

On the Power of Static Assignment Policies for Robust Facility Location Problems

Omar El Housni¹, Vineet Goyal², and David Shmoys³

¹ ORIE, Cornell Tech, New York, USA
oe46@cornell.edu

² IEOR, Columbia University, New York, USA
vg2277@columbia.edu

³ ORIE, Cornell University, Ithaca, USA
david.shmoys@cornell.edu

Abstract. We consider a two-stage robust facility location problem on a metric under an uncertain demand. The decision-maker needs to decide on the (integral) units of supply for each facility in the first stage to satisfy an uncertain second-stage demand, such that the sum of first stage supply cost and the worst-case cost of satisfying the second-stage demand over all scenarios is minimized. The second-stage decisions are only assignment decisions without the possibility of adding recourse supply capacity. This makes our model different from existing work on two-stage robust facility location and set covering problems. We consider an implicit model of uncertainty with an exponential number of demand scenarios specified by an upper bound k on the number of second-stage clients. In an optimal solution, the second-stage assignment decisions depend on the scenario; surprisingly, we show that restricting to a fixed (static) fractional assignment for each potential client irrespective of the scenario gives us an $O(\log k / \log \log k)$ -approximation for the problem. Moreover, the best such static assignment can be computed efficiently giving us the desired guarantee.

1 Introduction

We consider two-stage robust facility location problems under demand uncertainty where we are given a set of clients and a set of facilities in a common metric space. In the first stage, the decision-maker needs to select the (integral) units supply at each facility. The uncertain demand is then selected adversarially and needs to be satisfied by the existing supply with the minimum assignment cost in the second stage. The goal is to determine the first-stage supply such that the sum of first-stage supply cost and the worst-case assignment cost over all demand scenarios is minimized. Our problem is motivated by settings where the lead time to procure supply is large and obtaining additional units of supply in the second stage is not feasible. The uncertain second-stage demand must then be satisfied by supply units from the first stage, a common constraint in many applications. This is a departure from existing work on two-stage robust

facility location, network design and, more generally, robust covering problems that have been studied extensively in the literature [6,7,1,9] where the second-stage decisions allow for “adding more supply” (more specifically adding more sets/facilities to satisfy the requirement).

In this paper, we consider an implicit model of uncertainty with exponentially-many demand scenarios specified by an upper bound k on the number of second-stage demand clients. Since the number of scenarios is exponentially-many (specified compactly), we can not efficiently solve even the LP relaxation for the problem. In contrast, if the set of second-stage scenarios are explicitly specified (for the explicit scenario model, see for instance [6,1]), we can write a polynomially-sized LP relaxation with assignment decisions for each scenario. The main challenge then is related to obtaining an integral solution, which for the case of set covering and several network design problems can be reduced to deterministic versions (see, for instance, Dhamdhere et al. [6]).

In contrast, in an implicit model of uncertainty (with possibly exponentially-many scenarios), one of the fundamental challenges is to even approximately solve the linear relaxation of the problem efficiently. The implicit model of uncertainty with an upper bound on number of uncertain second-stage clients or elements has been studied extensively in the literature. Feige et al. [7] show under a reasonable complexity assumption that it is hard to solve the LP relaxation of a two-stage set covering problem within a factor better than $\Omega(\log n / \log \log n)$ under this implicit model of uncertainty. They also give an $O(\log^2 n)$ -approximation for the 0 – 1 two-stage robust set-covering problem. Gupta et al. [9] give an improved $O(\log n)$ -approximation for the set-covering problem, thereby matching the deterministic approximation guarantee. El Housni and Goyal [12] show that a static policy (that is, linear in the set of second-stage elements, also referred to as an affine policy) gives an $O(\log n / \log \log n)$ -approximation for the two-stage LP, thereby matching the hardness lower bound for the fractional problem. Gupta et al. [10] and Khandekar et al. [14] present approximations for several network design problems under the implicit model of uncertainty. Although there is a large body of work in this direction, as we mentioned earlier, our problem is different from set covering, since there is no possibility of adding recourse capacity; prior results do not imply an approximation for our model.

Other related work. Many variants of facility location problems have been studied extensively in the literature, including both deterministic versions as well as variants that address demand uncertainty. We refer the reader to [17] for a review of the many variants of deterministic facility location problems. Among the models that address demand uncertainty in the facility location problems, in addition to robust [2,4], there are also stochastic [11,13,20,16] and distributionally robust [15,3,5] models that have been studied extensively in the literature. In a stochastic model, there is a distribution on the second-stage demand scenarios and the goal is to minimize the total expected cost. A distributionally robust model can be thought of as a hybrid between stochastic and robust where the second-stage distribution is selected adversarially from a pre-specified set and the goal is to minimize the worst-case expected cost. We refer the reader to the

survey [19] for an extensive review of facility location problems under uncertainty.

1.1 Our Contributions

Let us begin with a formal problem definition. We are given a set of n facilities \mathcal{F} and m clients \mathcal{C} in a common metric d where d_{ij} denotes the distance between i and j . For each facility $i \in \mathcal{F}$, there is a cost c_i per unit of supply at i . The demand uncertainty is modeled by an implicit set of scenarios \mathcal{C}^k that includes all subsets of clients \mathcal{C} of size at most k . The decision-maker needs to select an (integral) number of units of supply x_i for each facility $i \in \mathcal{F}$ in the first-stage. An adversary observes the first stage decisions and selects a worst-case demand scenario $S \in \mathcal{C}^k$ that must be satisfied with the first-stage supply, where each client in the realized scenario needs one unit of supply. The goal is to minimize the sum of the first-stage supply cost and the worst-case assignment cost over all second-stage demand scenarios. We refer to this problem as *soft-capacitated robust facility location* (SCRFL). Typically in the literature, *soft-capacitated* refer to settings where violations of capacity upper bounds are allowed. The analogue here is that we can add any amount of supply in a facility without upper bounds ($x_i \in \mathbb{Z}_+$) but we pay a per unit cost of supply.

We consider a class of static assignment policies, where each of the m clients has a static fractional assignment to facilities that is independent of the scenario, leading to a feasible second-stage solution for each demand scenario, while respecting supply capacities. Note this is a restriction, since the optimal second-stage assignment decisions are scenario-dependent in general. As a warm-up, we show that static assignment policies are optimal for the uncapacitated case with unlimited supply at each open facility (i.e., there is a cost c_i to open facility i with unlimited supply). We refer to this problem as *uncapacitated robust facility location* (URFL). This is based on the intuition that each client can be assigned to the closest open facilities in an optimal solution in any scenario; this leads to optimality of a static assignment policy for the LP relaxation (Theorem 1).

Theorem 1. *A static assignment policy is optimal for the linear relaxation of (URFL).*

The optimality of static assignment is not true in general when the supply at facilities is constrained (or equivalently, there is a cost per unit of supply). The main contribution in this paper is to show that static assignment policies give an $O(\log k / \log \log k)$ -approximation for the LP relaxation of (SCRFL) (Theorem 2). We show this by constructing such a solution, starting from an optimal first-stage supply. The optimal static assignment policies can be computed efficiently by solving a compact LP.

Theorem 2. *A static assignment policy gives $O(\log k / \log \log k)$ -approximation for the linear relaxation of (SCRFL).*

Furthermore, the fractional supply in the first stage can be rounded to an integral supply using ideas similar to rounding algorithms for the deterministic

facility location [18]. In particular, the static assignment solution for the uncapacitated case can be rounded to give a 4-approximation algorithm for (URFL). The static assignment solution for the soft-capacitated case can be rounded within a constant factor, which results in an $O(\log k / \log \log k)$ -approximation algorithm for (SCRFL). We would like to note that while the fractional assignment is static in our approximate LP solution, our integral assignment for any client in the second-stage depends on the other demand clients in the scenario; thereby, making our static assignment policy adaptive in implementation.

2 Warm-up: uncapacitated robust facility location

2.1 Problem formulation

In this section, we consider the *uncapacitated robust facility location problem* (URFL) where for each $i \in \mathcal{F}$, there is a cost c_i to open facility i with unlimited supply. The problem can be stated as the following integer program, where each binary variable $x_i, i \in \mathcal{F}$ indicates if facility i is opened and each $y_{ij}^S, i \in \mathcal{F}, j \in S, S \in \mathcal{C}^k$ indicates the assignment of client j to facility i in scenario S .

$$\begin{aligned}
& \min \sum_{i \in \mathcal{F}} c_i x_i + \max_{S \in \mathcal{C}^k} \sum_{i \in \mathcal{F}} \sum_{j \in S} d_{ij} y_{ij}^S \\
& \text{s.t.} \quad \sum_{i \in \mathcal{F}} y_{ij}^S \geq 1, & \forall S \in \mathcal{C}^k, \forall j \in S, \\
& \quad x_i \geq y_{ij}^S, & \forall i \in \mathcal{F}, \forall S \in \mathcal{C}^k, \forall j \in S, \\
& \quad x_i \in \{0, 1\}, y_{ij}^S \geq 0, & \forall i \in \mathcal{F}, \forall S \in \mathcal{C}^k, \forall j \in S.
\end{aligned} \tag{URFL}$$

Note that the second-stage problem is a transportation problem and since the demand of a client is integral (0 or 1), the optimal solution y_{ij}^S is integral as well. The special case of (URFL) where the uncertainty set contains only a single scenario corresponds to the NP-hard classical and well-studied uncapacitated facility location problem, which is hard to approximate within a constant better than 1.463 unless NP has an $O(n^{O(\log \log n)})$ -time algorithm [8]. We let (LP-URFL) denote the linear relaxation of (URFL), where we replace $x_i \in \{0, 1\}$ by $x_i \geq 0$ for each $i \in \mathcal{F}$. We would like to note that, in our model, the number of scenarios could be exponential in the dimension of the problem. Hence, in general, even the linear relaxation of such a problem could be challenging. However, we show that (LP-URFL) can be solved in polynomial time using a *Static Assignment Policy* for the second-stage variables. Moreover, we can round the fractional solution losing only a constant factor, thereby getting a constant approximation for (URFL). This section serves as a warm-up for introducing and motivating static assignment policies before addressing the class of capacitated robust facility location problems.

2.2 Static assignment policy

Consider an optimal solution of (URFL). Since each open facility can have an unlimited amount of supply, each client in the realized scenario is assigned to the closest facility among the opened ones. Therefore, a client j is always assigned to the same open facility in all scenarios S where $j \in S$. The same observation holds as well for (LP-URFL) where each client is assigned to the same fractionally opened facilities independent of the realized scenario. Thus, the assignment of a client is static. This can be captured by the following policy.

Static assignment policy. There exists $y_{ij} \geq 0$ for each $i \in \mathcal{F}, j \in \mathcal{C}$ such that

$$\forall S \in \mathcal{C}^k, \forall i \in \mathcal{F}, \forall j \in S, \quad y_{ij}^S = y_{ij}. \quad (1)$$

Proof of Theorem 1. Let $(\mathbf{x}^*, \mathbf{y}^{*S}, S \in \mathcal{C}^k)$ be an optimal solution to (LP-URFL). Since there are no capacities on facilities, each client j is assigned to the closest fractionally opened facilities. In particular, for each $j \in \mathcal{C}$, let π_j be a permutation of $\mathcal{F} = \{1, \dots, n\}$ such that $d_{\pi_j(1)j} \leq d_{\pi_j(2)j} \leq \dots \leq d_{\pi_j(n)j}$, and let $\ell = \min\{p \mid x_{\pi_j(1)}^* + x_{\pi_j(2)}^* + \dots + x_{\pi_j(p)}^* \geq 1\}$. Denote $\hat{x}_{\pi_j(\ell)} = 1 - (x_{\pi_j(1)}^* + \dots + x_{\pi_j(\ell-1)}^*)$. The optimal solution can be written in the form (1) as follows: for $S \in \mathcal{C}^k$ and $j \in S$; $y_{ij}^S = x_i^*$ for $i \in \{\pi_j(1), \dots, \pi_j(\ell-1)\}$, $y_{ij}^S = \hat{x}_{\pi_j(\ell)}$ for $i = \pi_j(\ell)$, and $y_{ij}^S = 0$, otherwise. \square

Let (Static-URFL) denote the problem after restricting the second-stage variables y_{ij}^S in (LP-URFL) to a policy (1), which can then be reformulated as follows:

$$\begin{aligned} \min \quad & \sum_{i \in \mathcal{F}} c_i x_i + \max_{S \in \mathcal{C}^k} \sum_{i \in \mathcal{F}} \sum_{j \in \mathcal{C}} \mathbf{1}(j \in S) \cdot d_{ij} y_{ij} \\ \text{s.t.} \quad & \sum_{i \in \mathcal{F}} y_{ij} \geq 1, & \forall j \in \mathcal{C}, \quad (\text{Static-URFL}) \\ & x_i \geq y_{ij} \geq 0, & \forall i \in \mathcal{F}, \forall j \in \mathcal{C}. \end{aligned}$$

From Theorem 1, (Static-URFL) is equivalent to (LP-URFL). The number of variables in (Static-URFL) is reduced to a polynomial number since the y_{ij} no longer depend on the scenario S . The inner maximization problem is still taken over an exponential number of scenarios; however we can separate efficiently over these scenarios and write an efficient compact LP formulation for (Static-URFL):

$$\begin{aligned} \max_{S \in \mathcal{C}^k} \left\{ \sum_{i \in \mathcal{F}} \sum_{j \in \mathcal{C}} \mathbf{1}(j \in S) \cdot d_{ij} y_{ij} \right\} &= \max_{\mathbf{h} \in [0,1]^{|\mathcal{C}|}} \left\{ \sum_{i \in \mathcal{F}} \sum_{j \in \mathcal{C}} d_{ij} y_{ij} h_j \mid \sum_{j \in \mathcal{C}} h_j \leq k \right\} \\ &= \min_{\mu, \omega \geq 0} \left\{ k\mu + \sum_{j \in \mathcal{C}} \omega_j \mid \mu + \omega_j \geq \sum_{i \in \mathcal{F}} d_{ij} y_{ij}, \forall j \in \mathcal{C} \right\}, \end{aligned} \quad (2)$$

where the first equality holds because the optimal solution of the right maximization problem occurs at the extreme points of the k -ones polytope, which corresponds to the worst-case scenarios of \mathcal{C}^k and the second equality follows from strong duality. Therefore, by dropping the min and introducing μ and all ω_j as variables, we reformulate (Static-URFL) as the following linear program:

$$\begin{aligned}
& \min \sum_{i \in \mathcal{F}} c_i x_i + k\mu + \sum_{j \in \mathcal{C}} \omega_j \\
& \text{s.t. } \mu + \omega_j \geq \sum_{i \in \mathcal{F}} d_{ij} y_{ij}, & \forall j \in \mathcal{C}, \\
& \sum_{i \in \mathcal{F}} y_{ij} \geq 1, & \forall j \in \mathcal{C}, \\
& x_i \geq y_{ij}, & \forall i \in \mathcal{F}, \forall j \in \mathcal{C}, \\
& x_i \geq 0, y_{ij} \geq 0, \omega_j \geq 0, \mu \geq 0, & \forall i \in \mathcal{F}, \forall j \in \mathcal{C}.
\end{aligned} \tag{3}$$

Finally, we round the solution of (Static-URFL) to an integral solution for (URFL) while losing only a constant factor. This can be done using prior work on rounding techniques from the literature of deterministic facility location problems. In fact, the LP rounding technique in Shmoys et al. [18], which gives a 4-approximation algorithm to the deterministic uncapacitated problem also gives a 4-approximation algorithm for (Static-URFL). The idea is to define a ball around each client of radius equal to the fractional assignment cost of the client (which is independent of any scenario for our static policy). Then, we open facilities in non-intersecting balls of ascending radius. The result is given in the following theorem and for completeness, the details of the rounding are in Appendix A.

Theorem 3. (Static-URFL) *can be rounded to give a 4-approximation to (URFL).*

We would like to note that the focus in this section is not about finding the best constant approximation for (URFL), but we introduce it as a warm-up for motivating the static assignment policy before presenting our main result in the next section.

3 Soft-capacitated robust facility location

3.1 Problem formulation

In this section, we consider the *soft-capacitated robust facility location* (SCRFL) which is similar to (URFL) except that each facility i incurs a linear supply cost, where c_i is the cost per unit of supply. We refer to x_i as the supply (or capacity) in facility i . Each client in the realized scenario needs to be satisfied by one unit of supply. The problem is called soft-capacitated since there is no upper bound on x_i , ($x_i \in \mathbb{Z}_+$). The problem is given by the following integer program:

$$\begin{aligned}
& \min \sum_{i \in \mathcal{F}} c_i x_i + \max_{S \in \mathcal{C}^k} \sum_{i \in \mathcal{F}} \sum_{j \in S} d_{ij} y_{ij}^S \\
& \text{s.t. } \sum_{i \in \mathcal{F}} y_{ij}^S \geq 1, & \forall S \in \mathcal{C}^k, \forall j \in S, \\
& x_i \geq \sum_{j \in S} y_{ij}^S, & \forall i \in \mathcal{F}, \forall S \in \mathcal{C}^k, \forall j \in S, \\
& x_i \in \mathbb{N}, y_{ij}^S \geq 0, & \forall i \in \mathcal{F}, \forall S \in \mathcal{C}^k, \forall j \in S.
\end{aligned} \tag{SCRFL}$$

We let (LP-SCRFL) denote the linear relaxation of (SCRFL), where we replace $x_i \in \mathbb{N}$ by $x_i \geq 0$, for each $i \in \mathcal{F}$. We would like to note that even the linear relaxation (LP-SCRFL) is challenging to solve since it has exponentially-many variables (scenarios). Unlike the uncapacitated case, the static assignment policy (1) is not optimal for (LP-SCRFL) and the optimal assignment for each client depends, in general, on the realized scenario. In particular, the same client could be assigned to different facilities in different scenarios. In contrast, we show the surprising result that a static assignment policy gives $O(\log k / \log \log k)$ -approximation to (LP-SCRFL). Moreover, we can round the solution of the static assignment policy to an integral solution for (SCRFL) and only lose an additional constant factor. We let (Static-SCRFL) denote the problem when we restrict the second-stage variables y_{ij}^S in (LP-SCRFL) to static assignment policies (1). The problem can then be reformulated as follows:

$$\begin{aligned}
& \min \sum_{i \in \mathcal{F}} c_i x_i + \max_{S \in \mathcal{C}^k} \sum_{i \in \mathcal{F}} \sum_{j \in \mathcal{C}} \mathbf{1}(j \in S) \cdot d_{ij} y_{ij} \\
& \text{s.t. } \sum_{i \in \mathcal{F}} y_{ij} \geq 1, \quad \forall j \in \mathcal{C}, \\
& \quad x_i \geq \max_{S \in \mathcal{C}^k} \sum_{j \in \mathcal{C}} \mathbf{1}(j \in S) \cdot y_{ij}, \quad \forall i \in \mathcal{F}, \\
& \quad x_i \geq 0, y_{ij} \geq 0, \quad \forall i \in \mathcal{F}, \forall j \in \mathcal{C}.
\end{aligned} \tag{Static-SCRFL}$$

3.2 An $O(\frac{\log k}{\log \log k})$ -approximation algorithm

Our main contribution in this section is to show that a static assignment policy (1) gives $O(\log k / \log \log k)$ -approximation for (LP-SCRFL) (Theorem 2). To prove this theorem, we consider an optimal solution of (LP-SCRFL) and massage it to construct a solution of the form (1) while losing $O(\log k / \log \log k)$ factor. We first present our construction and several structural lemmas and then give the proof of Theorem 2.

Our construction. Let $\mathbf{x}^* : (x_i^*)_{i \in \mathcal{F}}$ be an optimal first-stage solution of (LP-SCRFL), let OPT_1 be the corresponding optimal first-stage cost and let OPT_2 be the corresponding optimal second-stage cost. We will classify the clients \mathcal{C} into three subsets C_1, C_2, C_3 using Procedure 1 (below) and then specify a static assignment policy for each subset. We use the following notation in the procedure. Let $\alpha > 1$ and $r = 5 \cdot \text{OPT}_2 / k$. For $\ell \geq 1$ and $j \in \mathcal{C}$, we let B_j^ℓ denote the ball centered at client j of radius ℓr . We initialize the sets $F \leftarrow \mathcal{F}$ and $C \leftarrow \mathcal{C}$ and update them at iteration, as explained in the procedure, until C becomes empty. We let $Cl(B)$ denote the set of clients in C that are inside the ball B and let $Sp(B)$ denote the total optimal supply of facilities F that are inside the ball B , i.e.,

$$Sp(B) = \sum_{i \in F} \mathbf{1}(i \in B) \cdot x_i^* \quad \text{and} \quad Cl(B) = \{j \in C \mid j \in B\}.$$

Procedure 1

```

1: Initialize  $C \leftarrow \mathcal{C}, C_1 \leftarrow \emptyset, C_2 \leftarrow \emptyset, C_3 \leftarrow \emptyset, F \leftarrow \mathcal{F}$ .
2: while  $C \neq \emptyset$  do
3:   Pick a client  $j \in C$ . Initialize  $\ell = 1$ 
4:   if  $|Cl(B_j^{2\ell-1})| \geq k$  then
5:      $C_1 \leftarrow C_1 \cup Cl(B_j^{2\ell-1}), C \leftarrow C \setminus Cl(B_j^{2\ell-1})$ 
6:     Stop, return to line 2
7:   if  $Sp(B_j^{2\ell}) < \frac{1}{2} \cdot |Cl(B_j^{2\ell-1})|$  then
8:      $C_2 \leftarrow C_2 \cup Cl(B_j^{2\ell-1}), C \leftarrow C \setminus Cl(B_j^{2\ell-1})$ 
9:     Stop, return to line 2
10:  if  $Sp(B_j^{2\ell}) \geq \frac{1}{2\alpha} \cdot |Cl(B_j^{2\ell+1})|$  then
11:     $C_3 \leftarrow C_3 \cup Cl(B_j^{2\ell+1}), C \leftarrow C \setminus Cl(B_j^{2\ell+1}),$ 
12:     $F \leftarrow F \setminus \{i \in F \mid i \in B_j^{2\ell}\}$ 
13:    Stop, return to line 2
14:  else  $\ell \leftarrow \ell + 1$ , return to line 4.

```

Note that both $Sp(B)$ and $Cl(B)$ depend on the current sets of facilities F and clients C , which we update at each iteration of the while loop in the procedure. But for the ease of notation, we do not refer to them with the indices F and C .

In the procedure, while the set C is not empty, we pick a client $j \in C$ and grow three balls around it: $B_j^{2\ell-1}$ (*internal* ball), $B_j^{2\ell}$ (*medium* ball) and $B_j^{2\ell+1}$ (*external* ball) starting with $\ell = 1$. For each ℓ , we check if the number of clients in the internal ball $B_j^{2\ell-1}$ is greater than k (line 4); if this is the case, we remove them from C , put them in C_1 and restart in line 2. If not, we check if the supply in the medium ball $B_j^{2\ell}$ is sufficient to satisfy half of the clients in the internal ball $B_j^{2\ell-1}$ (line 7); if that is not the case, we remove those clients from C , put them in C_2 , and restart in line 2. Otherwise, we finally check if the supply in the the medium ball $B_j^{2\ell}$ is sufficient to satisfy a fraction $1/2\alpha$ of the clients in the external ball $B_j^{2\ell+1}$ (line 10); if that is the case we remove all the clients in $B_j^{2\ell+1}$ and put them in C_3 , we also remove all the facilities in $B_j^{2\ell}$ and restart in line 2. If none of these three conditions holds, we increase ℓ to $\ell + 1$. First, we show that after at most $\log_\alpha k$ increments (i.e., $\ell \leq \log_\alpha k$), one of three conditions must hold and therefore we will remove some clients from C and restart in line 2. Which implies that after a finite number of iterations, the set C becomes empty. In particular, We have the following lemma.

Lemma 1. *In Procedure 1, after a finite number of iterations, the set C becomes empty and $C_1 \cup C_2 \cup C_3$ is equal to \mathcal{C} . Moreover, ℓ is always less than $\log_\alpha k$.*

Proof. Fix a client j and let $\ell \geq 1$. If none of the three conditions (“if” statements) holds then

$$\alpha \cdot |Cl(B_j^{2\ell-1})| \leq 2\alpha \cdot Sp(B_j^{2\ell}) < |Cl(B_j^{2\ell+1})|.$$

Therefore, the number of clients grows geometrically when we increase the radius of the balls and by induction, we have that

$$\alpha^\ell \leq \alpha^\ell \cdot |Cl(B_j^1)| < |Cl(B_j^{2\ell+1})|,$$

where $|Cl(B_j^1)| \geq 1$, since $Cl(B_j^1)$ contains at least the client j . Hence, after at most $\log_\alpha k$ increments, we will reach k clients, and must stop by the first condition, and return to line 2. Hence, we always have $\ell \leq \log_\alpha k$. Finally since, we remove at least one client at each iteration of the while loop, the set C becomes empty after at most $|C|$ iterations, and finally $C_1 \cup C_2 \cup C_3 = \mathcal{C}$. \square

Now we are ready to present our static assignment policy for (LP-SCRFL). The following three lemmas show our constructed static assignment for each client in the three subsets C_1, C_2, C_3 . Moreover, we specify the supply used to satisfy each subset of these clients and present the analysis for the assignment cost.

For each client in the set C_1 , we know that it belongs to a ball with at least k clients. By the feasibility of the optimal solution, this implies that there exists k units of supply in \mathbf{x}^* “close” to this ball. Hence, we satisfy this client by using the same fraction $1/k$ of \mathbf{x}^* (static assignment) while paying a small assignment cost (roughly a constant times the radius of the ball). Since, there are at most k clients in each scenario, and each one is using at most \mathbf{x}^*/k , we need to dedicate only one \mathbf{x}^* for all clients in C_1 . Formally, we have the following lemma.

Lemma 2. *There exists a static assignment policy for C_1 such that each client in C_1 is using at most the supply \mathbf{x}^*/k and has an assignment cost less than $O(\log_\alpha k/k) \cdot \text{OPT}_2$, i.e., there exists $(\tilde{y}_{ij})_{i \in \mathcal{F}, j \in C_1}$ such that for each $j \in C_1$:*

$$\sum_{i \in \mathcal{F}} \tilde{y}_{ij} \geq 1, \quad \frac{x_i^*}{k} \geq \tilde{y}_{ij} \geq 0, \quad \forall i \in \mathcal{F}, \quad \text{and} \quad \sum_{i \in \mathcal{F}} d_{ij} \tilde{y}_{ij} = O(\log_\alpha k) \cdot \frac{\text{OPT}_2}{k}.$$

Proof. Let j be a client of C_1 . It is sufficient to show that the following minimization problem is feasible and its optimal cost is $O(\log_\alpha k/k) \cdot \text{OPT}_2$. Consider

$$\min \left\{ \sum_{i \in \mathcal{F}} d_{ij} y_{ij} \mid \sum_{i \in \mathcal{F}} y_{ij} \geq 1, \quad \frac{x_i^*}{k} \geq y_{ij} \geq 0, \quad \forall i \in \mathcal{F} \right\}. \quad (4)$$

Problem (4) must be feasible since the total supply in \mathbf{x}^* is greater than the total demand in any scenario, i.e., $\sum_{i \in \mathcal{F}} x_i^* \geq k$. Recall that a client j in C_1 belongs to one of the sets $Cl(B_t^{2\ell-1})$ for some $t \in \mathcal{C}$ and $\ell \leq \log_\alpha k$ (Lemma 1) such that $|Cl(B_t^{2\ell-1})| \geq k$. Consider a scenario S formed by k clients from $Cl(B_t^{2\ell-1})$. Let denote \mathbf{y}^S the assignment of scenario S in the optimal solution. Consider the following candidate solution for (4):

$$y_{ij} = \frac{1}{k} \cdot \sum_{p \in S} y_{ip}^S, \quad \forall i \in \mathcal{F}.$$

We have, by the feasibility of the optimal solution, for each $i \in \mathcal{F}$, $0 \leq y_{ij} \leq \frac{1}{k} x_i^*$ and

$$\sum_{i \in \mathcal{F}} y_{ij} = \frac{1}{k} \cdot \sum_{i \in \mathcal{F}} \sum_{p \in S} y_{ip}^S \geq \frac{1}{k} \sum_{p \in S} 1 = 1.$$

Therefore, our solution is feasible for (4). Moreover, we have

$$\begin{aligned} \sum_{i \in \mathcal{F}} d_{ij} y_{ij} &= \frac{1}{k} \cdot \sum_{i \in \mathcal{F}} \sum_{p \in S} d_{ij} y_{ip}^S \\ &\leq \frac{1}{k} \cdot \sum_{i \in \mathcal{F}} \sum_{p \in S} d_{ip} y_{ip}^S + \frac{1}{k} \cdot \sum_{p \in S} d_{pj} \\ &\leq \frac{1}{k} \cdot \text{OPT}_2 + 2(2\ell - 1)r \\ &\leq \frac{\text{OPT}_2}{k} + 2(2\log_\alpha k - 1) \cdot 5 \cdot \frac{\text{OPT}_2}{k} = O(\log_\alpha k) \cdot \frac{\text{OPT}_2}{k}, \end{aligned}$$

where the first inequality follows from the triangle inequality and the fact that $\sum_{i \in \mathcal{F}} y_{ip}^S = 1$ for all $p \in S$ in the optimal solution. For the second inequality, we use the definition of OPT_2 to bound the first term, the second term d_{pj} is bounded by the diameter of the ball $B_j^{2\ell-1}$ which contains client j and all clients $p \in S$. \square

Now, consider the set C_2 . By construction, these are clients such that there is not enough supply within a distance $r = 5\text{OPT}_2/k$ to satisfy half of them. Therefore, intuitively they need to pay “large” distances in the optimal assignment cost if they show up all together in the same scenario. In the following lemma, we show that we can have no more than k of these clients. As we would show later, this would imply that we can dedicate a supply x^* to C_2 and make a static assignment of all the clients C_2 to this x^* .

Lemma 3. *The set C_2 has at most k clients.*

Proof. Suppose, for the sake of contradiction, that $|C_2| > k$. Let G_1, G_2, \dots, G_T be the disjoint subsets of clients added at each iteration in the construction of C_2 in Procedure 1. In particular, $C_2 = G_1 \cup G_2 \cup \dots \cup G_T$ for some T where:

- (i) for $t = 1, 2, \dots, T$, $G_t = Cl(B_{j_t}^{2\ell_t-1})$ for some client j_t and $1 \leq \ell_t \leq \log_\alpha k$.
- (ii) the supply $Sp(B_{j_t}^{2\ell_t})$ is less than half of the clients in G_t .
- (iii) each set G_t has strictly less than k clients, since the procedure has to fail the first “if” statement before adding G_t into C_2 .

Recall that

$$Sp(B_{j_t}^{2\ell_t}) = \sum_{i \in F} \mathbf{1}(i \in B_{j_t}^{2\ell_t}) \cdot x_i^*,$$

where F is the current set of facilities in the procedure (and is not all \mathcal{F} since some facilities have been removed in line 12 of the procedure). However, we would like to emphasize that when a facility has been removed (in line 12 of the

procedure), all clients within distance r from this facility were removed as well (line 11). This is true, since when we remove the facilities in a medium ball, say $B_j^{2\ell}$, (line 12), we remove all clients in the corresponding external ball $B_j^{2\ell+1}$ (line 11). Hence, the remaining clients in C are at least $(2\ell + 1)r - (2\ell)r = r$ away from the removed facilities. In particular, for the clients G_t , the supply that has been removed before they were added to C_2 is at least r away from them. Therefore, all the facility in \mathcal{F} that are within a distance r from a client in G_t belong to the set F that verifies $\sum_{i \in F} \mathbf{1}(i \in B_{j_t}^{2\ell_t}) \cdot x_i^* \leq 2 \cdot |G_t|$. This implies that the supply of all facilities within a distance r from G_t in the optimal solution, is less than half of the clients G_t . Hence, if all of the clients G_t show up in a scenario, the optimal second-stage solution needs to pay an assignment cost of at least $r \cdot |G_t|/2$.

Order the sets G_t according to their cardinalities: wlog assume that $|G_1| \geq |G_2| \geq \dots \geq |G_T|$. We construct a scenario \hat{S} by taking clients from the sets G_1, G_2, \dots until we hit k . This is possible since by assumption $|C_2| > k$. Assume that

$$|G_1| + |G_2| + \dots + |G_{p-1}| + |\bar{G}_p| = k,$$

for some p , where $2 \leq p \leq T$. Note that \bar{G}_p is a subset of G_p , since we can reach k before taking all the clients of the last set G_p . For each $t = 1, \dots, p-1$, the optimal second-stage decision needs to pay at least $r \cdot |G_t|/2$. Therefore,

$$\text{OPT}_2 \geq \frac{1}{2} \cdot r \cdot (|G_1| + |G_2| + \dots + |G_{p-1}|).$$

We did not include G_p , since not all these clients are necessary in the scenario \hat{S} , but only $|\bar{G}_p|$ of them. Since \bar{G}_p has the smallest cardinality

$$|G_1| + |G_2| + \dots + |G_{p-1}| \geq \frac{1}{2}(|G_1| + |G_2| + \dots + |G_{p-1}| + |\bar{G}_p|) = \frac{k}{2}.$$

Therefore,

$$\text{OPT}_2 \geq r \cdot \frac{k}{4} = 5 \cdot \frac{\text{OPT}_2}{4},$$

which is a contradiction. Therefore, $|C_2| \leq k$. \square

Finally, for clients C_3 , we show that there exists $|C_3|/2\alpha$ units of supply “close” to them. In particular, we can multiply these units by 2α , dedicate them to C_3 and make a static assignment for C_3 . We have the following lemma.

Lemma 4. *There exists a supply \hat{x} that has a cost at most $2\alpha \cdot \text{OPT}_1$ and there exists a static assignment policy such that all clients in C_3 are assigned to supply \hat{x} and each client in C_3 has an assignment cost that is $O(\log_\alpha k/k) \cdot \text{OPT}_2$, i.e., there exists $(\hat{y}_{ij})_{i \in \mathcal{F}, j \in C_3}$ and $(\hat{x}_i)_{i \in \mathcal{F}}$ such that for all $j \in C_3$:*

$$\begin{aligned} \sum_{i \in \mathcal{F}} \hat{y}_{ij} &\geq 1, & \hat{x}_i &\geq \sum_{j \in C_3} \hat{y}_{ij}, \quad \forall i \in \mathcal{F}, & \hat{x}_i &\geq 0, \hat{y}_{ij} \geq 0, \quad \forall i \in \mathcal{F}, \\ \sum_{i \in \mathcal{F}} c_i \hat{x}_i &\leq 2\alpha \cdot \text{OPT}_1 & \text{and} & & \sum_{i \in \mathcal{F}} d_{ij} \hat{y}_{ij} &= O(\log_\alpha k) \cdot \frac{\text{OPT}_2}{k}. \end{aligned}$$

Proof. Let G_1, G_2, \dots, G_T be the disjoint subsets of clients added at each iteration to construct C_3 in Procedure 1. In particular, $C_3 = G_1 \cup G_2 \cup \dots \cup G_T$ for some T such that: for all $t = 1, 2, \dots, T$, $G_t = Cl(B_{j_t}^{2\ell_t+1})$ for some client j_t and some integer ℓ_t with $1 \leq \ell_t \leq \log_\alpha k$. Moreover, the supply $Sp(B_{j_t}^{2\ell_t})$ is greater than a $1/2\alpha$ fraction of the clients in G_t . Hence, for each ball $B_{j_t}^{2\ell_t}$, we multiply the supply by 2α , move it to the cheapest facility in this ball and make a static assignment of all clients G_t to this cheapest facility. Since the supply in $B_{j_t}^{2\ell_t}$ is removed along with clients G_t , it will not be used by the other clients in C_3 .

Formally, let i_t be the cheapest facility in the ball $B_{j_t}^{2\ell_t}$. We define, for each facility i in $B_{j_t}^{2\ell_t}$, $\hat{x}_i = 2\alpha \sum_{i' \in F} \mathbf{1}(i' \in B_{j_t}^{2\ell_t}) \cdot x_{i'}^*$ if $i = i_t$ and $\hat{x}_i = 0$ otherwise. For each client $j \in C_3$, we let $\hat{y}_{ij} = 1$ for $i = i_t$ and $j \in G_t$, and let $\hat{y}_{ij} = 0$, otherwise. Therefore, the first desired constraints in the lemma are verified. Let us check the last one. The distance between a client and its assigned facility in our solution is at most r plus the diameter of the ball $B_{j_t}^{2\ell_t}$, i.e.,

$$r + 4\ell_t r \leq (4 \log_\alpha k + 1) \cdot 5 \cdot \text{OPT}_2/k = O(\log_\alpha k/k) \cdot \text{OPT}_2.$$

□

Proof of Theorem 2. Let $(\tilde{y}_{ij})_{i \in \mathcal{F}, j \in C_1}$ be the solution given in Lemma 2 for satisfying the clients in C_1 . We dedicate a supply \mathbf{x}^* to clients C_1 . Let $(\hat{y}_{ij})_{i \in \mathcal{F}, j \in C_3}$ and $(\hat{x}_i)_{i \in \mathcal{F}}$ be the solution given in Lemma 4 for satisfying the clients in C_3 . Finally, we know from Lemma 3 that C_2 has at most k clients, and therefore C_2 is a scenario. So we dedicate a supply \mathbf{x}^* to C_2 and let the optimal assignment $y_{ij}^{C_2}$ be our static assignment solution for C_2 . In particular, we give the following solution to (LP-SCRFL), where the first stage solution is $2\mathbf{x}^* + \hat{\mathbf{x}}$ and the static assignment policy is for all $i \in \mathcal{F}$: $y_{ij} = \tilde{y}_{ij}$ for $j \in C_1$, $y_{ij} = y_{ij}^{C_2}$ for $j \in C_2$, and $y_{ij} = \hat{y}_{ij}$ for $j \in C_3$. It is clear that $\sum_{i \in \mathcal{F}} y_{ij} \geq 1$, for each j in $C_1 \cup C_2 \cup C_3$. Moreover, for any scenario $S \in \mathcal{C}^k$ and $i \in \mathcal{F}$,

$$\begin{aligned} \sum_{j \in S} y_{ij} &= \sum_{j \in S \cap C_1} \tilde{y}_{ij} + \sum_{j \in S \cap C_2} y_{ij}^{C_2} + \sum_{j \in S \cap C_3} \hat{y}_{ij} \\ &\leq \sum_{j \in S \cap C_1} \frac{x_i^*}{k} + x_i^* + \hat{x}_i \leq 2x_i^* + \hat{x}_i. \end{aligned}$$

Therefore, our solution is feasible for (LP-SCRFL). Let us evaluate its cost. The cost of the first stage is at most $2\text{OPT}_1 + 2\alpha\text{OPT}_1 = O(\alpha) \cdot \text{OPT}_1$. For the second-stage cost, consider any scenario $S \in \mathcal{C}^k$. We have

$$\begin{aligned} \sum_{i \in \mathcal{F}} \sum_{j \in S} d_{ij} y_{ij} &= \sum_{j \in S \cap C_1} \sum_{i \in \mathcal{F}} d_{ij} \tilde{y}_{ij} + \sum_{j \in S \cap C_2} \sum_{i \in \mathcal{F}} d_{ij} y_{ij}^{C_2} + \sum_{j \in S \cap C_3} \sum_{i \in \mathcal{F}} d_{ij} \hat{y}_{ij} \\ &\leq \sum_{j \in S \cap C_1} O(\log_\alpha k) \cdot \frac{\text{OPT}_2}{k} + \text{OPT}_2 + \sum_{j \in S \cap C_3} O(\log_\alpha k) \cdot \frac{\text{OPT}_2}{k} \\ &\leq O(\log_\alpha k) \cdot \text{OPT}_2 + \text{OPT}_2 + O(\log_\alpha k) \cdot \text{OPT}_2 \\ &= O(\log_\alpha k) \cdot \text{OPT}_2. \end{aligned}$$

By balancing the terms α and $\log_\alpha k$, we choose $\alpha = \log k / \log \log k$ which gives $O(\log k / \log \log k)$ -approximation to (LP-SCRFL). \square

Similar to the uncapacitated problem, we can solve (Static-SCRFL) efficiently using a compact linear program. In fact, we dualize the inner maximization problem in the objective function of (Static-SCRFL) in the same way as (2). In addition to that, we reformulate the second constraint in (Static-SCRFL) using the same dualization technique as follows: for each $i \in \mathcal{F}$,

$$\begin{aligned} \max_{S \in \mathcal{C}^k} \left\{ \sum_{j \in S} y_{ij} \right\} &= \max_{\mathbf{h} \in [0,1]^{|\mathcal{C}|}} \left\{ \sum_{j \in \mathcal{C}} y_{ij} h_j \mid \sum_{j \in \mathcal{C}} h_j \leq k \right\} \\ &= \min_{\eta_i, \lambda_{ij} \geq 0} \left\{ k\eta_i + \sum_{j \in \mathcal{C}} \lambda_{ij} \mid \eta_i + \lambda_{ij} \geq y_{ij}, \forall j \in \mathcal{C} \right\}. \end{aligned}$$

The linear program is given by

$$\begin{aligned} \min \quad & \sum_{i \in \mathcal{F}} c_i x_i + k\mu + \sum_{j \in \mathcal{C}} \omega_j \\ \text{s.t.} \quad & \mu + \omega_j \geq \sum_{i \in \mathcal{F}} d_{ij} y_{ij}, & \forall j \in \mathcal{C}, \\ & \sum_{i \in \mathcal{F}} y_{ij} \geq 1, & \forall j \in \mathcal{C}, \\ & x_i \geq k\eta_i + \sum_{j \in \mathcal{C}} \lambda_{ij}, & \forall i \in \mathcal{F}, \\ & \eta_i + \lambda_{ij} \geq y_{ij}, & \forall i \in \mathcal{F}, \forall j \in \mathcal{C}, \\ & x_i, y_{ij}, \lambda_{ij}, \eta_i, \omega_j, \mu \geq 0, & \forall i \in \mathcal{F}, \forall j \in \mathcal{C}, \end{aligned} \tag{5}$$

which can be reduced, after removing the variables y_{ij} , to

$$\begin{aligned} \min \quad & \sum_{i \in \mathcal{F}} c_i x_i + k\mu + \sum_{j \in \mathcal{C}} \omega_j \\ \text{s.t.} \quad & \mu + \omega_j \geq \sum_{i \in \mathcal{F}} d_{ij} (\eta_i + \lambda_{ij}), & \forall j \in \mathcal{C}, \\ & \sum_{i \in \mathcal{F}} \eta_i + \lambda_{ij} \geq 1, & \forall j \in \mathcal{C}, \\ & x_i \geq k\eta_i + \sum_{j \in \mathcal{C}} \lambda_{ij}, & \forall i \in \mathcal{F}, \\ & x_i, \lambda_{ij}, \eta_i, \omega_j, \mu \geq 0, & \forall i \in \mathcal{F}, \forall j \in \mathcal{C}. \end{aligned} \tag{6}$$

Finally, we round the optimal solution of (Static-SCRFL) to an integral solution using the filtering and rounding techniques from Shmoys et al. [18] while losing only a factor of 12. Again, this rounding technique was designed for the deterministic facility location problem, but the same argument works as

well for (Static-SCRFL). Finally, since (Static-SCRFL) gives $O(\log k / \log \log k)$ -approximation to (LP-SCRFL) and we only loose a constant factor in the rounding, this results in $O(\log k / \log \log k)$ -approximation algorithm for (SCRFL). We state the result in the following theorem and, for completeness, we present the details of the rounding in Appendix B.

Theorem 4. *(Static-SCRFL) can be rounded to give $O(\frac{\log k}{\log \log k})$ -approximation algorithm to (SCRFL).*

Note that after rounding the supply in our solution of (Static-SCRFL), the integral second-stage assignment for each realized scenario is a transportation problem and therefore its optimal solution is integral. We would like to emphasize that while the fractional assignment in our solution is static, our integral assignment is not necessarily static. In fact, a policy with an integral static assignment could even be bad our model.

4 Conclusion.

In this paper, we give a $O(\log k / \log \log k)$ -approximation for soft-capacitated robust facility location problems with an implicit model of demand uncertainty. It is an interesting open question to study whether there exists a constant approximation algorithm for the problem, even in special cases such as the Euclidean metric. Our solution approach relies on static fractional assignment policies, which we show are optimal for the uncapacitated problem and give a strong theoretical guarantee for soft-capacitated case. Static assignment policies, while reasonable for the case of soft-capacities can be shown to be arbitrarily bad for the case of hard-capacities, where in addition to cost per unit, there is also an upper bound on supply at each facility. It is another interesting open direction to study any non-trivial approximation in this setting.

References

1. B. Anthony, V. Goyal, A. Gupta, and V. Nagarajan. A plant location guide for the unsure: Approximation algorithms for min-max location problems. *Mathematics of Operations Research*, 35(1):79–101, 2010.
2. A. Atamtürk and M. Zhang. Two-stage robust network flow and design under demand uncertainty. *Operations Research*, 55(4):662–673, 2007.
3. B. Basciftci, S. Ahmed, and S. Shen. Distributionally robust facility location problem under decision-dependent stochastic demand. *arXiv preprint arXiv:1912.05577*, 2019.
4. M. Charikar, S. Khuller, D. M. Mount, and G. Narasimhan. Algorithms for facility location problems with outliers. In *Proceedings of the Twelfth annual ACM-SIAM symposium on Discrete algorithms*, pages 642–651, 2001.
5. E. Delage and Y. Ye. Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations research*, 58(3):595–612, 2010.
6. K. Dhamdhere, V. Goyal, R. Ravi, and M. Singh. How to pay, come what may: approximation algorithms for demand-robust covering problems. In *46th Annual IEEE Symposium on Foundations of Computer Science (FOCS’05)*, pages 367–376, 2005.
7. U. Feige, K. Jain, M. Mahdian, and V. Mirrokni. Robust combinatorial optimization with exponential scenarios. In *International Conference on Integer Programming and Combinatorial Optimization*, pages 439–453. Springer, 2007.
8. S. Guha and S. Khuller. Greedy strikes back: Improved facility location algorithms. *Journal of Algorithms*, 31(1):228–248, 1999.
9. A. Gupta, V. Nagarajan, and R. Ravi. Thresholded covering algorithms for robust and max-min optimization. *Mathematical Programming*, 146(1-2):583–615, 2014.
10. A. Gupta, V. Nagarajan, and R. Ravi. Robust and maxmin optimization under matroid and knapsack uncertainty sets. *ACM Transactions on Algorithms (TALG)*, 12(1):10, 2016.
11. A. Gupta, M. Pál, R. Ravi, and A. Sinha. Boosted sampling: approximation algorithms for stochastic optimization. In *Proceedings of the Thirty-Sixth annual ACM Symposium on Theory of Computing*, pages 417–426, 2004.
12. O. E. Housni and V. Goyal. On the optimality of affine policies for budgeted uncertainty sets. *arXiv preprint arXiv:1807.00163*, 2018.
13. N. Immorlica, D. Karger, M. Minkoff, and V. S. Mirrokni. On the costs and benefits of procrastination: Approximation algorithms for stochastic combinatorial optimization problems. In *Proceedings of the Fifteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 691–700, 2004.
14. R. Khandekar, G. Kortsarz, V. Mirrokni, and M. R. Salavatipour. Two-stage robust network design with exponential scenarios. In *Proceedings of the 16th annual European Symposium on Algorithms*, pages 589–600, 2008.
15. A. Linhares and C. Swamy. Approximation algorithms for distributionally-robust stochastic optimization with black-box distributions. In *Proceedings of the 51st Annual ACM Symposium on Theory of Computing*, pages 768–779, 2019.
16. R. Ravi and A. Sinha. Hedging uncertainty: Approximation algorithms for stochastic optimization problems. In *International Conference on Integer Programming and Combinatorial Optimization*, pages 101–115. Springer, 2004.
17. D. B. Shmoys. Approximation algorithms for facility location problems. In *Proceedings of the Third International Workshop on Approximation Algorithms for Combinatorial Optimization*, pages 27–33, 2000.

18. D. B. Shmoys, É. Tardos, and K. Aardal. Approximation algorithms for facility location problems. In *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing*, pages 265–274, 1997.
19. L. V. Snyder. Facility location under uncertainty: a review. *IIE Transactions*, 38(7):547–564, 2006.
20. C. Swamy and D. B. Shmoys. Approximation algorithms for 2-stage stochastic optimization problems. *ACM SIGACT News*, 37(1):33–46, 2006.

A Rounding and proof of Theorem 3

Rounding (Shmoys et al.[18]). Recall that n is the number of facilities and m is the number of clients. Let $(\mathbf{x}^*, \mathbf{y}^*) = ((x_i^*)_{i \in \mathcal{F}}, (y_{ij}^*)_{i \in \mathcal{F}, j \in \mathcal{C}})$ be an optimal solution for (Static-SCRFL). Let OPT_1 and OPT_2 respectively be the corresponding optimal first-stage and second-stage cost. In particular, $\text{OPT}_2 = \max_{S \in \mathcal{C}^k} \sum_{i \in \mathcal{F}} \sum_{j \in \mathcal{C}} \mathbf{1}(j \in S) \cdot d_{ij} y_{ij}^*$. For each client $j \in \mathcal{C}$, let

$$L_j = \sum_{i \in \mathcal{F}} d_{ij} y_{ij}^*.$$

We sort L_j in ascending order. Suppose without loss of generality, that $L_1 \leq L_2 \leq \dots \leq L_m$. For each client $j \in \mathcal{C}$, we consider B_j the ball centered at the client j with radius αL_j (where $\alpha \geq 1$ will be defined later). We choose greedily the non-intersecting balls in ascending order of L_j and open the cheapest facility in each selected ball. That is, we consider every client in this sorted order, and only include its ball if it does not intersect any previously selected ball. Each client is assigned to its closest opened facility. For each client $j \in \mathcal{C}$,

$$L_j \geq \sum_{i \notin B_j} d_{ij} y_{ij}^* \geq \sum_{i \notin B_j} \alpha L_j y_{ij}^*.$$

Hence,

$$\frac{1}{\alpha} \geq \sum_{i \notin B_j} y_{ij}^*,$$

which implies

$$\sum_{i \in B_j} y_{ij}^* \geq 1 - \frac{1}{\alpha}.$$

Therefore,

$$\sum_{i \in B_j} x_i^* \geq 1 - \frac{1}{\alpha}.$$

By summing over the non-intersecting balls, the cost of our opened facilities is less than $\frac{1}{1-\frac{1}{\alpha}} \text{OPT}_1$. Now let us consider the assignment cost. For client j , if its ball B_j is chosen by the greedy procedure above, then the client is assigned to the opened facility i in B_j and $d_{ij} \leq \alpha L_j$. If not, the client j is assigned to a

facility i no further away than the one chosen in the ball B_k that intersected B_j . By the triangle inequality:

$$d_{ij} \leq d_{iv} + d_{jv} \leq \alpha L_j + 2\alpha L_k \leq 3\alpha L_j,$$

where $v \in B_j \cap B_k$ and $L_k \leq L_j$. Hence, for any scenario $S \in \mathcal{C}^k$, the second-stage cost is at most $3\alpha \cdot \sum_{j \in S} L_j$, which is at most $3\alpha \cdot \text{OPT}_2$. By optimizing over α , we take $\alpha = \frac{4}{3}$, which results in the 4-approximation with respect to $\text{OPT}_1 + \text{OPT}_2$. Moreover, since (Static-URFL) is equivalent to (LP-URFL) and (LP-URFL) is a relaxation of (URFL), the rounding above gives 4-approximation algorithm for (URFL).

B Rounding and proof of Theorem 4

Rounding (Shmoys et al.[18]). Let $(\mathbf{x}^*, \mathbf{y}^*) = ((x_i^*)_{i \in \mathcal{F}}, (y_{ij}^*)_{i \in \mathcal{F}, j \in \mathcal{C}})$ be an optimal solution of (Static-SCRFL). Let $z = \max_{S \in \mathcal{C}^k} \sum_{i \in \mathcal{F}, j \in S} d_{ij} y_{ij}^*$ and set $\alpha \in (0, 1)$. For each client $j \in \mathcal{C}$, let π be a permutation of facilities that serve j in the optimal solution of (Static-SCRFL) such that $d_{\pi(1)j} \leq d_{\pi(2)j} \leq \dots \leq d_{\pi(n)j}$. Define the radius $d_j(\alpha) = d_{\pi(i^*)j}$ where $i^* = \min\{i' \mid \sum_{i=1}^{i'} y_{\pi(i)j}^* \geq \alpha\}$. Following the same definitions in Shmoys et al. [18], we say that a solution (\mathbf{x}, \mathbf{y}) is g -close if for any $j \in \mathcal{C}$, $y_{ij} > 0$ implies that $d_{ij} \leq g_j$.

Claim. Given a feasible solution (\mathbf{x}, \mathbf{y}) , we can construct a g -close feasible solution $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ such that $\bar{\mathbf{x}} = \frac{1}{\alpha} \mathbf{x}$, $\bar{z} = \frac{1}{\alpha} z$ and $g_j \leq d_j(\alpha)$, $\forall j \in \mathcal{C}$ where $\bar{z} = \max_{S \in \mathcal{C}^k} \sum_{i \in \mathcal{F}, j \in S} d_{ij} \bar{y}_{ij}$.

Proof. In fact, for each j , we set $\bar{y}_{\pi(i)j} = y_{\pi(i)j}/\alpha$ for $1 \leq i \leq i^*$ and $\bar{y}_{\pi(i)j} = 0$ otherwise. We have

$$\sum_{i \in \mathcal{F}} \bar{y}_{ij} = \sum_{i=1}^{i^*} \frac{1}{\alpha} y_{\pi(i)j} \geq 1.$$

The second constraint in (Static-SCRFL) is trivially verified since x is multiplied by $\frac{1}{\alpha}$ and y_{ij} are either multiplied by $\frac{1}{\alpha}$ or set to 0. Moreover, all edges in our solution with a positive flow have a cost at most $d_j(\alpha)$. Hence, $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ is g -close with $g_j \leq d_j(\alpha)$ for all j . \square

Let $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ be the g -close solution corresponding to the optimal solution $(\mathbf{x}^*, \mathbf{y}^*)$. For the ease of notation, we just use the notation (\mathbf{x}, \mathbf{y}) instead of $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ in the rest of the proof. We round-up all $x_i \geq \frac{1}{2}$ to $\lceil x_i \rceil$. We pay at most a factor 2 in the first-stage cost. Let us focus on the remaining fractional facilities, say $\hat{\mathcal{F}}$, i.e.,

$$\hat{\mathcal{F}} = \{i \in \mathcal{F} \mid 0 < x_i < 1/2\}.$$

We let $\hat{\mathcal{C}}$ denote the set of clients such that half of their supply is coming from $\hat{\mathcal{F}}$, i.e.,

$$\hat{\mathcal{C}} = \{j \in \mathcal{C} \mid \sum_{i \in \hat{\mathcal{F}}} y_{ij} \geq \frac{1}{2}\}.$$

We sort the clients in $\hat{\mathcal{C}}$ in ascending order of g_j and do the following until $\hat{\mathcal{C}}$ becomes empty. We take the client j with the smallest g_j in $\hat{\mathcal{C}}$, let say j' . Let

$$V = \{i \in \hat{\mathcal{F}} : y_{ij'} > 0\},$$

and

$$T = \{j \in \mathcal{C} : \exists i \in V \text{ s.t. } y_{ij} > 0\}.$$

We put the $\lceil \sum_{i \in V} x_i \rceil$ units of supply in the cheapest facility in V which we denote f_c . We set each x_i in $V \setminus \{f_c\}$ to 0. The clients in T are the ones affected by this change. We will route all of their demand from V to f_c and make this assignment static. Note that, at this step, we have routed only their demand from V and not their entire demand. This is feasible because

$$\lceil \sum_{i \in V} x_i \rceil \geq \sum_{i \in V} x_i \geq \sum_{j \in \mathcal{C}} \mathbf{1}(j \in S) \cdot \sum_{i \in V} y_{ij}$$

for any scenario S . Since we choose the clients in $\hat{\mathcal{C}}$ in ascending order of g_j , the triangle inequality ensures that the solution is $3g$ -close. We keep doing this until $\hat{\mathcal{C}}$ becomes empty.

Let us analyze the final cost. We lost a factor $\frac{1}{\alpha}$ in both first-stage and second-stage cost. Then, we lost a factor 2 in the first-stage cost by rounding up the solution and moving the supply to the cheapest facilities after each rounding. We lost another factor 2 in the first stage to satisfy $\mathcal{C} \setminus \hat{\mathcal{C}}$. For the second stage, we have a solution that is $3g$ -close with $g_j \leq d_j(\alpha)$. Note that

$$d_j(\alpha) \leq \frac{1}{1-\alpha} \sum_{i \in \mathcal{F}} d_{ij} y_{ij}.$$

In particular, for any scenario S ,

$$\sum_{j \in S} g_j \leq \frac{1}{1-\alpha} \sum_{j \in S} \sum_{i \in \mathcal{F}} d_{ij} y_{ij}.$$

Hence, we loose $\frac{3}{1-\alpha}$ in the second stage. Overall, the factor is

$$\frac{4}{\alpha} \text{OPT}_1 + \frac{3}{\alpha(1-\alpha)} \text{OPT}_2.$$

We set $\alpha = \frac{1}{2}$, which gives 12 approximation to (LP-SCRFL). Finally, since (Static-SCRFL) gives $O(\log k / \log \log k)$ -approximation to (LP-SCRFL) and (LP-SCRFL) is a relaxation of (SCRFL), this results in $O(\log k / \log \log k)$ -approximation algorithm for (SCRFL).