Breaching the 2 LMP Approximation Barrier for Facility Location with Applications to k-Median[*]

Vincent Cohen-Addad Viallat[†]   Fabrizio Grandoni[‡]   Euiwoong Lee[§]

Chris Schwiegelshohn[¶]

Abstract

The Uncapacitated Facility Location (UFL) problem is one of the most fundamental clustering problems: Given a set of clients C and a set of facilities F in a metric space $(C \cup F, dist)$ with facility costs $\mathtt{open} : F \to \mathbb{R}^+$, the goal is to find a set of facilities $S \subseteq F$ to minimize the sum of the opening cost $\mathtt{open}(S)$ and the connection cost $d(S) := \sum_{p \in C} \min_{c \in S} dist(p, c)$. An algorithm for UFL is called a Lagrangian Multiplier Preserving (LMP) α approximation if it outputs a solution $S \subseteq F$ satisfying $\mathtt{open}(S) + d(S) \leq \mathtt{open}(S^*) + \alpha d(S^*)$ for any $S^* \subseteq F$. The best-known LMP approximation ratio for UFL is at most 2 by the JMS algorithm of Jain, Mahdian, and Saberi [STOC'02, J.ACM'03] based on the Dual-Fitting technique. The lack of progress on improving the upper bound on α_{LMP} in the last two decades raised the natural question whether $\alpha_{LMP} = 2$.

We answer this question negatively by presenting a (slightly) improved LMP approximation algorithm for UFL. This is achieved by combining the Dual-Fitting technique with Local Search, another popular technique to address clustering problems. In more detail, we use the LMP solution S produced by JMS to seed a local search algorithm. We show that local search substantially improves S unless a big fraction of the connection cost of S is associated with facilities of relatively small opening costs. In the latter case however the analysis of Jain, Mahdian, and Saberi can be improved (i.e., S is cheaper than expected). To summarize: Either S is close enough to the optimum, or it must belong to the local neighborhood of a good enough local optimum. From a conceptual viewpoint, our result gives a theoretical evidence that local search can be enhanced so as to avoid bad local optima by choosing the initial feasible solution with LP-based techniques.

Our result directly implies a (slightly) improved approximation for the related k-Median problem, another fundamental clustering problem: Given $(C \cup F, dist)$ as in a UFL instance and an integer $k \in \mathbb{N}$, find $S \subseteq F$ with $|S| = k$ that minimizes $d(S)$. The current best approximation algorithms for k-Median are based on the following framework: use an LMP α approximation algorithm for UFL to build an α approximate *bipoint solution* for k-Median, and then round it with a ρ_{BR} approximate bipoint rounding algorithm. This implies an $\alpha \cdot \rho_{BR}$ approximation. The current-best value of ρ_{BR} is 1.338 by Byrka, Pensyl, Rybicki, Srinivasan, and Trinh [SODA'15, TALG'17], which yields 2.6742-approximation. Combining their algorithm with our refined LMP algorithm for UFL (replacing JMS) gives a 2.67059-approximation.

---

1 Introduction

Given a set of points (clients) C and a set of facilities F in a metric space $(C \cup F, dist)$, the classic k-Median problem asks to find a subset $S \subseteq F$ of size k (the *centers* or *facilities*), such that the total distance $d(S) := \sum_{p \in C} \min_{c \in S} dist(p, c)$ of points in C to the closest facility is minimized. In the related Uncapacitated Facility Location problem (UFL), opening costs $\mathtt{open} : F \to \mathbb{R}^+$ replace k in the input, and we need to compute $S \subseteq F$ that minimizes $\mathtt{open}(S) + d(S)$, $\mathtt{open}(S) := \sum_{f \in S} \mathtt{open}(f)$. Initially motivated by operations research applications, the study of the complexity of these two problems dates back to the early 60s [KH63, Sto63] and has been the focus of a great deal of attention over the years. They are APX-hard in general metric spaces, and have been well-studied in terms of approximation algorithms. For UFL, Li [Li13] presented an algorithm with the currently best known approximation ratio of 1.488, while Guha and Khuller [GK99] show that it admits no better than an $\gamma_{GK} \approx 1.463$ approximation. For k-Median, while a hardness-of-approximation bound of $1 + 2/e \approx 1.73$ has been known for more than 20 years [GK99, JMS02], designing an algorithm matching this lower bound remains a major challenge despite ongoing efforts.

The most successful techniques to approximate k-Median and UFL are essentially of two kinds: Local search and LP-based techniques. Local search starts with a given feasible solution, and iteratively improves it by enumerating all "neighbouring" feasible solutions (obtained by swapping in and out a subset of at most $\Delta = O(1)$ centers) and switching to any such solution if the cost decreases. [AGK+04] showed that this approach achieves a $(3 + 2/\Delta)$ approximation for k-Median, and this factor is tight (more precisely, there exists a locally optimal solution which is $(3 + 2/\Delta)$ approximate). They also showed that local search achieves a $(1 + \sqrt{2}) \approx 2.41$ approximation for UFL. More recently, [CGH+22] showed that a non-oblivious local search can achieve a better than 3-approximation for k-median, yet the bound of $2.836 + \varepsilon$ achieved is larger than the 2.6742-approximation of [BPR+17].

There are two families of LP-based approximation algorithms. The first family is based on the direct rounding of a fractional solution to an LP relaxation (see e.g. [ARS03, CGTS02, BA10, Li13, CL12]). The second approach, which is the most relevant for this paper, is based on LMP approximations for UFL. Suppose that we are given a Lagrangian Multiplier Preserving (LMP) α_{LMP} approximation algorithm for UFL; the solution S produced by this algorithm satisfies $\mathtt{open}(S) + d(S) \leq \mathtt{open}(S^*) + \alpha_{LMP} \cdot d(S^*)$ for any feasible solution S^*.[1] For k-Median, there is a framework using *bipoint solutions* that takes any LMP α_{LMP} approximation algorithm for UFL and constructs an $\alpha_{LMP} \cdot \rho_{BR}$ approximation, where ρ_{BR} denotes the best known bipoint rounding ratio. (See Section 1.3 for details.) The fundamental paper by Jain and Vazirani [JV01] gives $\alpha_{LMP} = 3$ and $\rho_{BR} = 2$, leading to a 6 approximation for k-Median. The value of α_{LMP} was later improved to 2 in another seminal paper by Jain, Mahdian, and Saberi [JMS02, JMM+03], where the authors use the *Dual-Fitting* technique. While the lower bound remains $1 + 2/e \approx 1.73$ [JMS02], in the last two decades no progress was made on α_{LMP}. Recent progress was however made on ρ_{BR}: In a breakthrough result ρ_{BR} was improved to 1.3661 by Li and Svensson [LS16], and later refined to 1.3371 by [BPR+17]. [BPR+17] also shows that $\rho_{BR} \geq \frac{1+\sqrt{2}}{2} > 1.207$. Combining [JMS02] with [BPR+17] one obtains the current best $2 \cdot 1.3371 = 2.6742$ approximation for k-Median. There is also a simple reduction showing that an α LMP inapproximability for UFL with uniform facility costs implies an α inapproximability for k-Median for any $\alpha \geq 1$, so the question $\alpha_{LMP} < 2$ is directly related to whether k-Median would eventually admit a strictly better than 2 approximation. (The standard LP relaxation for k-Median has an integrality gap at least $2 - o(1)$ [JMS02].)

Bicriteria approximation algorithms for UFL, of which LMP approximation algorithms are an extreme case, play an important role for the approximation ratio of UFL. The best known 1.488 approximation algorithm for UFL by Li [Li13] uses the algorithms by Madhian, Ye, and Zhang [MYZ06] and Byrka and Aardal [BA10] that guarantee $\mathtt{open}(S) + d(S) \leq 1.11\mathtt{open}(S^*) + 1.78d(S^*)$ and $\mathtt{open}(S) + d(S) \leq 1.67\mathtt{open}(S^*) + 1.37d(S^*)$ for any S^* respectively. [MYZ06] also proves that the optimal LMP $(1 + 2/e)$ approximation automatically yields the optimal γ_{GK} approximation for UFL.

1.1 Our Results
Due to the lack of progress on improving α_{LMP} in the last two decades, it is natural to ask the following question:

---

[1]An equivalent definition also used in the literature is that $\alpha_{LMP}\mathtt{open}(S) + d(S) \leq \alpha_{LMP}(\mathtt{open}(S^*) + d(S^*))$. Two definitions are equivalent by scaling facility costs.

OPEN PROBLEM 1. *Is there an LMP $(2 - \eta)$ approximation for UFL for some constant $\eta > 0$?*

We answer the above question in the affirmative (though for a rather small $\eta > 2.25 \cdot 10^{-7}$), even for the general version of UFL where facilities can have different opening costs. The constant η can be substantially improved for the case of uniform opening costs, which is sufficient in our application to k-Median. Indeed, for k-Median some other technical refinements are possible, altogether leading to a more substantial (still small) improvement of k-Median approximation factor from 2.6742 to 2.67059[2].

While showing (though only qualitatively) that the LMP 2 approximation barrier can be breached is interesting in our opinion, probably the most interesting aspect of our work is at technical level. The dual-fitting LMP 2 approximation by [JMS02] is tight. In particular, a tight instance can be built via their factor-revealing LP. One might consider also a natural local-search-based algorithm for UFL, where one swaps in and out subsets of up to $\Delta = O(1)$ facilities: however (see Claim 3 in Section G) such an algorithm is not an LMP α_{LMP} approximation for any constant α_{LMP} (in the case of non-uniform opening costs).

Somewhat surprisingly, we show that a combination of dual-fitting and local search is better than each such approach in isolation. Consider first the case of uniform opening costs. We consider the UFL solution S produced by the algorithm JMS in [JMS02]. We seed the above mentioned local search UFL algorithm with S (i.e., S is used as the initial feasible solution), hence obtaining a refined solution S'. The latter solution turns out to be an LMP $2 - \eta < 2$ approximation for the problem. More precisely, we can show that either S is already a good enough solution, or it belongs to the local neighbourhood of a good enough solution. Intuitively, the worst case instances for JMS and local search are to a certain degree mutually exclusive. For general opening costs we augment the standard local-search-neighborhood by also considering the solutions obtained by setting to zero the cost of some facilities and running JMS itself on the obtained instance. We believe that our result is conceptually interesting: seeding a local search algorithm with an LP-based approximate solution is a natural heuristic; our result supports this heuristic from a theoretical viewpoint.

1.2 Related Work

k-Median. k-Median remains APX-hard [CK19, CCSL21, MS84] in Euclidean metrics, and here the best-known approximation is 2.406 [CEMN22]. The problem however admits a local-search-based PTAS for a constant number of dimensions [Coh18]. This extends to metrics with bounded doubling dimension [FRS19, CFS19], and graph metrics with an excluded minor [CKM19]. Local search computes optimal or nearly optimal solutions also assuming a variety of stability conditions [BW17, CS17] (see also [BBLM14, BT10, DGK02, HM01, CAD04, YSZ+08] for related clustering results). For fixed k, there are a few PTASs, typically using small coresets [BBC+19, FL11, HV20, KSS10]. There exists also a large body of bicriteria approximations, see [BV16, CG05, CM15, KPR00a, MMSW16].

A related result by Chen [Che08] also combines bipoint solutions and local-search-based approaches to obtain the first $O(1)$-approximation algorithm for k-Median with outliers (recently significantly improved by Krishnaswamy, Li, and Sandeep [KLS18] using a different method). In the k-Median with outliers problem the goal is to place k centers so as to minimize the sum of the distances from the closest $n - z$ points to their closest center, where z is an input integer designating the number of outliers. In his paper, Chen cleverly rounds a bipoint solution for the problem by using a local search algorithm for the k-Median problem with penalty (where the contribution of each point to the objective is the minimum between the distance to the closest center and the penalty) with increasing values of penalty. Though the focus of his local search is different from ours (we use local search to obtain better bipoint solutions instead of rounding them), we believe that these examples suggest that combining local search with other algorithms at various stages has a great potential to obtain improved results.

Facility Location. As mentioned previously, following the extensive literature on UFL [BA10, CG05, CS03, JMM+03, JMS02, KPR00a, MYZ06, STA97], the 1.488 approximation by Li [Li13] cleverly combines the algorithms of Mahdian, Ye, and Zhang [MYZ06] and Byrka and Aardal [BA10] that achieve different (α, β)-approximations; an algorithm for UFL (producing a solution S) is an (α, β) approximation if, for any feasible solution S', one has $\mathtt{open}(S) + d(S) \leq \alpha\mathtt{open}(S') + \beta d(S')$. In this language, [MYZ06] gives a $(1.11, 1.78)$ approximation and [BA10] gives a $(1.6774, 1.3738)$-approximation. Achieving better than $(\gamma, 1 + e^{-\gamma})$ is NP-hard [JMS02], and the latter result matches this lower bound for $\gamma = 1.6774$. Note that Guha and Khuller's lower bound $\gamma_{GK} \approx 1.463$ for UFL is the solution of $\gamma = 1 + e^{-\gamma}$.

---

[2]We did not insist too much on improving η, but we made a substantial effort to refine the approximation factor for k-Median.

There is also a large body of work in Capacitated Facility Location, where each facility location has a limited capacity [CW99, KPR00b, PTW01, ALB$^+$13, BGG12, ASS17]. After many constant factor approximation algorithms using local search, An, Singh, and Svensson [ASS17] presented the first constant factor approximation algorithm using a LP relaxation.

1.3 Preliminaries

Facility Location and JMS Algorithm. In the classical (metric uncapacitated) Facility Location problem (UFL) we are given a set of facilities F and a set of clients C together with a metric $dist(\cdot, \cdot)$ over $F \cup C$. Furthermore, we are given opening costs $\mathtt{open}(f) \geq 0$ over facilities $f \in F$. A feasible solution consists of a subset $S \subseteq F$ of *open* facilities. Let $\mathtt{open}(S) := \sum_{f \in S} \mathtt{open}(f)$ be the opening cost of the solution, and $d(S) := \sum_{j \in C} dist(j, S)$ be its connection cost, where $dist(j, S) := \min_{f \in S}\{dist(j, f)\}$. Intuitively, we open the facilities in S and connect each client j to the closest open facility $f \in S$ (paying the associated connection cost $dist(j, S)$). Our goal is to minimize the total cost $\mathtt{open}(S) + d(S)$. We will also consider the special case of UFL where all opening costs have uniform value λ. Recall that an algorithm for UFL (producing a solution S) is an (α, β) approximation if, for any feasible solution S', one has $\mathtt{open}(S) + d(S) \leq \alpha \mathtt{open}(S') + \beta d(S')$. If $\alpha = 1$, that algorithm is a *Lagrangian Multiplier Preserving* (LMP) β approximation for FL.

As mentioned in the introduction, our algorithm performs local search starting from the solution computed by the classical JMS algorithm [JMS02, JMM$^+$03]. Roughly speaking, the JMS algorithm works as follows. At any point of time we maintain a subset of *active* clients $A \subseteq C$ and a subset of open facilities $S \subseteq F$. Furthermore, for each $j \in C$, we have a variable $\alpha_j \geq 0$ and a facility $S(j) \in S$ which is defined only for *inactive* clients $j \in C \setminus A$ (j is *connected* or *assigned* to $S(j)$). Initially $A = C$, $S = \emptyset$ and $\alpha = 0$. At any point of time each client j makes an *offer* to each *closed* facility $f \in F \setminus S$. If j is active, this offer is $\max\{0, \alpha_j - dist(j, f)\}$. Otherwise ($j$ is inactive), the offer is $\max\{0, dist(j, S(j)) - dist(j, f)\}$. We increase at uniform rate the dual variables of active clients until one of the following events happens. If $\alpha_j = dist(j, f)$ for some $j \in A, f \in S$, we remove j from A and set $S(j) = f$ (j is connected to f). If the offers to a closed facility $f \in F \setminus S$ reach its opening cost $\mathtt{open}(f)$, we open f (i.e. add f to S), and set $S(j) = f$ for all the clients who made a strictly positive offer to f. Notice that $S(j)$ might change multiple times. We iterate the process until $A = \emptyset$, and then return S.

This algorithm is analyzed with the dual-fitting technique. We interpret the variables α as a dual solution for a proper standard LP (that we will omit). From the above construction it should be clear that the sum of the dual variables upper bounds the total cost of the final solution S. The dual solution turns out to be infeasible, but it can be made feasible by scaling down all the dual variables by a constant factor $\gamma > 1$ which also determines the approximation factor of the algorithm. The value of γ is determined by upper bounding the value of a properly defined *factor-revealing* LP, which is parameterized by an integer cluster size q. In particular the authors show that $\gamma < 1.61$. By modifying the objective function of such LP, one can also show that JMS is an (α, β) approximation algorithm for different pairs (α, β), and this is at the heart of the current best approximation algorithm for UFL [Li13]. In particular, we consider the following variant of a factor-revealing LP used in [JMS02]

to show that JMS is an LMP 2-approximation for UFL[3]:

$$\max \sum_{i=1}^{q} \alpha_i - \lambda \qquad\qquad (LP_{JMS(q)})$$

(1.1) $$\text{s.t.} \sum_{i=1}^{q} d_i = 1$$

(1.2) $$\alpha_i \leq \alpha_{i+1} \qquad\qquad \forall 1 \leq i < q$$

(1.3) $$r_{j,i+1} \leq r_{j,i} \qquad\qquad \forall 1 \leq j \leq i < q$$

(1.4) $$\alpha_i \leq r_{j,i} + d_i + d_j \qquad\qquad \forall 1 \leq j < i \leq q$$

(1.5) $$r_{j,j} \leq \alpha_j \qquad\qquad \forall 1 \leq j \leq q$$

(1.6) $$\sum_{j=1}^{i-1} \max\{r_{j,i} - d_j, 0\} + \sum_{j=i}^{q} \max\{\alpha_i - d_j, 0\} \leq \lambda \qquad\qquad \forall 1 \leq i \leq q$$

$$\alpha, d, \lambda, r \geq 0$$

In the above LP, for $1 \leq j < i \leq q$, $r_{j,i}$ is interpreted as the distance between client j and the facility it is connected to right before time α_i, assuming $\alpha_j < \alpha_i$. Otherwise, $r_{j,i} = \alpha_j$. Equation (1.1) is the normalizing constraint so that the objective value of the program directly becomes the LMP approximation ratio, (1.2) orders the clients in the increasing order of α_i's, and (1.3) indicates that a client will be connected to closer facilities as time progresses. W.r.t. the LP in [JMS02], we add the variables $r_{j,j}$ which denote the distance between client j and the first facility it is connected to. W.r.t. to the LP studied in [JMS02], we add the constraint $r_{j,j+1} \leq r_{j,j}$ in (1.3) and the new constraint (1.5), namely $r_{j,j} \leq \alpha_j$. Both constraints are clearly satisfied by the *original* algorithm in [JMS02]. Equation (1.4) uses the triangle inequality to show that client i can be connected to the facility client j is already connected to, and (1.6) shows that the current facility is never overpaid (i.e., it becomes open as soon as the equality holds). Let $opt_{JMS}(q)$ be the optimal solution to $LP_{JMS}(q)$, and $opt_{JMS} = \sup_q\{opt_{JMS}(q)\}$. The analysis in [JMS02] directly implies $opt_{JMS} \leq 2$ (since we only added constraints). We will use a variant of the above LP to prove our improved LMP bound for UFL[4].

Given a facility location solution S and a client j, we let $S(j)$ denote the facility serving client j in solution S. We also let $C_S(f)$ be the clients served by facility $f \in S$ in solution S. For $S' \subseteq S$, $C_S(S') = \cup_{f \in S'} C_S(f)$. For $f \in S$ and $S' \subseteq S$, we let $d_S(f) := \sum_{j \in C_S(f)} dist(j, f)$ and $d_S(S') = \sum_{f \in S'} d_S(f)$. In all the mentioned cases, we sometimes omit the subscript S when S is clear from the context.

Bipoint solution. Consider a k-Median instance $(C, F, dist, k)$ with an optimal solution OPT of cost opt. Let us assume w.l.o.g. that $k < |F|$, otherwise the problem can be solved optimally in polynomial time by opening the entire F. Given an LMP α_{LMP} approximation for UFL, a bipoint solution is constructed as follows. For a given parameter $\lambda \geq 0$, consider the UFL instance with clients C, facilities F, and uniform opening cost λ. Let $S(\lambda) \subseteq F$ be the solution produced by the considered UFL algorithm on this instance. We assume that $S(\lambda')$ consists of precisely one facility for a large enough λ', and that $S(0) = F$. These properties are either automatically satisfied by the UFL algorithm, or they can be easily enforced. We perform a binary search in the interval $[0, \lambda']$ until we find two values $0 \leq \lambda_2 < \lambda_1 \leq \lambda'$ such that $S_1 := S(\lambda_1)$ opens $k_1 \leq k$ facilities and $S_2 := S(\lambda_2)$ opens $k_2 > k$ facilities. Furthermore, $(\lambda_1 - \lambda_2)k \leq \varepsilon \cdot opt$. Interpreting S_i as an incidence vector over F, the bipoint solution $S_B = aS_1 + bS_2$ is a convex combination of S_1 and S_2 with coefficients $a = \frac{k_2 - k}{k_2 - k_1}$ and $b = 1 - a = \frac{k - k_1}{k_2 - k_1}$. Notice that S_B opens precisely k facilities in a fractional sense. In S_B each client c is connected by an amount a to $S_1(c)$ and by an amount b to $S_2(c)$, hence the connection cost of S_B is $a\, d(S_1) + b\, d(S_2)$. Recall that by assumption $\lambda_i |S_i| + d(S_i) \leq \lambda_i k + \alpha_{LMP} \cdot opt$. Thus

$$a\, d(S_1) + b\, d(S_2) \leq a\lambda_1(k - k_1) + b\lambda_2(k - k_2) + \alpha_{LMP} \cdot opt \leq (\alpha_{LMP} + \varepsilon)opt.$$

---

[3]The normalization constraint (1.1) can be omitted by dividing the objective function by $\sum_{j=1}^{q} d_j$. The maxima in (1.6) can be avoided by introducing auxiliary variables.

[4]We remark that JMS produces an LMP 2-approximation w.r.t. the value of a standard LP, while our approach provides an LMP $2 - \eta$ approximation but only w.r.t. the optimal integral solution (due to our use of local search). Finding an LMP $2 - \eta$ approximation w.r.t. a natural LP is an interesting open problem.

We will next neglect the term εopt since it is subsumed by our numerical overestimates. Furthermore, to lighten the notation, we will neglect the slight gap between λ_1 and λ_2 and simply assume $\lambda_1 = \lambda_2 = \lambda$.

Organization. To illustrate (some of) our main ideas in a simpler setting, in the main body we will focus on approximation algorithms for k-Median. To that aim, it is sufficient to consider LMP approximation algorithms for UFL *in the case of uniform opening cost λ* (which in turn is sufficient to build a convenient bipoint solution). As mentioned earlier, our UFL algorithm starts with the solution produced by JMS, and then refines it by means of local search. In Sections 2 and 3 (with some proofs in Sections A and C, resp.) we provide two distinct upper bounds on the cost of the produced UFL solution. The first such bound is based on a refinement of the analysis in [JMS02] which holds in certain special cases (see Section B for an alternative analytical approach). The second bound instead exploits the local optimality of the final solution. Combining these two bounds, we derive an improved approximation for k-Median in Section 4 (with some proofs in Section D). In Section E we describe how to obtain a better than 2 LMP approximation for UFL with uniform facility costs. A refined approximation for k-Median is sketched in Section F (using results from Section E.1). Finally our improved LMP approximation for UFL with arbitrary opening costs is given in Section G.

2 An Improved LMP Bound for Small Facility Cost

Given a Facility Location instance $(C, F, dist, \mathsf{open})$, let $OPT \subseteq F$ be some given solution to a UFL instance. Suppose that a constant fraction $\alpha > 0$ of the connection cost of OPT is associated with a subset $OPT' \subseteq OPT$ of facilities whose opening cost $\mathsf{open}(OPT')$ is at most some constant T times the respective connection cost $d(OPT')$. Then we can show that JMS on this instance produces a solution of cost at most $\mathsf{open}(OPT) + (2 - \eta)d(OPT)$, for some constant $\eta = \eta(T, \alpha) > 0$. Hence in some sense JMS is an LMP $(2 - \eta)$ approximation algorithm w.r.t. this type of solutions. This result will be used in our overall win-win strategy for k-Median; later in Theorem 3.1 and Lemma 4.2, we will show that if an outcome of local search with mild additional requirements (satisfied by starting local search from a JMS solution) does not achieve a $(2 - \eta')$ LMP approximation for some $\eta' > 0$, then OPT will have the above property, so JMS already guarantees a $(2 - \eta)$ LMP approximation.

In order to prove this claim, let us have a closer look at the analysis of JMS in [JMS02]. Using the dual-fitting technique, the authors show that, for any $f \in OPT$, $\sum_{j \in C(f)} \alpha_j \leq \mathsf{open}(f) + opt_{JMS} \cdot d(f) \leq \mathsf{open}(f) + 2d(f)$. The worst case value of opt_{JMS} in their analysis is achieved for $\lambda = \mathsf{open}(f)$ arbitrarily larger than $d(f) = \sum_{j=1}^{q} d_j = 1$. However for the facilities in OPT' this cannot happen on average since $\sum_{f \in OPT'} \mathsf{open}(f) = \mathsf{open}(OPT') \leq T \cdot d(OPT') = T \cdot \sum_{f \in OPT'} d(f)$. This motivated us to study a variant $LP_{JMS}(q, T)$ of $LP_{JMS}(q)$, for a parameter $T > 0$, where we add the following constraint that intuitively captures the condition $\mathsf{open}(f) \leq T \cdot d(f)$

$$(2.7) \qquad\qquad\qquad\qquad \lambda \leq T.$$

Let $opt_{JMS}(q, T)$ be the optimal value of $LP_{JMS}(q, T)$, and $opt_{JMS}(T) = \sup_q\{opt_{JMS}(q, T)\}$. Observe that $opt_{JMS}(q, T) \leq opt_{JMS}(q)$ where the equality holds for any fixed q and T large enough. A straightforward adaptation of the analysis in [JMS02] using our modified LP $LP_{JMS}(q, T)$ implies the following.

LEMMA 2.1. *Let OPT be a solution to a UFL instance and S be the JMS solution on the same instance. Then*

$$\mathbf{\textit{open}}(S) + d(S) \leq \mathbf{\textit{open}}(OPT) + \sum_{f^* \in OPT} opt_{JMS}\left(\frac{\mathbf{\textit{open}}(f^*)}{d(f^*)}\right) \cdot d(f^*).$$

Given Lemma 2.1 and the above discussion, it makes sense to study the behaviour of $opt_{JMS}(T)$ for bounded values of T. We can show that the supremum defining $opt_{JMS}(T)$ is achieved for q going to infinity, i.e. $opt_{JMS}(T) = \lim_{q \to +\infty} opt_{JMS}(q, T)$:

LEMMA 2.2. *Fix $T > 0$. For any positive integers q and c, $opt_{JMS}(q, T) \leq opt_{JMS}(cq, T)$.*

Proof. [Proof sketch] Given an optimal solution $S = (\alpha, d, r, \lambda)$ of $LP_{JMS}(q, T)$, consider the following solution $S' = (\alpha', d', r', \lambda')$ of $LP_{JMS}(cq, T)$:

$$\alpha'_i := \frac{\alpha_{\lceil i/c \rceil}}{c}, \quad d'_i = \frac{d_{\lceil i/c \rceil}}{c}, \quad r'_{j,i} := \begin{cases} \frac{r_{\lceil j/c \rceil, \lceil i/c \rceil}}{c} & \text{if } \lceil j/c \rceil < \lceil i/c \rceil; \\ \frac{\alpha_{\lceil j/c \rceil}}{c} & \text{if } \lceil j/c \rceil = \lceil i/c \rceil. \end{cases} \quad \lambda' = \lambda, \quad \forall 1 \leq j \leq i \leq cq.$$

Observe that S and S' have exactly the same objective value. The reader can easily verify that S' is feasible (details in Section A). The claim follows. □

We also observe that $opt_{JMS(T)}$ is a concave function.

LEMMA 2.3. $opt_{JMS}(T)$ *is a concave function of* T.

We next show how to compute an upper bound on $opt_{JMS}(T)$ for any given T. Here we slightly adapt an approach presented in [BPR$^+$15], which avoids to construct a family of feasible dual solutions as in [JMS02][5]. This leads to a simpler and easier to adapt analysis.

We define an alternative LP $LP_{JMS}^+(q, T)$ which is obtained from $LP_{JMS}(q, T)$ by replacing Constraint (1.6) with the following constraint

$$(2.8) \qquad \sum_{j=1}^{i} \max\{r_{j,i} - d_j, 0\} + \sum_{j=i+1}^{q} \max\{\alpha_i - d_j, 0\} \leq \lambda, \qquad \forall 1 \leq i \leq q.$$

In particular, there is a shift by one of two indexes in the two sums. Let $opt_{JMS}^+(q, T)$ be the optimal solution to $LP_{JMS}^+(q, T)$. The next lemma follows analogously to [BPR$^+$15].

LEMMA 2.4. *Fix* $T > 0$. *For any positive integers* q *and* c, $opt_{JMS}^+(q, T) \geq opt_{JMS}(cq, T)$.

Proof. [Proof sketch] Let $S := (\alpha, d, r, \lambda)$ be an optimal solution to $LP_{JMS}(cq, T)$. Consider the following solution $S^+ := (\alpha^+, d^+, r^+, \lambda^+)$ for $LP_{JMS}^+(q, T)$:

$$\alpha_i^+ := \sum_{\ell=(i-1)c+1}^{ic} \alpha_\ell, \quad d_i^+ = \sum_{\ell=(i-1)c+1}^{ic} d_\ell, \quad r_{j,i}^+ := \sum_{\ell=(j-1)c+1}^{jc} r_{\ell, ic}, \quad \lambda^+ = \lambda, \quad \forall 1 \leq j \leq i \leq q.$$

Clearly the objective values of S and S^+ are identical. The reader can easily check that S^+ is feasible (details in Section A). The claim follows. □

LEMMA 2.5. *Fix* $T > 0$. *For any positive integer* q, $opt_{JMS}^+(q, T) \geq opt_{JMS}(T)$.

Proof. Suppose by contradiction that the claim does not hold. In particular, for some constant $\varepsilon > 0$, one has $opt_{JMS}^+(q, T) \leq opt_{JMS}(T) - \varepsilon$. There must exist a finite integer value c such that $opt_{JMS}(c, T) > opt_{JMS}(T) - \varepsilon$. Hence we obtain the desired contradiction by Lemmas 2.2 and 2.4:

$$opt_{JMS}^+(q, T) \overset{\text{Lem. 2.4}}{\geq} opt_{JMS}(cq, T) \overset{\text{Lem. 2.2}}{\geq} opt_{JMS}(c, T) > opt_{JMS}(T) - \varepsilon \geq opt_{JMS}^+(q, T).$$

□

Lemma 2.5 provides a convenient way to upper bound $opt_{JMS}(T)$. Though we will not prove this formally, an empirical evaluation shows that $opt_{JMS}^+(q, T)$ is a decreasing function of q which (rather quickly) converges to $opt_{JMS}(T)$. Hence, for a given T, a very tight upper bound on $opt_{JMS}(T)$ is obtained by computing $opt_{JMS}^+(q, T)$ for a large enough but fixed value of q (in the order of a few hundreds). In the following we will use q as a parameter that we will fix only in the final numerical computation. We are also able to prove an analytical (though weaker) upper bound on $opt_{JMS}(T)$. Though this is not needed for the rest of our analysis, we present the latter upper bound in Section B.

The next corollary follows from Lemmas 2.1, 2.3, and 2.5 (proof in Section A).

COROLLARY 2.1. *Let* OPT *be a solution to a UFL instance and* S *be the JMS solution on the same instance. For a given constant* $T > 0$, *let* $OPT' \subseteq OPT$ *be a subset of facilities satisfying* $\mathbf{open}(OPT') \leq T \cdot d(OPT')$. *Then, for every integer* $q > 0$,

$$\mathbf{open}(S) + d(S) \leq \mathbf{open}(OPT) + 2d(OPT \setminus OPT') + opt_{JMS}^+(q, T) \cdot d(OPT').$$

---

[5][BPR$^+$15] contains a technical bug which was later fixed in the journal version [BPR$^+$17] (leading to a slightly worse approximation factor). Here we adapt a part of the analysis in [BPR$^+$15] which does not appear in [BPR$^+$17] (nor in the most recent arXiv version).

3 A Local-Search-Based Bound

In this section we present our local-search-based bound for UFL. Suppose that we start with a facility location solution S' for some given uniform facility opening cost λ. (In our case S' will be produced by the JMS algorithm, though the result of this section holds for an arbitrary starting solution S'.) We refine S' via local search as follows. Let $\varepsilon > 0$ be a constant parameter. A feasible *swap pair* (A, B) consists of $A \subseteq S'$ and $B \subseteq F$ such that $|A|, |B| \leq \Delta$. Here $\Delta \geq 1$ is a constant depending on ε to be fixed later. For any feasible swap pair, we consider the alternative solution $S'' = S' \setminus A \cup B$ and if S'' has total (i.e., facility plus connection) cost strictly smaller than S', we replace S' with S'' and repeat. Lemma C.1 shows that with a loss of $1 + \varepsilon$ in the approximation factor, we can guarantee that the local-search procedure ends after a polynomial number of rounds. For the sake of clarity we simply assume here and omit the extra factor $1 + \varepsilon$ since it is subsumed by similar factors in the rest of our analysis.

Let S' be the solution at the end of the process (i.e., S' is a *local optimum*), with $k' = |S'|$ and $d' = d(S')$. Let also OPT denote some given UFL solution, with $k = |OPT|$ and connection cost $opt = d(OPT)$. Intuitively, for us OPT will be the optimal solution to the underlying k-Median problem. We will next describe a bound on the total cost of S' which is based on a carefully chosen classification of the facilities in S' and OPT.

Matched vs. Lonely facilities: a key notion. Given two facility location solutions S and S' and $\alpha \geq 1/2$, we say that $f \in S$ α-captures $f' \in S'$ if more than an α fraction of the clients served by f' in S' are served by f in S, i.e. $|C_S(f) \cap C_{S'}(f')| > \alpha |C_{S'}(f')|$. Since $\alpha \geq 1/2$, $f' \in S'$ can be captured by at most one $f \in S$. Similarly, we say that $A \subseteq S$ α-captures $f' \in S'$ if more than an α fraction of the clients served by f' in S' are served by some $f \in A$ in S.

We classify the facilities in S' and OPT as lonely and matched as follows. Let $\delta \leq 1/2$ be a parameter to be fixed later. Suppose first that $|S'| = k' \leq k = |OPT|$ (intuitively corresponding to $S' = S_1$). In this case, consider $f^* \in OPT$, and let $M(f^*)$ be the facilities in S' which are $1/2$-captured by f^*. If $M(f^*)$ $(1-\delta)$-captures f^*, then we say that f^* and $M(f^*)$ are *matched*. In the complementary case $k' > k$ (intuitively corresponding to $S' = S_2$) we use a symmetric definition. Let $f' \in S'$, and let $M(f')$ be the facilities in OPT which are $(1-\delta)$-captured by f'. If $M(f')$ $1/2$ captures f' we say that f' and $M(f')$ are *matched*.

All the facilities which are not matched are *lonely*. We use OPT^M and OPT^L to denote the facilities in OPT which are respectively matched and lonely. We define similarly S^M and S^L w.r.t. to S'. By opt^X, $X \in \{M, L\}$, we denote the connection cost of clients served by OPT^X in OPT. Similarly, by opt^{XY}, $X, Y \in \{M, L\}$, we denote the connection cost of clients served by S^X in S' and by OPT^Y in OPT. We define d^{XY} analogously w.r.t. the connection cost of S'.

High-level overview. At a high-level, pairs of facilities of S' and OPT induce *well-served* clients in the solution output by local search. Indeed, think about the extreme scenario where there is one facility of $f' \in S'$ and one facility of $f^* \in OPT$ that serve the exact same group of clients. Then, by local optimality, one could argue that this group of clients is served optimally in S' since the swap that closes down f' and opens f^* does not decrease the cost. Thus generalizing this idea, our analysis shows that clients served by matched facilities of OPT are served nearly-optimally (i.e,: much better than the 2 approximation of JMS) while the remaining clients are served similarly than in the classic local search analysis (namely within a factor 3 of the optimum).

Our proof then decouples the matched and lonely facilities. We provide an analysis of the local search solution that is specific to the matched facilities, and use the analysis of Gupta and Tangwongsan [GT08]. We then show how to mix the two analysis. Concretely, the proof works as follows. We define feasible swap pairs (A_i, B_i) so that the A_i's and B_i's are carefully chosen subsets of S' and OPT. The goal is thus to bound the variations Δ_i defined by the cost of solution $S' - A_i \cup B_i$ minus the cost of S'. Local optimality of S' implies that Δ_i is non-negative for all i, providing us with an upper bound on the cost of S'. The challenge is thus to provide the best possible upper bound on the cost of $S' - A_i \cup B_i$, in terms of the cost of OPT and S'. The usual challenge here is to bound the *reassignment* cost of the clients served by a facility of A_i in S' but not served by a facility of B_i in OPT.

The swap pairs are defined so that matched facilities belong to the same swap pair (when one facility is matched to a large number of facilities, we carefully break them into multiple swap pairs to have the same effect). This way, for matched facilities $f' \in S'$ and $f^* \in OPT$, the clients served by f' in S' and f^* in OPT (call them *matched* clients) do not induce a reassignment cost – they are served optimally in the swap pair involving f' (and so f^*). The clients of f^* that are not served by f' still benefit from being served by a matched facility f^*. Any such client served by a facility $f'' \in S'$ can be reassigned to f' when f'' is swapped out. The reassignment cost

in this case can be charged to the matched clients; since the clients of f^* served by a facility different from f' are in minority, the charge is tiny.

The swap pairs are also built in a way that lonely facilities of OPT belong to the same swap with their closest facility in S', following the approach of Gupta and Tangwongsan [GT08]. This ensures that the reassignment cost is overall bounded by a factor 3 (as shown in [GT08]). Building the swap pairs satisfying the above two properties requires a more careful approach than in [GT08] and is an important step in our proof.

We are ready to state our local-search based bound (with proof in Section C).

THEOREM 3.1. *For* $\Delta = (2\varepsilon^{-2\varepsilon^{-7}})$, *one has*

$$\lambda k' + d' \leq \lambda k + 3opt^L + opt^M + \frac{\delta}{1-\delta}(d^{MM} + opt^{MM}) + O(\varepsilon(d' + opt)).$$

A high-level corollary of the above theorem is that, if most of the connection cost of OPT is due to clients served by matched facilities (and for δ small enough), then S' provides an LMP α_{LMP} approximation (w.r.t. to OPT) for some $\alpha_{LMP} < 2$ (hence improving on the JMS bound). Since this result is an intrinsic property of local search and does not use the properties of the starting solutions (e.g., those produced by the JMS algorithm), we hope that it may be useful in the future to show that local search can improve the quality of the solution of other algorithms.

Handling non-uniform facility opening costs via local search requires a new notion of local neighborhood because the disparity in facility costs forces some swap pairs (A_i, B_i) to contain a superconstant number of facilities to result in a meaningful guarantee. For general UFL, we introduce a new local search algorithm where each local move closes a constant number of open facilities and runs a variant of the JMS algorithm to find a new solution that can possibly open a large number of new facilities. Specifically, we set the opening cost of the currently open facilities to zero, and run the JMS algorithm to possibly open more facilities. Theorem G.2 provides a guarantee for this new local search; essentially the guarantee of Theorem 3.1 holds except $3opt^L$ is replaced by $4opt^L$ due to additional approximations created by the use of JMS. In our opinion defining a local neighbourhood via an LP-based algorithm is a promising approach which deserves further investigation. See Section G.1 for the new local search and its guarantee.

4 An Improved Approximation for k-Median

In this section we present the claimed improved approximation algorithm for k-Median. A refinement (for which we did not compute the explicit approximation factor) is sketched in Section F.

We consider the LMP UFL algorithm described in previous sections, where we start with the solution produced by the JMS algorithm and then apply local search to improve it. We use this algorithm to build a bipoint solution $S_B = aS_1 + bS_2$ via binary search as described in Section 1.3. Let $k_i = |S_i|$ and $d_i = d(S_i)$. Recall that $k_1 \leq k < k_2$, $1 - b = a = \frac{k_2 - k}{k_2 - k_1}$, and $ad_1 + bd_2 \leq 2opt$. Since S_1 and S_2 are the outcomes of local search starting from solutions produced by the JMS algorithm, their total costs ($\lambda k_1 + d_1$ for S_1 and $\lambda k_2 + d_2$ for S_2) are at most the total cost of the starting JMS solutions, which are at most $\lambda k + 2opt$ by the standard 2 LMP approximation guarantee. Furthermore, since $k_2 \geq k$, one can conclude that $d_2 \leq 2opt$.

Based on the results from Sections 2 and 3, we will prove the following improved upper bound on the total cost of S_2.

LEMMA 4.1. $\lambda k_2 + d_2 \leq \lambda k + (2 - \eta_2)opt$ *for some absolute constant* $\eta_2 > 0.00536$.

Using $\lambda k_1 + d_1 \leq \lambda k + 2opt$ with Lemma 4.1, we derive an improved bound on the connection cost of S_B:

$$ad_1 + bd_2 \leq (2 - (1-a)\eta_2)opt.$$

Thus combining the bipoint solution S_B with the $\rho_{BR} < 1.3371$ approximate bipoint rounding procedure in [BPR$^+$17], one obtains a feasible solution with the approximation factor

$$\rho_{BR}(2 - (1-a)\eta_2).$$

For a bounded away from 1, this is clearly an improvement on the $2\rho_{BR} < 2.6742$ approximation achieved in [BPR$^+$17]. (Indeed, in Section E, we also prove $\lambda k_1 + d_1 \leq \lambda k + (2 - \eta_1)opt$ for some $\eta_1 > 0$ as well, which

establishes $ad_1 + bd_2 < (2-\eta)opt$ for some $\eta > 0$ for all values of a.) To get an improvement only from Lemma 4.1, Li and Svensson [LS16] present a $2(1+2a) + \varepsilon$ approximation for any constant $\varepsilon > 0$ (with extra running time $n^{\mathrm{poly}(1/\varepsilon)}$), hence we can also assume that a is bounded away from 0 (since we aim at an approximation factor strictly larger than 2). For a bounded away from 0 and 1, the same authors present a $\frac{2(1+2a)}{(1+2a^2)} + \varepsilon$ approximation that we will next use. Taking the best of the two mentioned solutions, and neglecting the ε in the approximation factor, one obtains the following approximation factor for k-Median

$$\rho_{kMed} = \max_{a \in [0,1]} \min\{\frac{2(1+2a)}{1+2a^2}, \rho_{BR}(2 - (1-a)\eta_2)\}.$$

The first term in the minimum is decreasing for $a \geq \frac{\sqrt{3}-1}{2}$, hence $\rho_{kMed} < 2\rho_{BR}$, thus improving on [BPR$^+$17]. We numerically[6] obtained that the worst case is achieved for $a \simeq 0.4955391$, leading to $\rho_{kMed} < 2.67059$.

4.1 Proof of Lemma 4.1 We next use S_2^M instead of S^M and similarly for related quantities to stress that we are focusing on the locally optimal solution $S' = S_2$. We let $\alpha^L := opt^L/opt$, $\alpha^M := opt^M/opt$, $\alpha^{MM} = opt^{MM}/opt$, $\beta_2 = d_2/opt$, and $\beta_2^{MM} = d_2^{MM}/opt$. Let also $k^L = |OPT^L|$, $k^M = |OPT^M|$, $k_2^L = |S_2^L|$, and $k_2^M = |S_2^M|$.

We apply to S_2 the local-search-based bound from Theorem 3.1 in Section 3[7]. In the following we will neglect the term depending on ε: indeed the latter parameter can be chosen arbitrarily small, independently from the other parameters in the proof (at the cost of a larger, yet polynomial running time). Taking that term into account would involve an extra additive term $O(\varepsilon)$ in the approximation factor, which is however absorbed by the overestimations in our numerical analysis. Thus we obtain

$$(4.9) \qquad \lambda k_2 + d_2 \leq \lambda k + \big(1 + 2\alpha^L + \frac{\delta}{1-\delta}(\beta_2^{MM} + \alpha^{MM})\big) \cdot opt =: \lambda k + \rho^A \cdot opt.$$

Observe that the above bound already implies Lemma 4.1 when α^L and δ are sufficiently small (notice that $\beta_2^{MM} \leq \beta_2 \leq 2$). We next derive an alternative bound which implies the claim in the complementary case. Here we exploit the LMP result from Section 2. More specifically, we will show that the facility cost $\lambda|OPT^L|$ due to lonely facilities in OPT is upper bounded by $O(opt^L)$. This will allow us to exploit Corollary 2.1. By construction for each $f' \in S_2^M$, there is at least one distinct $f^* \in OPT^M$, hence $k_2^M \leq k^M$. Thus

$$(4.10) \qquad k_2^L = k_2 - k_2^M \geq k_2 - k^M = k_2 - k + k^L.$$

The local optimality of S_2 implies the following bound (proof in Section D):

LEMMA 4.2. *One has*

$$(4.11) \qquad \lambda k_2^L \leq \big(\frac{2}{\delta}(1 - \alpha^{MM}) + 2\frac{1-\delta}{\delta}(\beta_2 - \beta_2^{MM}) + 2\frac{\delta}{1-\delta}(\beta_2^{MM} + \alpha^{MM})\big) \cdot opt.$$

The idea of the proof is that removing any lonely facility f' from S_2 determines an increase of the connection cost of the clients $C_{S_2}(f')$ which is at least equal to the facility cost λ. It is therefore sufficient to upper bound the cost of connecting each $c \in C_{S_2}(f')$ to some facility $f'' \in S_2 \setminus \{f\}$. Here we exploit the fact that f' is lonely. This implies that if f' $(1-\delta)$-captures certain facilities $M(f')$ in OPT, the latter facilities do not $1/2$-capture f'. In particular, at least one half $C'(f')$ of the clients $C_{S_2}(f')$ are served in OPT by facilities not captured by f'. The idea is then to define a path that goes from c to some $c' \in C'(f')$, from there to the facility $f^* \in OPT$ serving c' and from f^* to the desired facility f'' (which in particular exists since f^* is not captured by f'). This idea is inspired by PTASes for k-Median and k-Means in constant-dimensional Euclidean spaces [CKM19] where the authors remove lonely facilities to convert a bicriteria approximation to a true approximation. Also note that while Lemma 4.2 uses local optimality as Theorem 3.1 and holds for any local optimum as well, the upper bound on λk_2^L (i.e., the RHS of (4.11)) depends on $\beta_2 = d_2/opt$, which can be upper bounded by 2 using the fact that the starting solution is already 2 LMP approximate and $k_2 \geq k$.

---

[6]The worst case can be computed analytically, but the formula is complicated and we omit it.

[7]Theorem 3.1 is proved in Section C. However for this part of the analysis it is sufficient to consider the simpler subcase $|S'| > k$ discussed in Section C.1.

We are ready to prove our second upper bound on the total cost of S_2. One has

$$\lambda k^L \overset{(4.10)}{\leq} \lambda k_2^L \overset{Lem.4.2}{\leq} \frac{2}{\delta \alpha^L}\Big(1 + (1-\delta)\beta_2 - (1 - \frac{\delta^2}{1-\delta})\alpha^{MM} - (1 - \frac{\delta}{1-\delta})\beta_2^{MM}\Big)opt^L =: T_L opt^L.$$

We can thus apply Corollary 2.1 with $OPT' = OPT^L$ and $T = T_L$ to infer

$$(4.12) \qquad \lambda k_2 + d_2 \leq \lambda k + (2 \cdot (1-\alpha^L) + opt_{JMS}^+(q, T_L) \cdot \alpha^L)opt =: \lambda k + \rho^B \cdot opt.$$

Combining (4.9) and (4.12) we obtain

$$\lambda k_2 + d_2 \leq \lambda k + (2 - \eta_2)opt =: \lambda k + \min\{\rho^A, \rho^B\}opt.$$

Notice that $opt_{JMS}^+(q, T_L)$ is a non-decreasing function of T_L. Since increasing $\beta_2 \leq 2$ makes T_L, hence ρ^B, bigger without decreasing ρ^A, we can pessimistically assume $\beta_2 = 2$. We can choose $\delta \in [0, 1/2]$ freely so as to minimize $\min\{\rho^A, \rho^B\}$. Given δ, we can pessimistically choose $\alpha^L \in [0,1]$, $\alpha^{MM} \in [0, 1-\alpha^L]$, and $\beta_2^{MM} \in [0, \beta_2] = [0, 2]$ so as to maximize $\min\{\rho^A, \rho^B\}$. Altogether we get:

$$2 - \eta_2 \leq \min_{\delta \in [0,1/2]} \max_{\alpha^L \in [0,1]} \max_{\alpha^{MM} \in [0, 1-\alpha^L]} \max_{\beta_2^{MM} \in [0, \beta_2]} \min\{\rho^A, \rho^B\}.$$

It should now be clear that $\eta_2 > 0$. Indeed, by choosing δ sufficiently close to zero (but positive), we achieve $\rho^A < 2$ for α^L small enough, say $\alpha^L < 1/2$. Otherwise we can assume that both α^L and δ are bounded away from zero. In that case T_L is upper bounded by some constant, and it turns out that $opt_{JMS}^+(q, T_L) < 2$ in that case (implying $\rho^B < 2$ since by assumption $\alpha^L > 0$). This concludes the proof of Lemma 4.1. We numerically obtained (setting $q = 400$) that $\eta_2 > 0.005360$. The worst case is achieved for $\delta \simeq 0.49777$, $\alpha^L \simeq 0.4948527$, $\alpha^{MM} \simeq 0.004005537$, $\beta_2^{MM} \simeq 0.00097229266$ (leading to $T_L \simeq 16.25852$).

References

[AGK⁺04] Vijay Arya, Naveen Garg, Rohit Khandekar, Adam Meyerson, Kamesh Munagala, and Vinayaka Pandit. Local search heuristics for k-median and facility location problems. *SIAM J. Comput.*, 33(3):544–562, 2004.

[ALB⁺13] Ankit Aggarwal, Anand Louis, Manisha Bansal, Naveen Garg, Neelima Gupta, Shubham Gupta, and Surabhi Jain. A 3-approximation algorithm for the facility location problem with uniform capacities. *Mathematical Programming*, 141(1):527–547, 2013.

[ARS03] Aaron Archer, Ranjithkumar Rajagopalan, and David B. Shmoys. Lagrangian relaxation for the k-median problem: New insights and continuity properties. In *Algorithms - ESA 2003, 11th Annual European Symposium, Budapest, Hungary, September 16-19, 2003, Proceedings*, pages 31–42, 2003.

[ASS17] Hyung-Chan An, Mohit Singh, and Ola Svensson. Lp-based algorithms for capacitated facility location. *SIAM Journal on Computing*, 46(1):272–306, 2017.

[BA10] Jaroslaw Byrka and Karen Aardal. An optimal bifactor approximation algorithm for the metric uncapacitated facility location problem. *SIAM J. Comput.*, 39(6):2212–2231, 2010.

[BBC⁺19] Luca Becchetti, Marc Bury, Vincent Cohen-Addad, Fabrizio Grandoni, and Chris Schwiegelshohn. Oblivious dimension reduction for k-means: beyond subspaces and the johnson-lindenstrauss lemma. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, STOC 2019, Phoenix, AZ, USA, June 23-26, 2019*, pages 1039–1050, 2019.

[BBLM14] MohammadHossein Bateni, Aditya Bhaskara, Silvio Lattanzi, and Vahab S. Mirrokni. Distributed balanced clustering via mapping coresets. In *Advances in Neural Information Processing Systems 27: Annual Conference on Neural Information Processing Systems 2014, December 8-13 2014, Montreal, Quebec, Canada*, pages 2591–2599, 2014.

[BGG12] Manisha Bansal, Naveen Garg, and Neelima Gupta. A 5-approximation for capacitated facility location. In *European Symposium on Algorithms*, pages 133–144. Springer, 2012.

[BPR⁺15] Jaroslaw Byrka, Thomas W. Pensyl, Bartosz Rybicki, Aravind Srinivasan, and Khoa Trinh. An improved approximation for k-median, and positive correlation in budgeted optimization. In Piotr Indyk, editor, *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2015, San Diego, CA, USA, January 4-6, 2015*, pages 737–756. SIAM, 2015.

[BPR⁺17] Jaroslaw Byrka, Thomas W. Pensyl, Bartosz Rybicki, Aravind Srinivasan, and Khoa Trinh. An improved approximation for k-median and positive correlation in budgeted optimization. *ACM Trans. Algorithms*, 13(2):23:1–23:31, 2017.

[BT10] Guy E. Blelloch and Kanat Tangwongsan. Parallel approximation algorithms for facility-location problems. In *SPAA 2010: Proceedings of the 22nd Annual ACM Symposium on Parallelism in Algorithms and Architectures, Thira, Santorini, Greece, June 13-15, 2010*, pages 315–324, 2010.

[BV16] Sayan Bandyapadhyay and Kasturi Varadarajan. On variants of k-means clustering. In *32nd International Symposium on Computational Geometry (SoCG 2016)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2016.

[BW17] Maria-Florina Balcan and Colin White. Clustering under local stability: Bridging the gap between worst-case and beyond worst-case analysis. *CoRR*, abs/1705.07157, 2017.

[CAD04] David Cohen-Steiner, Pierre Alliez, and Mathieu Desbrun. Variational shape approximation. *ACM Trans. Graph.*, 23(3):905–914, 2004.

[CCSL21] Vincent Cohen-Addad, Karthik C. S., and Euiwoong Lee. On approximability of clustering problems without candidate centers. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2635–2648. SIAM, 2021.

[CEMN22] Vincent Cohen-Addad, Hossein Esfandiari, Vahab Mirrokni, and Shyam Narayanan. Improved approximations for euclidean k-means and k-median, via nested quasi-independent sets. In *Proceedings of the 54th ACM Symposium on Theory of Computing, STOC 2022*, 2022.

[CFS19] Vincent Cohen-Addad, Andreas Emil Feldmann, and David Saulpic. Near-linear time approximations schemes for clustering in doubling metrics. In David Zuckerman, editor, *60th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2019, Baltimore, Maryland, USA, November 9-12, 2019*, pages 540–559. IEEE Computer Society, 2019.

[CG05] Moses Charikar and Sudipto Guha. Improved combinatorial algorithms for facility location problems. *SIAM J. Comput.*, 34(4):803–824, 2005.

[CGH⁺22] Vincent Cohen-Addad, Anupam Gupta, Lunjia Hu, Hoon Oh, and David Saulpic. An improved local search algorithm for k-median. In Joseph (Seffi) Naor and Niv Buchbinder, editors, *Proceedings of the 2022 ACM-SIAM Symposium on Discrete Algorithms, SODA 2022, Virtual Conference / Alexandria, VA, USA, January 9 - 12, 2022*, pages 1556–1612. SIAM, 2022.

[CGTS02] Moses Charikar, Sudipto Guha, Éva Tardos, and David B. Shmoys. A constant-factor approximation algorithm for the k-median problem. *J. Comput. Syst. Sci.*, 65(1):129–149, 2002.

[Che08] Ke Chen. A constant factor approximation algorithm for k-median clustering with outliers. In Shang-Hua Teng, editor, *Proceedings of the Nineteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2008, San Francisco, California, USA, January 20-22, 2008*, pages 826–835. SIAM, 2008.

[CK19] Vincent Cohen-Addad and Karthik C. S. Inapproximability of clustering in lp metrics. In David Zuckerman, editor, *60th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2019, Baltimore, Maryland, USA, November 9-12, 2019*, pages 519–539. IEEE Computer Society, 2019.

[CKM19] Vincent Cohen-Addad, Philip N. Klein, and Claire Mathieu. Local search yields approximation schemes for k-means and k-median in euclidean and minor-free metrics. *SIAM J. Comput.*, 48(2):644–667, 2019.

[CL12] Moses Charikar and Shi Li. A dependent lp-rounding approach for the k-median problem. In Artur Czumaj, Kurt Mehlhorn, Andrew M. Pitts, and Roger Wattenhofer, editors, *Automata, Languages, and Programming - 39th International Colloquium, ICALP 2012, Warwick, UK, July 9-13, 2012, Proceedings, Part I*, volume 7391 of *Lecture Notes in Computer Science*, pages 194–205. Springer, 2012.

[CM15] Vincent Cohen-Addad and Claire Mathieu. Effectiveness of local search for geometric optimization. In *31st International Symposium on Computational Geometry, SoCG 2015, June 22-25, 2015, Eindhoven, The Netherlands*, pages 329–343, 2015.

[Coh18] Vincent Cohen-Addad. A fast approximation scheme for low-dimensional k-means. In Artur Czumaj, editor, *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2018, New Orleans, LA, USA, January 7-10, 2018*, pages 430–440. SIAM, 2018.

[CS03] Fabián A. Chudak and David B. Shmoys. Improved approximation algorithms for the uncapacitated facility location problem. *SIAM J. Comput.*, 33(1):1–25, 2003.

[CS17] Vincent Cohen-Addad and Chris Schwiegelshohn. On the local structure of stable clustering instances. In Chris Umans, editor, *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*, pages 49–60. IEEE Computer Society, 2017.

[CW99] Fabián A Chudak and David P Williamson. Improved approximation algorithms for capacitated facility location problems. In *International Conference on Integer Programming and Combinatorial Optimization*, pages 99–113. Springer, 1999.

[DGK02] Inderjit S. Dhillon, Yuqiang Guan, and Jacob Kogan. Iterative clustering of high dimensional text data augmented by local search. In *Proceedings of the 2002 IEEE International Conference on Data Mining (ICDM 2002), 9-12 December 2002, Maebashi City, Japan*, pages 131–138, 2002.

[FL11] D. Feldman and M. Langberg. A unified framework for approximating and clustering data. In *STOC*, pages 569–578, 2011.

[FRS19] Zachary Friggstad, Mohsen Rezapour, and Mohammad R. Salavatipour. Local search yields a PTAS for k-means in doubling metrics. *SIAM J. Comput.*, 48(2):452–480, 2019.

[GK99] Sudipto Guha and Samir Khuller. Greedy strikes back: Improved facility location algorithms. *J. Algorithms*, 31(1):228–248, 1999.

[GT08] Anupam Gupta and Kanat Tangwongsan. Simpler analyses of local search algorithms for facility location. *CoRR*, abs/0809.2554, 2008.

[HM01] Pierre Hansen and Nenad Mladenovic. J-means: a new local search heuristic for minimum sum of squares clustering. *Pattern Recognition*, 34(2):405–413, 2001.

[HV20] Lingxiao Huang and Nisheeth K. Vishnoi. Coresets for clustering in euclidean spaces: importance sampling is nearly optimal. In Konstantin Makarychev, Yury Makarychev, Madhur Tulsiani, Gautam Kamath, and Julia Chuzhoy, editors, *Proccedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing, STOC 2020, Chicago, IL, USA, June 22-26, 2020*, pages 1416–1429. ACM, 2020.

[JMM$^+$03] 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, 2003.

[JMS02] Kamal Jain, Mohammad Mahdian, and Amin Saberi. A new greedy approach for facility location problems. In *Proceedings on 34th Annual ACM Symposium on Theory of Computing, May 19-21, 2002, Montréal, Québec, Canada*, pages 731–740, 2002.

[JV01] Kamal Jain and Vijay V. Vazirani. Approximation algorithms for metric facility location and k-median problems using the primal-dual schema and Lagrangian relaxation. *J. ACM*, 48(2):274–296, 2001.

[KH63] Alfred A Kuehn and Michael J Hamburger. A heuristic program for locating warehouses. *Management science*, 9(4):643–666, 1963.

[KLS18] Ravishankar Krishnaswamy, Shi Li, and Sai Sandeep. Constant approximation for k-median and k-means with outliers via iterative rounding. In Ilias Diakonikolas, David Kempe, and Monika Henzinger, editors, *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*, pages 646–659. ACM, 2018.

[KPR00a] Madhukar R. Korupolu, C. Greg Plaxton, and Rajmohan Rajaraman. Analysis of a local search heuristic for facility location problems. *J. Algorithms*, 37(1):146–188, 2000.

[KPR00b] Madhukar R Korupolu, C Greg Plaxton, and Rajmohan Rajaraman. Analysis of a local search heuristic for facility location problems. *Journal of algorithms*, 37(1):146–188, 2000.

[KSS10] Amit Kumar, Yogish Sabharwal, and Sandeep Sen. Linear-time approximation schemes for clustering problems in any dimensions. *J. ACM*, 57(2):5:1–5:32, 2010.

[Li13] Shi Li. A 1.488 approximation algorithm for the uncapacitated facility location problem. *Inf. Comput.*, 222:45–58, 2013.

[LS16] Shi Li and Ola Svensson. Approximating k-median via pseudo-approximation. *SIAM J. Comput.*, 45(2):530–547, 2016.

[MMSW16] Konstantin Makarychev, Yury Makarychev, Maxim Sviridenko, and Justin Ward. A bi-criteria approximation algorithm for k-means. In Klaus Jansen, Claire Mathieu, José D. P. Rolim, and Chris Umans, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2016, September 7-9, 2016, Paris, France*, volume 60 of *LIPIcs*, pages 14:1–14:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2016.

[MS84] Nimrod Megiddo and Kenneth J. Supowit. On the complexity of some common geometric location problems. *SIAM J. Comput.*, 13(1):182–196, 1984.

[MYZ06] Mohammad Mahdian, Yinyu Ye, and Jiawei Zhang. Approximation algorithms for metric facility location problems. *SIAM J. Comput.*, 36(2):411–432, 2006.

[PTW01] Martin Pal, T Tardos, and Tom Wexler. Facility location with nonuniform hard capacities. In *Proceedings 42nd IEEE symposium on foundations of computer science*, pages 329–338. IEEE, 2001.

[STA97] David B. Shmoys, Éva Tardos, and Karen Aardal. Approximation algorithms for facility location problems (extended abstract). In Frank Thomson Leighton and Peter W. Shor, editors, *Proceedings of the Twenty-Ninth Annual ACM Symposium on the Theory of Computing, El Paso, Texas, USA, May 4-6, 1997*, pages 265–274. ACM, 1997.

[Sto63] John F Stollsteimer. A working model for plant numbers and locations. *Journal of Farm Economics*, 45(3):631–645, 1963.

[YSZ+08] Yi Yang, Min Shao, Sencun Zhu, Bhuvan Urgaonkar, and Guohong Cao. Towards event source unobservability with minimum network traffic in sensor networks. In *Proceedings of the First ACM Conference on Wireless Network Security, WISEC 2008, Alexandria, VA, USA, March 31 - April 02, 2008*, pages 77–88, 2008.

A Omitted proofs from Section 2

LEMMA 2.2. *Fix $T > 0$. For any positive integers q and c, $opt_{JMS}(q, T) \leq opt_{JMS}(cq, T)$.*

Proof. Given an optimal solution $S = (\alpha, d, r, \lambda)$ of $LP_{JMS}(q, T)$, we build a feasible solution $S' = (\alpha', d', r', \lambda')$ of $LP_{JMS}(cq, T)$ with the same objective value. We set

$$\alpha'_i := \frac{\alpha_{\lceil i/c \rceil}}{c}, \quad d'_i = \frac{d_{\lceil i/c \rceil}}{c}, \quad r'_{j,i} := \begin{cases} \frac{r_{\lceil j/c \rceil, \lceil i/c \rceil}}{c} & \text{if } \lceil j/c \rceil < \lceil i/c \rceil; \\ \frac{\alpha_{\lceil j/c \rceil}}{c} & \text{if } \lceil j/c \rceil = \lceil i/c \rceil. \end{cases} \quad \lambda' = \lambda, \quad \forall 1 \leq j \leq i \leq cq.$$

Observe that S and S' have exactly the same objective value. Let us show that S' is feasible. Clearly all the variables in S' are non-negative since the same holds for S. To avoid possible confusion, we use $(x)'$ to denote constraint (x) in $LP_{JMS}(cq, T)$. The following inequalities show that Constraints $(1.1)'$, $(1.2)'$, $(1.5)'$, and $(2.7)'$, resp., hold for every feasible choice of the parameters:

$$\sum_{i=1}^{cq} d'_i = \sum_{\ell=1}^{q} d_i \overset{(1.1)}{=} 1; \qquad \alpha'_i = \frac{\alpha_{\lceil i/c \rceil}}{c} \overset{(1.2)}{\leq} \frac{\alpha_{\lceil (i+1)/c \rceil}}{c} = \alpha'_{i+1}; \qquad r'_{j,j} = \frac{\alpha_{\lceil j/c \rceil}}{c} = \alpha'_j; \qquad \lambda' = \lambda \overset{(2.7)}{\leq} T.$$

Consider next Constraint $(1.3)'$. We distinguish three subcases. If $\lceil j/c \rceil < \lceil i/c \rceil$, one has

$$r'_{j,i+1} = \frac{r_{\lceil j/c \rceil, \lceil (i+1)/c \rceil}}{c} \overset{(1.3)}{\leq} \frac{r_{\lceil j/c \rceil, \lceil i/c \rceil}}{c} = r'_{j,i}.$$

If $\lceil j/c \rceil = \lceil i/c \rceil < \lceil (i+1)/c \rceil$, one has

$$r'_{j,i+1} = \frac{r_{\lceil j/c \rceil, \lceil (i+1)/c \rceil}}{c} \overset{(1.3)}{\leq} \frac{r_{\lceil j/c \rceil, \lceil i/c \rceil}}{c} = \frac{r_{\lceil j/c \rceil, \lceil j/c \rceil}}{c} \overset{(1.5)}{\leq} \frac{\alpha_{\lceil j/c \rceil}}{c} = r'_{j,i}.$$

Finally if $\lceil j/c \rceil = \lceil i/c \rceil = \lceil (i+1)/c \rceil$, one has

$$r'_{j,i+1} = \frac{\alpha_{\lceil j/c \rceil}}{c} = r'_{j,i}.$$

Consider now Constraint $(1.4)'$. We distinguish two subcases. If $\lceil j/c \rceil < \lceil i/c \rceil$,

$$\alpha'_i = \frac{\alpha_{\lceil i/c \rceil}}{c} \overset{(1.4)}{\leq} \frac{r_{\lceil j/c \rceil, \lceil i/c \rceil} + d_{\lceil i/c \rceil} + d_{\lceil j/c \rceil}}{c} = r'_{j,i} + d'_i + d'_j.$$

Otherwise (i.e., $\lceil j/c \rceil = \lceil i/c \rceil$)

$$\alpha'_i = \frac{\alpha_{\lceil i/c \rceil}}{c} = \frac{\alpha_{\lceil j/c \rceil}}{c} = r'_{j,i} \leq r'_{j,i} + d'_i + d'_j.$$

It remains to consider Constraint $(1.6)'$. For the considered index i, let $i = c\ell + h$ with $\ell \in \{0, q-1\}$ and $h \in \{1, \ldots, c\}$. Notice that

$$(A.1) \qquad\qquad r'_{j,i} = \frac{\alpha_{\lceil i/c \rceil}}{c} = \alpha'_i, \qquad \forall c\ell + 1 \leq j \leq i = c\ell + h.$$

Then

$$\sum_{j=1}^{i-1} \max\{r'_{j,i} - d'_j, 0\} + \sum_{j=i}^{cq} \max\{\alpha'_i - d'_j, 0\}$$

$$= \sum_{j=1}^{c\ell} \max\{r'_{j,i} - d'_j, 0\} + \sum_{j=c\ell+1}^{c\ell+h} \max\{r'_{j,i} - d'_j, 0\} + \sum_{j=c\ell+h+1}^{cq} \max\{\alpha'_i - d'_j, 0\}$$

$$\overset{(A.1)}{\leq} \sum_{j=1}^{c\ell} \max\{r'_{j,i} - d'_j, 0\} + \sum_{j=c\ell+1}^{cq} \max\{\alpha'_i - d'_j, 0\}$$

$$= \sum_{h=1}^{\ell} c \cdot \max\{\frac{r_{h,\ell+1}}{c} - \frac{d_h}{c}, 0\} + \sum_{h=\ell+1}^{q} c \cdot \max\{\frac{\alpha_{\ell+1}}{c} - \frac{d_h}{c}, 0\} \overset{(1.6)}{\leq} \lambda = \lambda'.$$

$\square$

LEMMA 2.3. $opt_{JMS}(T)$ *is a concave function of* T.

Proof. We will show that, for any fixed q, $opt_{JMS}(q, T)$ is concave. The claim then follows. Indeed, assume by contradiction that there exist values $0 < T_1 < T_2$ and $0 < \alpha < 1$ such that

$$\alpha opt_{JMS}(T_1) + (1 - \alpha) opt_{JMS}(T_2) \geq opt_{JMS}(T) + \delta$$

where $T = \alpha T_1 + (1 - \alpha) T_2$ and $\delta > 0$. There must exist a finite q such that $opt_{JMS}(q, T_i) \geq opt_{JMS}(T_i) - \delta/2$ for $i \in \{1, 2\}$. Then we get the contradiction

$$
\begin{aligned}
opt_{JMS}(T) &\geq opt_{JMS}(q, T) \geq \alpha opt_{JMS}(q, T_1) + (1 - \alpha) opt_{JMS}(q, T_2) \\
&\geq \alpha(opt_{JMS}(T_1) - \frac{\delta}{2}) + (1 - \alpha)(opt_{JMS}(T_2) - \frac{\delta}{2}) \\
&\geq opt_{JMS}(T) + \delta - \frac{\delta}{2} > opt_{JMS}(T).
\end{aligned}
$$

It remains to show that $opt_{JMS}(q, T)$ is concave for any given $q \geq 1$. Consider any values $0 < T_1 < T_2$ and $\alpha \in (0, 1)$. Let $T = \alpha T_1 + (1 - \alpha) T_2$ and $OPT_{JMS}(q, T_i)$ be some optimal solution for $LP_{JMS}(q, T_i)$. We use $\lambda(q, T_i)$ to refer to the value of λ in $OPT_{JMS}(q, T_i)$ and similarly for the other variables. Consider the solution

$$APX_{JMS}(q, T) = \alpha OPT_{JMS}(q, T_1) + (1 - \alpha) OPT_{JMS}(q, T_2)$$

for $LP_{JMS}(q, T)$. Trivially $APX_{JMS}(q, T)$ if feasible. Indeed, it is the convex combination of two feasible solutions restricted to Constraints (1.1)-(1.6). Furthermore its value of λ satisfies Constraint (2.7) since, by the feasibility of $OPT_{JMS}(q, T_i)$,

$$\lambda = \alpha \lambda(q, T_1) + (1 - \alpha) \lambda(q, T_2) \leq \alpha T_1 + (1 - \alpha) T_2 = T.$$

The value of $APX_{JMS}(q, T)$ is $\alpha opt_{JMS}(q, T_1) + (1 - \alpha) opt_{JMS}(q, T_2)$, hence $opt_{JMS}(q, T)$ is at least the latter amount (since $LP_{JMS}(q, T)$ is a maximization LP). □

LEMMA 2.4. *Fix* $T > 0$. *For any positive integers* q *and* c, $opt_{JMS}^+(q, T) \geq opt_{JMS}(cq, T)$.

Proof. Let $S := (\alpha, d, r, \lambda)$ be an optimal solution to $LP_{JMS}(cq, T)$. We show how to build a feasible solution $S^+ := (\alpha^+, d^+, r^+, \lambda^+)$ for $LP_{JMS}^+(q, T)$ with the same objective value. We set

$$\alpha_i^+ := \sum_{\ell=(i-1)c+1}^{ic} \alpha_\ell, \quad d_i^+ = \sum_{\ell=(i-1)c+1}^{ic} d_\ell, \quad r_{j,i}^+ := \sum_{\ell=(j-1)c+1}^{jc} r_{\ell, ic}, \quad \lambda^+ = \lambda, \quad \forall 1 \leq j \leq i \leq q.$$

Clearly the objective values of S and S^+ are identical. Let us show that S^+ is a feasible solution. Clearly all the variables in S^+ are non-negative since the same holds for S. In order to avoid possible confusion, we will use $(x)^+$ instead of (x) to denote the (x)-constraint in $LP_{JMS}^+(q, T)$. The following inequalities show that Constraints $(1.1)^+$, $(1.2)^+$, $(1.3)^+$, $(1.4)^+$, $(1.5)^+$, and $(2.7)^+$, resp., hold for every valid choice of the indexes:

$$\sum_{i=1}^{q} d_i^+ = \sum_{j=1}^{cq} d_i \overset{(1.1)}{=} 1.$$

$$\alpha_i^+ = \sum_{\ell=(i-1)c+1}^{ic} \alpha_\ell \overset{(1.2)}{\leq} \sum_{\ell=(i-1)c+1}^{ic} \alpha_{\ell+c} = \sum_{\ell=ic+1}^{(i+1)c} \alpha_\ell = \alpha_{i+1}^+.$$

$$r_{j,i+1}^+ = \sum_{\ell=(j-1)c+1}^{jc} r_{\ell,(i+1)c} \overset{(1.3)}{\leq} \sum_{\ell=(j-1)c+1}^{jc} r_{\ell, ic} = r_{j,i}^+.$$

$$\alpha_i^+ = \sum_{h=1}^{c} \alpha_{(i-1)c+h} \overset{(1.4)}{\leq} \sum_{h=1}^{c} (r_{(j-1)c+h,(i-1)c+h} + d_{(i-1)c+h} + d_{(j-1)c+h})$$

$$= d_i^+ + d_j^+ + \sum_{h=1}^{c} r_{(j-1)c+h,(i-1)c+h} \overset{(1.3)}{\leq} d_i^+ + d_j^+ + \sum_{h=1}^{c} r_{(j-1)c+h,ic} = d_i^+ + d_j^+ + r_{j,i}^+.$$

$$r_{j,j}^+ = \sum_{h=1}^{c} r_{(j-1)c+h,jc} \overset{(1.3)}{\leq} \sum_{h=1}^{c} r_{(j-1)c+h,(j-1)c+h} \overset{(1.5)}{\leq} \sum_{h=1}^{c} \alpha_{(j-1)c+h} = \alpha_j^+.$$

$$\lambda^+ = \lambda \overset{(2.7)}{\leq} T.$$

It remains to consider Constraint (2.8) that replaces (1.6). Let us prove some intermediate inequalities

$$\sum_{j=1}^{t} \max\{r_{j,t}^+ - d_j^+, 0\} = \sum_{j=1}^{t} \max\left\{ \sum_{\ell=(j-1)c+1}^{jc} r_{\ell,tc} - \sum_{\ell=(j-1)c+1}^{jc} d_\ell, 0 \right\}$$

$$\text{(A.2)} \qquad \leq \sum_{j=1}^{t} \sum_{\ell=(j-1)c+1}^{jc} \max\{r_{\ell,tc} - d_\ell, 0\} = \sum_{\ell=1}^{tc} \max\{r_{\ell,tc} - d_\ell, 0\}.$$

$$\sum_{j=t+1}^{q} \max\{\alpha_t^+ - d_j^+, 0\} = \sum_{j=t+1}^{q} \max\left\{ \sum_{\ell=(j-1)c+1}^{jc} \alpha_\ell - \sum_{\ell=(j-1)c+1}^{jc} d_\ell, 0 \right\}$$

$$\text{(A.3)} \qquad \leq \sum_{j=t+1}^{q} \sum_{\ell=(j-1)c+1}^{jc} \max\{\alpha_\ell - d_\ell, 0\} = \sum_{\ell=tc+1}^{cq} \max\{\alpha_\ell - d_\ell, 0\}.$$

Constraint (2.8) follows since:

$$\sum_{j=1}^{i-1} \max\{r_{j,i}^+ - d_j^+, 0\} + \sum_{j=i}^{q} \max\{\alpha_i^+ - d_j^+, 0\} \overset{(A.2)+(A.3)}{\leq} \sum_{\ell=1}^{tc} \max\{r_{\ell,tc} - d_\ell, 0\} + \sum_{\ell=tc+1}^{cq} \max\{\alpha_\ell - d_\ell, 0\}$$

$$\overset{(1.5)}{\leq} \sum_{\ell=1}^{tc-1} \max\{r_{\ell,tc} - d_\ell, 0\} + \max\{\alpha_{tc} - d_{tc}, 0\} + \sum_{\ell=tc+1}^{cq} \max\{\alpha_\ell - d_\ell, 0\} \overset{(1.6)}{\leq} \lambda = \lambda^+.$$

□

COROLLARY 2.1. *Let OPT be a solution to a UFL instance and S be the JMS solution on the same instance. For a given constant $T > 0$, let $OPT' \subseteq OPT$ be a subset of facilities satisfying $\mathbf{open}(OPT') \leq T \cdot d(OPT')$. Then, for every integer $q > 0$,*

$$\mathbf{open}(S) + d(S) \leq \mathbf{open}(OPT) + 2d(OPT \setminus OPT') + opt_{JMS}^+(q, T) \cdot d(OPT').$$

Proof. Recall that $opt_{JMS}(T) \leq opt_{JMS} \leq 2$. By Lemma 2.3

$$\sum_{f^* \in OPT'} opt_{JMS}\left(\frac{\mathbf{open}(f^*)}{d(f^*)}\right) \cdot \frac{d(f^*)}{d(OPT')} \leq opt_{JMS}\left(\sum_{f^* \in OPT'} \frac{d(f^*)}{d(OPT')} \cdot \frac{\mathbf{open}(f^*)}{d(f^*)} \right)$$

$$\text{(A.4)} \qquad = opt_{JMS}\left(\frac{\mathbf{open}(OPT')}{d(OPT')}\right) \leq opt_{JMS}(T) \overset{Lem.2.5}{\leq} opt_{JMS}^+(q, T),$$

where the second-last inequality follows since $opt_{JMS}(T)$ is non-decreasing in T. Then

$$\mathbf{open}(S) + d(S) \overset{Lem.2.1}{\leq} \mathbf{open}(OPT) + \sum_{f^* \in OPT \setminus OPT'} 2d(f^*) + \sum_{f^* \in OPT'} opt_{JMS}\left(\frac{\mathbf{open}(f^*)}{d(f^*)}\right) d(f^*)$$

$$\overset{(A.4)}{\leq} \mathbf{open}(OPT) + 2d(OPT \setminus OPT') + opt_{JMS}^+(q, T) \cdot d(OPT').$$

□

B An Alternative LMP Bound for Small Facility Cost

Here we provide an alternative, analytical, way to upper bound the approximation factor of JMS. We define a variant of $LP_{JMS}(q)$ as follows:

$$\max \sum_{i=1}^{q} \alpha_i - \lambda$$

(B.5)
$$\text{s.t.} \sum_{i=1}^{q} d_i = 1$$

(B.6)
$$\forall 1 \leq i < q : \alpha_i - \alpha_{i+1} \leq 0$$

(B.7)
$$\forall 1 \leq j < i < q : r_{j,i+1} - r_{j,i} \leq 0$$

(B.8)
$$\forall 1 \leq j < i \leq q : \alpha_i - r_{j,i} - d_i - d_j \leq 0$$

(B.9)
$$\forall 1 \leq j < i \leq q : r_{j,i} - d_j - g_{i,j} \leq 0$$

(B.10)
$$\forall 1 \leq i \leq j \leq q : \alpha_i - d_j - h_{i,j} \leq 0$$

(B.11)
$$\forall 1 \leq i \leq q : \sum_{j=1}^{i-1} g_{i,j} + \sum_{j=i}^{q} h_{i,j} - \lambda \leq 0$$

(B.12)
$$\forall 1 \leq j < q : r_{j,j+1} \leq \alpha_j$$
$$\alpha, d, f, r, g, h \geq 0$$

In particular notice that we do not use the variables $r_{j,j}$. Furthermore, we replace $r_{j,j} \leq \alpha_j$ with $r_{j,j+1} \leq \alpha_j$. Clearly the latter constraint is also guaranteed by JMS. Indeed, since α_j is an upper bound on the distance between client j and its first facility and client j only switches to closer facilities, $r_{j,i} \leq \alpha_j$ holds for all possible LP values that correspond to a run of the JMS algorithm.

Consider the variant $LP_{JMS}(q, T)$ of $LP_{JMS}(q)$, for a parameter $T > 0$, where we add the following constraint that intuitively captures the condition $\mathtt{open}(f) \leq T \cdot d(f)$

(B.13)
$$\lambda \leq T.$$

Let $opt_{JMS}(q, T)$ be the optimal value of $LP_{JMS}(q, T)$, and $opt_{JMS}(T) = \sup_q \{opt_{JMS}(q, T)\}$. Observe that $opt_{JMS}(q, T) \leq opt_{JMS}(q)$ where the equality holds for any fixed q and T large enough. We are able to show that, for any finite T, $opt_{JMS}(T) < 2$. More precisely, we obtained the following result.

LEMMA B.1. *For any $T > 0$, $opt_{JMS}(T) \leq \min_{z \in [0, 1/3]}(V(z) + T(M(z) - 1))$, where*

$$V(z) = \max\{2 - \frac{z}{1-z}, 2 - \frac{2z}{1-z} + \ln(1 + \frac{z}{1-2z}) + \frac{4z^2}{(1-z)(1-2z)}\}, \ and$$

$$M(z) - 1 = \ln(1 + \frac{z}{1-2z}) - \frac{z}{1-z} + \frac{2z^2}{(1-z)(1-2z)}.$$

The following simple corollary shows that the improvement with respect to 2 is $\Theta(\frac{1}{c+T})$ for some constant c. Note that this is asymptotically tight because [JMS02] proved that $opt_{JMS}(T) \geq 2 - \frac{2}{T+2}$ for any positive even integer T.

COROLLARY B.1. *For any $T > 0$, $opt_{JMS}(T) \leq 2 - \frac{1}{4(7+3T)}$.*

Proof. Using the approximation $\ln(1 + \frac{z}{1-2z}) \leq \frac{z}{1-2z}$, simple calculations show that $V(z) \leq 2 - \frac{z-7z^2}{(1-z)(1-2z)}$ and $M(z) - 1 \leq \frac{z^3}{(1-z)(1-2z)}$. Therefore, $V(z) + T(M(z) - 1) \leq 2 - \frac{z-(7+3T)z^2}{(1-z)(1-2z)} \leq 2 - (z - (7 + 3T)z^2)$. Taking $z = \frac{1}{2(7+3T)}$ yields the claim. □

It remains to prove Lemma B.1. To that aim we first show that $opt_{JMS}(T) = \lim_{q \to +\infty} opt_{JMS}(q, T)$. This is a straightforward consequence of the following lemma.

LEMMA B.2. *Fix $T > 0$. For any positive integers q and m, $opt_{JMS}(q, T) \leq opt_{JMS}(mq, T)$.*

Proof. Create qm clients $(i, a) \in [q] \times [m]$ with the lexicographical ordering (first coordinate first). Define the values as follows.

- $\alpha_{i,a} = \alpha_i / m$.

- $d_{i,a} = d_i / m$.

- $r_{(j,b),(i,a)}$: if $j < i$, then $r_{(j,b),(i,a)} = r_{j,i}/m$. Otherwise, $i = j$ and $r_{(j,b),(i,a)} = \alpha_j/m$.

- As usual, $g_{(i,a),(j,b)} = \max(r_{(j,b),(i,a)} - d_{(j,b)}, 0)$ and $h_{(i,a),(j,b)} = \max(\alpha_{i,a} - d_{(j,b)}, 0)$.

- f stays the same.

Equations (B.5) and (B.6) are easy to check. (B.7) is also true because new $r_{(j,b),(j,a)} = \alpha_j/m \geq r_{(j,b),(i,c)}$ if $i > j$ (by the assumption). Consider (B.8) for $(j,b) < (i,a)$, if $j < i$, then it just follows from the old inequality with i and j. If $i = j$ and $b < a$, since $r_{(j,b),(i,a)} = \alpha_j/m$,

$$\alpha_{(i,a)} - r_{(j,b),(i,a)} - d_{i,a} - d_{j,b} = -d_{i,a} - d_{j,b} \leq 0.$$

Finally, let us look at (B.11). Fix (i, a). We first compute relevant g and h.

- For $(j,b) < (i,a)$, if $j < i$, then $g_{(i,a),(j,b)} = g_{i,j}/m$. Otherwise, $i = j$ and $g_{(i,a),(j,b)} = \max(r_{(j,b),(i,a)} - d_j, 0) = \max(\alpha_j - d_j, 0)/m$.

- For $(j,b) \geq (i,a)$, $h_{(i,a),(j,b)} = \max(\alpha_{i,a} - d_{j,b}, 0) = h_{i,j}/m$. Note that if $i = j$, it is again exactly $\max(\alpha_j - d_j, 0)/m$.

Since $g_{(i,a),(i,b)} = h_{i,i}/m = h_{(i,a),(i,c)}$ and $b < a \leq c$, so

$$
\begin{aligned}
&\sum_{(j,b) < (i,a)} g_{(i,a),(j,b)} + \sum_{(j,b) \geq (i,a)} h_{(i,a),(j,b)} \\
&= \sum_{(j,b): j < i} g_{(i,a),(j,b)} + \sum_{(j,b): j \geq i} h_{(i,a),(j,b)} \\
&= \sum_{j < i} g_{i,j} + \sum_{j \geq i} h_{i,j} \leq f.
\end{aligned}
$$

□

    The above lemma motivated us to study the continuous version of the factor-revealing LP and its dual. [JMS02] showed that $opt_{JMS} \leq 2$ by constructing a feasible solution of value $2 - 1/q$ for the dual of $LP_{JMS}(q)$ for every $q \geq 1$. For $LP_{JMS}(q, T)$, letting $\{A_{i,j}\}_{1 \leq j < i \leq k}$, $\{B_{i,j}\}_{1 \leq j < i \leq k}$, $\{C_{i,j}\}_{1 \leq i \leq j \leq k}$, $\{N_i\}_{1 \leq i \leq q}$, V, and M be the dual variables corresponding to (B.8), (B.9), (B.10), (B.11), (B.5), and (B.13) and optimizing the other dual variables depending on them, we obtain the dual LP called $DP_{JMS}(q, T)$ shown below. The *Continuous Dual LP*, dubbed $CDP_{JMS}(T)$, is obtained by letting $q \to \infty$ and replacing the counting measure on $\{1, \ldots, q\}$ by the Lebesgue measure on $[0, 1]$; so variables are represented as functions and sums become integrals. In particular, we have $A, B : L \to \mathbb{R}^{\geq 0}$, $C : U \to \mathbb{R}^{\geq 0}$ where $L := \{(i, j) \in [0, 1]^2 : j \leq i\}$, and $U := \{(i, j) \in [0, 1]^2 : j \geq i\}$. We use subscripts to denote function arguments (e.g., $A_{i,j} = A(i, j)$). See Appendix B.1 for a full derivation of both $DP_{JMS}(q, T)$ and $CDP_{JMS}(T)$.

$$\begin{aligned}
&\min \quad V - (M-1)T && (DP_{JMS}(q,T)) \\
&\text{s.t.} \quad \forall i \in [q]: \sum_{j=1}^{i-1} A_{i,j} + \sum_{j=i}^{q} C_{i,j} = 1 \\
&\qquad \forall j \in [q]: \\
&\qquad \sum_{i=j+1}^{q} A_{i,j} + \sum_{i=1}^{j-1} A_{j,i} + \sum_{i=j+1}^{q} B_{i,j} + \sum_{i=1}^{j} C_{i,j} \leq V \\
&\qquad \forall\, 1 \leq j \leq i \leq q: \sum_{\ell=j+1}^{i} B_{\ell,j} \geq \sum_{\ell=j+1}^{i} A_{\ell,j}, \\
&\qquad N_i = \max\left(\max_{j\in[1,i-1]} B_{i,j}, \quad \max_{j\in[i+1,q]} C_{i,j}\right), \\
&\qquad M = \sum_{i=1}^{q} N_i
\end{aligned}$$

$$\begin{aligned}
&\min \quad V - (M-1)T && (CDP_{JMS}(T)) \\
&\text{s.t.} \quad \forall i \in [0,1]: \int_{j=0}^{i} A_{i,j} + \int_{j=i}^{1} C_{i,j} = 1 \\
&\qquad \forall j \in [0,1]: \\
&\qquad \int_{i=j}^{1} A_{i,j} + \int_{i=0}^{j} A_{j,i} + \int_{i=j}^{1} B_{i,j} + \int_{i=0}^{j} C_{i,j} \leq V \\
&\qquad \forall\, 0 \leq j \leq i \leq 1: \int_{\ell=j}^{i} B_{\ell,j} \geq \int_{\ell=j}^{i} A_{\ell,j}, \\
&\qquad N_i = \max\left(\max_{j\in[0,i]} B_{i,j}, \quad \max_{j\in[i,1]} C_{i,j}\right), \\
&\qquad M = \int_{i=0}^{1} N_i
\end{aligned}$$

Let $cdopt_{JMS}(T)$ be the optimal value of $CDP_{JMS}(T)$. Note that setting $A(\cdot)$, $B(\cdot)$, $C(\cdot)$, and $N(\cdot)$ to the constant 1 function and letting $V = 2$ and $M = 1$, one obtains a feasible solution to $CDP_{JMS}(T)$, showing that $cdopt_{JMS}(T) \leq 2$ for any T. We are able to construct a better dual solution for finite T, showing that

$$cdopt_{JMS}(T) \leq \min_{z\in[0,1/3]} (V(z) + (M(z)-1)T).$$

The mentioned solution to $CDP_{JMS}(T)$, parameterized by z, is constructed by dividing the overall domain $[0,1] \times [0,1]$ into several regions defined by the lines including $x = z, y = z, x = 1 - z, y = 1 - z, x = y$ and by assigning different values to A, B, and C for each such region. In order to obtain a feasible solution for $DP_{JMS}(q,T)$, we discretize our solution for $CDP_{JMS}(T)$. In more detail, we use the natural strategy of partitioning $[0,1] \times [0,1]$ into squares of side length $1/q$ and integrating the continuous variables over each square to obtain the appropriate discrete variables. (E.g., $A_{q,1}$ of $DP_{JMS}(q,T)$ is defined to be $\int_{i=(q-1)/q}^{1} \int_{j=0}^{1/q} A_{i,j}$ where $A_{i,j}$ is taken from our $CDP_{JMS}(T)$ solution). When z is an integer multiple of $1/q$, the lines defining the small squares "align nicely" with the lines defining the continuous dual solution, so one can obtain the following lemma.

LEMMA B.3. *For any $T > 0$, $q \in \mathbb{N}$ and $z \in [0,1/3]$ where $z = d/q$ for some integer d, $opt_{JMS}(q,T) \leq V(z) + T(M(z)-1)$.*

We now have all the ingredients to prove Lemma B.1.

Proof. [Proof of Lemma B.1] Let $z^* = \mathrm{argmin}_{z\in[0,1/3]}(V(z)+T(M(z)-1))$. For any $\varepsilon > 0$, one can choose $c,d \in \mathbb{N}$ such that $z' := \frac{d}{cq} \in [0,1/3]$ is close enough to z^* to satisfy $V(z') + T(M(z')-1) < V(z^*) + T(M(z^*)-1) + \varepsilon$.

$$opt_{JMS}(q,T) \overset{Lem.B.2}{\leq} opt_{JMS}(cq,T) \overset{Lem.B.3}{\leq} V(z') + T(M(z')-1) < V(z^*) + T(M(z^*)-1) + \varepsilon.$$

Since this holds for any ε, we have $opt_{JMS}(q,T) \leq \min_{z\in[0,1/3]}(V(z) + T(M(z) - 1))$, hence the claim. □

It remains to prove Lemma B.3. As mentioned before, the proof of this lemma first considers the *continuous case* where the set of clients becomes a continuous set $[0,1]$ followed by a discretization step. Section B.1 and B.2 will discuss the factor-revealing LP for the continuous case.

B.1 Reduction to dual continuous LP In this subsection, we drive the dual LP $DP_{JMS}(q,T)$ and the continuous dual LP $CDP_{JMS}(T)$ from the factor-revealing LP $LP_{JMS}(q,T)$ introduced before.

We will primarily use dual variables for (B.8), (B.9), and (B.10). (Dual variables for other constraints will be implicitly defined depending on them, but we will not use (B.6) at all.) Let $\{A_{i,j}\}_{1\leq j < i \leq q}$, $\{B_{i,j}\}_{1\leq j < i \leq q}$, $\{C_{i,j}\}_{1\leq i \leq j \leq q}$ be the dual variables corresponding to (B.8), (B.9), and (B.10). Note that these are the only variables that contain α_i, d_i primal variables apart from (B.5) and (B.6). Next, we will go over each dual constraint and give some conditions that our dual solutions need to satisfy:

- α and d: In our dual solution, we will satisfy the constraints induced by α_i's with equality, which implies

$$(B.14) \qquad \forall i : \sum_{j=1}^{i-1} A_{i,j} + \sum_{j=i}^{q} C_{i,j} = 1$$

For d_i's, we have

$$(B.15) \qquad \forall j : \sum_{i=j+1}^{q} A_{i,j} + \sum_{i=1}^{j-1} A_{j,i} + \sum_{i=j+1}^{q} B_{i,j} + \sum_{i=1}^{j} C_{i,j} \leq V$$

where V will be the dual value corresponding to (B.5). (Note that i, j is swapped only for the second term because (B.8) has both d_i and d_j.)

- r: The coefficient of $r_{j,i}$ is increased by $B_{i,j}$ and decreased by $A_{i,j}$. Also, (B.7) increases the coefficient of $r_{j,i+1}$ and decrease the coefficient of $r_{j,i}$. (Intuitively, for same j, we can increase the coefficient for higher i and decrease the coefficient for smaller i.) Therefore, as long as we satisfy

$$\forall 1 \leq j < q : \text{ the sequence } (B_{j+1,j}, ..., B_{q,j}) \text{ "dominates" } (A_{j+1,j}, ..., A_{q,j})$$

$$(B.16) \qquad \Leftrightarrow \quad \forall 1 \leq j < i \leq q : \sum_{\ell=j+1}^{i} B_{\ell,j} \geq \sum_{\ell=j+1}^{i} A_{\ell,j},$$

then we can make sure that the dual constraints for $r_{j,i}$'s are satisfied.

- g and h: Once we fix A, B, C, then for each $i \in [q]$, the dual variable corresponding to the ith equation of (B.11) should be

$$N_i := \max \left(\max_{j=1}^{i-1} B_{i,j}, \quad \max_{j=i}^{q} C_{i,j} \right).$$

due to the constraints corresponding to g, h.

- f: Let $M := \sum_i N_i$. (We will make sure $M \geq 1$ always.) Then from the dual constraint for f, we will have the dual variable for $f \leq T$ exactly $(M-1)$. And the dual objective function value is $V + T(M-1)$.

So that's the rule of the game: find A, B, C that satisfy (B.14), (B.15), and (B.16) to minimize the value $V + T(M-1)$. The JMS gave every A, B, C variable $1/q$, so that $V = 2 - 1/q$ and $M = 1$.

To ignore the diagonal issues, let us consider the *continuous setting* where for A, B, C, we divide each index by q and multiply each value by q. As q goes to infinity, then we are basically in the setting where we want to give a map $A, B : L \to \mathbb{R}^{\geq 0}$ and $C : U \to \mathbb{R}^{\geq 0}$ where

$$L := \{(i, j) \in [0, 1]^2 : j \leq i\}$$
$$U := \{(i, j) \in [0, 1]^2 : j \geq i\}$$

Then the three conditions (B.14), (B.15), and (B.16) will be translated as

$$(B.17) \qquad \forall i \in [0, 1] : \int_{j=0}^{i} A_{i,j} + \int_{j=i}^{1} C_{i,j} = 1$$

$$(B.18) \qquad \forall j \in [0, 1] : \int_{i=j}^{1} A_{i,j} + \int_{i=0}^{j} A_{j,i} + \int_{i=j}^{1} B_{i,j} + \int_{i=0}^{j} C_{i,j} \leq V$$

$$\forall j \in [0, 1] : \text{ the function } B(\cdot, j) \text{ "dominates" } A(\cdot, j)$$

$$(B.19) \qquad \Leftrightarrow \quad \forall 0 \leq j \leq i \leq 1 : \int_{\ell=j}^{i} B_{\ell,j} \geq \int_{\ell=j}^{i} A_{\ell,j},$$

and

$$N_i := \max\left(\max_{j\in[0,i]} B_{i,j}, \quad \max_{j\in[i,1]} C_{i,j}\right), \qquad M := \int_{i=0}^{1} N_i,$$

and the objective value is $V - (M-1)T$. Note that making A, B, C the constant 1 functions satisfy every constraint and give the value 2 for any T. In Section B.2, we construct a dual solution for this continuous dual program. In Section B.3, we will show how to convert it back to a dual solution for the discrete dual LP.

B.2 Solution to dual continuous LP Let $z \in [0, 1/3]$ be a small constant, $y := z^2/(1-z)$, and

$$V(z) = \max\{2 - \frac{z}{1-z}, 2 - \frac{2z}{1-z} + \ln(1 + \frac{z}{1-2z}) + \frac{4z^2}{(1-z)(1-2z)}\}, \text{ and}$$

$$M(z) - 1 = \ln(1 + \frac{z}{1-2z}) - \frac{z}{1-z} + \frac{2z^2}{(1-z)(1-2z)}.$$

We are ready to construct a slightly better dual solution parameterized by z.

- $A(i, j)$ is equal to

  - $1/(1-z)$ when $i \in [z, 1-z], j \in [z, i]$.
  - $1/z$ when $i \in [1-z, 1-y], j \in [0, z]$.
  - $1/(0.5 - z)$ when $i \in [1-y, 1], j \in [z, 0.5]$.
  - 0 otherwise.

- $B(i, j)$ is equal to

  - $1/(1-z)$ when $i \in [z, 1-z], j \in [0, i]$.
    * But, when $i \in [1-z-y, 1-z]$ and $j \in [z, 0.5]$, let $B(i,j) = 1/(1-z) + 1/(0.5-z)$.
  - 0 otherwise.

- $C(i, j)$ is equal to

  - $1/(1-z-i)$ when $i \in [0, z]$ and $j \in [i, 1-z]$.
  - $1/(1-z)$ when $i \in [z, 1-z]$ and $j \in [i, 1]$.
  - 0 otherwise.

Let us first check (B.17)

$$\forall i \in [0, 1] : \int_{j=0}^{i} A_{i,j} + \int_{j=i}^{1} C_{i,j} = 1.$$

- $i \in [0, z]$: 0 from A, 1 from C, for each i, $C(i,j) = 1/(1-z-i)$ for $j \in [i, 1-z]$.

- $i \in [z, 1-z]$: $(i-z)/(1-z)$ from A and $(1-i)/(1-z)$ from C.

- $i \in [1-z, 1-y]$: 1 from A (value $1/z$, interval length z).

- $i \in [1-y, 1]$: 1 from A (value $1/(0.5-z)$, interval length $0.5 - z$).

Let us check (B.19)

$$\forall j \in [0, 1] : B(\cdot, j) \text{ "dominates" } A(\cdot, j),$$

- $j \in [0, z]$: $A_{i,j} = 1/z$ when $i \in [1-z, 1-y]$. $B(i,j) = 1/(1-z)$ when $i \in [z, 1-z]$: For both A and B, the integral is

(B.20)
$$\frac{1}{z} \cdot (z - y) = \frac{1}{z} \cdot (z - \frac{z^2}{1-z}) = \frac{1-2z}{1-z},$$

and clearly $B(\cdot, j)$ dominates $A(\cdot, j)$.

- $j \in [z, 0.5]$: $A_{i,j} = 1/(1-z)$ when $i \in [j, 1-z]$ and $1/(0.5-z)$ when $i \in [1-y, 1]$. $B_{i,j} = 1/(1-z)$ when $i \in [j, 1-z-y]$ and $1/(1-z) + 1/(0.5-z)$ when $i \in [1-z-y, 1-z]$. For both A and B, the integral is

(B.21)
$$\frac{1-z-j}{1-z} + \frac{y}{0.5-z}$$

and clearly $B(\cdot, j)$ dominates $A(\cdot, j)$.

- $j \in [0.5, 1-z]$: $A_{i,j} = B_{i,j} = 1/(1-z)$ for $i \in [j, 1-z]$ and 0 otherwise. The sum is

(B.22)
$$(1-z-j)/(1-z)$$

for both A and B.

- $j \in [1-z, 1]$: Both A and B are all zero here.

We finally check (B.18).

$$\forall j \in [0,1] : \int_{i=j}^{1} A_{i,j} + \int_{i=0}^{j} A_{j,i} + \int_{i=j}^{1} B_{i,j} + \int_{i=0}^{j} C_{i,j} \leq V$$

- For $j \in [1-z, 1]$: Both $A(\cdot, j)$ and $B(\cdot, j)$ are zero here. Since $C(j, \cdot) = 0$ here, $\int_i A(j, i) = 1$, $C(i,j)$ is $1/(1-z)$ when $i \in [z, 1-z]$. So the total sum is

$$\int_{i=0}^{j} A_{j,i} + \int_{i=0}^{j} C_{i,j} = 1 + (1-2z)/(1-z) = 2 - \frac{z}{1-z} \leq V(z).$$

- For $j \in [0.5, 1-z]$: By (B.22),

$$\int_{i=j}^{1} A_{i,j} + \int_{i=j}^{1} B_{i,j} = 2(1-z-j)/(1-z).$$

$C(i,j) = 1/(1-z)$ for $i \in [z, j]$ and $1/(1-z-i)$ for $i \in [0, z]$. Finally, $A(j, i) = 1/(1-z)$ for $i \in [z, j]$. Therefore, the total sum is

$$\int_{i=j}^{1} A_{i,j} + \int_{i=0}^{j} A_{j,i} + \int_{i=j}^{1} B_{i,j} + \int_{i=0}^{j} C_{i,j} = \frac{2(1-z-j)}{1-z} + \frac{2(j-z)}{1-z} + \int_{i=0}^{z} \frac{1}{1-z-i}$$

$$= \frac{2(1-2z)}{1-z} + (\log(1-z) - \log(1-2z)) = \frac{2-4z}{1-z} + (\log(1 + \frac{z}{1-2z})) \leq V(z).$$

- For $j \in [z, 0.5]$: By (B.21),

$$\int_{i=j}^{1} A_{i,j} + \int_{i=j}^{1} B_{i,j} = \frac{2(1-z-j)}{1-z} + \frac{2y}{0.5-z},$$

which is $\frac{2y}{0.5-z} = \frac{4z^2}{(1-z)(1-2z)}$ more than the $j \in [0.5, 1-z]$ case. Everything else is identical. It is still at most $V(z)$.

- For $j \in [0, z]$: From (B.20),

$$\int_{i=j}^{1} A_{i,j} + \int_{i=j}^{1} B_{i,j} = \frac{2(1-2z)}{1-z} = 2 - \frac{2z}{1-z}.$$

And $A_{j,i} = 0$. Therefore,

$$\int_{i=j}^{1} A_{i,j} + \int_{i=0}^{j} A_{j,i} + \int_{i=j}^{1} B_{i,j} + \int_{i=0}^{j} C_{i,j}$$

$$= 2 - \frac{2z}{1-z} + \int_{i=0}^{j} \frac{1}{1-z-i} = 2 - \frac{2z}{1-z} + (\log(1-z) - \log(1-z-y))$$

$$\leq 2 - \frac{2z}{1-z} + (\log(1-z) - \log(1-2z)) \leq V(z).$$

Finally, let us compute N_i and M.

- For $i \in [0, z]$, $N_i := 1/(1 - z - i)$.

- For $i \in [z, 1 - z - y]$, $N_i := 1/(1 - z)$.

- For $i \in [1 - z - y, 1 - z]$, $N_i := 1/(1 - z) + 1/(0.5 - z)$.

Therefore,

$$M = \left(\log(1 - z) - \log(1 - 2z) \right) + \frac{1 - 2z - y}{1 - z} + y\left(\frac{1}{1 - z} + \frac{1}{0.5 - z} \right) = M(z).$$

B.3 Conversion to Discrete Dual Given the solutions to the continuous LP, the discretization step shown in the following claim proves Lemma B.3.

CLAIM 1. *Fix $T > 0$, and consider the solution A, B, C, V, N, M to the continuous dual program constructed in Section B.2, which is parameterized by $z \in [0, 1/3]$. If $z = d/q$ for some integer d, the value of the discrete factor-revealing LP with q clients is at most the dual value for the continuous program, which is $V + T(M - 1)$.*

Proof. We first construct dual solutions $\{a_{i,j}\}_{1 \leq j < i \leq q}$, $\{b_{i,j}\}_{1 \leq j < i \leq q}$, $\{c_{i,j}\}_{1 \leq i \leq j \leq q}$ to the discrete factor-revealing LP; V remains the same.

They are defined as follows. For $1 \leq i \leq q$, let $s_i := (i-1)/q$ and $t_i := i/q$. To simplify notations, we allow integrating A, B, or C over regions they are not defined; their value is 0 in those regions. Also, we always use the standard Lebesgue measure on $\mathbb{R}$ and $\mathbb{R}^2$ and omit $dxdy$ notations. (Variables y and x correspond to i and j respectively.)

- For all $1 \leq j < i \leq q$: $a_{i,j} = q \cdot \left(\int_{y=s_i}^{t_i} \int_{x=s_j}^{t_j} A_{y,x} \right)$.

- For all $1 \leq j < i \leq q$: $b_{i,j} = q \cdot \left(\int_{y=s_i}^{t_i} \int_{x=s_j}^{t_j} B_{y,x} \right)$.

- For all $1 \leq i \leq j \leq q$, $c_{i,j} = q \cdot \left(\int_{y=s_i}^{t_i} \int_{x=s_j}^{t_j} C_{y,x} \right)$.

  - If $i = j$, additionally increase $c_{i,j}$ by $q \cdot \left(\int_{y=s_i}^{t_i} \int_{x=s_i}^{t_i} A_{y,x} \right)$.

We first check they satisfy (B.14), (B.15), and (B.16).

- (B.14): for every $i \in [q]$,

$$\sum_{j=1}^{i-1} a_{i,j} + \sum_{j=i}^{q} c_{i,j}$$

$$= q \cdot \left(\sum_{j=1}^{i-1} \int_{s_i}^{t_i} \int_{s_j}^{t_j} A_{y,x} + \sum_{j=i}^{q} \int_{s_i}^{t_i} \int_{s_j}^{t_j} C_{y,x} + \int_{s_i}^{t_i} \int_{s_i}^{t_i} A_{y,x} \right)$$

$$= q \cdot \int_{y=s_i}^{t_i} \left(\int_{x=0}^{y} A_{y,x} + \int_{x=y}^{1} C_{y,x} \right)$$

$$= 1,$$

where the last equality follows from (B.17).

- (B.15): for all $j \in [q]$, first note that

$$\frac{1}{q} \cdot \left(\sum_{i=j+1}^{q} a_{i,j} + \sum_{i=1}^{j-1} a_{j,i} + \sum_{i=j+1}^{q} b_{i,j} + \sum_{i=1}^{j} c_{i,j} \right)$$

$$= \sum_{i=j+1}^{q} \int_{s_i}^{t_i} \int_{s_j}^{t_j} A_{y,x} + \sum_{i=1}^{j-1} \int_{s_i}^{t_i} \int_{s_j}^{t_j} A_{x,y} + \sum_{i=j+1}^{q} \int_{s_i}^{t_i} \int_{s_j}^{t_j} B_{y,x} + \sum_{i=1}^{j} \int_{s_i}^{t_i} \int_{s_j}^{t_j} C_{y,x} + \int_{s_j}^{t_j} \int_{s_j}^{t_j} A_{y,x},$$

where the last $\int_{s_i}^{t_i} \int_{s_j}^{t_j} A_{y,x}$ term appears because of the additional increases for $c_{j,j}$. This is at most

$$\int_{x=s_j}^{t_j} \int_{y=x}^{1} A_{y,x} + \int_{x=s_j}^{t_j} \int_{y=0}^{x} A_{x,y} + \int_{x=s_j}^{t_j} \int_{y=x}^{1} B_{y,x} + \int_{x=s_j}^{t_j} \int_{y=0}^{x} C_{y,x}$$

$$\leq \int_{x=s_j}^{t_j} \left(\int_{y=x}^{1} A_{y,x} + \int_{y=0}^{x} A_{x,y} + \int_{y=x}^{1} B_{y,x} + \int_{y=0}^{x} C_{y,x} \right) \leq \frac{V}{q}.$$

where the last inequality follows from (B.18).

- (B.16): we use (B.19) that proved $B(\cdot, x)$ dominates $A(\cdot, x)$ for every $x \in [0,1]$, which almost immediately implies (B.16). The only worry is that we *ignored* diagonals; for $j \in [q]$, since $a_{j,j}$ and $b_{j,j}$ are not variables, the integrals $a'_j := \int_{s_j}^{t_j} \int_{s_j}^{t_j} A_{y,x}$ and $b'_j := \int_{s_j}^{t_j} \int_{s_j}^{t_j} B_{y,x}$ were not added to any discrete variable, and it can possibly violate (B.16) if the latter is strictly larger than the former. So below we check that a'_j and b'_j are exactly same for every $j \in [q]$. Here we use that $z = d/q$ for some integer d.

  - $1 \leq j \leq d$: both a'_j and b'_j are zero.
  - $d < j \leq q - d$: both a'_j and b'_j are $(1/q^2) \cdot (1/(1 - z))$.
  - $q - d < j \leq q$: both a'_j and b'_j are zero.

Let $m_i = \max\left(\max_{j=1}^{i-1} b_{i,j}, \max_{j=i}^{q} c_{i,j} \right)$ for $i \in [q]$ and $N_y = \max\left(\max_{x \in [0,y]} B_{y,x}, \max_{x \in [y,1]} C_{y,x} \right)$ for $y \in [0,1]$. We show that

(B.23)
$$m_i \leq \int_{y=s_i}^{t_i} N_y,$$

which implies that the same $M = \int_{y=0}^{1} N_y$ is an upper bound of $\sum_{i=1}^{q} m_i$ and finishes the proof. By definition, for $i \in [q]$, $m_i = b_{i,j}$ for $j < i$ or $c_{j,i}$ for $j \geq i$. If $m_i = b_{i,j}$ for $j < i$,

$$m_i = b_{i,j} = q \cdot \int_{y=s_i}^{t_i} \int_{x=s_j}^{t_j} B_{y,x} \leq \int_{y=s_i}^{t_i} N_y.$$

The other case $m_i = c_{i,j}$ can be shown similarly. The only possible worry is that $c_{i,i}$ for $i \in [d+1, q-d]$ added an additional integral $\int_{s_i}^{t_i} \int_{s_i}^{t_i} A_{y,x}$. But in those regions all A, B, and C have the same value $1/(1-z)$, and A and C are defined only in the lower triangle and upper triangle respectively, $c_{i,i} = c_{i,i+1} = \ldots c_{i,q} = 1/(1-z)$ for $i \in [d+1, q-d]$. Therefore, adding an additional integral to $c_{j,j}$ does not increase m_i and (B.23) still holds. □

C Proof of Theorem 3.1

In this section we prove Theorem 3.1. We assume that the size of the swaps done by the algorithm is $\Delta = (2/\varepsilon)^{2/\varepsilon^7}$. We split the proof in two parts. In Section C.1 we will consider the simpler case $|S'| > k$, which also assumes swaps of smaller size. This helps to introduce part of the ideas. In Section C.2 we will consider the more complex case $|S'| \leq k$. We start with the following lemma showing that local search can be made polynomial-time.

LEMMA C.1. *For any $\varepsilon > 0$, the results of Lemmas C.2 and E.1 can be obtained in polynomial time up to losing an additive factor of $O(\varepsilon(opt + d'))$.*

Proof. We apply the following standard preprocessing of the graph. Let η be the connection cost of the optimum solution. Our algorithm partitions the input metric by creating a graph whose vertex set is the set of points of the input metric space and whose edge set is the set of all pairs of points at distance at most η; and applies the local search algorithm independently in each subinstance defined by the set of points that are in the same connected component.

Since the optimum connection cost is at most η, it is immediate that any optimum clustering is such that no two points of different connected components are in the same cluster, and so, obtaining the approximation

guarantees of Lemmas C.2E.1 for each subinstance is enough to recover the same approximation guarantees for the whole instance.

Next, for each subinstance of size n', let us assume that the minimum distance between any pair of points is $\varepsilon\eta/n'^2$ – this can be achieved up to losing a factor $(1+\varepsilon)$ in the connection cost of the optimum solution by "contracting" each pair of points at distance smaller than $\varepsilon\eta/n'^2$. Moreover, the maximum distance between any two points of the subinstance is at most $n' \cdot \eta$. Finally, we can assume without loss of generality that the opening cost is at most $2n'^2 \cdot \eta$ (since otherwise we know that the optimum solution opens at most 1 center and we can solve the instance optimally) and at least $\varepsilon^2\eta/n'^2$ (otherwise, since the minimum distance between a pair of points is $\varepsilon\eta/n'^2$, we can obtain a $(1+\varepsilon)$-approximate solution by opening a facility for each input client, at the candidate center location that is the closest).

We thus have that for each subinstance on n' points, the minimum cost is at least $\varepsilon^2\eta/n'^2$ and the maximum is $\eta n'^2$. The local search is run such that a new solution is picked if its cost improves the cost of the current solution by a factor at least $(1+\varepsilon^3/n'^5)$. Since the initial solution has cost at most $\eta n'^2$ and the final solution at most $\varepsilon^2\eta/n'^2$, the number of steps is at most $O(n'^9\varepsilon^{-5})$. Finally, two solutions of cost at least $\varepsilon^2\eta/n'^2$ and at most $\eta n'^2$ differ in cost by a multiplicative factor at most $(1+\varepsilon^3/n'^5)$ which is an additive factor of at most $\varepsilon^3/n'^5 \cdot \eta n'^2 = \varepsilon^3/n'^3 \leq \varepsilon/n \cdot \varepsilon^2\eta/n'^2$. Since the analysis of Lemmas C.2E.1 uses at most k potential swaps, the difference in cost between the analysis of the algorithm assuming it reaches a local optimum or a solution which is within a factor $(1+\varepsilon^3/n'^5)$ is indeed at most $\varepsilon\eta$ as desired.

To conclude, the algorithm is run with $\varepsilon^{-1}\log\frac{\text{maximum distance}}{\text{minimum distance}}$ different estimates of η, namely, all the powers of $(1+\varepsilon)$ times the minimum distance between any pair of points, and the best solution is returned. $\quad\square$

C.1 Case $|S'| > k$. We first describe a family of feasible swap pairs (A_i, B_i) for S', with $A_i \subseteq S'$ and $B_i \subseteq OPT$. Let $k' = S'$ and d' be the service cost of S'. The upper bound derives from the fact that, by the local optimality of S', for each such pair (A_i, B_i) one has

$$\lambda k' + d' \leq \lambda|S' \setminus A_i \cup B_i| + d(S' \setminus A_i \cup B_i).$$

We start by creating (possibly infeasible) swap pairs as follows. Each matched facility $f_2 \in S'^M$ and the corresponding matched facilities $M(f_2) \in OPT^M$ define a swap pair $(\{f_2\}, M(f_2))$. Each $f^* \in OPT^L$ is added to the swap pair containing the closest facility $cl(f^*) \in S'$, creating one such pair $(\{cl(f^*)\}, \{f^*\})$ if $cl(f^*)$ is not already contained in some swap pair. Notice that each swap pair now contains exactly one facility $f_2 \in S'$ and one or more facilities in OPT: such facility f_2 is the *leader* of the corresponding pair. Notice also that each $f^* \in OPT$ belongs to precisely one pair. Now we add the remaining lonely facilities S'^L in S' to the swap pairs, enforcing that each pair (A, B) satisfies $|A| \geq |B|$. Notice that this is always possible since $|S'| \geq |OPT|$. We also observe that the sum of the discrepancies $disc(A, B) = |B| - |A|$ is initially exactly $k - k'$. Note that these swap pairs may be such that $|A| > 1/\varepsilon$, and so may be infeasible. If this happens, we say that the pair is large. We break large pairs into smaller ones as follows. We take any subset $A' \subseteq A$ of $-disc(A, B) = |A| - |B|$ facilities and let $A'' = A \setminus A'$. We partition A'' and B into sets $A''_1, \ldots, A''_q$ and $B_1, \ldots, B_q$ such that $|A''_i| = |B_i|$ for all i and $|A''_i| = \frac{1}{4\varepsilon}$ for $i < q$. We partition A' into sets $A'_1, \ldots, A'_\ell$ where $|A'_i| = 1/(4\varepsilon)$ for $i < \ell$. We replace (A, B) with the swap pairs $(A''_1, B_1), \ldots, (A''_q, B_q), (A'_1, \emptyset), \ldots, (A'_\ell, \emptyset)$. We observe that all these swap pairs are obviously feasible.

However, we still need to perform one last modification to the swap pairs. Let a be the leader in A, and assume w.l.o.g. that $a \in A''_1$. For $q \geq 2$, when performing the swap (A''_1, B_1), we remove the leader of A without including in the solution all the facilities in B. This is problematic in the rest of our analysis. To resolve this issue, if $q \geq 3$, we replace a with a random facility $a' \in A''_2$. Otherwise, we simply remove a from A''_1. We remark that in both cases we keep the feasibility of the pair (A''_1, B_1).

Let $(A_1, B_1), \ldots, (A_h, B_h)$ be the resulting overall set of feasible swap pairs. The following three properties, which will be helpful in the following, trivially hold. From the local optimality of S' one has:

$$(C.24) \qquad \forall i : \lambda k' + d' \leq \lambda(k_2 + |B_i| - |A_i|) + d(S' \setminus A_i \cup B_i).$$

Each time we decrease by one the cardinality of some set A_1 derived from a large pair (A, B) to remove the leader a of that pair, this means that $\ell \geq 2$, and in particular the corresponding set A' satisfies $|A'| \geq \frac{1}{2\varepsilon}$. Notice that the sum of the cardinalities of the sets A' is precisely $k' - k$, and such sets are disjoint. Hence this event can

happen at most $2\varepsilon(k' - k)$ times. Thus:

$$(C.25) \qquad k - k' \leq \sum_i disc(A_i, B_i) = \sum_i (|B_i| - |A_i|) \leq k - k' + 2\varepsilon(k' - k).$$

Finally, by construction each facility $f_2 \in S'$ is contained in at most 2 pairs. When f_2 belongs to two pairs, this is because f_2 was sampled from a set A_2 of size $\frac{1}{4\varepsilon}$. Thus

$$(C.26) \qquad \forall f_2 \in S' : Pr[f_2 \text{ is contained in 2 swap pairs}] \leq 4\varepsilon.$$

We are now ready to prove the following upper bound.

Recall that given S' and opt, we let $\alpha^L := opt^L/opt$, $\alpha^M := opt^M/opt$, $\alpha^{MM} = opt^{MM}/opt$, $\beta = d'/opt$, and $\beta^{MM} = d'^{MM}/opt$. Let also $k^L = |OPT^L|$, $k^M = |OPT^M|$, $k'^L = |S'^L|$, and $k'^M = |S'^M|$.

LEMMA C.2. *One has*

$$\lambda k' + d' \leq \lambda k + 2\varepsilon\lambda(k' - k) + \frac{\delta(1 + 4\varepsilon)}{1 - \delta}(d'^{MM} + opt^{MM}) + 3(1 + 4\varepsilon)opt^L + (1 + 4\varepsilon)opt^M$$

$$= \lambda k + \varepsilon\left(2\lambda(k' - k) + \frac{4\delta}{1 - \delta}(d'^{MM} + opt^{MM}) + 12opt^L + 4opt^M\right) + \rho^A(\delta, \alpha^L, \beta, \alpha^{MM}, \beta^{MM})opt.$$

Proof. Let us consider the swap pairs $(A_1, B_1), \ldots, (A_h, B_h)$ as described before. Let $S(i) = S' \setminus A_i \cup B_i$. One has

$$(C.27) \qquad hd' \overset{(C.24)}{\leq} \sum_i \lambda(k' + |B_i| - |A_i|) + \sum_i d(S(i)) \overset{(C.25)}{\leq} \lambda(k - k') + 2\varepsilon\lambda(k' - k) + \sum_i d(S(i)).$$

Therefore it is sufficient to show that

$$(C.28) \qquad \mathbb{E}[\sum_i d(S(i))] \leq (h - 1)d' + \frac{\delta(1 + 4\varepsilon)}{1 - \delta}(d'^{MM} + opt^{MM}) + 3(1 + 4\varepsilon)opt^L + (1 + 4\varepsilon)opt^M.$$

We next focus on a specific client c. Let $f^*(c)$ and $f'(c)$ be the facilities serving c in OPT and S', resp. We use the shortcuts $d'(c) = d(c, S')$ and $opt(c) = d(c, OPT)$. The contribution of c to hd' is precisely $h\, dist(c, S') = h\, d'(c)$. Let us bound the expected contribution $\sum_i dist(c, S(i))$ of c to $\sum_i d(S(i))$. Define $j^*(c)$ such that $f^*(c) \in B_{j^*(c)}$. Define also $J'(c)$ as the set of indexes i such that $f'(c) \in A_i$. Notice that $J'(c)$ is empty only if $f'(c)$ is the leader of a large group. Otherwise, $J'(c)$ has cardinality 2 with probability at most 4ε, and otherwise cardinality 1.

We next assume that $J'(c)$ is empty or $J'(c) = \{j''(c)\}$. We later show how to fix the analysis for the case $|J'(c)| = 2$. We distinguish a few cases (for each case, we assume that the conditions of previous cases do not apply):

1. $f'(c)$ is the leader of a large group. Then

$$\sum_i dist(c, S(i)) = \sum_{i \neq j^*(c)} dist(c, S(i)) + dist(c, S(j^*(c))) \leq (h - 1)d'(c) + opt(c).$$

From now on we can assume that $j''(c)$ is defined.

2. $j^*(c) = j''(c)$ (i.e., both $f^*(c)$ and $f'(c)$ belong to the same swap pair). This is similar to the previous case

$$\sum_i dist(c, S(i)) = \sum_{i \neq j^*(c)} dist(c, S(i)) + dist(c, S(j^*(c))) \leq (h - 1)d'(c) + opt(c).$$

3. $f^*(c) \in OPT^M$. Let $f^*(c)$ be matched with $f_1(c) \in S'^M$, and $C'(f^*(c))$ be the clients served by $f^*(c)$ which are also served by $f_1(c)$. Let also $C(f^*(c))$ be all the clients served by $f^*(c)$. We observe that

$|C'(f^*(c))| \geq (1-\delta)|C(f^*(c))|$ by the definition of matching, and $c \in C(f^*(c)) \setminus C'(f^*(c))$ since we are not in case 2. We have the upper bound

$$dist(c, S(j''(c))) \leq dist(c, f_1(c)) \leq dist(c, f^*(c)) + dist(f^*(c), f_1(c))$$

$$\leq opt(c) + \frac{1}{|C'(f^*(c))|} \sum_{c' \in C'(f^*(c))} (dist(f^*(c), c') + dist(c', f_1(c)))$$

$$= opt(c) + \frac{1}{|C'(f^*(c))|} \sum_{c' \in C'(f^*(c))} (opt(c') + d'(c'))$$

Let us define $\Delta(c) = \frac{1}{|C'(f^*(c))|} \sum_{c' \in C'(f^*(c))} (opt(c') + d'(c'))$. Then

$$\sum_i dist(c, S(i)) = \sum_{i \neq j''(c)} dist(c, S(i)) + dist(c, S(j''(c))) \leq (h-1)d'(c) + opt(c) + \Delta(c).$$

4. $f^*(c) \in OPT^L$. Let us upper bound $dist(c, S(j''(c)))$. Let $f_1(c) \in S'$ be the closest facility to $f^*(c)$ (i.e. the leader of the group of $f^*(c)$). By the previous cases it must happen that $f_1(c) \notin A_{j''(c)}$, hence $f_1(c) \in S(j''(c))$. Thus

$$dist(c, S(j''(c))) \leq dist(c, f_1(c)) \leq dist(c, f^*(c)) + dist(f^*(c), f_1(c)) \leq opt(c) + dist(f^*(c), f'(c))$$

$$\leq opt(c) + dist(f^*(c), c) + dist(c, f'(c)) = d'(c) + 2opt(c).$$

Thus

$$\sum_i dist(c, S(i)) = \sum_{i \neq j''(c), j^*(c)} dist(c, S(i)) + dist(c, S(j''(c))) + dist(c, j^*(c))$$

$$\leq (h-2)d'(c) + opt(c) + d'(c) + 2opt(c) = (h-1)d'(c) + 3opt(c).$$

We observe that

$$\sum_{c \text{ in case } 3} \Delta(c) \leq \sum_{c' \in C^{MM}} \frac{|C(f^*(c)) \setminus C'(f^*(c))|}{|C'(f^*(c))|} (opt(c') + d'(c'))$$

$$\leq \frac{\delta}{1-\delta} \sum_{c' \in C^{MM}} (opt(c') + d'(c')) \leq \frac{\delta}{1-\delta} (opt^{MM} + d'^{MM}).$$

Summing the different cases over the different clients one obtains

$$\sum_i d(S(i)) \leq (h-1)d' + opt + \sum_{c \text{ in case } 4} 2opt(c) + \sum_{c \text{ in case } 3} \Delta(c)$$

$$\leq (h-1)d' + opt + 2opt^L + \frac{\delta}{1-\delta}(opt^{MM} + d'^{MM})$$

$$= (h-1)d' + \frac{\delta}{1-\delta}(d'^{MM} + opt^{MM}) + 3opt^L + opt^M$$

Recall that with probability 4ε for each client not in case 1 above it might happen that there exists a second index $j_2''(c)$ such that $f'(c) \in A_{j_2''(c)}$. In that case we might need to add some extra terms in cases 3 and 4 above, and also add a term $\min\{0, opt(c) - d'(c)\}$. We pessimistically assume that the latter term is 0, and that the mentioned probability is precisely 4ε for each such client c. Summing over all clients one gets in expectation an extra additive term of value at most

$$4\varepsilon(\frac{\delta}{1-\delta}(d'^{MM} + opt^{MM}) + 3opt^L + opt^M).$$

This gives the desired inequality (C.28). □

C.2 Case $|S'| \leq k$. We complete the proof of Theorem 3.1. In fact we are going to prove a slightly stronger theorem that implies Theorem 3.1. Let $\delta', \delta^* \in (0, 1/2)$ be parameters to be fixed later. We say that $f^* \in OPT$ *captures* $f' \in S'$ if more than a $1 - \delta^*$ fraction of the clients served by f' in S' are served by f^* in OPT. We say that $f' \in S'$ captures $f^* \in OPT$ if more than a $1 - \delta'$ fraction of the clients served by f^* in OPT are served by f' in S'. Notice that the definition is *not* symmetric. We moreover use a generalization of the definition of isolated regions from [CKM19]: We say that a set of facilities $F' \subseteq S'$ and a set of facilities $F^* \subseteq OPT$ such that $|F^*| \leq 1/\epsilon$ are in *isolation* if at least a $1 - \delta'$ fraction of the clients served by F^* in OPT are served by F' in S' and if at least a $1 - \delta^*$ fraction of the clients served by F' in S' are served by F^* in OPT. In the following, we consider an arbitrary set of disjoint pairs of facilities (F', F^*) in isolation.

We let OPT^M be an arbitrary set of facilities of OPT that satisfy either of the following: (1) they are captured by some facility in S', or (2) they are in at least one chosen group that is in isolation. We let $OPT^L = OPT \setminus OPT^M$. We let S^M be the set of facilities of S' that are captured by some facility in OPT or in a chosen group that is in isolation, and set $S^L = S' \setminus S^M$. We also let $opt^M = d(OPT^M)$, $opt^L = d(OPT^L)$, $d^M = d(S^M)$ and $d^L = d(S^L)$. We let $C^{XY} = \{c \in C \mid S'(c) \in S^X \text{ and } OPT(c) \in OPT^Y\}$, e.g.: C^{LM} is the set of clients served by a lonely facility in S' and by a matched facility in OPT. We let $C^{\star M} = C^{LM} \cup C^{MM}$, etc.. We let d^{XY} and opt^{XY} be the connection cost in S' and OPT, resp., of the clients in C^{XY}. We let $d^{\star X} = d^{LX} + d^{MX}$, etc..

We are ready to state our local-search based bound.

THEOREM C.1. *We have*

$$\lambda k' + d' \leq$$

$$k\lambda + 3opt^{LL} + opt^M + \frac{\delta'}{1 - \delta'}(d^{MM} + opt^{MM}) + O(\varepsilon(d' + opt)) + \min \begin{cases} 3opt^{ML} \\ opt^{ML} + \frac{\delta^*}{1 - \delta^*}(d^M + opt^M) \end{cases}$$

Assuming the above theorem, we can conclude the proof of Theorem 3.1.

Proof. [Proof of Theorem 3.1] In the case where the solution S' produced by the algorithm is such that $|S'| > k$, then we invoke Lemma C.2 and obtain the desired bounds.

In the case where $|S'| \leq k$, then we have that the matching of facilities defined in Section 3 slightly differs from the more general one we have in this section; it is the following: Given the parameter δ of Theorem 3.1, consider $f^* \in OPT$, and let $M(f^*)$ be the facilities in S' which are $1/2$-captured by f^*. If $M(f^*)$ $(1-\delta)$-captures f^*, then we say that f^* and $M(f^*)$ are *matched*. If $M(f^*)$ $(1 - \delta)$-captures f^*, then we say that f^* and $M(f^*)$ are *matched*.

We now translate the above matching in the matching defined for Theorem C.1. For each such pair above, we create a pair of facilities in isolation $(f^*, M(f^*))$, note that by definition these pairs are pairwise disjoint (i.e.: for all pairs $(f^*, M(f^*)), (f_1^*, M(f_1^*))$, we have $f_1^* \neq f^*$ and $M(f_1^*) \cap M(f^*) = \emptyset$). We can thus apply Theorem C.1, with $\delta' := \delta$ and $\delta^* := 1/2$, and define OPT^M to be the set containing the chosen groups of facilities in isolation described above. It follows that the set of facilities of OPT and the set of facilities of S' that are matched according to the definition of Theorem 3.1 and matched according to the definition of Theorem C.1 are the same. We can thus write:

$$\lambda k' + d' \leq k\lambda + 3opt^L + opt^M + \frac{\delta'}{1 - \delta'}(d^{MM} + opt^{MM}) + O(\varepsilon(d' + opt)),$$

as claimed. □

We complete the proof by showing Theorem C.1.

Proof. [Proof of Theorem C.1] We form swap sets (A_i, B_i), $A_i \subseteq S'$ and $B_i \subseteq OPT$, using the following directed graph G. For each chosen group (F_1, F^*) of facilities in isolation, create a vertex v_f for each $f \in F_1$ and a vertex v_{F^*} for F^* and add edges $\langle v_f, v_{F^*} \rangle$ then repeat on the remaining facilities. For each remaining facility f of OPT or S' create a vertex v_f. Then, for each $f' \in OPT^M$ and $f \in S'$, create an edge from v_f to $v_{f'}$ if f' captures f. Namely, if f' is in OPT^M and f is in S' then at least a $1 - \delta^*$ fraction of the clients served by f in S' are served by f' in OPT. Also create an edge $v_{f'}$ to v_f if f' is in S' and $f \in OPT$ and at least a $1 - \delta'$ fraction of the clients

served by f in OPT are served by f' in S'. Moreover, for each facility f of $OPT - OPT^M$, create an edge from f to the vertex corresponding to the closest facility of S'.

Note that the out-degree of each vertex of the graph is at most 1 and the out-degree of any vertex v_{F^*} is 0. For each vertex v_f with outgoing edge $\langle v_f, v_{f'} \rangle$, we define $\phi(f) = f'$. If v_f has out-degree 0 we let $\phi(f) = \emptyset$.

We let $\tilde{F}^*$ be the set of facilities f of OPT that satisfy (1) $\phi(f)$ has indegree at least 2; or (2) f is part of a chosen group of facilities in isolation (F_1, F^*) such that $|F^*| > |F_1|$.

We now choose a collection of subsets $\{(F_i, F_i^*)\}$, where $F_i \subseteq S'$ and $F_i^* \subseteq OPT$ and modify OPT to obtain a solution $\widetilde{OPT} = OPT \bigcup_i F_i \setminus \bigcup_i F_i^*$ that satisfies the following:

- Each client c served by a facility in F_i^* in OPT it is served by a facility in F_i in $\widetilde{OPT}$.

- The service cost of $\widetilde{OPT}$ is at most the service cost of OPT plus $O(\varepsilon(d' + opt))$.

- $\widetilde{OPT}$ opens at most $|OPT| - 2\varepsilon^2 |\tilde{F}^*|$ centers.

To define a solution $\widetilde{OPT}$ that satisfies the above, we proceed as follows. Consider each facility $f \in S'$ with indegree at least 2, and let $OPT(f)$ be the set of facilities of OPT with an outgoing edge toward f. Consider the group $f, OPT(f)$. We bound the change in cost induced by the following *group exchange*: add f and remove $OPT(f)$ from OPT and serve all the clients served by a facility in $OPT(f)$ by f. Since for each f^* we have that either f captures f^* or f is the closest facility to f^*, we have that the change in cost is at most

$$\frac{1}{1 - \delta'} \sum_{c:\ \text{served by } f^* \text{ in } OPT} dist(c, S') + dist(c, OPT).$$

Consider also any chosen group of facilities in isolations F_1, F^* and define a group exchange which consists in replacing F^* with F_1 in OPT. Replacing F^* with F_1 and serving the clients served by F^* in OPT with F_1 also increases the cost for these clients by at most the same amount as above.

Therefore, consider all the chosen groups of facilities in isolation (F_1, F^*), and all pairs (f, F^*) where $f \in S'$ and $F^* \subseteq OPT$, $|F^*| > 1$, and where each facilities $f^* \in F^*$ has an edge toward f. Let $\tilde{k}$ be the number of groups. Pick a random subset of $2\varepsilon\tilde{k}$ groups and perform the group exchange for each picked group. The expected cost increase is thus at most $O(\varepsilon(d' + opt))$. Moreover, each group exchange reduces the number of centers in the solution by at least 1 and since for each group $|F^*| \leq 1/\varepsilon$, we have that the number of groups is at least $\varepsilon|\tilde{F}^*|$. Therefore, the resulting solution contains at most $|OPT| - 2\varepsilon^2|\tilde{F}^*|$ centers. Denote by $I = S' \cap \widetilde{OPT}$. Observe that the service cost for the clients served by any $f^* \in S' \cap \widetilde{OPT}$ in $\widetilde{OPT}$ is as good in S' since $f^* \in S'$. Importantly, we keep the chosen group of facilities in isolation unchanged in spirit: if a group of a facilities in isolation (F^*, F) has been chosen above and so $F \subseteq \widetilde{OPT}$, we replace (F^*, F) by (F, F).

We now describe how to adapt the graph G. Each vertex representing a facility or a group of facilities that is both in S' and $\widetilde{OPT}$ is removed from the G and so the resulting graph only contains facilities of $S' - I$ and $\widetilde{OPT} - I$. Moreover, for each facility f and group of facilities F^* of OPT with outgoing edges towards f for which the exchange was performed, we remove f and F^* from the graph (and all the adjacent edges).

We now want to build a set of *swaps* $\{(A_i, B_i)\}$, namely a collection of pairs $A_i \subset S' \setminus I$, $B_i \subset \widetilde{OPT} - I$, where $|A_i| + |B_i| \leq \Delta$. From there we can analyze the cost of the solutions $S' \setminus A_i \cup B_i$ and obtain a bound on the cost of S'. We will give a randomized procedure that produces a set of feasible swaps that satisfies the following property.

Swap consistency: For any feasible swap (A_i, B_i), we have:

1. Property 1: For any facility $f \in S'$, if $f \in A_i$, we have $\phi(f) \in B_i$ with probability 1 if f is in a chosen group of facilities in isolation, and at least $(1 - \varepsilon)$ otherwise. In other words, if the swap *closes* f then it must open $\phi(f)$ with probability at least $(1 - \varepsilon)$ if f is not in an isolated group and with probability 1 otherwise;

2. Property 2: For any facility $f^* \in \widetilde{OPT} \setminus I$, if $f^* \notin B_i$, then $\phi(f^*) \notin A_i$. In other words, if the swap does not *open* f^* then it must leave $\phi(f^*)$ open;

3. Property 3: Each $f^* \in \widetilde{OPT} \setminus I$ belongs to at least one group and $\sum_i |B_i| \leq k + \varepsilon^2 |\tilde{F}^*|$;

4. Property 4: Each facility of S' belongs to at most one group. The number of facilities of $S' \setminus I$ that do not belong to any group is at most $\varepsilon^2 |\tilde{F}^*|$. The facilities of I do not belong to any group.

CLAIM 2. *There exists a set of feasible swaps that satisfies the swap consistency properties.*

Let's first assume that the claim holds and proceed to the analysis. Let h be the number of pairs (A_i, B_i) in the set of swaps prescribed by Claim 2.

For each swap set (A_i, B_i) we consider the local update where the algorithm opens B_i and removes A_i from S'. Note this is a feasible local update since we assume $|B_i| + |A_i| \leq \Delta$. Therefore, by the local optimality of S', it must be the case that, for $S(i) = S' \cup B_i - A_i$, one has

(C.29)
$$\lambda k' + d' \leq \lambda(k' + |B_i| - |A_i|) + d(S(i)).$$

Summing over i, from (C.29) one gets

(C.30)
$$\lambda h k' + h d' \leq \sum_i \lambda(k' + |B_i| - |A_i|) + \sum_i d(S(i)) = h\lambda k' + \lambda(k - k') + \sum_i d(S(i)),$$

where we have used Property 3 of the swap consistency, the definition of $\widetilde{\text{OPT}}$, and the fact that the swaps are defined on $S' \setminus I$ and $\widetilde{\text{OPT}} \setminus I$. We next focus on a specific client c. Let $f^*(c)$ and $f_1(c)$ be the facilities serving c in $\widetilde{\text{OPT}}$ and S', respectively. We use the shortcuts $d'(c) = dist(c, S')$ and $\widetilde{\text{opt}}(c) = dist(c, \widetilde{\text{OPT}})$. The contribution of c to $h d'$ is precisely $h\, dist(c, S') = h\, d'(c)$. Let us bound the contribution $\sum_i dist(c, S(i))$ of c to $\sum_i d(S(i))$.

First consider a client c served by a facility in I in $\widetilde{\text{OPT}}$. Then, for the swap that closes the facility serving it in S' (if there is such, and there can be at most one by Property 3), we can always reassign c to the facility serving it in I. In this case, the cost is $\widetilde{\text{opt}}(c)$. In any other swap it is at most $d'(c)$. Hence, the contribution is always at most $(h-1)d'(c) + \widetilde{\text{opt}}(c)$. Additionally, for any client c such that there is no swap containing the facility serving c in S', then since there must be a swap containing f^* (by Property 3), we have that the cost contribution of c is at most $(h-1)d'(c) + \widetilde{\text{opt}}(c)$. We will see in the remaining that the bound provided dominates this one and we henceforth focus on the case where there is at least one swap containing the facility serving c in S'.

We now turn to the rest of clients; We have 3 types of clients among the rest:

- $c \in C^{\star M}$: an *OPT-matched* client.

- $c \in C^{ML}$: an *S'-matched* client.

- $c \in C^{LL}$: a *lonely* client.

OPT-**matched clients.** Consider a client c served in *OPT* by facility f^* that is captured or in a chosen isolation group. Let f be the facility serving c in S'. Consider the swap (A_i, B_i) where $f \in A_i$. We bound the cost paid by c in $S(i)$ by one of the following

1. $\widetilde{\text{opt}}(c)$ if $f^* \in B_i$,

2. Otherwise, the distance from c to the facility f' of S' that is captured by f^* that is the closest to f^* if f^* is in a chosen isolation group, or

3. $\phi(f^*)$ if f^* is captured.

We argue that these are valid reassignments. (1) is obviously valid since it assumes that $f^* \in B_i$. For (2), observe that by Property 1, if $f^* \notin B_i$ then, $f' \notin A_i$ (since f' is captured by f^* and f^* is in isolation, f' is in an isolated group with f^* and so this holds with probability 1). For (3), by Property 2, if $f^* \notin B_i$ then $\phi(f^*) \notin A_i$. It follows that $dist(c, f')$ is a valid upper bound on the cost of c in $S(i)$ if f^* is in a chosen isolation group and $dist(c, \phi(f^*))$ is a valid upper bound on the cost of c if f^* is captured.

In the case of (2), let us now provide an upper bound on $dist(c, f')$, and therefore on the cost of c in $S(i)$. We write $dist(c, f') \leq dist(c, f^*) + dist(f^*, f')$. Then, denote by $U(f^*)$ the set of clients that are served by f^* in *OPT*

and by a facility captured by f^* in S'. Thus, we have that $dist(f^*, f') \leq \frac{1}{|U(f^*)|} \sum_{c' \in U(f^*)} dist(c', f^*) + dist(c', S')$. Hence

$$dist(c, f') \leq dist(c, f^*) + \frac{1}{|U(f^*)|} \sum_{c' \in U(f^*)} dist(c', f^*) + dist(c', S')$$

Finally, for each such f^*, the number of clients served by f^* in OPT that are not served by a facility of S' captured by f^* is at most $\frac{\delta'}{1-\delta'}|U(f^*)|$.

In the case of (3), let's then provide an upper bound on the $dist(c, \phi(f^*)) \leq dist(c, f^*) + dist(f^*, \phi(f^*))$. Denote by $U(f^*)$ the set of clients that are served by f^* in OPT and by a $\phi(f^*)$ in S'. We have that $dist(f^*, \phi(f^*)) \leq \frac{1}{|U(f^*)|} \sum_{c' \in U(f^*)} dist(c', f^*) + dist(c', S')$. Here again, for each such f^*, the number of clients served by f^* in OPT that are not served by $\phi(f^*)$ in S' is at most $\frac{\delta'}{1-\delta'}|U(f^*)|$.

Hence, the total service cost contribution of the OPT-matched clients to $\sum_i d(S(i))$ is at most

$$(h-1)d'^{\star M} + opt^{\star M} + \frac{\delta'}{1-\delta'}\left(opt^{MM} + d'^{MM}\right) + O(\varepsilon(d' + opt)).$$

S'-matched clients. Consider a client c served in S' by a facility f that is captured by some facility of $\widetilde{OPT}$ and served in $\widetilde{OPT}$ by a facility f^* not in isolation. Let (A_i, B_i) be the swap such that $f^* \in B_i$, and (A_j, B_j) be the swap such that $f \in A_j$. We have two options: (1) $i = j$ or (2) $i \neq j$. In the first case, we have that the cost for the swap (A_i, B_i) is at most $\widetilde{opt}(c)$, and $d'(c)$ for any other swap. In the second case, we bound the cost as follows. For the swap (A_i, B_i) the cost is at most $\widetilde{opt}(c)$. For any swap (A_ℓ, B_ℓ), $\ell \notin \{i, j\}$, we have that the service cost for c is at most $d'(c)$. Now, for the swap (A_j, B_j) we have two options which we will bound by the minimum of two quantities since the local search equation holds for both cases.

1. Case 1: We bound the cost of c by the distance from c to the facility of S' that is the closest to f^*. Since f^* is not in isolation, it has an edge to the closest facility f' in S'. Since $f^* \notin B_j$ it must be that $f' \notin A_j$. Hence this is a valid reassignment. In this case, the cost is by triangle inequality at most $dist(c, f^*) + dist(f^*, f') \leq dist(c, f^*) + dist(c, f^*) + dist(c, f)$, since $dist(f^*, f') \leq dist(f^*, f)$, and so at most $2\widetilde{opt}(c) + d'(c)$. Summing over all such clients, we have a total cost contribution of at most $(h-1)d'^{ML} + 3opt^{ML} + O(\varepsilon d')$. Note that this bound holds with probability 1.

2. Case 2: Recall that f is captured (possibly in isolation). Let F^* denote the set of facilities capturing f (it is the set of facilities of $\widetilde{OPT}$ if f is in isolation or a single facility otherwise). We bound the cost of c by the distance from c to the facility $\tilde{f} \in F^*$ that is the closest to f. By Property 1, $f \in A_j$ implies that $F^* \subseteq B_j$ with probability $1 - \varepsilon$ and so this is a valid upper bound on the cost only with this probability and we use Case 1 to provide an upper bound otherwise. Thus, this cost is at most $dist(c, f) + dist(f, F^*) + \varepsilon 3\widetilde{opt}(c)$. Let $U(f)$ be the set of clients served by f in S' and F^* in OPT. Again, by triangle inequality we have that $dist(f, F^*) \leq \frac{1}{|U(f)|} \sum_{c' \in U(f)} dist(c', f) + dist(c', F^*)$. It follows that the cost for c is at most $dist(c, f) + \frac{1}{|U(f)|} \sum_{c' \in U(f)} dist(c', f) + dist(c', F^*)$. Now since f is captured by F^*, we have that the number of clients served by f and not in $U(f)$ is at most $\frac{\delta^*}{1-\delta^*}|U(f)|$. Therefore, summing over all such f, the service cost contribution for the S'-isolated clients served by f is at most $(h-1)d'^{ML} + opt^{ML} + \frac{\delta^*}{1-\delta^*}(d'^{M\star} + opt^{M\star}) + O(\varepsilon d')$.

Lonely clients. Consider a client c served in S' by facility f that is not captured and served in OPT by a facility f^* that is not captured. Let (A_i, B_i) be the swap such that $f^* \in B_i$. For the swap (A_i, B_i) the cost is at most $\widetilde{opt}(c)$ since $f^* \in B_i$. Now, for any swap (A_j, B_j) where $f \in A_j$. We bound the cost of c by the distance from c to the facility of S' that is the closest to f^*. Since f^* is not in isolation, it has an edge to the closest facility f' in S'. By Property 2, $f^* \notin B_j$ implies that $f' \notin A_j$. Hence this is a valid reassignment. In this case, the cost is by triangle inequality at most $dist(c, f^*) + dist(f^*, f') \leq dist(c, f^*) + dist(c, f^*) + dist(c, f)$, since $dist(f^*, f') \leq dist(f^*, f)$, and so at most $2\widetilde{opt}(c) + d'(c)$. Thus, summing over all such clients, we have a total cost contribution of at most $(h-1)d'^{LL} + 3(1+\varepsilon)opt^{LL} + \varepsilon d'$. Combining the three above bounds yield the bound of the theorem.

We now turn to the proof of Claim 2.

Proof. [Proof of Claim 2] We will work with the graph $G = (A, B, E)$ defined above, where A, B are the vertices representing the facilities of $S' - I$ and $\widetilde{\text{OPT}} - I$ respectively. We first preprocess the graph and then partition its vertices (hence partitioning the associated facilities). For each part P_i of the partition we will create a swap (A_i, B_i) by setting A_i (resp. B_i) to be the set of facilities corresponding to vertices in $P_i \cap A$ (resp. $P_i \cap B$). Before defining the partition of G, we make the following modifications to the graph.

Degree reduction. For each vertex $v \in B$ with indegree $d > 1/\varepsilon^2$, called a *heavy vertex* create $\lfloor d\varepsilon^2 \rfloor$ new vertices $v_1, \ldots, v_{\lfloor d\varepsilon^2 \rfloor}$ that will go in the B side of the graph and remove v. We say that the new vertices also represent the facilities of $\widetilde{\text{OPT}}$ that v was representing (the facilites are now represented $\lfloor d\varepsilon^2 \rfloor$ times in the graph). Let us now bound the number of vertices in B at the end of this procedure. Let $B_{>\varepsilon^{-2}}$ be the set of vertices of B with indegree greater than $1/\varepsilon^2$. For each vertex $v \in B_{>\varepsilon^{-2}}$ with indegree d the contribution to the number of new vertices is $\lfloor d\varepsilon^2 \rfloor - 1$. Recall that $|A| \le |B| + \varepsilon^2 |\tilde{F}^*|$ and $\tilde{F}^*$ is the set of facilities f of $\widetilde{\text{OPT}}$ that satisfy (1) $\phi(f)$ has indegree at least 2; or (2) f is part of a chosen group of facilities in isolation (F_1, F^*) such that $|F^*| > |F_1|$. Thus the number of added vertices is at most $(1 + \varepsilon)\varepsilon^2 |\tilde{F}^*|$. Moreover, any vertex v_f where $f \in S'$ that has indegree larger than $1/\varepsilon^2$ is placed into a new set $\bar{A}$, we immediately infer that $|\bar{A}| \le \varepsilon^2 |\tilde{F}^*|$.

We now define edges adjacent to the new vertices. For each new vertex, the outgoing edge of v is present in each new vertex as well. The incoming edges are arbitrarily partitioned into $\lfloor d\varepsilon^2 \rfloor$ groups of size at most $2/\varepsilon^2$ and each new vertex receives a different group of incoming edges. The graph consisting of G plus the new vertices and the new edges, and resulting from deleting vertices of $\bar{A}$, and deleting the heavy vertices has thus degree at most $1/\varepsilon^2$.

Defining U and $\mathcal{P}$. Consider the graph where the above degree reduction procedure has been applied and the vertices of $\bar{A}$ have been deleted. We will now provide a partitioning of the above graph into parts (A_i, B_i) such that $|A_i| + |B_i| \le (1/\varepsilon + 1)^{O(1/\varepsilon^7)}$.

Then, contract all edges from $f^* \in \widetilde{\text{OPT}}$ to $f \in S'$ and replace the vertex by one vertex representing f and all the facilities of $\widetilde{\text{OPT}}$ now contracted to f. Moreover, for each pair of isolated facilities F_1, F^*, contract the edges $\langle v_f, v_{F^*} \rangle$. Since the outdegree of v_{F^*} is by definition 0, the degree of this graph is thus at most $1/\varepsilon^3$. Note that the only remaining edges in this graph are of the form $\langle f, f^* \rangle$, where $f \in S', f^* \in \widetilde{\text{OPT}}$. Let G' be the resulting graph.

Therefore, consider the following randomized procedure on G'. For each connected component of G', start at an arbitrary vertex of $A \setminus \bar{A}$ and performs a BFS on the undirected version of the graph, labeling the vertices of $A \setminus \bar{A}$ at hop-distance i with label $i \mod 1/\varepsilon^4 + 1$. Then pick a random integer i^* in $[0, 1/\varepsilon^4]$ and remove all the outgoing edges of the vertices with label equal to i^* – the probability of any edge of being removed is thus at most ε^4.

Since the degree of each vertex is at most $1/\varepsilon^3$, the connected components of the graph have thus size at most $(1/\varepsilon + 1)^{3/\varepsilon^4}$. We now create a part in $\mathcal{P}$ for each of them. Since each vertex may represent up to $1/\varepsilon^3$ vertices, we have that the total number of vertices represented in a given connected component is at most $(1/\varepsilon + 1)^{3/\varepsilon^7}$. For each connected component (A_i, B_i, E_i) we create a pair (U_i, V_i) where U_i is the set of facilities represented by vertices in A_i and V_i is the set of facilities represented by vertices in B_i. We now show that this set of swaps satisfies the swap consistency properties.

Let us start with Property 3. Observe that each vertex of G representing a facility of $\widetilde{\text{OPT}}$ and of degree at most $1/\varepsilon^2$ is also in G' and each vertex of G representing a facility of $\widetilde{\text{OPT}}$ and of degree $> 1/\varepsilon^2$ appears with multiplicity $\lfloor d\varepsilon^2 \rfloor$. So since each facility of $\widetilde{\text{OPT}}$ is initially represented in G it is also represented in G', and since we take a partition of the vertices of G' it must appear in at least one swap created. Moreover, since the number of new vertices in G' is at most $\varepsilon^2 |\tilde{F}^*|$, we have that $\sum_i |B_i| \le k + \varepsilon^2 |\tilde{F}^*|$, as desired.

Property 4 follow immediately from the fact that we have a partition of G' which contains vertices that represent all the facilities of S' with multiplicity 1 (note that we did not create copies of vertices of A) except for the facilities of S' whose corresponding vertices is in $\bar{A}$, in which case the facility is not present in any swap. Recall moreover that $|\bar{A}| \le \varepsilon^2 |\tilde{F}^*|$.

We now turn to Property 1. First, any edge of G that has been contracted is satisfied (since all the facilities are represented by one single vertex in G' and so necessarily in the same swap). Then, as discussed above, each edge of G' is cut with probability at most ε^4. If an edge $\langle f, f^* \rangle$ is not cut then f, f^* are in the same component and so for any swap (A_i, B_i), if $f \in A_i$ then $f^* = \phi(f) \in B_i$ as desired. Moreover, since G does not contain the

facilities of I anymore, we have that no facility in I is in any swap.

We finally turn to Property 2. Observe that all edges from f^* to f have been contracted, or the vertex corresponding to f has been deleted. In the first case, we have that f^* and f are represented by the same vertex in G' and so are necessarily in the same swap. In the latter case, f does not belong to any swap. Hence, for any swap (A_i, B_i), if $f^* \notin B_i$ then $f \notin A_i$. $\quad\square$

$\quad\square$

D Omitted Proofs from Section 4

Proof. [Proof of Lemma 4.2] Consider any $f' \in S_2^L$. Let $\Delta(f')$ be the increase of the connection cost of S_2 due to the removal of f'. The local optimality of S_2 implies $\lambda \leq \Delta(f')$, hence

$$\lambda k_2^L \leq \sum_{f' \in S_2^L} \Delta(f').$$

Thus it is sufficient to upper bound the righthand side of the above inequality. In order to do that, we consider each client c in the set $C_{S_2}(f')$ of clients served by f' in S_2^L, and describe a random path $P(c)$ from c to some facility in S_2 distinct from f'. Let $\ell(P)$ be the length of a path P. The sum of values $\ell(P(c)) - d_2(c)$ over clients $c \in C_{S_2}(f')$ gives a valid upper bound on $\Delta(f')$, and we will sum these values over all clients served by any $f' \in S_2^L$. The path $P(c)$ is defined as follows. Let $f^*(c) = OPT(c)$ be the facility serving c in OPT. We distinguish two cases: (a) $f^*(c)$ is not $(1-\delta)$-captured by f', and (b) the complementary case. In case (a), let $rand(c)$ be a random client served by $f^*(c)$ which is not served by f' in S_2. The random path $P(c)$ is given by the concatenation of the shortest paths from c to $f^*(c)$, from $f^*(c)$ to $rand(c)$, and from $rand(c)$ to the facility $f'(rand(c))$ serving $rand(c)$ in S_2. Notice that

$$\ell(P(c)) = opt(c) + opt(rand(c)) + d_2(rand(c)).$$

Case (b) is slightly more complex. Let $C'(f') \subseteq C_{S_2}(f')$ be the clients which in OPT are served by a facility not captured by f'. We remark that $|C'(f')| \geq |C_{S_2}(f')|/2$ since otherwise f' would be matched. Let $next(c)$ be a random client in $C'(f')$. The path $P(c)$ is the concatenation of the shortest path from c to f', the shortest path from f' to $next(c)$, and the random path $P(next(c))$ already defined earlier. In this case we have

$$\ell(P(c)) = d_2(c) + d_2(next(c)) + \ell(P(next(c)))$$
$$= d_2(c) + d_2(next(c)) + opt(next(c)) + opt(rand(next(c))) + d_2(rand(next(c)).$$

Let $next^{-1}(c)$ be the clients c' such that $c = next(c')$. We have that, for $c \in C'(f')$, $Ex[|next^{-1}(c)|] \leq 1$ (and $next^-1(c) = \emptyset$ otherwise). Let $\ell'(P(c))$ be the part of the lengths $\ell(P(c))$ not involving $rand(\cdot)$. Let also $\ell''(P(c))$ be the remaining part of each such length. One has

$$Ex[\sum_{f' \in S_2^L} \sum_{c \in C_{S_2}(f')} (\ell'(P(c)) - d_2(c)] = \sum_{f' \in S_2^L} \sum_{c \in C'(f')} (opt(c) - d_2(c) + Ex[|next^{-1}(c)|](d_2(c) + opt(c)))$$
$$\leq \sum_{f' \in S_2^L} \sum_{c \in C'(f')} 2opt(c) \leq 2(opt^{LM} + opt^{LL}).$$

In order to upper bound the sum of the cost $\ell''(P(c))$, let us define $rand^{-1}(c)$ as the clients c' in some set $C'(f')$ such that $c = rand(c')$. We wish to upper bound $Ex[|rand^{-1}(c)|]$. Consider the facility f^* serving c in OPT, and let us partition the clients $C_{OPT}(f^*)$ served by f^* into subsets $C^1, \ldots, C^h$ depending on the facility $f^1, \ldots, f^h$ serving them in S_2. Assume w.l.o.g. that $c \in C^1$. For each $c' \in C^j$, $j > 1$, c' belongs to $rand^{-1}(c)$ with probability $\frac{1}{|C_{OPT}(f^*)| - |C^j|}$. Thus

$$Ex[|rand^{-1}(c)|] = \sum_{i \geq 2} \frac{|C^i|}{|C_{OPT}(f^*)| - |C^i|}.$$

Observe that for $x, y \geq 0$ and $x + y \leq z > 0$, $\frac{x}{z-x} + \frac{y}{z-y} \leq \frac{x+y}{z-x-y}$. By repeatedly applying this inequality, one obtains that the largest value of the above sum is achieved when $h = 2$, being in particular $|C^2|/|C^1|$. Notice

however that f^2 does not $(1-\delta)$-capture f^* by construction. Indeed otherwise clients $c' \in C^2$ would not consider selecting $rand(c)$ in C^1. This implies $|C^2| \leq (1-\delta)|C(f^*)|$, hence $|C^2|/|C^1| \leq \frac{1-\delta}{\delta}$. Thus

$$Ex[|rand^{-1}(c)|] \leq \frac{1-\delta}{\delta}$$

Now let us define $\widetilde{rand}^{-1}(c)$ as the set of clients in $rand^{-1}(c)$ plus any client c' in some set $C_{S_2}(f') \setminus C'(f')$ such that $next(c') \in rand^{-1}(c)$. By the above discussion and the independence of the sampling processes defining $next()$ and $rand()$, one has that

$$Ex[|\widetilde{rand}^{-1}(c)|] \leq 2Ex[|rand^{-1}(c)|] \leq 2\frac{1-\delta}{\delta}$$

We can conclude that

$$Ex[\sum_{f' \in S_2^L} \sum_{c \in C_{S_2}(f')} (\ell''(P(c))] \leq \sum_c Ex[|\widetilde{rand}^{-1}(c)|](d_2(c) + opt(c)) \leq 2\frac{1-\delta}{\delta}(d_2 + opt).$$

We can refine the above analysis for the clients $C^{LM} \cup C^{MM}$ as follows. Suppose that $f^* \in OPT^M$, which implies that $c \in C^{LM} \cup C^{MM}$. In this case at least a $1-\delta$ fraction of the clients $C_{OPT}(f^*)$ are served by matched facilities $M(f^*) \subseteq S_2^M$. Those clients cannot belong to $rand^{-1}(c)$. We can adapt the above analysis by letting the sets C^i not include clients in C^{MM} (while possibly $c \in C^{MM}$). This gives

$$Ex[|rand^{-1}(c)|] \leq \sum_{i \geq 2} \frac{|C^i|}{|C_{OPT}(f^*)| - |C^i|} \leq \frac{1}{(1-\delta)|C_{OPT}(f^*)|} \sum_{i \geq 2} |C^i| \leq \frac{\delta}{1-\delta},$$

hence

$$Ex[|\widetilde{rand}^{-1}(c)|] \leq 2Ex[|rand^{-1}(c)|] \leq 2\frac{\delta}{1-\delta}.$$

This gives

$$Ex[\sum_{f' \in S_2^L} \sum_{c \in C_{S_2}(f')} (\ell''(P(c))] \leq \sum_c Ex[|\widetilde{rand}^{-1}(c)|](d_2(c) + opt(c))$$

$$\leq 2\frac{1-\delta}{\delta}(d_2^{LL} + opt^{LL} + d_2^{ML} + opt^{ML}) + 2\frac{\delta}{1-\delta}(d_2^{LM} + opt^{LM} + d_2^{MM} + opt^{MM}).$$

Altogether

$$Ex[\sum_{f' \in S_2^L} \sum_{c \in C_{S_2}(f')} (\ell(P(c)) - d_2(c)]$$

$$\leq 2(opt^{LM} + opt^{LL}) + 2\frac{1-\delta}{\delta}(d_2^{LL} + opt^{LL} + d_2^{ML} + opt^{ML}) + 2\frac{\delta}{1-\delta}(d_2^{LM} + opt^{LM} + d_2^{MM} + opt^{MM})$$

$$= \frac{2}{\delta}opt^{LL} + 2\frac{1-\delta}{\delta}(d_2^{LL} + d_2^{ML} + opt^{ML}) + \frac{2}{1-\delta}opt^{LM} + 2\frac{\delta}{1-\delta}(d_2^{MM} + opt^{MM})$$

$$\leq \frac{2}{\delta}(opt - opt^{MM}) + 2\frac{1-\delta}{\delta}(d_2 - d_2^{MM}) + 2\frac{\delta}{1-\delta}(d_2^{MM} + opt^{MM}).$$

$\square$

E An Improved LMP Approximation for Facility Location: Uniform Opening Costs

In this section we describe how to derive a better than 2 LMP approximation for UFL in the case of uniform facility costs (without giving an explicit bound on the approximation factor). Suppose we are given a bipoint solution $S_B = aS_1 + (1-a)S_2$ for k-Median as described in Section 4, for some given value of λ. Recall that $d_i = d(S_i)$, $k_i = |S_i|$, $k_1 \leq k < k_2$ and $k = ak_1 + (1-a)k_2$. Lemma 4.1 shows that $\lambda k_2 + d_2 \leq \lambda k + (2 - \eta_2)opt$ for some fixed constant $\eta_2 > 0$. In Section E.1 we will prove the following (weaker) version of Lemma 4.1 which applies to S_1.

LEMMA E.1. *For some function $\eta_1(a)$ which is strictly positive for $a > 0$, one has $\lambda k_1 + d_1 \leq \lambda k + (2 - \eta_1(a))opt$.*

As a corollary we immediately obtain that S_B is strictly better than 2 approximate.

COROLLARY E.1. *For some absolute constant $\eta > 0$, one has $ad_1 + bd_2 \leq (2 - \eta)opt$.*

Proof. From Lemmas 4.1 and E.1, a valid choice for η is

$$\eta := \min_{a \in [0,1]} \{a\eta_1(a) + (1-a)\eta_2\}.$$

It is clear that $\eta > 0$. Indeed, for $a \leq 1/2$ this is at least $\eta_2/2 > 0$ while for $a \geq 1/2$ it is at least $\frac{1}{2}\min_{a \in [1/2,1]}\{\eta_1(a)\} > 0$. $\square$

We can use Corollary E.1 to achieve an LMP $(2 - \eta)$-approximation for UFL with uniform facility costs λ as follows. We consider any possible number k of facilities in a feasible solution. For each such k, we use the construction from Section 4 to obtain a bipoint solution $a(k)S_1(k) + (1 - a(k))S_2(k)$ for k-Median, for some value $\lambda(k)$ of the Lagrangian multiplier. We return the cheapest solution $S_i(k)$.

THEOREM E.1. *There is a deterministic LMP $(2 - \eta)$-approximation for facility location with uniform facility costs, where $\eta > 0$ is the constant from Corollary E.1.*

Proof. Consider the above algorithm, and let OPT be any feasible solution with total cost $\lambda k + opt$, where $opt = d(OPT)$ and $k = |OPT|$. Consider the random solution S' which is equal to $S_1(k)$ with probability $a(k)$ and $S_2(k)$ otherwise. Clearly the cost $\lambda|S'| + d(S')$ of S' (notice that here we consider the original facility cost λ rather than $\lambda(k)$) is not worse than the cost of the solution returned by our algorithm. The claim follows since the expected cost of S' is

$$a(k)(\lambda|S_1(k)| + d(S_1(k)) + (1 - a(k))(\lambda|S_2(k)| + d(S_2(k)) \leq \lambda k + (2 - \eta)opt.$$

$\square$

E.1 Improved Upper Bounds on S_1 We next prove Lemma E.1. To that aim, we next assume $a > 0$.

Recall that, for a parameter $\delta \in [0, 1/2]$ to be fixed later, we defined a many-to-one matching between S_1 and OPT, where a facility $f^* \in OPT$ is matched with a set $M(f^*)$ of facilities in S_1 which are $1/2$-captured by f^* and which together $(1 - \delta)$-capture f^*. We next use S_1^M instead of S^M and similarly for related quantities to stress that we are focusing on the locally optimal solution $S' = S_1$. With the usual notation, we let $\alpha^L := opt^L/opt$, $\alpha^M := opt^M/opt$, $\alpha^{MM} = opt^{MM}/opt$, $\beta_1 = d_1/opt$, and $\beta_1^{MM} = d_2^{MM}/opt$. Let also $k^L = |OPT^L|$, $k^M = |OPT^M|$, $k_1^L = |S_1^L|$, and $k_1^M = |S_1^M|$. Finally, we let d_1^L be the connection cost in S_1 related to the clients served by lonely facilities, and set $\beta_1^L = d_1^L/opt$.

We next apply Theorem 3.1 to S_1, neglecting the term depending on ε for the usual reasons. Hence we get

(E.31) $\qquad \lambda k_1 + d_1 \leq \lambda k + (1 + 2\alpha^L + \frac{\delta}{1-\delta}(\beta_1^{MM} + \alpha^{MM})) \cdot opt =: \lambda k + \rho_1^A(\delta, \alpha^L, \alpha^{MM}, \beta_1^{MM})opt.$

Observe that the above bound already implies an improved upper bound on the cost of S_1 (w.r.t. $\lambda k + 2opt$) when α^L and δ are sufficiently small and $a > 0$. Indeed, from $ad_1 + (1 - a)d_2 \leq 2opt$ and recalling that $\beta = d_2/opt$, we derive:

(E.32) $\qquad \beta_1^{MM} \leq \beta_1 = \frac{d_1}{opt} \leq \frac{2opt - (1-a)d_2}{a \cdot opt} = \frac{2 - (1-a)\beta_2}{a} \leq \frac{2}{a}.$

We next derive an alternative bound which implies the claim in the complementary case. Similarly to Section 4.1, we show that the facility cost $\lambda|OPT^L|$ due to lonely facilities in OPT is upper bounded by $O(opt^L)$. This will allow us to exploit Corollary 2.1.

Let us partition OPT^L in two sets D_A, D_B as follows. We define a directed graph G^L over the node set OPT^L. Let $cl(f)$ be the closest facility to f in OPT. For each $f^* \in OPT^L$, we add a directed edge $(f^*, cl(f^*))$ iff $cl(f^*) \in OPT^L$. In particular each node f^* in G^L has outdegree 1 or 0, the latter case happening if $cl(f^*) \in OPT^M$.

LEMMA E.2. *By breaking ties properly, one can guarantee that cycles of G^L have length at most 2 (in particular, G^L is bipartite).*

Proof. Suppose that we sort the nodes arbitrarily, and in case of ties we direct the edge leaving v to the nearest node u that appears first in this sorted list. Assume by contradiction that there exists a directed cycle $C = (v_1, v_2, \ldots, v_k, v_1)$ with $k \geq 3$. Assume w.l.o.g. that v_1 appears last in the sorted list among nodes of C. Let $d_{i,j}$ denote the distance between v_i and v_j. By the minimality of the distances corresponding to the arcs of G^L and by the symmetry of the distances between nodes in the original graph we have:

$$d_{1,2} \leq d_{1,k} \leq d_{k-1,k} \leq d_{k-2,k-1} \leq \ldots \leq d_{2,3} \leq d_{1,2}.$$

This implies that all the distances corresponding to the arcs in C are identical. This is however a contradiction since node v_k should have its outgoing arc directed towards v_{k-1} rather than towards v_1 since v_{k-1} appears earlier than v_1 in the sorted list and $d_{1,k} = d_{k-1,k}$. ☐

The sets D_A and D_B are defined by computing a bipartition of G^L. Let $k_A = |D_A|$ and $k_B = |D_B|$. Notice that $k^L := k_A + k_B$.

REMARK 1. k_A *and* k_B *are potentially larger than any constant, hence removing* D_A *or* D_B *from* S_1 *might not be an option considered by the local search algorithm. Nonetheless, we can to upper bound the increase of the connection cost of* S_1 *due to the removal of such sets.*

The proof of the following lemma is analogous to the proof of Lemma 4.2, with some subtle technical differences due to the fact that, unlike S_2, OPT is not a locally optimal facility location solution (hence leading to a slightly weaker bound).

LEMMA E.3. *Let* $OPT_X = OPT \setminus D_X$ *and* $d_X = d(OPT_X)$ *for* $X \in \{A, B\}$*. Then*

$$d_A + d_B \leq 2opt + \frac{1}{\delta}(opt + d_1 + opt^L + d_1^L).$$

Proof. Consider a given D_X, and any client c served by $f^* \in D_X$ in OPT. We upper bound the connection cost of c in $OPT_X = OPT \setminus D_X$ by upper bounding the distance between c and $cl(f^*)$. Notice that $cl(f^*) \in OPT_X$ by construction. In order to do that we describe a random path $P(c)$ between c and some facility $f' \in OPT$ distinct from f^* (possibly $f' \in D_X$). This path $P(c)$ always includes f^* (this is a technical difference w.r.t. Lemma 4.2). Since

$$dist(c, cl(f^*)) \leq dist(c, f^*) + dist(f^*, cl(f^*)) \leq dist(c, f^*) + dist(f^*, f'),$$

the expected length $\ell(P(c))$ of $P(c)$ is a valid upper bound on the connection cost of c in OPT_X.

The path $P(c)$ is defined as follows. Let $C'(f^*)$ be the clients served by f^* in OPT which in S_1 are served by a facility which is not 1/2-captured by f^*. Notice that $|C'(f^*)| \geq \delta|C_{OPT}(f^*)|$ since otherwise f^* would be matched. Let $next(f^*)$ be a random client in $C'(f^*)$. Let $f_1 = S_1(next(f^*))$ be the facility serving $next(f^*)$ in S_1. Let $rand(f^*)$ be a random client served by f_1 in S_1 but not by f^* in OPT. Finally let $f' \neq f^*$ be the facility serving $rand(f^*)$ in OPT. Then $P(c)$ is the concatenation of the shortest paths from c to f^*, from f^* to $next(f^*)$, from $next(f^*)$ to f_1, from f_1 to $rand(f^*)$, and from $rand(f^*)$ to f'.

By the above discussion, the increase of the connection cost of OPT_A and OPT_B together w.r.t $2opt$ is upper bounded by

$$\sum_{f^* \in OPT^L} \sum_{c \in C_{OPT}(f^*)} \ell(P(c)) - opt(c)$$
$$= \sum_{f^* \in OPT^L} |C_{OPT}(f^*)| \left(opt(next(f^*)) + d_1(next(f^*)) + d_1(rand(f^*)) + opt(rand(f^*))\right).$$

Let us bound the terms in the above sum involving $next(f^*)$ and $rand(f^*)$ separately. For the part involving $next(f^*)$, we observe that each $c' \in C'(f^*)$ is selected as $next(f^*)$ with probability $\frac{1}{|C'(f^*)|} \leq \frac{1}{\delta|C_{OPT}(f^*)|}$, and it

that case contributes to the sum with $|C_{OPT}(f^*)|(opt(c') + d_1(c'))$. Thus

$$Ex[\sum_{f^* \in OPT^L} |C_{OPT}(f^*)|opt(next(f^*)) + d_1(next(f^*))] \leq \sum_{f^* \in OPT^L} \sum_{c' \in C'(f^*)} \frac{1}{\delta}(opt(c') + d_1(c'))$$

$$\leq \frac{1}{\delta} \sum_{f^* \in OPT^L} \sum_{c \in C(f^*)} opt(c) + d_1(c) \leq \frac{1}{\delta}(opt^L + d_1^L).$$

Consider next the terms depending on $rand(f^*)$. For a client c, define $rand^{-1}(c)$ as the facilities f^* with $rand(f^*) = c$. Then the expected contribution of c to the considered terms in the sum is

$$(opt(c) + d_1(c)) \cdot Ex[\sum_{f^* \in rand^{-1}(c)} |C_{OPT}(f^*)|]$$

Let $f_1 = S_1(c)$ be the facility serving c in S_1. Let us partition $C_{S_1}(f_1)$ in subsets $C^1, \ldots, C^h$ depending on the facilities $f^1, \ldots, f^h$ serving them in OPT. W.l.o.g., assume $c \in C_1$. Notice that for each f^i, $i \geq 2$, f^i selects c as $rand(f^i)$ with probability $\frac{1}{|C_{S_1}(f_1)| - |C^i|}$. Then

$$Ex[\sum_{f^* \in rand^{-1}(c)} |C_{OPT}(f^*)|] = \sum_{i \geq 2} \frac{|C(f^i)|}{|C^i|} \frac{|C^i|}{|C_{S_1}(f_1)| - |C^i|} \leq \frac{1}{\delta} \sum_{i \geq 2} \frac{|C^i|}{|C_{S_1}(f_1)| - |C^i|}.$$

As in the proof of Lemma 4.2, the righthand side of the above inequality is maximized for $h = 2$, in particular being $|C^2|/|C^1|$. Since f^2 does not $1/2$-capture f_1 by construction, we have that $|C^2| \leq \frac{1}{2}|C_{S_1}(f_1)|$ and hence $|C^2|/|C^1| \leq 1$. Altogether

$$Ex[\sum_{f^* \in rand^{-1}(c)} |C_{OPT}(f^*)|] \leq \frac{1}{\delta}.$$

Thus the terms involving $rand(f^*)$ altogether contribute in expectation with at most

$$\sum_{c \in C} \frac{1}{\delta}(opt(c) + d_1(c)) = \frac{1}{\delta}(opt + d_1).$$

The claim follows. □

As a corollary, we achieve the desired alternative bound on the total cost of S_1.

COROLLARY E.2. *One has*

$$\lambda k_1 + d_1 \leq \lambda(k - \frac{k^L}{2}) + (2 + \frac{1 + \beta_1 + \beta_1^L + \alpha^L}{\delta}) \cdot opt.$$

Proof. Since S_1 is LMP 2-approximate, from the existence of the solutions OPT_A and OPT_B one gets $\lambda k_1 + d_1 \leq \lambda(k - k_A) + 2d_A$ and $\lambda k_1 + d_1 \leq \lambda(k - k_B) + 2d_B$. The claim follows by averaging and Lemma E.3. □

Let $\eta \in [0, 1]$ be a parameter to be fixed later. Next we distinguish two cases. If $\frac{\lambda k^L}{2opt} \geq \frac{1+\beta_1+\beta_1^L+\alpha^L}{\delta} + \eta$, then

(E.33)
$$\lambda k_1 + d_1 \overset{Cor. (E.2)}{\leq} \lambda k + (2 - \eta)opt.$$

Otherwise, we have

$$\lambda k^L \leq 2(\frac{1 + \beta_1 + \beta_1^L + \alpha^L}{\delta} + \eta)opt = 2(\frac{1 + \beta_1 + \beta_1^L}{\alpha^L \delta} + \frac{1}{\delta} + \eta)opt^L =: T_1 opt^L.$$

In this case we can apply Corollary 2.1 with $OPT' = OPT^L$ and $T = T_1$ to infer

(E.34)
$$\lambda k_1 + d_1 \leq \lambda k + (2 \cdot (1 - \alpha^L) + opt_{JMS}^+(q, T_1) \cdot \alpha^L)opt =: \lambda k + \rho_1^B(\delta, \alpha^L, \beta_1, \beta_1^L, \eta).$$

From the above discussion we get

$$\lambda k_1 + d_1 \le \lambda k + (2 - \eta_1(a))opt$$

where

$$2 - \eta_1(a) \ge \min\{\rho_1^A(\delta, \alpha^L, \alpha^{MM}, \beta_1^{MM})), \max\{2 - \eta, \rho_1^B(\delta, \alpha^L, \beta_1, \beta_1^L, \eta)\}\}.$$

We can simplify the above formula by making some pessimistic choices. From (E.32) and assuming pessimistically that $\beta_2 = 0$, we can set $\beta_1 = 2/a$. We can pessimistically upper bound α^{MM} with $\alpha^M = 1 - \alpha^L$ and β_1^{MM} with $\beta_1^{MM} = \beta_1 - \beta_1^L = \frac{2}{a} - \beta_1^L$. Then we can choose freely the value of δ to minimize the lower bound on $2 - \eta_1(a)$. For a given δ, we can pessimistically choose the values of $\alpha^L \in [0, 1]$ and $\beta_1^L \in [0, \frac{2}{a}]$ that maximize the minimum. Given that, we can freely choose η to minimize the maximum. Altogether we obtain

$$2 - \eta_1(a) = \min_{\delta \in [0, 1/2]} \max_{\alpha^L \in [0,1], \beta_1^L \in [0, \frac{2}{a}]} \min_{\eta \in [0,1]} \min\{\rho_1^A(\delta, \alpha^L, 1 - \alpha^L, \frac{2}{a} - \beta_1^L), \max\{2 - \eta, \rho_1^B(\delta, \alpha^L, \frac{2}{a}, \beta_1^L, \eta)\}\}$$

It is not hard to see that for $a > 0$ one has $\eta_1(a) > 0$. Indeed, for positive and small enough δ and α^L, one has $\rho_1^A(\delta, \alpha^L, 1 - \alpha^L, \frac{2}{a} - \beta_1^L) < 2$. Otherwise, we can assume that δ and α^L are both bounded away from zero. For any positive η trivially $2 - \eta < 2$. Finally, for η small enough one has that T_1 is upper bounded by a constant. It turns out that $opt_{JMS}^+(q, T_1)$ is strictly smaller than 2 for any constant T_1 and large enough q. This implies $\rho_1^B(\delta, \alpha^L, \frac{2}{a}, \beta_1^L, \eta) < 2$. More formally, one can replace $opt_{JMS}^+(q, T_1)$ in the above analysis with the analytical upper bound $opt'_{JMS}(T_1)$ on $opt_{JMS}(T_1)$ obtained in Section B. It is not hard to show analytically that $opt'_{JMS}(T_1) < 2$ for any constant T_1, hence the claim.

F A Refined Approximation Factor for k-Median

In this section we sketch how to refine the approximation factor for k-Median from Section 4. Consider a bipoint solution $S_B = aS_1 + (1 - a)S_2$ computed as in Section 4. With the usual notation, $d_i = d(S_i)$ and $\beta_i = d_i/opt$. Recall that Li and Svensson [LS16] present a $2(1+2a)+\varepsilon$ approximation for any constant $\varepsilon > 0$ (with extra running time $n^{\text{poly}(1/\varepsilon)}$), hence we can also assume that a is bounded away from 0 since we aim at an approximation factor larger than 2. Notice also that $a\beta_1 + (1 - a)\beta_2 \le 2$, hence $\beta_1 \le \frac{2}{a}$. This means that S_1, which is a feasible solution, provides a $2/a$ approximation. Hence we can also assume that a is bounded away from 1 for similar reasons. Recall that, for a bounded away from 0 and 1, Li and Svensson [LS16] present a $\frac{2(1+2a)}{(1+2a^2)} + \varepsilon$ approximation that we will next use.

We will consider the best of 3 approximate solutions:

1. The solution S_1 alone, which provides a β_1 approximation.

2. The solution S_{LS} computed with the mentioned $\frac{2(1+2a)}{(1+2a^2)} + \varepsilon$ approximation in [LS16].

3. The solution obtained by rounding our bipoint solution (S_1, S_2) with the $\rho_{BR} < 1.3371$ rounding algorithm in [BPR+17].

Let us focus on the latter solution. The discussion from Section E directly implies that this solution has approximation ratio at most

$$\rho_{BR}(a\beta_1 + (1 - a)\beta_2) \le \rho_{BR}(a(2 - \eta_1(a)) + (1 - a)(2 - \eta_2)) = \rho_{BR}(2 - a\eta_1(a) - (1 - a)\eta_2).$$

This already gives an improvement on the approximation factor from Section 4 due to the extra term $-a\eta_1(a)$ which is strictly positive for $a > 0$. We can further refine this factor as follows. In the computation of $\eta_1(a)$ we assumed pessimistically that $\beta_1 = 2/a$, and in the computation of η_2 that $\beta_2 = 2$. However these two conditions cannot simultaneously hold unless $a = 1$. In the latter case however S_1 provides a 2 approximation already. Under the assumption $a < 1$ and fixing a value $\beta_1 \in [2, 2/a]$, we can rather use $\frac{2 - a\beta_1}{1 - a}$ as a pessimistic upper bound on β_2. Hence with essentially the same analysis as before we get two new functions $\eta_1(a, \beta_1) \ge \eta_1(a)$ and $\eta_2(a, \beta_1) \ge \eta_2$ for a fixed β_1, leading to the refined approximation factor

$$\rho_{BR}(2 - a\eta_1(a, \beta_1) - (1 - a)\eta_2(a, \beta_1)).$$

Taking the best of the 3 mentioned approximate solutions, and choosing the worst possible parameter $a \in [0,1]$ and $\beta_1 \in [2, 2/a]$, we obtain the approximation factor:

$$\rho_{kMed} = \max_{a \in [0,1], \beta_1 \in [2, \frac{2}{a}]} \min\{\beta_1, \frac{2(1+2a)}{1+2a^2} + \varepsilon, \rho_{BR}\left(2 - a\eta_1(a, \beta_1) - (1-a)\eta_2(a, \beta_1)\right)\}$$

G   An Improved LMP Approximation for Facility Location: General Opening Costs

Consider a general facility location instance where each facility $f \in F$ has a cost $\mathsf{open}(f)$. In this section, we prove the following theorem for Facility Location with general facility cost.

THEOREM G.1. *There exists an absolute constant $\eta > 0$ such that there is a polynomial time LMP $(2 - \eta)$ approximation algorithm for Facility Location with general facility costs.*

The current (unoptimized) lower bound for η is $2.25 \cdot 10^{-7}$.

G.1   New Local Search and Facility Deletion Given an instance of Facility Location, our algorithm runs the JMS algorithm to get the initial solution S' and improves it via local search until no local improvement is possible. In this subsection, we prove that when the final solution S' has a relatively small connection cost, we already improve on an LMP 2 approximation. The proof is similar to the one that we presented for uniform facility costs, but has additional ideas that may be of independent interest. In fact, the following example shows that the standard local search where one swaps a constant number of facilities will not give any LMP approximation for a large class of objective functions.

CLAIM 3. *For any constants $\Delta \in \mathbb{N}$ and $\alpha, \beta \in \mathbb{R}^+$, there is an instance of general Facility Location $(C, F, dist, o)$ that admits feasible solutions $OPT, S \subseteq F$ such that (1) S is locally optimal for the objective function $\alpha \cdot \mathsf{open}(S) + \beta \cdot d(S)$ and the width-Δ local search, but (2) $\mathsf{open}(OPT) + t \cdot d(OPT) < \mathsf{open}(S) + d(S)$ for any $t \in \mathbb{R}$.*

Proof. For some n to be determined, consider the instance of UFL where $F = \{f_0, \ldots, f_n\}$, $C = \{c_1, \ldots, c_n\}$ where $dist(f_i, c_i) = 0$ and $dist(f_0, c_i) = 1$ for every $i \in [n]$. We set $\mathsf{open}(f_0) = x$ and $\mathsf{open}(f_i) = y$ for $i \in [n]$ for some $x, y > 0$ such that $y > \beta/\alpha$, $\alpha(x - ry) < \beta(n - 2r)$ for every $r \in [\Delta]$, and $ny < x + n$. Such x and y exist for large enough n depending on Δ, α, β (e.g., take $y \approx \beta/\alpha$ for the first inequality, $x \approx n \cdot \min(\beta/\alpha - 1, 0)$ for the third, and n large enough for the second). This finishes the description of the instance. Let $OPT = \{f_1, \ldots, f_n\}$ and $S = \{f_0\}$. Note that OPT has $\mathsf{open}(OPT) = n \cdot y$ and $d(OPT) = 0$, and $\mathsf{open}(S) = x$ and $d(S) = n$.

From S, the width-Δ local search algorithm may close f_0 or not. Suppose that it considers a new solution S' where f_0 is closed and $r \leq \Delta$ facilities from OPT are open. Then $\mathsf{open}(S') = r \cdot y$ and $d(S') = 2(n - r)$. Since $\alpha x + \beta n < \alpha r \cdot y + \beta \cdot 2(n - r) \Leftrightarrow \alpha(x - ry) < \beta(n - 2r)$, S is better than S' in terms of the objective function $\alpha \cdot \mathsf{open}(S) + \beta \cdot d(S)$. For the other case, when the local search considers a new solution S' where f_0 is kept open and $r \leq \Delta$ facilities from OPT are open, $\mathsf{open}(S') = x + r \cdot y$ and $d(S') = (n - r)$. Since $\alpha x + \beta n < \alpha(x + r \cdot y) + \beta \cdot (n - r) \Leftrightarrow \beta/\alpha < y$, S is better than S' in terms of the objective function $\alpha \cdot \mathsf{open}(S) + \beta \cdot d(S)$. Therefore, one can conclude that S is locally optimal with respect to Δ-local moves and the objective function $\alpha \cdot \mathsf{open}(S) + \beta \cdot d(S)$, but since $ny < x + n$, $\mathsf{open}(OPT) + t \cdot d(OPT) < \mathsf{open}(S) + d(S)$ for any finite t.   □

Therefore, our local search considers a new kind of local moves where we remove one facility from S' and augment S' by running the JMS algorithm on top of it. More concretely, our local search algorithm `LocalSearch-JMS` goes as follows.

1. Start with an LMP 2 approximate solution S' (e.g., as returned by JMS).

2. While there exists a solution S'' such that either
   (1) $|S'' - S'| + |S' - S''| \leq 1/\varepsilon + 1$ or
   (2) $S'' = \mathtt{Extend\text{-}JMS}(S' - \{f\} \cup \{f'\})$ for any $f \in S'$, $f' \notin S'$,
   and such that $\mathsf{cost}(S'') < \mathsf{cost}(S')$.

(a) Do $S' \leftarrow S''$.

3. Return S'.

The procedure `Extend-JMS`(X) is simply

1. Modify the instance so that all facilities in X have opening cost 0.

2. Run the standard JMS on the new instance.

Let S' be the solution at the end of `LocalSearch-JMS` (i.e., S' is a local optimum). Let OPT denote some given facility location solution. Let d' and $d^* = opt$ be the total connection cost of S' and OPT respectively. The main result of this subsection is the following lemma.

LEMMA G.1. *There exists an absolute constant $\eta \geq 4.5 \cdot 10^{-7}$ such that if S' is a local optimum of* `LocalSearch-JMS` *and $d' \leq 4opt$, $\boldsymbol{open}(S') + d' \leq \boldsymbol{open}(OPT) + (2 - \eta)d^*$.*

Matching and Local Search. We classify the facilities in S' and OPT as lonely and matched as follows. Let $0 < \delta_1' < \delta_1 \leq 1/2$, $0 < \delta_2' < \delta_2 \leq 1/2$ be parameters to be determined later. Say $f' \in S'$ and $f^* \in OPT$ are *matched* if f' $(1 - \delta_1)$-captures f^* and f^* $(1 - \delta_2)$-captures f'. All the facilities which are not matched are *lonely*. As usual, let OPT^M, OPT^L, S^M, and S^L denote the set of matched facilities in OPT, lonely facilities in OPT, matched facilities in S', and lonely facilities in S' respectively. Let opt^M (resp. opt^L) be the total connection cost of the clients served by a matched facility in OPT^M (resp. OPT^L). Also for a client $c \in C$, let $d'(c)$ and $d^*(c)$ of the connection cost of c in S' and OPT respectively.

The following theorem proved in Section G.1.1 shows that `LocalSearch-JMS` already gives the desired guarantee when opt^L is small and d' is not too large compared to opt.

THEOREM G.2. *Fix $\delta_1 = \delta$ and $\delta_2 = 1/2$ in the above definition of matching. Let S' be the solution at the end of* `LocalSearch-JMS`. *Then one has $\boldsymbol{open}(S') + d' \leq \boldsymbol{open}(OPT) + \frac{\delta}{1-\delta}d' + \frac{1}{1-\delta}opt^M + 4opt^L + O(\varepsilon(d' + opt))$.*

Bounding $\boldsymbol{open}(S^L)$. Using the local optimality of S', one can prove the following lemma.

LEMMA G.2. *There exists a constant $t = t(\delta_1, \delta_2')$ such that*

(G.35)
$$\boldsymbol{open}(S^L) \leq \frac{1 - \delta_2}{1 - \delta_2'} \cdot \boldsymbol{open}(OPT^L) + t \cdot (opt + d').$$

Proof. Consider a $f' \in S^L$. We will consider a (randomized) local move that closes f' and opens at most one facility in OPT. Let $M(f') \subseteq OPT$ be the facilities that are $(1 - \delta_1)$-captured by f'. We consider the following cases. For a client c, let $f^*(c)$ (resp. $f'(c)$) be the facility serving c in OPT (resp. S').

1. $M(f')$ does not $(1 - \delta_2')$ capture f': Here we close f' without opening any facility, leading to a decrease of the opening cost by $\boldsymbol{open}(f')$. Consider the following randomized rerouting of the clients in $C(f')$ to $S' \setminus \{f'\}$.

    - If $c \in C(f') \setminus C(M(f'))$: Choose a random client $rand(c)$ from $C(f^*(c)) \setminus C(f')$. Reroute c to $f'(rand(c))$. The expected increase in the connection cost for c is at most

    (G.36)
    $$d^*(c) + \mathbb{E}_{rand(c)}[d^*(rand(c)) + d'(rand(c))] - d'(c).$$

    - If $c \in C(f') \cap C(M(f'))$: Choose a random client $next(c)$ from $C(f') \setminus C(M(f'))$. As in the previous case, choose a random $rand(next(c))$ from $C(f^*(next(c))) \setminus C(f')$ and reroute c to $f'(rand(next(c)))$. The increase in the connection cost for c is at most

    (G.37) $$\mathbb{E}_{next(c)}[d'(next(c)) + d^*(next(c))] + \mathbb{E}_{rand(next(c))}[d^*(rand(next(c))) + d'(rand(next(c)))].$$

2. If $M(f')$ does $(1-\delta_2')$ captures f': Here we open one facility $f^* \in M(f')$ with probability $|C(f^*) \cap C(f')|/|C(M(f')) \cap C(f')|$ (note that the sum of the probabilities is 1). Each $c \in C(f')$ is rerouted to the open facility f^*. Note that for each $c \in C(f')$, this is equivalent to choosing a random client $rand'(c)$ uniformly from $C(M(f')) \cap C(f')$ and rerouting c to $f^*(rand'(c))$. Therefore, in expectation, the connection cost for c is increased by

$$\text{(G.38)} \qquad \mathbb{E}_{rand'(c) \in C(M(f')) \cap C(f')}[d'(rand'(c)) + d^*(rand'(c))].$$

Since S' is local optimal, all the (randomized) local moves considered above do not improve the cost of S'. Consider the sum of all the (expected) increased costs, considering both opening and connection cost. This sum has the following terms.

1. $-\mathbf{open}(f')$ for every $f' \in S^L$.

2. At most $\frac{1-\delta_2}{1-\delta_2'}\mathbf{open}(f^*)$ for every $f^* \in OPT^L$. Note that $M(f') \subseteq OPT^L$ for every $f' \in S^L$, and $M(f') \cap M(f'') = \emptyset$ for $f' \neq f'' \in S^L$. When $M(f')$ $(1-\delta_2')$-captures f' and $f^* \in M(f')$ does not $(1-\delta_2)$-capture f^*, f^* is open with probability at most $\frac{1-\delta_2}{1-\delta_2'}$. (If f^* $(1-\delta_2)$-captured f', then they should be matched.)

3. At most $d^*(c)$ for every $c \in C$. (From the sum of $d^*(c)$ in (G.36)).

4. For every c' such that $f'(c') \in S^L$ and $f^*(c') \notin M(f'(c'))$:

   (a) In the sum of $\mathbb{E}_{next(c)}[d'(next(c)) + d^*(next(c))]$ in (G.37), $(d^*(c') + d'(c'))$ will appear with coefficient at most $\frac{1-\delta_2'}{\delta_2'}$; c' can be sampled by $c \in C(f') \cap c(M(f'))$, and each such c samples $next(c)$ from $|C(f') \setminus c(M(f'))|$ clients. In (G.37), we have $|C(f') \setminus c(M(f'))| \geq \frac{\delta_2'}{1-\delta_2'}|C(f') \cap C(M(f'))|$.

5. For every $c' \in C$:

   (a) In the sum of $\mathbb{E}_{rand(c)}[d^*(rand(c)) + d'(rand(c))]$ in (G.36), $(d^*(c') + d'(c'))$ will appear with coefficient at most $\frac{1}{\delta_1}$, because c' can be possibly sampled by $c \in C(f^*(c)) \cap C(f'')$ where f'' does not capture $f^*(c)$ but each such c samples c' from at least $|C(f^*(c)) \setminus C(f'')| \geq \delta_1|C(f^*(c))|$ clients.

   (b) In the sum of $\mathbb{E}_{rand(next(c))}[d^*(rand(next(c))) + d'(rand(next(c)))]$ in (G.37), $(d^*(c') + d'(c'))$ will appear with coefficient at most $\frac{1-\delta_2'}{\delta_1\delta_2'}$; each $c'' \in C$ will be $next(c)$ at most $\frac{1-\delta_2'}{\delta_2'}$ times in expectation by the same argument as in item 4.(a), and using the same argument as in item 5.(a), the expected number of times c' is chosen as $rand(next(c))$ is at most $\frac{1}{\delta_1}$ times the maximum expected number of times any $c'' \in C$ is chosen as $next(c)$.

6. For every c' such that $f' \in S^L$ and $f^*(c') \in M(f'(c'))$:

   (a) In the sum of $\mathbb{E}_{rand'(c) \in C(M(f')) \cap C(f')}[d'(rand'(c)) + d^*(rand'(c))]$ in (G.38), $(d^*(c) + d'(c))$ will appear with coefficient at most $\frac{1}{1-\delta_2'}$; c' can be possibly sampled by $c \in C(f')$ and each such c samples from at least $|C(M(f')) \cap C(f')| \geq (1-\delta_2')|C(f')|$ clients.

Therefore, if we add 3, 4, 5, 6 above over each client, we have

$$\text{(G.39)} \qquad \sum_{f' \in S^L} \mathbf{open}(f') \leq \frac{1-\delta_2}{1-\delta_2'} \cdot \sum_{f^* \in OPT^L} \mathbf{open}(f^*) + t(d^* + d'),$$

where $t = 1 + \frac{1}{\delta_1\delta_2'} + \max(\frac{1-\delta_2'}{\delta_2'}, \frac{1}{1-\delta_2'})$ (note that 4 and 6 are mutually exclusive so we have the maximum instead of the sum). $\square$

Bounding open(OPT^L). Similarly to Section E.1, partition OPT^L into D_A and D_B such that for every $f \in D_A$ (resp. $f \in D_B$), one of the facilities closest to f in OPT is in $OPT \setminus D_A$ (resp. $OPT \setminus D_B$).

LEMMA G.3. *Let $t = t(\delta_1, \delta_2')$ be the constant determined in Lemma G.2. There exists a constant $t' = t'(\delta_2, \delta_1')$ and a randomized solution $OPT^\dagger$ such that*

- *The expected facility cost of $OPT^\dagger$ is at most $\mathtt{open}(OPT) - \left(\frac{1}{2} - \frac{(1-\delta_1)(1-\delta_2)}{2(1-\delta_1')(1-\delta_2')}\right)\mathtt{open}(OPT^L) + t \cdot (opt + d')$.*

- *The expected connection cost of $OPT^\dagger$ is at most $t' \cdot (opt + d')$.*

Proof. For each $X \in \{A, B\}$, let OPT_X be the randomized solution achieved by applying the proof of the above Lemma G.2 to delete each of D_X and possibly reopen some facilities in S^L. $OPT^\dagger$ samples $X \in \{A, B\}$ with probability $1/2$ each and samples OPT_X.

In $OPT^\dagger$, each $f^* \in OPT^L$ is deleted exactly with probability $1/2$, and each $f' \in S^L$ is reopen with probability at most $\frac{1-\delta_1}{2(1-\delta_1')}$. (Note that if $f' \in S^L$ reopens only when some $f^* \in OPT^L$ capturing f' is deleted, so it can possibly reopen in at most one of OPT_A and OPT_B.) Therefore, the expected facility cost is at most

$$\mathtt{open}(OPT) - \frac{\mathtt{open}(OPT^L)}{2} + \frac{1-\delta_1}{2(1-\delta_1')}\mathtt{open}(S^L)$$

$$\leq \mathtt{open}(OPT) - \frac{\mathtt{open}(OPT^L)}{2} + \frac{1-\delta_1}{2(1-\delta_1')}\left(\frac{1-\delta_2}{1-\delta_2'} \cdot \mathtt{open}(OPT^L) + t(opt + d')\right) \quad \text{(Lemma G.2)}$$

$$\leq \mathtt{open}(OPT) - \left(\frac{1}{2} - \frac{(1-\delta_1)(1-\delta_2)}{2(1-\delta_1')(1-\delta_2')}\right)\mathtt{open}(OPT^L) + t(opt + d').$$

For the connection cost, since each $f' \in D_X$ has one of its closest facilities in OPT in $OPT \setminus D_X$, the proof of Lemma G.2 can be applied verbatim (switching the roles of $\delta_1 \leftrightarrow \delta_2$ and $\delta_1' \leftrightarrow \delta_2'$) to show that the expected connection cost is $t'(opt + d')$ for $t' = \frac{1}{2}(1 + \frac{1}{\delta_2 \delta_1'} + \max(\frac{1-\delta_1'}{\delta_1'}, \frac{1}{1-\delta_1'}))$. (The additional factor $\frac{1}{2}$ comes from the fact that each $f^* \in OPT^L$ is deleted with probability exactly $1/2$.) $\quad\square$

Set $\delta_1 = \delta, \delta_2 = 1/2$, and assume that $d' \leq 4opt$. By Theorem G.2,

$$\mathtt{open}(S') + d' \leq \mathtt{open}(OPT) + \frac{\delta_1}{1-\delta_1}d' + \frac{1}{1-\delta_1}opt^M + 4opt^L.$$

Let $\alpha^L = opt^L/opt$. Since

$$\frac{\delta_1}{1-\delta_1}d' + \frac{1}{1-\delta_1}opt^M + 4opt^L \leq opt(\frac{1+4\delta_1}{1-\delta_1} + 4\alpha^L),$$

if $\alpha^L \leq (1.9 - \frac{1+4\delta_1}{1-\delta_1})/4$, S' is already an LMP 1.9 approximate solution. Therefore, we can assume that $\alpha^L \geq (1.9 - \frac{1+4\delta_1}{1-\delta_1})/4$. (We will choose $\delta_1 > 0$ such that the RHS is strictly positive.)

Also by Lemma G.3, if we let $d^\dagger$ be the connection cost of $OPT^\dagger$,

$$\mathtt{open}(S') + d' \leq \mathbb{E}[\mathtt{open}(OPT^\dagger)] + 2\mathbb{E}[d^\dagger] \leq \mathtt{open}(OPT) - \zeta\mathtt{open}(OPT^L) + t''(opt + d'),$$

where $\zeta = \zeta(\delta_1, \delta_2, \delta_1', \delta_2') = \left(\frac{1}{2} - \frac{(1-\delta_1)(1-\delta_2)}{2(1-\delta_1')(1-\delta_2')}\right) > 0$ and $t'' = t''(\delta_1, \delta_2, \delta_1', \delta_2') = t + 2t' < \infty$.

When $\zeta \cdot \mathtt{open}(OPT^L) \geq t''(opt + d')$ then S' is already an LMP 1 approximate solution. Therefore, one can assume $\zeta \cdot \mathtt{open}(OPT^L) < t''(opt + d') \leq 3t''opt$. Together with $opt^L \geq opt \cdot (1.9 - \frac{1+4\delta_1}{1-\delta_1})/4$, it upper bounds the ratio between the facility cost and the connection cost of OPT^L by some constant depending only on $\delta_1, \delta_2, \delta_1', \delta_2'$. Therefore, we can apply an analogue of Corollary 2.1 to conclude that S' is an LMP $(2 - \eta)$ approximation for some $\eta = \eta(\delta_1, \delta_2, \delta_1', \delta_2') > 0$. More specifically, a lower bound of η can be obtained as follows.

- Let δ be a free parameter to be determined and $\delta_1 = \delta, \delta_1' = \delta, \delta_2 = 1/2, \delta_2' = 1/4$.

- $t = 1 + \frac{1}{\delta_1 \delta_2'} + \max(\frac{1-\delta_2'}{\delta_2'}, \frac{1}{1-\delta_2'}) = 4 + \frac{4}{\delta} \leq \frac{5}{\delta}$ for $\delta \leq \frac{1}{4}$.

- $t' = \frac{1}{2}(1 + \frac{1}{\delta_2 \delta_1'} + \max(\frac{1-\delta_1'}{\delta_1'}, \frac{1}{1-\delta_1'})) \le \frac{1}{2} + \frac{3}{\delta} \le \frac{4}{\delta}$.

- $\zeta = \left(\frac{1}{2} - \frac{(1-\delta_1)(1-\delta_2)}{2(1-\delta_1')(1-\delta_2')}\right) \ge \frac{1}{2} - \frac{1-\delta_2}{2(1-\delta_2')} = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}$.

- $opt^L \ge opt \cdot (1.9 - \frac{1+4\delta}{1-\delta})/4$.

- $\mathtt{open}(OPT^L) \le \frac{3(t+2t')}{\zeta} opt \le 3 \cdot 6 \cdot (\frac{5}{\delta} + \frac{8}{\delta}) = \frac{234}{\delta}$.

- By Corollary 2.1 and the analytic bound on $opt_{JMS}(T)$ proved in Corollary B.1 with $OPT' \leftarrow OPT^L$ and $T \leftarrow \frac{234/\delta}{(1.9 - \frac{1+4\delta}{1-\delta})/4}$, we have $\eta \ge \frac{1.9 - \frac{1+4\delta}{1-\delta}}{4} \cdot \frac{1}{4(7 + 3\frac{234/\delta}{(1.9 - \frac{1+4\delta}{1-\delta})/4})}$.

- For $\delta = 0.05$, the lower bound becomes at least $4.5 \cdot 10^{-7}$.

G.1.1 Extend-JMS Local Search – Proof of Theorem G.2

We let S_1 denote the solution output by Extend-JMS local search and OPT be the optimum solution. We will need the following lemma to analyse the `Extend-JMS` swap.

LEMMA G.4. *Consider a set of facilities S_0, S^*, such that $S_0 \subseteq S^*$. Consider a not-necessarily-optimal assignment μ of clients to facilities in S^*, let $C_\mu(f)$ be the set of clients assigned to facility $f \in S^*$ in assignment μ. The cost of the solution S'' (with optimal assignment) produced by $\mathtt{Extend\text{-}JMS(S_0)}$ is at most*

$$\mathit{open}(S^*) + \sum_{f \in S_0} \sum_{c \in C_\mu(f)} dist(c, f) + 2 \sum_{f \in S^* - S_0} \sum_{c \in C_\mu(f)} dist(c, f).$$

Proof. We follow the standard JMS analysis. The only difference between $\mathtt{Extend\text{-}JMS}(S_0)$ and the standard JMS algorithm is that we pretend that the opening cost of each $f \in S_0$ is zero, while paying the original cost once f is open. We conservatively assume that each $f \in S_0$ is open, paying $\mathtt{open}(S_0)$ from the beginning. Then the cost of S'' is at most $\mathtt{open}(S_0)$ plus the cost of S'' in the facility location instance when each $f \in S_0$ has the opening cost zero.

Let $\alpha : C \to \mathbb{R}$ be the dual variables during the JMS algorithm. For each $f \in S^* - S_0$, the standard LMP 2 approximation guarantee ensures that $\sum_{c \in C_\mu^f} \alpha_c \le \mathtt{open}(f) + 2d(c, f)$. For each $f \in S_0$, we claim that $\sum_{c \in C_\mu^f} \alpha_c \le d(c, f)$. This follows from the fact that α_c is exactly the time at which c is connected to some facility for the first time; since f has the opening cost 0 and is at distance $d(c, f)$ from c, c is surely connected to some facility no later than $d(c, f)$.

Combining the two bounds, we conclude that the cost of S'' is at most

$$\mathtt{open}(S_0) + \left(\sum_{f \in S_0} \sum_{c \in C_\mu(f)} dist(c, f)\right) + \left(\mathtt{open}(S^* - S_0) + 2 \sum_{f \in S^* - S_0} \sum_{c \in C_\mu(f)} dist(c, f)\right),$$

which proves the lemma. $\square$

Proof. [Proof of Theorem G.2] Let S' be the solution at the end of `LocalSearch-JMS` (i.e., a local optimum). Let $d' = d(S')$ be the total connection cost of S', and $d'(c) := dist(c, S')$ be the connection cost of c. We analyse the algorithm by forming *swap* pairs (A_i, B_i), $A_i \subseteq S'$ and $B_i \subseteq OPT$. At the intuition level, each swap leads to a solution $S' - A_i \cup B_i$ whose cost is thus at least as high as S'. Start by creating a pair $(\{f'\}, \emptyset)$ for each facility of S'. Then, each matched facility $f^* \in OPT^M$ is added to the group of the corresponding (unique!) facility $f' \in OPT^M$ it is matched to. Finally, each $f^* \in OPT^L$ is added to the group of the closest $f' \in S'$. Notice that each group now consists of exactly one facility $f' \in S'$ and that each facility $f' \in S'$ appears in exactly one group. Notice also that each $f^* \in OPT$ belongs to exactly one group.

We say that a group containing more than $1/\varepsilon$ facilities of OPT is a *heavy* group. A group that is not heavy is *light*. For each heavy group $g = (\{f'\}, B)$, designate a special facility $g(B) \in B$ that is the facility that is matched to f' if $f' \in S^M$ and an arbitrary facility otherwise. For each group $g = (\{f'\}, B)$, we will consider a specific solution S''_g. Let $s''_g = d(S''_g)$ and $s''_g(c) := dist(c, S''_g)$.

1. If $(\{f'\}, B)$ is light, we consider the solution $S''_g = S' - \{f'\} \cup B$. Since g is light, we have by local optimality that $\mathsf{cost}(S') \leq \mathsf{cost}(S''_g)$.

2. If $(\{f'\}, B)$ is heavy, we proceed as follows. If $\{f'\}$ is matched, then the corresponding facility f^* is part of B and we consider the solution S''_g produced by $\texttt{Extend-JMS}(S' - \{f'\} \cup \{g(B)\})$. We have by local optimality that $\mathsf{cost}(S') \leq \mathsf{cost}(S''_g)$.

We now aim at bounding the cost of each solution S''_g defined by the above swaps. Given a facility $f^* \in OPT$, recall that $C_{OPT}(f^*)$ is the set of clients served by f^* in solution OPT. We first consider the case where $g = (\{f'\}, B)$ is light. We then partition the clients into four categories.

- Let C_1 be the set of clients that are either (1) served by a lonely facility in B in OPT, or (2) served by f' in S' and the facility of OPT that is matched to f'.

- Let C_2 be the set of clients not in C_1 that are served by $\{f'\}$ in S' and by a matched facility in OPT.

- Let C_3 be the set of clients not in C_1 that are served by $\{f'\}$ in S' and by a lonely facility in OPT.

- Finally let C_4 be the remaining clients.

We let $s''_g(c)$ denote the cost of client c in solution S''_g. We have

- If $c \in C_1$, $s''_g(c) \leq opt(c)$.

- If $c \in C_2$, we consider the facility f^* serving c in OPT. Let $(\{f_2\}, B')$ be the group such that $f^* \in B'$. Since $B \neq B'$ we have that $f_2 \in S''_g$ and so $s''_g(c) \leq dist(c, f_2)$, which by triangle inequality is at most $opt(c) + dist(f_2, f^*)$. We have that

$$dist(f_2, f^*) \leq \frac{1}{(1-\delta)|C_{OPT}(f^*)|} \sum_{c' \in C_{OPT}(f^*)} (opt(c') + d'(c')).$$

- If $c \in C_3$, we consider the facility f^* serving c in OPT. Let $(\{f_2\}, B')$ be the group such that $f^* \in B'$. Since $B \neq B'$ we have that $f_2 \in S''_g$ and so $s''_g(c) \leq dist(c, f_2)$, which by triangle inequality is at most $opt(c) + dist(f_2, f^*)$. We have that

$$dist(f^*, f_2) \leq dist(f^*, f') \leq opt(c) + d'(c).$$

- If $c \in C_4$, we have $s''_g(c) \leq d'(c)$.

Recall that for each client c, let $OPT(c)$ be the facility of OPT that serves it in OPT. Since $\mathsf{cost}(S') \leq \mathsf{cost}(S''_g)$, we have $\mathsf{cost}(S') - \mathsf{cost}(S''_g) \leq 0$, and so

$$
\begin{aligned}
& \mathsf{open}(f') - \mathsf{open}(B) + \sum_{c \in C_1} \big(opt(c) - d'(c)\big) \\
\text{(G.40)} \quad & + \sum_{c \in C_2} \left(opt(c) - d'(c) + \frac{\sum_{c' \in C_{OPT}(OPT(C))} opt(c') + d'(c')}{(1-\delta)|C_{OPT}(OPT(c))|}\right) + \sum_{c \in C_3} 2opt(c) \quad \leq 0
\end{aligned}
$$

We now turn to the case where $(\{f'\}, B)$ is heavy. Recall $S''_g = \texttt{Extend-JMS}(S' - \{f'\} \cup \{g(B)\})$ and let $S'_g = S' - \{f'\} \cup B$. We consider the following assignment of points to centers. We then partition the clients into four categories and define a not-necessarily-optimal assignment μ of points to center in S'_g.

- If $g(B)$ and f' are matched, let C_1^a be the set of clients that are served by $g(B)$ in OPT and by f' in S'. Otherwise let C_1^a be the whole set of clients served by $g(B)$ in OPT. Let the clients in C_1^a be assigned to $g(B)$ in μ.

- Let C_1^b be the set of clients that are served by a in $B - g(B)$ in OPT. This facility is necessarily a lonely facility. These clients are assigned to the closest facility in $B - g(B)$ in μ.

- Let C_2 be the set of clients not in $C_1^a \cup C_1^b$ that are served by $\{f'\}$ in S' and by a matched facility f^* in OPT. These clients are assigned to the facility of S' that is matched to f^* in μ.

- Let C_3 be the set of clients not in $C_1^a \cup C_1^b$ that are served by $\{f'\}$ in S' and by a lonely facility f^* in OPT. These clients are assigned to the facility of S' that is the closest to f^* in μ.

- Finally let C_4 be the remaining clients. These clients are assigned to the closest facility of $S' - \{f'\}$ in μ.

We now bound the cost of solution S_g' with the assignment μ, which will allow us to bound the cost of S_g'' via Lemma G.4. Let s_g' and $s_g'(c)$ be the total connection cost and the connection cost of c in S_g', *according to the assignment* μ.

- If $c \in C_1^a$, $s_g'(c) \leq opt(c)$.

- If $c \in C_1^b$, $s_g'(c) \leq opt(c)$.

- If $c \in C_2$, we consider the facility f^* serving c in OPT. Let $(\{f_2\}, B')$ be the group such that $f^* \in B'$. Since $B \neq B'$ we have that $f_2 \in S_g''$ and so $s_g'(c) \leq dist(c, f_2)$, which by triangle inequality is at most $opt(c) + dist(f_2, f^*)$. We have that

$$dist(f_2, f^*) \leq \frac{1}{(1-\delta)|C_{OPT}(f^*)|} \sum_{c' \in C_{OPT}(f^*)} (opt(c') + d'(c')).$$

- If $c \in C_3$, we consider the facility f^* serving c in OPT. Let $(\{f_2\}, B')$ be the group such that $f^* \in B'$. Since $B \neq B'$ we have that $f_2 \in S_g''$ and so $s_g'(c) \leq dist(c, f_2)$, which by triangle inequality is at most $opt(c) + dist(f_2, f^*)$. We have that

$$dist(f^*, f_2) \leq dist(f^*, f') \leq opt(c) + d'(c).$$

- If $c \in C_4$, we have $s_g'(c) \leq d'(c)$.

We then apply Lemma G.4 ($S_0 \leftarrow S' - \{f'\} \cup \{g(B)\}$, $S^* \leftarrow S_g'$ and $\mu \leftarrow \mu$) to get a guarantee on $S_g'' = \texttt{Extend-JMS}(S' - \{f'\} \cup \{g(B)\})$. Since clients of type C_1^a, C_2, C_3 and C_4 are served by a facility in $S' - \{f'\} \cup \{g(B)\}$ in assignment μ, Lemma G.4 implies

$$\texttt{cost}(S_g'') \leq \texttt{open}(S' - \{f'\} \cup B) + \sum_{c \in C_1^a} opt(c) + \sum_{c \in C_2} \left(opt(c) + \frac{\sum_{c' \in C_{OPT}(OPT(C))} opt(c') + d'(c')}{(1-\delta)|C_{OPT}(OPT(c))|} \right)$$

$$+ \sum_{c \in C_3} \left(2opt(c) + d'(c) \right) + \sum_{c \in C_4} d'(c) + \sum_{c \in C_1^b} 2opt(c).$$

It follows that, since $\texttt{cost}(S') - \texttt{cost}(S_g'') \leq 0$, we have

$$\texttt{open}(f') - \texttt{open}(B) + \sum_{c \in C_1^a} \left(opt(c) - d'(c) \right)$$

$$+ \sum_{c \in C_2} \left(opt(c) - d'(c) + \frac{\sum_{c' \in C_{OPT}(OPT(C))} opt(c') + d'(c')}{(1-\delta)|C_{OPT}(OPT(c))|} \right) + \sum_{c \in C_3} 2opt(c) + \sum_{c \in C_1^b} \left(2opt(c) - d'(c) \right)$$

(G.41) $\quad \leq 0$

We now conclude the proof of the theorem. Consider summing up the inequalities (G.40) and (G.41) over all groups $(\{f'\}, B)$. First, observe that each facility of OPT and S' appears in exactly one group. Hence, the sum of the quantities $\texttt{open}(f') - \texttt{open}(B)$ is exactly $\texttt{open}(S') - \texttt{open}(OPT)$.

Moreover, each client c served by a facility $f^* \in OPT^L$ (hence not matched) appears as a client of type C_1, C_1^a or C_1^b in exactly one swap (i.e.: the swap $(\{f'\}, B)$ where $f^* \in B$), as a client of type C_3 in at most one swap

(i.e.: the swap $(\{f'\}, B)$ where f' is the facility serving it in S') and as a type C_4 client in all remaining swaps. Therefore, its total contribution to LHS of the sum is at most $4opt(c) - d'(c)$.

In addition, each client c served by a facility $f^* \in OPT^M$ and served by a facility $f' \in S'^M$ such that f^* and f' are matched appears as a client of type C_1 or C_1^a in exactly one swap and as a client of type C_4 otherwise. Its total contribution is thus at most $opt(c) - d'(c)$.

Each client c served by a facility $f^* \in OPT^M$ and served by a facility $f' \in S'$ such that f^* and f' are not matched appears as a client of type C_2 in exactly one swap and as a client of type C_4 otherwise. Its total contribution is thus at most $opt(c) - d'(c) + \frac{\sum_{c' \in C_{OPT}(OPT(c))} opt(c') + d'(c')}{(1-\delta)|C_{OPT}(OPT(c))|}$. Observe that for each matched facility f^* of OPT^M the number of such clients is at most $\delta|C_{OPT}(OPT(c))|$ by the definition of matched. Hence, summing up over all clients served by f^* in OPT, the above bound is at most

$$\sum_{c \in C_{OPT}(f^*)} \big(opt(c) - d'(c)\big) + \frac{\delta}{1-\delta} \sum_{c \in C_{OPT}(f^*)} \big(opt(c) + d'(c)\big).$$

Therefore, the total cost of S' is as claimed at most

$$\mathtt{open}(OPT) + \frac{\delta}{1-\delta}d' + \frac{1}{1-\delta}opt^M + 4opt^L.$$

□

G.2   Final Result by Cost Scaling The result in the above subsection shows that for any Facility Location instance, if OPT is the solution minimizing $\mathtt{open}(OPT) + (2-\eta)opt$, S' is a local optimum of the extended local search starting from a JMS solution, and $d' \leq 4opt$, then it is already an LMP $(2-\eta)$ approximate solution for some $\eta > 0$ depending only on our choices of $\delta_1, \delta_2, \delta_1', \delta_2'$. Though ensuring $d' \leq 4opt$ in general is a nontrivial task, we show how to get around this bottleneck and get an LMP $(2 - \eta/2)$ approximation. Let OPT be the feasible solution that minimizes $\mathtt{open}(OPT) + (2 - \eta/2)opt$.

The idea is to use the concept of bipoint solutions again. In particular, we multiply each facility cost by a factor λ, and apply the JMS algorithm and the extended local search to the resulting Facility Location instance. Let $S(\lambda)$ be the corresponding solution.

If $\lambda = 0$, $S(0)$ connects all clients to their closest facilities. If the original facility cost of $S(0)$ is at most that of OPT, then $S(0)$ is already an LMP 1 approximation, so we can assume otherwise. On the other hand, let $\gamma = \min(\min_{f \in F} \mathtt{open}(f), \min_{f, f' \in F: \mathtt{open}(f) \neq \mathtt{open}(f')} |\mathtt{open}(f) - \mathtt{open}(f')|)$, and $M = \sum_{c \in C, f \in F} d(c, f)$. Suppose we scale facility costs by $\lambda = 3M/\gamma$, and consider an arbitrary solution S' that consists of one facility with the lowest opening cost. By the definition of λ, for any solution S'' that strictly spends more on the facility cost than S', whether it opens two cheapest facilities or opens a facility which is not the cheapest, the difference between the opening costs between S' and S'' is at least $3M$, which is at least 3 times larger than the connection cost of any solution. The fact that $S(\lambda)$ is also LMP 2 approximate w.r.t. S' (with respect to scaled facility costs) implies that $S(\lambda)$ also opens exactly one cheapest facility, so its facility cost must be at most that of OPT (both scaled and original).

We perform a binary search over λ until we find a value λ^* such that $S_1 = S(\lambda^* + \delta)$ has facility cost (with respect to original facility costs) $\mathtt{open}(S_1) \leq \mathtt{open}(OPT)$ and $S_2 = S(\lambda^*)$ has facility cost (with respect to original facility costs) $\mathtt{open}(S_2) \geq \mathtt{open}(OPT)$. Here $\delta > 0$ can be chosen to be exponentially small in the input size and we will next neglect it at the cost of an extra factor $(1 + \varepsilon)$ in the approximation for arbitrarily small constant $\varepsilon > 0$. Also note that the value of $\mathtt{open}(OPT)$ can be guessed at the cost of an extra $(1 + \varepsilon)$ in the approximation.

Define $a \in [0, 1]$ such that $a\mathtt{open}(S_1) + (1 - a)\mathtt{open}(S_2) = \mathtt{open}(OPT)$. Let $d_i = d(S_i)$. Since S_1 and S_2 are LMP 2 approximate solutions with respect to scaled facility costs, we have

$$\lambda^*\mathtt{open}(S_1) + d_1 \leq \lambda^*\mathtt{open}(OPT) + 2opt, \quad \lambda^*\mathtt{open}(S_2) + d_2 \leq \lambda^*\mathtt{open}(OPT) + 2opt,$$

which implies

$$ad_1 + (1 - a)d_2 \leq 2opt.$$

Suppose that $a \geq 1/2$. (The other case is symmetric.) This implies that $d_1 \leq 4opt$, so we can Lemma G.1 so that S_1 is indeed LMP $(2 - \eta)$ approximate with respect to scaled facility costs for some $\eta \geq 4.5 \cdot 10^{-7}$;

$\lambda^* \mathsf{open}(S_1) + d_1 \leq \lambda^* \mathsf{open}(OPT) + (2 - \eta)opt$. This implies that a random solution choosing S_1 with probability a and S_2 with probability $(1 - a)$ has the expected (original) facility cost $\mathsf{open}(OPT)$ and the connection cost at most $a(2 - \eta)opt + (1 - a)2opt \leq (2 - \eta/2)opt$. Therefore, one of S_1 and S_2 is an LMP $(2 - \eta/2)$ approximate solution. This proves Theorem G.1.