Sherali-Adams Gaps, Flow-cover Inequalities and Generalized Configurations for Capacity-constrained Facility Location*

Stavros G. Kolliopoulos and Yannis Moysoglou

Department of Informatics and Telecommunications
National and Kapodistrian University of Athens
Panepistimiopolis Ilissia, Athens 157 84, Greece
{sgk,gmoys}@di.uoa.gr

Abstract

Metric facility location is a well-studied problem for which linear programming methods have been used with great success in deriving approximation algorithms. The capacity-constrained generalizations, such as capacitated facility location (Cfl) and lower-bounded facility location (Lbfl), have proved notorious as far as LP-based approximation is concerned: while there are local-search-based constant-factor approximations, there is no known linear relaxation with constant integrality gap. According to Williamson and Shmoys devising a relaxation-based approximation for Cfl is among the top 10 open problems in approximation algorithms.

This paper advances significantly the state-of-the-art on the effectiveness of linear programming for capacity-constrained facility location through a host of impossibility results for both Cfl and Lbfl. We show that the relaxations obtained from the natural LP at $\Omega(n)$ levels of the Sherali-Adams hierarchy have an unbounded gap, partially answering an open question of [27, 6]. Here, n denotes the number of facilities in the instance. Building on the ideas for this result, we prove that the standard Cfl relaxation enriched with the generalized flow-cover valid inequalities [1] has also an unbounded gap. This disproves a long-standing conjecture of [25]. 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 and Lbfl through a sharp threshold phenomenon.

1998 ACM Subject Classification G.1.6 Optimization

Keywords and phrases Approximation Algorithms, Linear Programming, Facility Location

Digital Object Identifier 10.4230/LIPIcs.APPROX-RANDOM.2014.297

1 Introduction

Facility location is one of the most well-studied problems in combinatorial optimization. In the *uncapacitated* version *(*Ufl*)* we are given a set F of facilities and 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

* This research has been 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 Social Fund".

17th Int'l Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX'14) /
18th Int'l Workshop on Randomization and Computation (RANDOM'14).
Editors: Klaus Jansen, José Rolim, Nikhil Devanur, and Cristopher Moore; pp. 297–312
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

solution with cost at most ρ times the optimum. The approximability of general UFL is settled by an $O(\log |C|)$-approximation [18] which is asymptotically best possible, unless P = NP. 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 very 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., [33]). After a long series of papers the currently best approximation ratio for metric UFL is 1.488 [26], while the best known lower bound is 1.463, unless P = NP ([17] and Sviridenko [32]). In this paper we focus on two generalizations of metric UFL: the *capacitated facility location (*CFL*)* and the *lower-bounded facility location (*LBFL*)*.

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, or even proving a constant integrality gap for an efficient LP relaxation, are notorious open problems. Intriguingly, the following rare phenomenon occurs. The natural LP relaxations have an unbounded integrality gap and the only known $O(1)$-approximation algorithms are based on local search, with the currently best ratios being 5 [9] 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 approximation [25]. 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. According to Williamson and Shmoys [33] devising a relaxation-based algorithm for CFL is one of the top 10 open problems in approximation algorithms.

A lot of effort has been devoted to understanding the quality of relaxations obtained by an iterative lift-and-project procedure. Such procedures define hierarchies of successively stronger relaxations, where valid inequalities are added at each level. After at most n levels, where n 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. [8], Lovász and Schrijver [28] (for linear and semidefinite programs), Sherali and Adams [3], Lasserre [22] (for semidefinite programs). See [23] for a comparative discussion.

The seminal work of Arora et al. [7], 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 after 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 after a linear number of levels, one can take this as evidence that polynomially-sized relaxations are unlikely to yield approximations better than g (see also [29]). In fact, the former belief is now a theorem for maximum constraint satisfaction problems: in terms of approximation, LPs of size n^k, are exactly as powerful as $O(k)$-level Sherali-Adams relaxations [11].

LBFL is in a sense the opposite problem to CFL. In an LBFL instance every facility i comes with a lower bound b_i which is the minimum number of clients that must be assigned to i if we open it. In *uniform* LBFL all the lower bounds have the same value B. LBFL is even less well-understood than CFL. The first approximation algorithm for the uniform case had a performance guarantee of 448 [31], which has been improved to 82.6 [5]. Both use local search.

Apart from some work of the authors [21, 20] there has been no systematic theoretical study of the power of linear programming for approximating Cfl. In [21] we show an unbounded gap for Cfl at $\Omega(n)$ levels of the LS and the semidefinite mixed-LS$_+$ hierarchies, n being the number of facilities. In [20] we show that linear relaxations in 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 after one round. For related problems there are some recent interesting results. Improved approximations were given for k-median [27] and capacitated k-center [14, 6], 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 [27] 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. In [6] the authors raise as an important question to understand the power of lift-and-project methods for capacitated location problems, including whether they automatically capture relevant preprocessing steps.

Our results. We give impossibility results on arguably the most promising directions for strengthening linear relaxations for Cfl and Lbfl and in doing so we answer open problems from the literature. Our contribution is threefold.

First, we show that the LPs obtained from the natural relaxations for Cfl and Lbfl at $\Omega(n)$ levels of the SA hierarchy have an unbounded gap on an instance where $|F| = \Theta(n)$ and $|C| = \Theta(n^3)$. This result answers the questions of [27] and [6] stated above as far as the natural LP is concerned and moreover it is asymptotically tight. In the instances we consider clients have unit demands and it is well known that in this case the integer polytope and the mixed-integer (where fractional client assignments are allowed) polytope are the same. Since SA extends to mixed-integer programs as well [13, 8], the mixed-integer polytope is obtained after at most n levels. Thus at most that many levels are needed also by the stronger, full-integer, SA procedure we employ, which in the lifting stage multiplies also with assignment variables. From a qualitative aspect, we give the first, to our knowledge, SA 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. The lifting of the assignment variables raises obstacles in the proof that we managed to overcome as discussed in Section 3.

We use the *local-to-global* method which was implicit in [7] for local-constraint relaxations and was then extended to the SA hierarchy in [15]. See also [16] for an explicit description and [12] 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 Theorem 4 for Cfl we devise first in Lemma 3 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, explained in Propositions 3.1–3.3, produces the final set of distributions. Theorem 4 extends significantly our earlier result on the LS hierarchy for Cfl [21] to the stronger SA hierarchy. It turns out that in both cases we can start from the same bad instance. It should be noted that the methodology in the two proofs is completely different – in [21] the result was obtained via an inductive construction of protection matrices.

Our second contribution (cf. Theorem 6) is that the *effective capacity* 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 [25] 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 can be extended to a solution s^π that is a convex combination of feasible integer solutions. Our ideas can be extended to even more general families such as the *submodular inequalities* [1], cf. Theorem 7. All results in this paper make no time-complexity assumptions. To our knowledge no efficient separation algorithm for the effective capacity inequalities is known.

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 [10] for the Santa Claus problem and has yielded valuable insights, mostly for resource allocation and scheduling problems (e.g., [30]). The analogue of the Configuration LP for facility location already exists, it is the *star relaxation* (see, e.g., [19]). 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 12 we characterize the behavior of proper relaxations for Cfl and Lbfl through a threshold result: anything less than maximum complexity results in unboundedness of the integrality gap, while there are proper relaxations of maximum complexity with a gap of 1.

Our results disqualify the so far most promising approaches for an efficient LP relaxation for Cfl. Moreover, we advance drastically the state-of-the-art for the little understood Lbfl. Whether a fundamentally new approach may succeed for either problem remains as an open question.

2 Preliminaries

Given an instance $I(F,C)$ of Cfl or Lbfl, we use n, m to denote $|F|$ and $|C|$ respectively. We will show our negative results for uniform, integer, capacities and lower bounds. 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.

$$\min\{\textstyle\sum_{i\in F} f_i y_i + \sum_{i\in F}\sum_{j\in C} x_{ij}c_{ij} \;\; | \quad x_{ij} \le y_i \;\; \forall i\in F, \forall j \in C,$$
$$\textstyle\sum_{i\in F} x_{ij} = 1 \;\; \forall j\in C, \qquad y_i, x_{ij} \in \{0,1\} \;\; \forall i \in F, \forall j \in C\}$$

The linear relaxation results from the above IP by replacing the integrality constraints with: $0 \le y_i \le 1,\ 0 \le x_{ij} \le 1,\ \forall i \in F, \forall j \in C$. To obtain the standard LP relaxations for uniform Cfl (and Lbfl) with capacity U (lower bound B) the following constraints are added respectively:

$$\textstyle\sum_j x_{ij} \le U y_i \;\; \forall i \in F \quad \text{and} \quad \sum_j x_{ij} \ge B y_i \;\; \forall i \in F.$$

We will slightly abuse terminology by using the term *(LP-classic)* for both LPs. It will be clear from the context to which problem, Cfl or Lbfl, we refer.

We proceed to define the Sherali-Adams hierarchy [3]. Consider a polytope $P \subseteq \mathbb{R}^d$ defined by the linear constraints $Ax - b \leq 0$, $0 \leq x_i \leq 1$, $i = 1, \ldots, d$. 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 $U \subseteq \{x_i \mid i = 1, \ldots, d\}$ such that $|U| \leq k$, and for every $W \subseteq U$, consider the valid constraint: $\pi(x) \prod_{x_i \in U-W} x_i \prod_{x_i \in W}(1 - x_i) \leq 0$. Linearize the system obtained this way by replacing (i) x_i^2 with x_i for all i and (ii) $\prod_{x_i \in I} x_i$ with x_I for each set $I \subseteq \{x_i | i = 1, \ldots, d\}$. $\text{SA}^k(P)$ is the projection of the resulting linear system onto the singleton variables. 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 $\min\{c^T x \mid x \in \text{SA}^k(P)\}$.

3 Sherali-Adams Gap for CFL

Consider an instance of metric CFL with a total of $2n$ facilities, n with opening cost 0 which we call cheap (and denote the corresponding set by $Cheap$) and n with opening cost 1 which we call costly (and denote by $Costly$). The capacity $U = n^3$ and we have a total of $nU + 1$ clients. All connection costs are 0. We will show that the following bad solution s to the instance[1] survives a number of SA levels, which is linear in the number $2n$ of facilities, more specifically for $n/10$ levels. On the other hand, it is known that at level $2n$ the relaxation obtained expresses the integral polytope. Let $\alpha = n^{-2}$. For all $i \in Cheap$ and for all $j \in C$, $y_i = 1$ and $x_{ij} = \frac{1-\alpha}{n}$, and for all $i \in Costly$ and for all $j \in C$ $y_i = \frac{10}{n^2}$ and $x_{ij} = \frac{\alpha}{n}$. Theorem 4 below indicates that, as often with hierarchies, simple valid inequalities are generated after many rounds. The reader who is further interested in the robustness of SA for CFL may consult Section 3.2.

The following lemma, which is implicit in previous work [15, 16] gives sufficient conditions for a solution to be feasible at level k of the SA hierarchy.

▶ **Lemma 1** ([15, 16]). *Let s be a feasible solution to the relaxation and let $v(\pi, z)$ be the set of variables appearing in a lifted constraint obtained from π multiplied by z. Solution s survives k levels of SA if for every constraint π and each multiplier z with at most k distinct variables there is:*

1. *A solution $s' = s_{\pi,z}$ which agrees with s on $v(\pi, z)$ such that s' is a convex combination E_d of integer solutions (and thus E_d defines a distribution on integer solutions) and*
2. *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 lifted constraints obtained from π_1 and π_2 multiplied 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.*

First consider a constraint π: $\sum_j x_{i^\pi j} \leq U y_{i^\pi}$ and a multiplier z. 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 letting all other variables the same as in the original bad solution s. We say that facility 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 constraint $\sum_{i \in Costly} y_i = 1$. While $s_{\pi,z}$ can be obtained as a convex combination

[1] The reader should notice that any similarity with Knapsack is superficial. Theorem 4 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.

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 $p = z_1 z_2 \dots z_l$ as an abbreviation of the event $\mathcal{E}_p := \bigwedge_{i=1}^{l}(z_i = 1)$.

▶ **Definition 2.** Let i_b be the facility that takes the blame. We say that a distribution E_d is *assignment-symmetric* if the following are true:

1. $P_{E_d}[x_{i_{a_1} j_{b_1}} \dots x_{i_{a_t} j_{b_t}} y_{i_{a_{t+1}}} \dots y_{i_{a_l}}]$, with $t + l \le k + 1$ is the same if we exchange all occurrences of cheap facility i_r by cheap facility $i_{r'}$ (in other words relabeling facilities). Note that we allow repetitions of facilities and clients in the description of the event.
2. $P_{E_d}[x_{i_{a_1} j_{b_1}} \dots x_{i_{a_t} j_{b_t}} y_{i_{a_{t+1}}} \dots y_{i_{a_l}}]$ is the same if we exchange all occurrences of client j_q by client $j_{q'}$.
3. $P_{E_d}[x_{i_{a_1} j_{b_1}} \dots x_{i_{a_t} j_{b_t}} y_{i_{a_{t+1}}} \dots y_{i_{a_l}}]$ is the same if we exchange all occurrences of costly facility i_1 by costly facility i_2, $i_1, i_2 \neq i_b$.

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

▶ **Lemma 3.** *Solution $s_{\pi,z}$ is a convex combination E_d of integer solutions which defines an assignment-symmetric distribution.*

Proof. We describe a probabilistic experiment which induces an assignment-symmetric distribution E_d over integer solutions satisfying $\sum_{i \in Costly} y_i = 1$.

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

Pick costly facility i_c with probability y'_{i_c}. If $i_c = i_b$ ($i_c \neq i_b$) 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 $w^1_{i_b}$ (w^2_{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. Moreover the induced distribution E_d is assignment-symmetric and the expected (y, x) vector with respect to E_d is solution $s_{\pi,z}$.

Clearly, $s_{\pi,z}$ is the convex combination induced by E_d and E_d is assignment-symmetric: 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 $y'_{i_{co}} w_{co}$ 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 $s_{\pi,z}$.

As for the assignments to the cheap facilities, observe that in every outcome of the experiment the demand not assigned to costly facilities is exactly the demand assigned to cheap. Since we have proved that the expected assignments to the costly facilities are those of the bad solution, 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 add up to 1 by the constraints of the LP). By the symmetric way the cheap are handled in the experiment we have that the total expected demand assigned to each $i \in Cheap$ is $\sum_j x_{ij}$ and by the symmetric way the clients are assigned to i through the experiment we get that the expected assignment of each j to i is x_{ij}.

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

We set the product-variables x_I appearing in constraint π multiplied by multiplier z to $P_{E_d}[I]$. 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.

The second and more challenging case is when constraint π is $\sum_i x_{ij^\pi} = 1$ for some client j^π. Let again z be a multiplier of level k. Observe now that all facilities in F appear in $v(\pi, z)$ as indexes of at least the x_{ij} variables. We select one facility i_b not appearing in z to *take the blame*. Let $s_{\pi,z} = (y', x')$ be the corresponding extended solution that can be written as a convex combination/assignment–symmetric distribution E_d of integer solutions; the existence of E_d is ensured by Lemma 3. In this case there is a major obstacle to the agreement of the products x_I: conditioning on the event $x_{i_b j}$ the probability of an event $x_{ij'}, i \in Cheap$ for some $j' \neq j$ is higher than it would be if we were to condition on the event $x_{i'j}, i' \in Costly - \{i_b\}$. 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. This can be problematic since facility i_b takes the blame in some distributions but does not in some others and thus there is the danger of violating the consistency required by the 2nd condition of Lemma 1. We overcome this difficulty by making alterations to E_d and constructing a distribution E_f where the probabilities of the aforementioned events are the same.

We now devise the altered distribution E_f. We first display the intuition in 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 call *transfusion of probability*. We pick some measure of an integer solution s_1 for which $x_{i_{ch}j'} = 1 \wedge x_{i_b j} = 1 \wedge x_{i_b j''} = 0$ for some client j''. We pick the same quantity 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. Let that quantity 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 E_f now has $P[A] = P[B]$. 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.

Now let p be a product appearing in constraint π after having multiplied by multiplier z. We only consider products where exactly one variable $x_{i_b j}$ appears. Recall 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 a product p. We may also assume that there is no y_i variable in p, since if there is for some $i \in \mathit{Costly} - \{i_b\}$ the probability of $\mathcal{E}_p$ is simply 0 and if $i \in \mathit{Cheap}$ the we can ignore the effect of $y_i = 1$ since it is always true. Likewise we assume that there is no assignment variable of another costly facility. We shall make corrections of the probability of all such events $\mathcal{E}_p$ in a top-down manner: at step i 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. In other words, we fix the probabilities in decreasing order of the cardinality of the set of variables appearing in p. The following proposition relates the probability of $\mathcal{E}_p$ with that of $\mathcal{E}_{p'} = \mathcal{E}_{p x_{ij}}$, an event with the additional requirement that $x_{ij} = 1$.

▶ **Proposition 3.1.** *Let $p = x_{i_b j} x_{i_{a_1} j_{b_1}} x_{i_{a_2} j_{b_2}} \ldots x_{i_{a_l} j_{b_l}}$ and let $p' = p x_{i_{a_{l+1}} j_{b_{l+1}}}$. Then in E_d, $(1 - o(1))P[\mathcal{E}_p]/n \le P[\mathcal{E}_{p'}] \le (1 + o(1))P[\mathcal{E}_p]/n$.*

Consider step i of the above iterative construction of E_f. Let $p = x_{i_b j} x_{i_{a_1} j_{b_1}} \ldots x_{i_{a_{k-i+1}} j_{b_{k-i+1}}}$ and the event $\mathcal{E}_p$: $x_{i_b j} = 1 \wedge x_{i_{a_1} j_{b_1}} = 1 \wedge \ldots \wedge x_{i_{a_{k-i+1}} j_{b_{k-i+1}}} = 1$. We wish in E_f the probability $P[\mathcal{E}_p]$ to be equal to $P[\mathcal{E}_{p/fixed}] = P[x_{i^* j} = 1 \wedge x_{i_{a_1} j_{b_1}} = 1 \wedge \ldots \wedge x_{i_{a_{k-i+1}} j_{b_{k-i+1}}} = 1]$ in E_d for $i^* \in \mathit{Costly} - \{i_b\}$. We bound the ratio $\frac{P[\mathcal{E}_p]}{P[\mathcal{E}_{p/fixed}]}$:

▶ **Proposition 3.2.** *Let $\mathcal{E}_p$ and $\mathcal{E}_{p/fixed}$ be defined as above. Then* $(1 + (1 - o(1))1/n)^{k-i+1} \le \frac{P[\mathcal{E}_p]}{P[\mathcal{E}_{p/fixed}]} \le (1 + (1 + o(1))1/n)^{k-i+1}$.

Now we describe in detail the alterations of the probabilities in each iteration. The corrections of the probabilities of events of previous iterations affect the probabilities of the events of the current iteration of the procedure that constructs E_f. We bound this effect on the probability of an event $\mathcal{E}_p$ of the current iteration i by considering the corrections 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 iteration 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. The amount of the effect of the correction of the previous iteration is by Proposition 3.2 at most $i((1+(1+o(1))1/n)^{k-i+2} - 1)P[\mathcal{E}_{p'/fixed}]$ while the measure of the needed correction for $\mathcal{E}_p$ is at least $((1+(1-o(1))1/n)^{k-i+1} - 1)P[\mathcal{E}_{p/fixed}]$ which by Proposition 3.1 and by the number of rounds we consider is higher, in particular $((1 + (1 - o(1))1/n)^{k-i+1} - 1)P[\mathcal{E}_{p/fixed}] \ge n(1-o(1))((1+(1-o(1))1/n)^{k-i+1}-1)P[\mathcal{E}_{p'/fixed}] > i((1+(1+o(1))1/n)^{k-i+2}-1)P[\mathcal{E}_{p'/fixed}]$. To subtract from $P[\mathcal{E}_p]$ the rest of the probabilistic measure required from the correction at

[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.

step i, say a measure of μ, we do the following transfusion step: pick a measure μ of solutions from distribution E_d 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 of solutions from E_d such that $x_{i_b j} = 1$, $x_{i_b j'} = 0$, all the other events of $\mathcal{E}_p$ are 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. The resulting distribution has the probability of $\mathcal{E}_p$ fixed to the desired value and moreover, by the choice of the sets of solutions on which we perform the transfusion step, the probability of the events fixed in previous iterations was not altered and neither was the probability of events containing only assignments of cheap facilities. Clearly, the solution $s_{\pi,z}$ is still obtained in expectation. It remains to show that the transfusion step can be performed, i.e., that there is enough measure μ in the involved sets of integer solutions.

▶ **Proposition 3.3.** *The probabilistic transfusion step of the above iterative procedure can always be performed.*

Proof. The intuition behind the proof is that the "donor" event that supplies the required measure is 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 at most $(e^2 - 1)P[\mathcal{E}_{p/fixed}] = \Theta(1/n^i)$. 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, so we can always pick the required amount of measure. ◀

▶ **Theorem 4.** *There is a family of* CFL *instances with* $2n$ *facilities and* $n^4 + 1$ *clients such that 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 lifted 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 1 is satisfied. For the second condition, 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. 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 iterative procedure of probabilistic transfusion, 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 1 holds. It is easy to see that bad solution has cost $\Theta(n^{-1})$ while any feasible solution to the instance has cost $\Omega(1)$. ◀

3.1 SA Gap for LBFL

A similar result to Theorem 4 can be proved for LBFL. Consider an instance with n facilities, lower bound $B = n^3$ and a total of $n(B-1)$ clients. The metric space here is more intriguing than the one for the CFL case. Consider a regular $(n-1)$-dimensional simplex with edge length 1. On each of the n vertices of the simplex a facility along with some $B-1$ clients are located. All opening costs are 0. Clearly every integer solution has a cost of at least $B-1$ since we can open at most $n-1$ of the facilities, and so at least $B-1$ clients will have to be assigned to some facility other than the one on the same vertex. We call a client j that is located on the same vertex with facility i, *exclusive* client of i. We denote by $Exclusive(i)$ the set of clients that are exclusive to facility i. On the other hand we can show that the following bad solution s is feasible at $\Omega(n)$ levels of the SA hierarchy. For all $i \in F$, $y_i = 1 - n^{-2}$; for a client $j \in C$, $x_{ij} = 1 - 10n^{-2}$, if $j \in Exclusive(i)$, and $x_{ij} = \frac{10n^{-2}}{n-1}$ for all other facilities. Solution s incurs a cost of $o(B)$.

▶ **Theorem 5.** *There is a family of* LBFL *instances with n facilities and $n^4 - n$ clients such that the relaxations obtained from (LP-classic) at $\Omega(n)$ levels of the Sherali-Adams hierarchy have an integrality gap of $\Omega(n)$.*

The proof is similar to that of CFL and is thus omitted. Here the reader can find a sketch of the necessary changes to the proof of Theorem 4.

Proof sketch of Theorem 5. Consider a constraint $\pi : \sum_j x_{i^\pi j} \geq B y_{i^\pi}$ and a multiplier z at level k and let $v(\pi, z)$ be the set of variables appearing in the multiplied constraint. We pick a facility i_b not in $v(\pi, z)$ to take the blame. We construct a solution s' where we set $y'_{i_b} = n - 1 - \sum_{i \neq i_b} y_i$ and for each $j \in Exclusive(i_b)$ we set $x'_{i_b j} = y'_{i_b} = \frac{1-1/n}{n}$ and we distribute the remaining demand that was assigned to i_b to each facility from a constant-size set I_b of facilities not appearing in $v(\pi, z)$. Solution s' can be obtained as a convex combination of integer solutions by constructing a distribution similarly to Lemma 3. This time the distribution satisfies that exactly $n-1$ facilities are opened in each outcome of the experiment. Note that we do not require the underlying distribution to be assignment symmetric, because facilities have to treat differently their exclusive clients. We set the values of the linearized products appearing in the multiplied constraint equal to the probability of the corresponding events with respect to the aforementioned distribution. No product involving variables of $i_b \cup I_b$ appear in the constraint. For constraints $0 \leq x_{ij}, y_i \leq 1$ and $x_{ij} \leq y_i$ the construction of the distribution is the same. The distributions constructed so far are locally consistent as required by Lemma 1.

The case where the constraint is $\pi : \sum_i x_{ij^\pi} = 1$ is once again more complicated. We choose a facility $i_b \notin z$ and moreover $j^\pi \notin Exclusive(i_b)$ to take the blame and the set I_b is defined as before except we also require that j^π is not exclusive to any of them. Solution s' is constructed like in the previous case. All products take the value of the corresponding events in the distribution except those in which the unique variable involving i_b appears, namely $x_{i_b j}$ and those involving facilities in I_b. We perform a transfusion step so that the probabilities of all the events whose corresponding products appear in the lifted constraint become consistent with the distributions of the previous case: this time we need to fix the probabilities of the events involving facility i_b or some facility $i \in I_b$. ◀

3.2 Robustness of the SA Gap

In this section we explain how adding simple valid inequalities does not affect our arguments on the SA hierarchy.

As an example we address the valid inequality $\sum_i y_i \geq \lceil D/U \rceil$, where D is the total amount of demand. This is a well-known facet-inducing constraint for our instance, see, e.g., [24, p. 283]. 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 is impossible without some yet unknown form of preprocessing. In fact part of the motivation behind Theorem 4 is to demonstrate that the SA hierarcy is inadequate for such preprocessing purposes.

We modify the family of "bad" instances by using the same trick we used in the proof of Theorem 6: we have n cheap and n costly facilities and $Un + 1$ clients, and the bad solution in which for every $ch \in Cheap$, $co \in Costly$, and client j, $y_{ch} = 1, x_{chj} = \frac{1-\alpha}{n}$, $y_{co} = 10/n^2, x_{coj} = \frac{\alpha}{n}$ with $\alpha = n^{-2}$, and additionally we add a set of n *dummy* facilities a_i, $1 \leq i \leq n$, all with 0 opening costs, on the same point at distance 1 from the rest. In the bad solution s we additionally set $y_{a_i} = 1$ and $x_{a_ij} = 0$ for all i and for all clients j. The inequality is obviously satisfied.

In the design of the locally consistent distributions, now we must give a distribution for the case where the constraint π is the new one $\sum_i y_i \geq \lceil D/U \rceil$, and verify that the "visible" part of the distribution agrees with the visible part of all other distributions of the proof. In this case there must be some dummy facility a_d not appearing as an index in the multiplier z of the constraint (although its y variable does appear in π). Additionally there must be a costly facility i' for which the assignments of clients to i' do not appear in $v(\pi, z)$ – this is ensured by the number of rounds we consider. We modify the solution (y, x) to obtain (y', x') where the facilities i' and a_d exchange the values of their corresponding assignments. We define now the random experiment similarly to the proof of Lemma 3 with facility a_d taking the blame. The only difference is that while a_d is opened 100% of the time, it is not assigned any demand when a costly facility other than i' is opened. In the terminology of Theorem 6 that follows, a_d is always open but it is inactive when some $i \in Costly$, $i \neq i'$, is opened. It is easy to see that the distribution obtained is consistent with all the other distributions defined for this modified instance, as required by Lemma 1.

4 Fooling the Effective Capacity Inequalities for CFL

In this section we show that the (LP-classic) for CFL with the addition of the effective capacity inequalities proposed in [1] has unbounded gap.

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].

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

The proof of the following theorem uses some of the ideas we introduced earlier for Theorem 4.

▶ **Theorem 6.** *The integrality gap of the relaxation obtained from (LP-classic) with the addition of the effective capacity inequalities is* $\Omega(n)$*, where* n *is the number of facilities in the instance.*

Proof. Consider an instance with n cheap and $n+2$ costly facilities and $Un+1$ clients, $U = n^3$. Define the bad solution s, similarly to Section 3, s.t. for every $ch \in Cheap$, $co \in Costly$, and client j, $y_{ch} = 1, x_{chj} = \frac{1-\alpha}{n}$, $y_{co} = 10/n^2, x_{coj} = \frac{\alpha}{n+2}$. Recall that $\alpha = n^{-2}$. We add a set of $n+2$ facilities a_i, $1 \le i \le n+2$, all with 0 opening costs, on the same point at distance 1 from the rest (an instance of the so-called *facility location on a line*). In the bad solution s we additionally set $y_{a_i} = 1$ and $x_{a_i j} = 0$ for all i and for all clients j.

We will prove that in every cover I with respect to some client set J and to the J_i client sets for each i, there must always be a number of at least $2n^3$ clients whose assignment variables to some costly and to some a_i do not appear in the constraint. This is because if, $\bar{u}_i = U$ for each $i \in Costly$, or, $\bar{u}_{a_i} = U$ for each $i \le n+2$, then the excess capacity $\lambda > U$ since $d(J) \le Un+1$. This contradicts the requirement that $\lambda < U$. So there must be a costly facility $i_{co'}$ and some facility $a_{i'}$ such that for the corresponding sets we have $|J_{i_{co'}}|, |J_{a_{i'}}| < U$, and so there is a set J^* of $2n^3$ clients whose assignments to those two facilities do not appear in the constraint. We exchange the values of $x_{i_{co'}j}$ and $x_{a_{i'}j}$ for all $j \in J^*$, leaving everything else the same, and we obtain a solution $s' = (y', x')$. We can prove similarly to the proof of Lemma 3 that s' is a convex combination of integer solutions and thus solution s satisfies the inequality since the parts of s and s' visible to that inequality are the same.

We modify the construction of Lemma 3 in the following way: facility $a_{i'}$ is opened 100% of the time but is active $1 - \sum_{i\in Costly} y'_i$ of the time, when none of the costly facilities are opened. When it is not active, the capacity of its corresponding bin is 0. When a costly other than $i_{co'}$ is opened the experiment is the same as in Lemma 3. If costly facility $i_{co'}$ is opened the capacity of the corresponding bin is $w^2_{co'} = \frac{\sum_j x'_{co'j}}{y'_{i_{co'}}}$ and the capacity of the cheap is $\frac{|C|-w^2_{co'}}{n}$. We randomly select some $w^2_{co'}$ clients that do not belong to J^* to be tossed in the bin of $i_{co'}$; we randomly distribute the balls corresponding to the remaining clients to the slots of the cheap facilities. When $a_{i'}$ is active, and thus no costly facility is opened, the capacity of the corresponding bin is $w^1_{a_{i'}} = \frac{\sum_j x'_{a_{i'}j}}{1-\sum_{i\in Costly} y'_i}$ and the capacity of the cheap is $\frac{|C|-w^1_{a_{i'}}}{n}$. We select randomly some $w^1_{a_{i'}}$ clients in J^* and we toss the corresponding balls in the bin of $a_{i'}$. We randomly toss the remaining balls to the slots of the bins of the cheap facilities.

Note that the above experiment induces a distribution over feasible integer solutions since all the defined bin capacities are less than U (this is by the choice of the size of J^*) and every client is assigned to exactly one opened facility in each outcome. We do not need this distribution to be assignment-symmetric. Observe that the expected vector with respect to the latter distribution is solution s'. Finally, note that we once again treated the capacities w of the bins as being integral. For fractional bin capacities (which is actually always the case for the defined w's) we can define the experiment in a similar way to the proof of Lemma 3. ◀

The submodular inequalities introduced in [1] are even stronger than the effective capacity inequalities. We limit our discussion to uniform CFL where all clients have 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 [1]. The name reflects the fact that the function $f(I)$ is submodular.

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

Theorem 7 below strictly generalizes Theorem 6 to the submodular inequalities.

▶ **Theorem 7.** *The integrality gap of the relaxation obtained from (LP-classic) with the addition of the submodular inequalities is $\Omega(n)$, where n is the number of facilities in the instance.*

5 Proper Relaxations

In this section we present the family of proper relaxations and characterize their strength. Consider a 0-1 (y, x) vector on the set of variables of the classic relaxation (LP-classic) such that $y_i \geq x_{ij}$ for all $i \in F, j \in C$. The meaning of $y_i = 1$ is the usual one that we open facility i. Likewise, the meaning of $x_{ij} = 1$ is that we assign client j to facility i. We call such a vector a *class*. Note that the definition is quite general and a class can be defined from any such (y, x), which may or may not have a relationship to a feasible integer solution. We denote the vector corresponding to a class cl as $(y, x)_{cl}$. We associate with class cl the *cost of the class* $c_{cl} = \sum_{i | y_i = 1 \in (y,x)_{cl}} f_i + \sum_{i,j | x_{ij} = 1 \in (y,x)_{cl}} c_{ij}$. Let the *assignments of class cl* be defined as $Agn_{cl} = \{(i, j) \in F \times C \mid x_{ij} = 1 \text{ in } (y, x)_{cl}\}$. We say that cl *contains* facility i, if the corresponding entry y_i in the vector $(y, x)_{cl}$ equals 1. The set of facilities contained in cl is denoted by $F(cl)$.

▶ **Definition 8** (Constellation LPs)**.** Let $\mathcal{C}$ be a set of classes defined for an instance $I(F, C)$ of CFL or LBFL. Let x_{cl} be a variable associated with class $cl \in \mathcal{C}$. The *constellation LP with class set* $\mathcal{C}$, denoted LP($\mathcal{C}$), is defined as $\min\{\sum_{cl \in \mathcal{C}} c_{cl} x_{cl} \mid \sum_{cl | \exists i : (i,j) \in Agn_{cl}} x_{cl} = 1 \ \forall j \in C, \quad \sum_{cl | i \in F(cl)} x_{cl} \leq 1 \ \forall i \in F, \quad x_{cl} \geq 0 \ \forall cl \in \mathcal{C}\}$.

We refer simply to a *constellation LP* when $\mathcal{C}$ is implied from the context. We define the *projection* $s' = (y^{s'}, x^{s'})$ of solution $s = (x^s_{cl})_{cl \in \mathcal{C}}$ of LP($\mathcal{C}$) to the facility opening and assignment variables (y, x) as $y^{s'}_i = \sum_{cl | i \in cl} x^s_{cl}$ and $x^{s'}_{ij} = \sum_{cl | (i,j) \in Agn_{cl}} x^s_{cl}$. We restrict our attention to constellation LPs that satisfy a symmetry property that is very natural for uniform capacities and unit demands.

▶ **Definition 9** (P_1: Symmetry)**.** We say that property P_1 holds for the constellation linear program LP($\mathcal{C}$) if for every class $cl \in \mathcal{C}$, all classes resulting from a permutation that relabels the facilities and/or the clients of cl are also in $\mathcal{C}$.

▶ **Definition 10** (Proper Relaxations)**.** We call *proper relaxation* for CFL (LBFL) 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., [19]) where $\mathcal{C}$ corresponds to the set of all *stars*: a facility and a set of at most U (or at least B for LBFL) 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 and LBFL. In order to characterize the strength of a proper LP we need the notion of complexity.

▶ **Definition 11** (Complexity of proper relaxations)**.** Given an instance $I(F, C)$ of CFL (LBFL) let F' be a maximum-cardinality set of open facilities in an integral feasible solution. The *complexity* α of a proper relaxation $LP(\mathcal{C})$ for I is defined as the $\sup_{cl \in \mathcal{C}}(|F(cl)|/|F'|)$.

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 take each 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.

▶ **Theorem 12.** *Every proper relaxation for uniform* CFL *(*LBFL*) with complexity* $\alpha < 1$ *has an unbounded integrality gap. There is a proper relaxation for* CFL *(*LBFL*) of complexity* 1 *whose projection to* (y, x) *expresses the integral polytope.*

Proof sketch of Theorem 12 for LBFL**.** We are given an arbitrary proper relaxation $LP(\mathcal{C})$ of complexity $\alpha < 1$, for an instance with $n + 1$ facilities, n^3 clients and $B = n^2$, and the following metric distances: put every facility i, $i \le n - 1$, together with $B - 1$ clients, which we call *exclusive* clients of i, on a distinct vertex of an $(n - 2)$-dimensional regular simplex in $\mathbb{R}^{n-2}$ with edge length D. Put facilities $n, n + 1$ together with their exclusive clients, which are all the $B + n - 1$ remaining clients, to a point far away from the simplex, so the minimum distance from a vertex is $D' = \Omega(nD)$. We set all the facility costs to 0.

A major challenge is that we have no a priori knowledge of $\mathcal{C}$. We use the validity of $LP(\mathcal{C})$ and the fact that $\alpha < 1$, to prove that there is a class cl_0 with some desired properties that must belong to $\mathcal{C}$. Using classes that are symmetric to cl_0, which also must belong to $\mathcal{C}$, we construct a vector s that is feasible for $LP(\mathcal{C})$ and whose projection on the classic variables is the following (y^*, x^*): for each facility $i \le n - 1$, its exclusive clients are assigned to it with a fraction of $\frac{n^2-1}{n^2}$ each, while they are assigned with a fraction of $\frac{1}{(n^2)(n-2)}$ to each other facility $i' \le n - 1$. As for facilities $n, n + 1$, all of their exclusive clients are assigned with a fraction of $1/2$ to each. Moreover $y^*_i = \frac{n^2-1}{n^2}$, for $i \le n - 1$, and $y^*_n = y^*_{n+1} = \frac{n^2+n-1}{2n^2}$.

The cost of the fractional solution we constructed is $O(nD)$ due to the assignments of exclusive clients of facility i, $i \le n - 1$, to facilities i' with $i' \ne i$, $i' \le n - 1$. As for the cost of an arbitrary integral solution, observe that since the $n^2 + n - 1$ exclusive clients of $n, n + 1$ are very far from the rest of the facilities, using n of them to satisfy some demand of those facilities and help to open all of them, incurs a cost of $\Omega(nD') = \Omega(n^2 D)$. On the other hand, if we do not open all of the $n - 1$ facilities on the vertices of the simplex (since they have in total $(n - 1)(B - 1)$ exclusive clients which is not enough to open all of them), there must be at least one such facility not opened in the solution, thus its $B - 1 = \Theta(n^2)$ exclusive clients must be assigned elsewhere, incurring a cost of $\Omega(n^2 D)$. ◀

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 Warren P. Adams and Hanif D. Sherali. Linearization strategies for a class of zero-one mixed integer programming problems. *Oper. Res.*, 38(2):217–226, April 1990.

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. *Math. Program.*, 141(1-2):527–547, 2013.

5 Sara Ahmadian and Chaitanya Swamy. Improved approximation guarantees for lower-bounded facility location. In *Proceedings of the 10th International Workshop on Approximation and Online Algorithms* , WAOA'12, pages 257–271, Ljubljana, Slovenia, 2012.

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

7 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.

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

9 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.

10 Nikhil Bansal and Maxim Sviridenko. The Santa Claus problem. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, STOC'06, pages 31–40, New York, NY, USA, 2006. ACM.

11 Siu On Chan, James R. Lee, Prasad Raghavendra, and David Steurer. Approximate constraint satisfaction requires large lp relaxations. *CoRR*, abs/1309.0563, 2013.

12 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.

13 Gérard Cornuéjols. Valid inequalities for mixed integer linear programs. *Math. Program.*, 112(1):3–44, 2008.

14 Marek Cygan, MohammadTaghi Hajiaghayi, and Samir Khuller. LP rounding for k-centers with non-uniform hard capacities. In *FOCS*, pages 273–282, 2012.

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

16 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.

17 S. Guha and S. Khuller. Greedy strikes back: improved facility location algorithms. *Journal of Algorithms*, 31:228–248, 1999.

18 D. S. Hochbaum. Heuristics for the fixed cost median problem. *Mathematical Programming*, 22:148–162, 1982.

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

20 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.

21 Stavros G. Kolliopoulos and Yannis Moysoglou. Tight bounds on the Lovász-Schrijver rank for approximate capacitated facility location. *Manuscript*, 2013.

22 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, UK, 2001. Springer-Verlag.

23 Monique Laurent. A Comparison of the Sherali-Adams, Lovász-Schrijver, and Lasserre relaxations for 0–1 Programming. *Math. Oper. Res.*, 28(3):470–496, July 2003.

24 J. M. Y. Leung and T. L. Magnanti. Valid inequalities and facets of the capacitated plant location problem. *Math. Program.*, 44(1-3):271–291, 1989.

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

26 Shi Li. A 1.488 approximation algorithm for the uncapacitated facility location problem. *Inf. Comput.*, 222:45–58, January 2013.

27 Shi Li and Ola Svensson. Approximating k-median via pseudo-approximation. In Dan Boneh, Tim Roughgarden, and Joan Feigenbaum, editors, *Proc. 45th STOC*, pages 901–910. ACM, 2013.

28 L. Lovász and A. Schrijver. Cones of matrices and set-functions and 0-1 optimization. *SIAM Journal on Optimization*, 1:166–190, 1991.

29 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 thirty-ninth annual ACM symposium on Theory of computing*, STOC'07, pages 302–310, New York, NY, USA, 2007. ACM.

30 Ola Svensson. Santa claus schedules jobs on unrelated machines. *SIAM J. Comput.*, 41(5):1318–1341, 2012.

31 Z. Svitkina. Lower-bounded facility location. In *Proceedings of the 19th ACM-SIAM Symposium on Discrete Algorithms*, pages 1154–1163, 2008.

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

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