get $z_{I\cup\{x_i\}} \leq z_I$.

The proof is by induction on $|I|$. For $|I| = 0$ the claim obviously holds. Now, if $|I| > 0$ is even, consider the constraint of the lifted polytope obtained by multiplying the constraint $x_j \geq 0$ for some $x_j \in I$ with the multiplier $(1 - x_i)\prod_{x_l \in I-\{x_j\}}(1 - x_l)$. The obtained constraint is $\sum_{I'\subseteq I\cup\{x_i\}|x_j\in I'}(-1)^{|I'|-1} z_{I'}$. By the inductive hypothesis the term $z_{I'}$ with $|I'| < |I|$ and $x_i \notin I'$ cancels the term $z_{I'\cup\{x_i\}}$. So what is left is $z_{I\cup\{x_i\}} - z_I \geq 0$ which, together with $z_I \geq z_{I\cup\{x_i\}}$, gives the claim. If $|I| > 0$ is odd the proof is similar, we just start from $(1 - x_j) \geq 0$ and we multiply by $(1 - x_i)\prod_{x_l\in I-\{x_j\}}(1 - x_l)$. For $x_i = 0$ the lemma is proved again similarly. ∎

Lemma 4.5 *Consider some polytope $P \subseteq [0,1]^d$, and $\bar{x} \in P$ with $\bar{x}_i = \beta$, where $\beta \in \{0,1\}$. Let $P' \subseteq [0,1]^{d-1}$ be a polytope defined by all constraints*

$$\sum_{j:j\neq i} \alpha_j x_j + \alpha_i \cdot \beta \leq b$$

where $\alpha^T x \leq b$ is a constraint of P. Let also x' be the projection of $\bar{x}$ onto coordinates $j \neq i$. Then $\bar{x} \in S^k(P)$ if and only if $x' \in S^k(P')$.

Proof. Let z' be a solution of the lifted P' that projects to x'. Consider the solution $\overline{z}$ for the lifted P obtained by setting for all sets $I \not\ni i$, $\overline{z}_{I\cup\{i\}} = \beta \cdot z'_I$, and $\overline{z}_I = z'_I$. Notice that $\overline{z}$ projects to $\overline{x}$. By Lemma 4.4 it is immediate that $\overline{z}$ is feasible for the lifted P if and only if z' is feasible for the lifted P'. ∎

Given the above facts, it is now easy to conclude the argument.

Theorem 4.2 *Let P be the polytope defined on instance I_a of family $\mathcal{I}_a$ by (LP-classic) augmented by constraint (9). For sufficiently large n and $k = \Omega(n)$, vector $s^a \in \mathrm{SA}^k(P)$.*

Proof. Let P_1 be the polytope defined by (LP-classic) where variable y_a has been replaced by the constant 1 and for all j x_{aj} has been replaced by the constant 0. Observe that the system defining P_1 is simply (LP-classic) on instance I augmented by the redundant inequality $0 \leq 1$. By Theorem 4.1, $s \in SA^k(P_1)$.

Let P' be the polytope whose description results from the system defining P after replacing y_a and x_{aj} for all j, with their values in s^a, namely, 1 and 0 respectively. By an inductive application of Lemma 4.5, $s^a \in \mathrm{SA}^k(P)$ iff $s \in \mathrm{SA}^k(P')$. Constraint (9) appears in the description of P' as

$$\sum_{i \neq a} y_i + 1 \geq \lceil \frac{n^4+1}{n^3} \rceil \Leftrightarrow \sum_{i \neq a} y_i \geq n \qquad (10)$$

Inequality (10) is redundant for P' as it results from the summation of the $2n$ capacity constraints for the cheap and costly facilities. By Lemma 4.3, $\mathrm{SA}^k(P') = \mathrm{SA}^k(P_1)$ and since $s \in \mathrm{SA}^k(P_1)$ it follows that $s^a \in \mathrm{SA}^k(P)$. ∎

5 Fooling the submodular inequalities for CFL

In this section we show that the relaxation obtained from (LP-classic) with the addition of the submodular inequalities proposed by Aardal et al. [1] has a gap not bounded by any constant. Levi et al. [39] conjectured that the addition of a subset of those inequalities, called the flow-cover inequalities, to the classic relaxation would improve the integrality gap to constant. We note that it is not known how to separate in polynomial time even the flow-cover inequalities. We wish to clarify, that the inequalities introduced by Carr et al. [14] and extended by Carnes and Shmoys [13] are referred to also as "flow-cover inequalities". Our proof does not apply to the latter.

We proceed to define first the flow-cover inequalities and then generalize the definition to the submodular ones. Consider the general case where facility i has capacity u_i and client j has demand d_j. For a set J of clients, we denote their total demand by $d(J) = \sum_{j \in J} d_j$. Let $J \subseteq C$ be a set of clients, let $I \subseteq F$ be a set of facilities, and let $J_i \subseteq J$ be a set of clients for each facility $i \in I$. Given a facility i, we denote the *effective capacity* of i with respect to J_i by $\bar{u}_i = \min\{u_i, d(J_i)\}$. I is a *cover* with respect to J if $\sum_{i \in I} \bar{u}_i = d(J) + \lambda$ with $\lambda > 0$. λ is called the *excess capacity.* Let $(x)^+ = \max\{x, 0\}$. In the case where $J_i = J$ for all $i \in I$ the following inequalities called *flow-cover* inequalities were introduced for CFL in [1].

$$\sum_{i \in I} \sum_{j \in J} d_j x_{ij} + \sum_{i \in I} (u_i - \lambda)^+ (1 - y_i) \leq d(J)$$

If $\max_{i \in I}(\bar{u}_i) > \lambda$, the following inequalities, called the *effective capacity inequalities* are valid and strengthen the flow-cover inequalities [1]. Note that we no longer assume that $J_i = J$.

$$\sum_{i \in I} \sum_{j \in J_i} d_j x_{ij} + \sum_{i \in I} (\bar{u}_i - \lambda)^+ (1 - y_i) \leq d(J)$$

The submodular inequalities introduced in [1] are even stronger than the effective capacity inequalities. From now on we limit our discussion to uniform CFL with *all clients having unit demands.*

Choose a subset $J \subseteq C$ of clients, and let $I \subseteq F$ be a subset of facilities. For each facility $i \in I$ choose a subset $J_i \subseteq J$. Consider a 3-level network G with a source s, a set of nodes corresponding to the facilities, a set of nodes corresponding to the clients and a sink t. The source s is connected by an edge of capacity $\min\{U, |J_i|\}$ to each facility node i. That node is connected by an edge of unit capacity to each node corresponding to client j, $j \in J_i$. Each node corresponding to some client is connected by an edge of unit capacity to the sink t.

Define $f(I)$ as the maximum s-t flow value in G. Define $f(I \setminus \{i\})$ as the maximum flow when facility i is closed, i.e., when the capacity of edge (s, i) is set to zero. The difference in maximum flow when all facilities in I are open, and when all facilities except facility i are open, is called the *increment* function and is defined as $\rho_i(I \setminus \{i\}) = f(I) - f(I \setminus \{i\})$.

For any choice of $I \subseteq F$, $J \subseteq C$, and $J_i \subseteq J$, for all i, the following inequalities, called *the submodular inequalities,* are valid for CFL [1]. The name reflects the fact that the function $f(I)$ is submodular.

$$\sum_{i\in I}\sum_{j\in J_i} x_{ij} + \sum_{i\in I} \rho_i(I \setminus \{i\})(1 - y_i) \leq f(I)$$

The intuitive explanation for an integer solution is the following: if some set $S \subseteq I$ of facilities is closed then the loss in the total flow through G is at least $\sum_{i\in S} \rho_i(I \setminus \{i\})$. Thus the total assignments wrt the selected client sets J_i's cannot be greater than the maximum possible flow $f(I)$ minus the flow loss due to the closed facility of S which is at least $\sum_{i\in S} \rho_i(I \setminus \{i\})$. The proof of the following theorem uses some of the ideas we introduced earlier for Theorem 4.1.

Theorem 5.1 *There is a family of instances $\mathcal{I}$ of uniform* CFL, *such that for every sufficiently large n, there is an instance in the family with $2n$ facilities and $n^4 + 1$ clients for which the integrality gap of the relaxation produced from (LP-classic) with the addition of the submodular inequalities is $\Omega(n)$.*

Proof. Fix n. Consider the resulting instance $\mathcal{I}(n)$ from the family $\mathcal{I}$ and the bad solution s that we used in Theorem 4.1 for the SA result. To prove that s is feasible for the classic relaxation strengthened by the submodular inequalities we take the idea of fooling local constraints a little further: either the constraint is local enough that we can use the ideas from our previous proofs (define s' that is a convex combination of integer solutions and the values of the variables in the support of the constraint agree with s), or we can define another instance $\mathcal{I}'(n)$ and solution s' for which the inequality in question is true with respect to s' and again s' has the same visible part as s with respect to the constraint. Note that our arguments include two different instances as opposed to all our other proofs so far.

Consider the submodular inequality π for some I, J and some selection of J_i's. If not all the costly facilities appear in the constraint the proof is similar to that of Lemma 4.2. If at least n assignment variables to cheap facilities do not appear in π we do the following:

we add one more facility b to the instance. We construct a solution s' for the new instance $\mathcal{I}'(n)$ as follows. We transfer the demand corresponding to the missing assignments of the cheap to b, and we set $y'_b = 1$. Observe that π is valid for $\mathcal{I}'(n)$. Now we can show that s' is a convex combination of integer solutions, again similarly to the proof of Lemma 4.2. Again, we call facility b *active* if it is assigned some nonzero demand. Facility b will be open 100% of the time but it will be active only when no costly facilities are open, i.e., a fraction $1 - \sum_{i \in Costly} y'_i$ of the time. When it is active, it is assigned $\frac{\sum_j x'_{bj}}{1 - \sum_{i \in Costly} y'_i}$ amount of demand. For simplicity, assume that this quantity is an integer. If it is not, we can take the remedial action described in detail in the proof of Lemma 4.2.

In the fractional solution s', b is assigned a total demand of at least $1 - 1/n^2$, therefore, in each outcome of the random experiment in which b is active, it will be assigned at least one client. Thus, in each such outcome we obtain a feasible integer solution. By the convex combination produced, the inequality is satisfied by s'. Thus the same inequality for the original instance is satisfied by s, since the exact same values of variables are involved in both cases.

Now consider the case where less than n assignments to cheap facilities are missing from π. We will show that it cannot be the case that all y_i variables of costly facilities appear in the constraint as well. Consider the quantity $\rho_i(I \setminus \{i\})$ for some costly facility i. If $\rho_i(I \setminus \{i\}) > 0$, then J_i is not empty. We will show that the set of nodes $(Cheap \cap I) \cup \{i\}$ in G has enough incident edges so that the flow originating from them is equal to the total client demand $|J|$ in G. We first give some properties of graph G.

Claim 5.1 *If less than n assignments to cheap facilities are missing from π, then $(Cheap \cap I) = Cheap$ and $J = C$.*

Proof of Claim. To see that $(Cheap \cap I) = Cheap$, notice that if a cheap facility is missing from I, at least $|C| = n^4 + 1$ assignment variables will be missing from π, a contradiction. For the second part of the claim, if a client j is missing from J, then all the corresponding n edges that would connect j to a cheap facility cannot be in G. Therefore at least n assignment-to-cheap variables are missing from π, a contradiction. The proof of the claim is complete.

We return to proving that $Cheap \cup \{i\}$ has enough incident edges so that the flow originating from them is equal to the total client demand $|C|$ in G. "Assign" one client $j \in J_i$ to facility i and for the remaining $|C| - 1$ clients do the following: assign each client j' involved in the set of variables of assignments-to-cheap that are missing from π to a cheap facility i' such that $j' \in J_{i'}$. There is always such a cheap facility i' since the missing edges from the client-nodes in G to the cheap-facility nodes are less than n. Assign the remaining clients arbitrarily to the cheap facilities respecting the capacities, since all the edges from cheap to those clients are included in the network. Thus it must be the case that $\rho_{i'}(I \setminus \{i'\}) = 0$ for any other costly facility $i' \neq i$. Since the $y_{i'}$ variable of such a facility i' has 0 coefficient in the constraint, it can take the blame and the proof is similar to that of Lemma 4.2. ∎

6 Proper Relaxations

In this section we present the family of proper relaxations and characterize their strength. Given an instance of CFL a *class* is defined as a 0-1 vector $\eta = (y^\eta, x^\eta)$ which satisfies constraint (2) of (LP-classic). Note that the definition of a class η is quite general as the support of the vector (y^η, x^η) is not necessarily a subset of the support of some feasible integer solution for CFL.

We associate with class η the *cost of the class*

$$c_\eta = \sum_{i|y_i^\eta=1} f_i + \sum_{i,j|x_{ij}^\eta=1} c_{ij}.$$

Let the *pairings of class* η be defined as

$$\mathrm{Pairs}(\eta) = \{(i,j) \in F \times C \mid x_{ij}^\eta = 1\}.$$

We say that class η *opens* facility i, if $y_i^\eta = 1$. The set of facilities opened in η is denoted by $F(\eta)$.

Definition 6.1 (Constellation LPs) *Let $\mathcal{C}$ be a set of classes defined for an instance $I(F,C)$ of* CFL*. Let z_η be a variable associated with class $\eta \in \mathcal{C}$. The* constellation LP with class set $\mathcal{C}$ *is defined as*

$$\min \sum_{\eta\in\mathcal{C}} c_\eta z_\eta \qquad (\mathrm{LP}(\mathcal{C}))$$

$$\sum\nolimits_{\eta|\exists i:(i,j)\in \mathrm{Pairs}(\eta)} z_\eta = 1 \qquad \forall j \in C$$

$$\sum\nolimits_{\eta|i\in F(\eta)} z_\eta \leq 1 \qquad \forall i \in F$$

$$z_\eta \geq 0 \qquad \forall \eta \in \mathcal{C}$$

We refer simply to a *constellation LP* when $\mathcal{C}$ is implied from the context. Given a solution z of LP($\mathcal{C}$), the induced solution to (LP-classic) is $(y,x) = \sum_{\eta\in\mathcal{C}} z_\eta (y^\eta, x^\eta)$.

We restrict our attention to constellation LPs that satisfy a symmetry property that is very natural for uniform capacities and unit demands.

Definition 6.2 (P_1: Symmetry) *We say that property P_1 holds for the constellation linear program LP($\mathcal{C}$) if for every class $\eta \in \mathcal{C}$, all classes resulting from a permutation that relabels the facilities and/or the clients of η are also in $\mathcal{C}$.*

Definition 6.3 (Proper Relaxations) *We call* proper relaxation *for* CFL *a constellation LP that is valid and satisfies property P_1.*

A simple example of a constellation LP is the well-known *(LP-star)* (see, e.g., [28]) where $\mathcal{C}$ corresponds to the set of all *stars*: a facility and a set of at most U clients assigned to it. Obviously (LP-star) is a proper relaxation, while (LP-classic) is equivalent to (LP-star). Therefore proper relaxations generalize the known natural relaxations for CFL. In order to characterize the strength of a proper LP we need the notion of complexity.

Definition 6.4 (Complexity of proper relaxations) *Given an instance $I(F, C)$ of* CFL *the* complexity α *of a proper relaxation $LP(\mathcal{C})$ for I is defined as the* $\max_{\eta \in \mathcal{C}}(|F(\eta)|/|F|)$.

The complexity of a proper LP represents the maximum fraction of the total number of feasibly openable facilities that is allowed in a single class. A complexity of nearly 1 means that there are classes that each take into consideration almost the whole instance at once. Low complexity means that all classes consider the assignments of a small fraction of the instance at a time. It is easy to create families of instances for which by increasing the complexity one obtains strictly better lower bounds. The following theorem characterizes the strength of proper relaxations with regard to their complexity.

Theorem 6.1 *The quality of solutions of proper relaxations is characterized by the following:*

(i) There is a family of instances $\mathcal{J}$ of uniform CFL, *such that for every sufficiently large n, there is an instance in the family with n facilities and $n^2(n-1)+1$ clients for which every proper relaxation with complexity $\alpha < 1$ has an integrality gap of $\Omega(n^2)$.*

(ii) There is a proper relaxation of complexity 1 for which the set of induced (y, x) vectors of its solutions is the integral CFL *polytope.*

6.1 Proof of (ii) of Theorem 6.1

We first prove the easy part, that there are proper relaxations for CFL with complexity 1 that express the integral polytope. For any instance, let $\mathcal{C}$ consist of the vectors of feasible 0-1 integer solutions. The resulting $LP(\mathcal{C})$ is clearly proper. Let z be any feasible solution of $LP(\mathcal{C})$ and let S be the support of the solution. For every $\eta \in S$, and for every client $j \in C$, there is an $i \in F$, such that $(i, j) \in \text{Pairs}(\eta)$. From this observation and the feasibility of z, for any $j \in C$ we have:

$$\sum\nolimits_{\eta | \exists i : (i,j) \in \text{Pairs}(\eta)} z_\eta = 1 \Rightarrow \sum_{\eta \in S} z_\eta = 1.$$

This implies that z is a convex combination of integral solutions. By the boundedness of the feasible region of $LP(\mathcal{C})$, the corresponding polytope is integral. Clearly, not every LP with complexity 1 has an integrality gap of 1 since it might contain "strong" classes which correspond to feasible integer solutions together with "weak" ones, such as stars, whose conic combinations yield fractional solutions not in the integer CFL polytope.

6.2 Proof of (i) of Theorem 6.1

In the next two subsections, we prove the first part of Theorem 6.1. In Subsection 6.2.1, we show that, for any proper relaxation $LP(\mathcal{C})$ of complexity strictly less than 1, there must be a specific set of classes contained in $\mathcal{C}$. In Subsection 6.2.2, we use these classes to construct a desired feasible fractional solution. By defining appropriately the cost structure of the instance, we establish an unbounded integrality gap.

6.2.1 Existence of a specific type of classes

Consider an instance I with n facilities, where n is sufficiently large to ensure that $\alpha n \leq n - c_0$ where c_0, is a constant greater than or equal to 1. Let the capacity be $U = n^2$, and let the number of clients be $(n-1)U + 1$. Notice that in every integer solution of the instance we must open all n facilities. The facility costs and the assignment costs will be defined later.

We assume, like before, that the facilities are numbered $1, 2, \ldots, n$. Consider an integral solution $s = (y^s, x^s)$ for I where all the facilities are opened, and furthermore facilities $1, \ldots, n-1$ are assigned U clients each and facility n is assigned one client. Since our proper relaxation is valid, there must be a solution s' in the space of feasible solutions of the proper relaxation whose induced (y, x) vector equals s. By Definition 6.1, it is easy to see that s' can only be obtained as a positive combination of classes η such that for every facility i and client j we have $x^{\eta}_{ij} \leq x^{s}_{ij}$. Recall that since the complexity of our relaxation is α, the classes in the support of any solution have at most $n - c_0 \leq n-1$ facilities.

Now consider the support of s'. We will distinguish the classes η for which variable z_η is in the support of s' into 2 sets. The first set consists of the classes that assign exactly one client to facility n; call them *type A* classes. The second set consists of the classes that do not assign any client to facility n; call those *type B* classes. By the discussion above those sets form a partition of the classes in the support of s', and moreover they are both non-empty: this is by the fact that at most $n - c_0$ facilities are in any class, and by the fact that in s all n facilities are opened integrally. Notice also that no class of type B can open facility n even though the definition of a class does not exclude the possibility that a class opens a facility to which no clients are assigned.

We call *density* of a class η the ratio $d(\eta) = \frac{\sum_{i \neq n} \sum_{j \in C} x^{\eta}_{ij}}{\sum_{i \neq n} y^{\eta}_i}$. By the discussion above we have that $d(\eta) \leq U$ for all η in the support of s'. The following holds:

Lemma 6.1 *All classes in the support of s' have density U.*

Proof. The amount of demand that a class η contributes to the demand assigned to the set of the first $n-1$ facilities by s' is $d(\eta) z_\eta \sum_{i \neq n} y^{\eta}_i$. We have $\sum_\eta d(\eta) \sum_{i \neq n} y^{\eta}_i z_\eta = (n-1)U$. By the fact that in the induced (y, x) solution of s, for $i = 1, \ldots, n-1$, $y_i = \sum_\eta z_\eta y^{\eta}_i = 1$, we obtain that $\sum_\eta (z_\eta \sum_{i \neq n} y^{\eta}_i) = n - 1$. Setting $m_\eta = \frac{z_\eta \sum_{i \neq n} y^{\eta}_i}{n-1}$ we have from the above $\sum_\eta m_\eta = 1$ and $\sum_\eta m_\eta d(\eta) = U$. The latter equations together with the fact that $d(\eta) \leq U$ imply that $d(\eta) = U$ for all classes η in the support of s'. ∎

The following lemma is immediate from the above:

Lemma 6.2 *The support of s' contains a variable corresponding to a type B class, denoted η_0, that has density U.*

So far we have proved that in the class set of any proper relaxation for I, there is a class η_0 of type B with density $d(\eta_0) = U$. Let $|F(\eta_0)| = t \leq n - 1$.

6.2.2 Construction of a bad solution

Consider the symmetric classes of η_0 for all permutations of the n facilities and for all permutations of the clients. Those classes are not necessarily in the support of s'. Take a quantity of measure ϵ and distribute it equally among all those classes. Since class η_0 has density U, all those symmetric classes assign on average U clients to each of their facilities. Due to symmetry, each facility is in a class $\epsilon\frac{t}{n}$ of the time and is assigned $\epsilon\frac{t}{n}U$ demand. Each client is assigned to each facility $\epsilon\frac{tU}{((n-1)U+1)n}$ of the time. We call that step of our construction *round A*.

Now consider the symmetric classes of η_0 for all permutations of the first $n-1$ facilities and for all permutations of the clients (those classes are well defined since $t \leq n-1$). Again distribute a quantity of measure ϵ equally among all those classes. Similarly to the previous, each facility is in a class $\epsilon\frac{t}{n-1}$ of the time and is assigned $\epsilon\frac{t}{n-1}U$ demand. Each client is assigned to each facility $\epsilon\frac{tU}{((n-1)U+1)(n-1)}$ of the time. We call that step of our construction *round B*.

Spending $\phi = \frac{1}{nt}$ measure in round A and $\xi = \frac{(n-1)(1-1/n^2)}{t}$ measure in round B we construct a solution s_b whose corresponding (y, x) vector is the following (y^*, x^*): $y_i^* = 1$ for $i = 1, \ldots, n-1$, $y_n^* = \frac{1}{n^2}$, and for every client j, $x_{nj}^* = \frac{U/n^2}{(n-1)U+1}$ and $x_{ij}^* = \frac{1-x_{nj}^*}{n-1}$ for $i = 1, \ldots, n-1$. It is easy to see that s_b is a feasible solution for our proper relaxation.

Now simply set all distances to 0, and define the facility opening costs as $f_n = 1$ and $f_i = 0$ for $i \leq n-1$. Simple calculations show that the integrality gap of the proper relaxation is $\Omega(n^2)$. The proof of Theorem 6.1 is now complete.

7 Conclusion

In this work we showed that the unboundedness of the integrality gap for CFL persists even after applying the tightenings of the LS and SA hierarchies on the natural relaxation.

In particular, we proved the feasibility of a bad fractional solution for an asymptotically tight number of levels. We also proved that the submodular inequalities do not reduce the integrality gap to constant. Lastly, we proved similar negative results for families of proper relaxations that capture general configuration LPs.

In the recent work of An et al. [6] the first constant factor LP-based approximation algorithm for CFL was given. However, the proposed relaxation is exponential in size and, according to the authors, it is not known to be separable in polynomial time. A natural question that arises is whether there is a polynomially-sized relaxation achieving a constant integrality gap. An interesting direction is that of determining the minimum size of an approximate extended formulation of the CFL polytope, which our results arguably suggest to be exponential. We present progress towards answering the latter question in an upcoming paper [30].

Acknowledgment. We wish to thank the anonymous reviewers of this article for many valuable and substantial comments. We particularly thank the reviewer who suggested a proof for Theorem 4.2 which was more instructive than ours.

References

[1] Karen Aardal, Yves Pochet, and Laurence A. Wolsey. Capacitated facility location: Valid inequalities and facets. *Mathematics of Operations Research*, 20:562–582, 1995.

[2] Karen Aardal, Yves Pochet, and Laurence A. Wolsey. Erratum: Capacitated facility location: Valid inequalities and facets. *Mathematics of Operations Research*, 21:253–256, 1996.

[3] Zoë Abrams, Adam Meyerson, Kamesh Munagala, and Serge Plotkin. The integrality gap of capacitated facility location. Technical Report CMU-CS-02-199, Carnegie Mellon University, 2002.

[4] Ankit Aggarwal, Anand Louis, Manisha Bansal, Naveen Garg, Neelima Gupta, Shubham Gupta, and Surabhi Jain. A 3-approximation algorithm for the facility location problem with uniform capacities. *Mathematical Programming*, 141(1-2):527–547, 2013.

[5] Hyung-Chan An, Aditya Bhaskara, and Ola Svensson. Centrality of trees for capacitated k-center. *CoRR*, abs/1304.2983, 2013.

[6] Hyung-Chan An, Mohit Singh, and Ola Svensson. LP-based algorithms for capacitated facility location. In *55th IEEE Annual Symposium on Foundations of Computer Science, FOCS '14, Philadelphia, PA, USA,* pages 256–265. IEEE Computer Society, 2014.

[7] Sanjeev Arora, Béla Bollobás, and László Lovász. Proving integrality gaps without knowing the linear program. In *Proceedings of the 43rd Symposium on Foundations of Computer Science*, FOCS '02, pages 313–322, Washington, DC, USA, 2002. IEEE Computer Society.

[8] Sanjeev Arora, Béla Bollobás, László Lovász, and Iannis Tourlakis. Proving integrality gaps without knowing the linear program. *Theory of Computing*, 2(1):19–51, 2006.

[9] Egon Balas, Sebastián Ceria, and Gérard Cornuéjols. A lift-and-project cutting plane algorithm for mixed 0-1 programs. *Mathematical Programming*, 58(3):295–324, February 1993.

[10] Manisha Bansal, Naveen Garg, and Neelima Gupta. A 5-approximation for capacitated facility location. In Leah Epstein and Paolo Ferragina, editors, *Algorithms ESA 2012*, volume 7501 of *Lecture Notes in Computer Science*, pages 133–144. Springer Berlin Heidelberg, 2012.

[11] Nikhil Bansal and Maxim Sviridenko. The Santa Claus problem. In *Proceedings of the 38th annual ACM Symposium on Theory of computing*, STOC '06, pages 31–40, New York, NY, USA, 2006. ACM.

[12] MohammadHossein Bateni and MohammadTaghi Hajiaghayi. Assignment problem in content distribution networks: Unsplittable hard-capacitated facility location. *ACM Transactions on Algorithms*, 8(3):20:1–20:19, July 2012.

[13] Tim Carnes and David Shmoys. Primal-Dual Schema for Capacitated Covering Problems In *Proceedings of 13th International IPCO Conference on Integer Programming and Combinatorial Optimization*, pages 288–302, 2008.

[14] Robert D. Carr, Lisa Fleischer, Vitus J. Leung, and Cynthia A. Phillips. Strengthening integrality gaps for capacitated network design and covering problems. In *Proceedings of the Eleventh Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2000, San Francisco, CA, USA*, pages 106–115, 2000.

[15] Siu On Chan, James R. Lee, Prasad Raghavendra, and David Steurer. Approximate constraint satisfaction requires large LP relaxations. In *54th Annual IEEE Symposium on Foundations of Computer Science, FOCS '13, 26-29 October, 2013, Berkeley, CA, USA*, pages 350–359, 2013.

[16] Moses Charikar and Sudipto Guha. Improved combinatorial algorithms for the facility location and k-median problems. In *Proceedings of the 40th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '99, pages 378–388, 1999.

[17] Moses Charikar, Konstantin Makarychev, and Yury Makarychev. Integrality gaps for Sherali-Adams relaxations. In *Proceedings of the 41st annual ACM Symposium on Theory of Computing*, STOC '09, pages 283–292, New York, NY, USA, 2009. ACM.

[18] Fabián A. Chudak and David B. Shmoys. Improved approximation algorithms for a capacitated facility location problem. In *Proceedings of the 10th annual ACM-SIAM symposium on Discrete algorithms*, SODA '99, pages 875–876, Philadelphia, PA, USA, 1999. Society for Industrial and Applied Mathematics.

[19] Fabián A. Chudak and David P. Williamson. Improved approximation algorithms for capacitated facility location problems. *Mathematical Programming*, 102(2):207–222, March 2005.

[20] Michele Conforti, Gérard Cornuéjols and Giacomo Zambelli. Extended formulations in combinatorial optimization. *Annals of Operations Research*, 204(1):97–143, 2013.

[21] Gérard Cornuéjols. Valid inequalities for mixed integer linear programs. *Mathematical Programming*, 112(1):3–44, 2008.

[22] Marek Cygan, MohammadTaghi Hajiaghayi, and Samir Khuller. LP rounding for k-centers with non-uniform hard capacities. In *53rd Annual IEEE Symposium on Foundations of Computer Science, FOCS '12, New Brunswick, NJ, USA, October 20-23, 2012*, pages 273–282, 2012.

[23] Wenceslas Fernandez de la Vega and Claire Kenyon-Mathieu. Linear programming relaxations of maxcut. In *Proceedings of the 18th annual ACM-SIAM symposium on Discrete algorithms*, SODA '07, pages 53–61, Philadelphia, PA, USA, 2007. Society for Industrial and Applied Mathematics.

[24] Konstantinos Georgiou and Avner Magen. Limitations of the Sherali-Adams lift and project system: Compromising local and global arguments. Technical Report CSRG-587, University of Toronto, 2008.

[25] Michel X. Goemans and Levent Tunçel. When does the positive semidefiniteness constraint help in lifting procedures? *Mathematics of Operations Research*, 26(4):796–815, 2001.

[26] Sudipto Guha and Samir Khuller. Greedy strikes back: improved facility location algorithms. *Journal of Algorithms*, 31:228–248, 1999.

[27] Dorit S. Hochbaum. Heuristics for the fixed cost median problem. *Mathematical Programming*, 22:148–162, 1982.

[28] Kamal Jain, Mohammad Mahdian, Evangelos Markakis, Amin Saberi, and Vijay V. Vazirani. Greedy facility location algorithms analyzed using dual fitting with factor-revealing LP. *Journal of the ACM*, 50(6):795–824, November 2003.

[29] Kamal Jain and Vijay V. Vazirani. Approximation algorithms for metric facility location and k-median problems using the primal-dual schema and Lagrangian relaxation. *Journal of the ACM*, 48(2):274–296, March 2001.

[30] Stavros G. Kolliopoulos and Yannis Moysoglou. Extended formulation lower bounds via hypergraph coloring? To appear *in Proc. 32nd International Symposium on Theoretical Aspects of Computer Science, STACS 2015, Munich, Germany*, 2015.

[31] Stavros G. Kolliopoulos and Yannis Moysoglou. Exponential lower bounds on the size of approximate formulations in the natural encoding for capacitated facility location. *CoRR*, abs/1312.1819, 2013.

[32] Stavros G. Kolliopoulos and Yannis Moysoglou. Integrality gaps for strengthened LP relaxations of capacitated and lower-bounded facility location. *CoRR*, abs/1305.5998, 2013.

[33] Stavros G. Kolliopoulos and Yannis Moysoglou. Sherali-Adams gaps, flow-cover inequalities and generalized configurations for capacity-constrained Facility Location *CoRR*, abs/1312.0722, 2013.

[34] Stavros G. Kolliopoulos and Yannis Moysoglou. Sherali-Adams gaps, flow-cover inequalities and generalized configurations for capacity-constrained Facility Location. *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2014, Barcelona, Spain*, pages 297–312, 2014.

[35] Madhukar R. Korupolu, C. Greg Plaxton, and Rajmohan Rajaraman. Analysis of a local search heuristic for facility location problems. In *Proceedings of the 9th annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '98, pages 1–10, Philadelphia, PA, USA, 1998. Society for Industrial and Applied Mathematics.

[36] Jean B. Lasserre. An explicit exact SDP relaxation for nonlinear 0-1 programs. In *Proceedings of the 8th International IPCO Conference on Integer Programming and Combinatorial Optimization*, pages 293–303, London, UK, 2001.

[37] Monique Laurent. A Comparison of the Sherali-Adams, Lovász-Schrijver, and Lasserre relaxations for 0–1 Programming. *Mathematics of Operations Research*, 28(3):470–496, July 2003.

[38] Janny M. Y. Leung and Thomas L. Magnanti. Valid inequalities and facets of the capacitated plant location problem. *Mathematical Programming*, 44(1-3):271–291, 1989.

[39] Retsef Levi, David B. Shmoys, and Chaitanya Swamy. LP-based approximation algorithms for capacitated facility location. *Mathematical Programming*, 131(1-2):365–379, 2012. Preliminary version in Proc. IPCO 2004.

[40] Shi Li. A 1.488 approximation algorithm for the uncapacitated facility location problem. *Information and Computation*, 222:45–58, January 2013.

[41] Shi Li and Ola Svensson. Approximating k-median via pseudo-approximation. In Dan Boneh, Tim Roughgarden, and Joan Feigenbaum, editors, *Proc. 45th ACM Symposium on Theory of Computing*, STOC '13, pages 901–910. ACM, 2013.

[42] Lazlo Lovász and Alexander Schrijver. Cones of matrices and set-functions and 0-1 optimization. *SIAM Journal on Optimization*, 1:166–190, 1991.

[43] Mohammad Mahdian and Martin Pál. Universal facility location. In Giuseppe Battista and Uri Zwick, editors, *Algorithms - ESA 2003*, volume 2832 of *Lecture Notes in Computer Science*, pages 409–421. Springer Berlin Heidelberg, 2003.

[44] Mohammad Mahdian, Yinyu Ye, and Jiawei Zhang. Approximation algorithms for metric facility location problems. *SIAM Journal on Computing*, 36(2):411–432, August 2006.

[45] Éva Tardos Martin Pál and Tom Wexler. Facility location with nonuniform hard capacities. In *Proceedings of the 42nd IEEE symposium on Foundations of Computer Science*, FOCS '01, pages 329–, Washington, DC, USA, 2001. IEEE Computer Society.

[46] Claire Mathieu and Alistair Sinclair. Sherali-Adams relaxations of the matching polytope. In *Proceedings of the 41st annual ACM symposium on Theory of computing*, STOC '09, pages 293–302, New York, NY, USA, 2009. ACM.

[47] Ran Raz and Shmuel Safra. A sub-constant error-probability low-degree test, and a sub-constant error-probability PCP characterization of NP. In *Proceedings of the 29th Annual ACM Symposium on Theory of Computing*, STOC '97, pages 475–484, 1997.

[48] Thomas Rothvoß. The matching polytope has exponential extension complexity. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*, STOC '14, pages 263–272, 2014.

[49] Grant Schoenebeck, Luca Trevisan, and Madhur Tulsiani. Tight integrality gaps for Lovasz-Schrijver LP relaxations of Vertex Cover and Max Cut. In *Proceedings of the 39th annual ACM Symposium on Theory of Computing*, STOC '07, pages 302–310, New York, NY, USA, 2007. ACM.

[50] Hanif D. Sherali and Warren P. Adams. A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. *SIAM Journal of Discrete Mathematics*, 3(3):411–430, 1990.

[51] David B. Shmoys, Éva Tardos, and Karen I. Aardal. Approximation algorithms for facility location problems. In *Proceedings of the 29th Annual ACM Symposium on Theory of Computing*, STOC '97, pages 265–274, 1997.

[52] Ola Svensson. Santa Claus schedules jobs on unrelated machines. In *Proceedings of the 43rd annual ACM symposium on Theory of computing*, STOC '11, pages 617–626, New York, NY, USA, 2011. ACM.

[53] Madhur Tulsiani. Lovász-Schrijver reformulation. In J. J. Cochran, Jr. L. Anthony Cox, P. Keskinocak, J. P. Kharoufeh, and J. Cole Smith, editors, *Encyclopedia of Operations Research and Management Science*. John Wiley and Sons, 2011.

[54] Jens Vygen. Approximation algorithms for facility location problems (Lecture Notes). Report 05950-OR, Research Institute for Discrete Mathematics, University of Bonn, 2005. `URL: www.or.uni-bonn.de/~vygen/files/fl.pdf`.

[55] David P. Williamson and David B. Shmoys. *The Design of Approximation Algorithms*. Cambridge University Press, New York, NY, USA, 1st edition, 2011.

[56] Mihalis Yannakakis. Expressing combinatorial optimization problems by linear programs. *Journal of Computer and System Sciences*, 43(3):441 – 466, 1991.

[57] Jiawei Zhang, Bo Chen, and Yinyu Ye. A multi-exchange local search algorithm for the capacitated facility location problem. In *Proceedings of the 10th International IPCO Conference on Integer Programming and Combinatorial Optimization,* pages 219–233, 2004.