An Improved Local Search Algorithm for k-Median

Vincent Cohen-Addad, Anupam Gupta, Lunjia Hu, Hoon Oh, David Saulpic

An Improved Local Search Algorithm for k-Median

Vincent Cohen-Addad* Anupam Gupta† Lunjia Hu‡ Hoon Oh†

David Saulpic§

Abstract

We present a new local-search algorithm for the k-median clustering problem. We show that local optima for this algorithm give a $(2.836+\epsilon)$-approximation; our result improves upon the $(3+\epsilon)$-approximate local-search algorithm of Arya et al. [AGK+01]. Moreover, a computer-aided analysis of a natural extension suggests that this approach may lead to an improvement over the best-known approximation guarantee for the problem.

The new ingredient in our algorithm is the use of a potential function based on both the closest and second-closest facilities to each client. Specifically, the potential is the sum over all clients, of the distance of the client to its closest facility, plus (a small constant times) the truncated distance to its second-closest facility. We move from one solution to another only if the latter can be obtained by swapping a constant number of facilities, and has a smaller potential than the former. This refined potential allows us to avoid the bad local optima given by Arya et al. for the local-search algorithm based only on the cost of the solution.

*Google Research, Zurich and Sorbonne Université, Paris.

†Carnegie Mellon University, Pittsburgh PA 15217.

‡Stanford University.

§Sorbonne Université, Paris.

1 Introduction

The *k-median* problem is a classic optimization problem for metric spaces, and has been widely studied by the algorithm-design community with a two-fold motivation: on the one hand getting good algorithms for the k-median problem immediately yields important practical implications in operations research, bioinformatics, or data analysis. On the other hand, the study of the approximability of k-median has given us a deeper understanding of key algorithmic ideas like primal-dual techniques and Lagrangian-multiplier preserving algorithms, sophisticated dependent LP roundings, local search, iterative rounding, and algorithmic notions of stability.

Concretely, given a finite metric space $(\mathcal{X}, d)$, where the point set $\mathcal{X}$ is partitioned into *client* locations $\mathcal{C}$ and possible *facility* locations $\mathcal{F}$, with $\mathcal{X} := \mathcal{C} \cup \mathcal{F}$, and a parameter k, the *k-median* problem asks to pick k "medians" $F \subseteq \mathcal{F}$ to minimize

$$\mathsf{kmed}(F) := \sum_{c \in \mathcal{C}} d(c, F). \tag{1.1}$$

Throughout the paper, given a set $S \subseteq \mathcal{X}$, and point $x \in \mathcal{X}$ we let $d(x, S)$ denote $\min_{s \in S} d(x, s)$.

An interesting perspective on the k-median problem is to view it as a "metric set cover" problem, where one needs to find k medians (seen as "sets") to cover the clients (seen as the universe) – with the relaxation that each client pays a cost that is a function of how well it is covered and this cost function is a metric. This perspective has long been known (see e.g. [GK99, JMS02]), but although the complexity of the classic set cover problem is well-understood since the 90s, the approximability of this metric variant is still quite open.

The current-best result is the 2.675-approximation of Byrka et al. [BPR+15], improving on a breakthrough 2.732-factor of Li and Svensson [LS16]. These papers use the clever idea of finding pseudo-approximations (i.e., solutions with good cost but opening a few extra facilities) by first giving bi-point solutions (i.e., a feasible fractional solution that is the convex combination of two integer solutions) using the primal-dual framework, and then rounding these bi-point solutions carefully into integer solutions. Nevertheless, the gap between these results and the current best hardness bound of $1 + 2/e$ remains large. While various techniques can give good approximations for k-median in specific metrics, the current arsenal for getting a better approximation bound for the general case is not very rich. E.g., a significant improvement using the bi-point rounding approach seems challenging, since it requires either improving the quality of the bi-point solution computed (on which no progress has been made over the last 20 years), or improving on the rounding scheme. Other techniques to obtain $O(1)$-approximations are primal-dual, or greedy-plus-pruning, but the best bounds using these techniques do not even give a 3-approximation. Finally, the best result before [LS16] was an analysis of the p-swap local-search algorithm that tries to improve the current solution by closing some p facilities and opening p others. Arya et al. [AGK+01] showed that any local optimum was a $(3+2/p)$-approximation. However, they also showed instances with a matching "locality gap" for this algorithm (see §A.2 for a simple example showing a gap arbitrarily close to 3). In summary, the only known way to do better than a factor of 3 remains bi-point rounding.

In this paper, we draw on parallels with set cover and submodular optimization problems and propose an extension of the simple local-search paradigm that has the potential to improve the current best-known approximation factor. While our current analysis does not improve the best approximation it provides the first alternative to bi-point solutions to go below a 3-approximation—namely, to 2.836—and offers the possibility of better results. The new idea is to perform the local search with respect to some other "surrogate" potential $\Phi(F)$ instead of the k-median objective function. This allows us to avoid the bad local minima present in the standard local search. Of course, this Φ needs to be easily computable, and also to be close enough to the original objective function so that finding a local-optimum with respect to Φ implies a good approximation for k-median objective as well. Such local-search procedures are called *non-oblivious* in the literature, and have been successful in several settings [Ali94, KMSV98, FW12, FW14, GGK$^+$18].

1.1 Our Approach and Results

Let $d_i(c, F)$ be the distance between the client c and the facility in F that is i^{th}-closest to it, so that $d_1(c, F) = d(c, F)$ as defined above. Define the potential function

$$\Phi(F) := \sum_{c \in \mathcal{C}} \Big(\underbrace{d_1(c,F)}_{\text{closest}} + \beta \min\{ \underbrace{d_2(c,F), \alpha\, d_1(c,F)}_{\text{truncated second-closest}} \} \Big). \tag{1.2}$$

For almost all of the paper, we choose $\alpha = 3$ and $\beta = 1/5$. While we motivate the potential in detail in §1.2, consider two clients whose closest facilities are both at distance D: one with its second-closest facility at the same distance D pays $(1+\beta)D \approx 1.2\,D$, whereas another whose second-closest facility is much farther away pays $(1 + \alpha\beta)D \approx 1.6\,D$. Hence a lower potential prefers solutions with good "backup" facilities, so that local moves can then explore a richer space. Our main result is the following:

Theorem 1.1 (Pseudo-approximation). *Let $\alpha = 3, \beta = 1/5$, and let $p(\varepsilon), r(\varepsilon)$ be sufficiently large constants that depend only on ε. If F is a local minimum of our non-oblivious local-search procedure with $|F| = k$ facilities and swap size $p(\varepsilon)$, then*

$$\mathsf{kmed}(F) \le (2.836 + \varepsilon) \cdot \mathsf{kmed}(F^*)$$

for any solution F^* with $k - r(\varepsilon)$ facilities.

We can convert this pseudo-approximation into a regular approximation using ideas from [LS16, ABS10]. Indeed, if the original instance is "stable" (i.e., if reducing the number of facilities by $r(\varepsilon)$) causes the optimal cost to increase by more than $(1 + \varepsilon)$), we can get a PTAS [ABS10] in time $\text{poly}(|\mathcal{X}|^{r(\varepsilon)})$. Hence, this reduction of the number of facilities does not change the optimal cost much, and then the pseudo-approximation of Theorem 1.1 is also a true approximation.

We are yet to understand the limitations of this specific potential function, and of this general approach. The best lower bound for this potential function we currently know is the following:

Theorem 1.2 (Lower Bound for Φ). *There exists $\varepsilon > 0$ and an infinite family of instances on which the local-minimum F of our non-oblivious local-search function with constant-sized swaps satisfies*

$$\mathsf{kmed}(F) \ge \min\{\max\{(3 - 2\beta - \varepsilon, 1 + 4\beta - \varepsilon)\}, \max\{2, \alpha - \varepsilon\}\} \cdot \mathsf{kmed}(F^*).$$

Balancing the two terms gives us a locality gap lower bound of $2 \cdot \mathsf{kmed}(F^*)$ for all values of α, β.

This lower bound holds even if F is allowed to have more facilities than F^*. The gap between the two results above suggests that local-search with respect to Φ still has the possibility of beating the current-best approximation bounds.

Extending our Potential Function. We consider extending this non-oblivious approach using more expressive potentials. E.g., we can look at the $q = 3$ closest facilities, as follows: (we use d_i as shorthand for $d_i(c, F)$, and $(a \wedge b) := \min(a, b)$)

$$\Phi_3(F) = \sum_{c \in \mathcal{C}} \Big(d_1 + \beta_2 \underbrace{(\alpha_2 d_1 \wedge d_2)}_{\text{truncated second-closest}} + \beta_3 \underbrace{(\alpha_3 d_1 \wedge d_3)}_{\text{truncated third-closest}} \Big).$$

Again α_i, β_i are constants, discussed in §7. A preliminary implementation of this LP discussed in that section suggests that we can get an approximation ratio of 2.69. However, these are based on experiments, and since we do not have a formal proof, computer-assisted or otherwise, these should just be considered circumstantial evidence and promising first steps. We hope that we (or others) will be able to extend these to a formal proof.

1.2 Our Techniques

Since the algorithm is just the p-swap local search algorithm, all the work is in the analysis of the local optima.

The choice of the objective function. Our potential function is inspired by the work of Filmus and Ward [FW12, FW14], who improved the local-search algorithm for submodular maximization from a $1/2$-factor to the optimal $(1-1/e)$-factor. We describe their idea in the context of max-k-coverage: the potential gets a bonus if it covers elements multiple times. I.e., for each element, we get a value of 1 if we cover it once, a small bonus β_2 if we cover it at least twice, a smaller additional bonus β_3 if we cover it at least thrice, etc. The total overall bonus is small compared to the gain in covering it once (so that the potential remains close to the true objective), but enough to evade the bad local minima. Indeed, if an element is covered twice, the algorithm has more flexibility in choosing local-search steps, since any single-set swap will leave this element still covered.

The k-median problem is a minimization problem, so the natural objective is $\sum_c d_1(c)+\sum_{i\geq 2}\beta_i d_i(c)$, where $d_i(c)$ is the distance from c to its i^{th}-closest facility: this penalty term can incentivize each facility to have "backup" facilities close to it. Indeed, just using $d_1+\beta_2 d_2$ (for small constant $\beta_2>0$) side-steps the standard bad examples with respect to the objective function d_1. However, this potential penalizes us too heavily for not having backups. So if the instance has k widely-separated clusters, the penalty term overwhelms the original cost. This suggests the potential (1.2) we eventually use:

$$\sum_c d_1(c)\left[1+(\text{small constant})\times\min\left(1,\frac{d_2(c)}{(\text{large constant})\times\ d_1(c)}\right)\right].$$

However, the introduction of the minimum in the objective function makes the analysis more involved, since it forces a case distinction between clients which pay the truncated and untruncated values.

Important Swaps. The standard approach to analyze the quality of local optima for clustering problems is to define a subset of swaps we call *important*. Since all swaps are non-improving, these important ones are too. This non-improvement gives linear inequalities that relate the cost of the solution F_{new} after the swap to the cost of the local optimum F. To relate F_{new} to the optimal solution F^*, we define important swaps to be ones that replace a constant number of local facilities $P\subseteq F$ with the same number of optimal facilities $Q\subseteq F^*$. Hence, the cost of F_{new} is the sum of the costs for (1) "happy" clients that are now served optimally (or even better) in F_{new} because their optimal facility is in Q, (2) the "sad" clients which were previously assigned to the facilities in P that were swapped out, but which are not happy and hence require *reassignment*, and (3) the remaining "indifferent" clients. The art in these proofs is to define the important swaps to control the reassignment cost for the sad clients.

For example, we can pair each optimal facility with its closest local facility (assume for now this is a bijection), and form the important swaps by swapping some constant-sized subset of these pairs.

This ensures:

$$\sum_{c \text{ happy}} d(c, F^*) + \sum_{c \text{ sad}} (d(c, F) + 2d(c, F^*)) + \sum_{c \text{ indifferent}} d(c, F) \geq \mathsf{cost}(F_{\text{new}}) \geq \mathsf{cost}(F) = \sum_{c} d(c, F).$$

(see [GT08] for details). Simplifying gives

$$\sum_{c \text{ happy}} d(c, F^*) + \sum_{c \text{ sad}} 2d(c, F^*) \geq \sum_{c \text{ happy}} d(c, F).$$

Summing over important swaps (one per local facility) means each client appears on the left at most twice (once when happy, and once when sad) and on the right exactly once, which means $ALG \leq 3OPT$. Handling the non-bijective case loses another ε factor, so the local optimum is at most $(3 + \varepsilon)$ times the global optimum. The important lessons are that (a) important swaps need to be "rich" enough to infer the small locality gap, and (b) "simple" enough to be able to reason about.

However, the important swaps used in past works [AGK$^+$01, GT08] do not work with the new potential: Figures B.11 and B.12 in Appendix B show instances and local solutions that cost three times the optimum but are not locally optimal with respect to the new objective function. Yet previously-used important swaps are not rich/expressive enough to deduce non-local-optimality, and only prove a 3-approximation.

New Swaps. Given a local solution F, we distinguish the *far* clients c with $d_2(c, F) \geq \alpha d(c, F)$ from the *close* ones with $d_2(c, F) < \alpha d(c, F)$. The type of a client determines which value attains the minimum in the potential function (1.2): a far client c pays $(1 + \alpha\beta)d(c, F)$ while a close one pays $d(c, F) + \beta d_2(c, F)$. The two types of clients require different analysis.

Far Clients. Consider a facility ℓ_2 of F closest to the optimal facility f^* for far client c. If ℓ_2 is also the local facility that is closest to c, and if we pair it with f^*, client c is a happy client (as described above) and we get a good bound on the cost of client c (so we should always associate f^* with ℓ_2). Else if ℓ_2 is not a facility that is the closest to c, then a simple argument using the triangle-inequality shows there exists a second facility in the local solution at distance $2d(c, f^*) + d(c, F)$ to c. But c is a far client, so this facility cannot be too close: $2d(c, f^*) + d(c, F) \geq \alpha d(c, F)$, and so $d(c, F) \leq \frac{2}{\alpha - 1} d(c, f^*)$, which is an excellent bound.

Close Clients. On the other hand, the close clients, may now be sad both when their closest facility closes, and also *when their second-closest closes.* E.g., consider a client whose closest optimal facility is far from the rest of the instance, but which has two local facilities at the same distance to it (with $d_1 \approx d_2$). (See Figure 1.1.) In this case, moving from two facilities to one in the local solution without opening the optimal facility incurs a large reassignment cost. Hence, such clients want the swap which opens the optimal facility to also close both local facilities close to them. If not, closing any one of these close local facilities would mean reassigning them to the other, and suffering a cost of $(1 + \alpha\beta)d_1$. These woud be *very sad* clients. So we would like to close both the facilities for the close clients at the same time. Else the potential that was helping the far clients now hurts these close ones when they become very sad.

Our approach mitigates the risks: we define *two different swap structures* and take a linear combination of the inequalities obtained from these. Since the local-search algorithm tries all possible swaps, the resulting inequalities remain valid. The two swaps structures can be viewed as follows. One of them, referred to as *simple swaps*, is similar to the one described by [GT08], where each facility of F^* is mapped to its closest facility in F. The other one, which resolves the "bad example" described in Figure B.11 for single swaps, is to also consider the reverse map: i.e., to map each

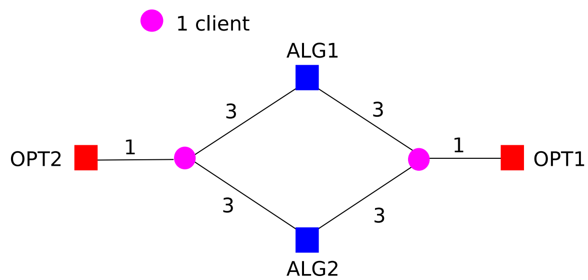


Figure 1.1: Illustration of the tension between clients for defining the swap structure. In order to get a good bound for the right client, we need to open OPT1 and close both local facilities ALG1 and ALG2. However, closing both facilities and opening OPT1 increases potential value of the left client to $(1+\alpha\beta)7$ *from* $(1+\beta)3$.

facility of F to its closest one in F^*. These two maps induce a directed graph G where the vertices are $F^* \cup F$, with an arc from f_1 to f_2 if f_1 is mapped to f_2 in the appropriate map. This graph G has outdegree-1 and hence has a nice structure. We show how to break it into pieces of bounded size; these define *tree swaps*. We then work with all the inequalities coming from these two families of swaps.

A final ingredient is randomization: instead of always mapping each facility f in one of the solutions to its closest facility f' in the other solution, we randomize these maps—we map f to its second-closest facility in the other solution with some probability that depends on their relative distances. This allows us to again mitigate bad and good scenarios for different types of clients that are in tension.

In summary, here's what we do: we flip a coin to either consider simple swaps or tree swaps. In either case, we randomly map some facilities to the closest or second-closest facilities in the other solution, and use this to build a set of important swaps. Since all these are non-improving, this gives us linear inequalities that relate the local cost to the optimum. Finally, we deduce the approximation ratio from these linear inequalities.

1.3 Related Work

The first $O(1)$-approximation for the k-median problem was given by Charikar et al. [CGTS99]. After many developments using, e.g., the primal-dual schema [JV01, CG99], greedy algorithms (and dual fitting) [JMM$^+$03], improved LP rounding [CL12], local-search [AGK$^+$01], and pseudo-approximations [LS16], the current best approximation guarantee is 2.675 [BPR$^+$15]. The best hardness result is $(1+2/e)$ [GK99, JMS02]. Local-search algorithms have been widely used for clustering problems. Despite their simplicity, they often give good theoretical guarantee: the $(3+\varepsilon)$-approximation result of [AGK$^+$01] was the best factor for some time; a simplified proof is given in [GT08]. The best results for the closely related k-means problem are by Ahmadian et al. [ANSW17], who give a $6.35+\varepsilon$-approximation for Euclidean metrics and $9+\varepsilon$ for general metrics, both using the primal-dual method: these improve on results of Kanungo et al. [KMN$^+$02] who show that the simple local-search with respect to the objective function gives a $(9+\varepsilon)$-approximation for Euclidean k-means.

Ahmadian et al. [AFS13] give a local-search algorithm for mobile k-median, where they also construct a 1-tree using the optimal and algorithm's centers (and the original centers, which play a role for that problem), and consider swaps based on its subtrees. However, the details of the analysis seem to be different from ours, since the concerns in the two problems are quite different.

The use of an alternate potential function instead of the objective function in local-search was termed non-oblivious by [Ali94, KMSV98]. Filmus and Ward [FW12, FW14] used non-oblivious local-search for the *maximum coverage* and *submodular maximization* problems, getting $1 - 1/e$-approximations in both cases. (A further simplification of the submodular algorithm/analysis appears in [FFSW17].)

1.4 Paper Outline

We formally define the algorithm in §2, and the set of important swaps in §3. We classify the clients into types in §4, and bound the expected change in potential for each client type in §6; combining them proves Theorem 1.1. In §7, we present how to construct a linear program that mimics our analysis. In Appendix A, we prove the lower bound from Theorem 1.2. Details of calculations, as well as deferred proofs, appear in the appendix.

2 The Local Search Algorithm

The algorithm performs swaps of constant size $p = p(\varepsilon) > 1/\varepsilon$: given any solution F (initially arbitrarily chosen) of k facilities from $\mathcal{F}$, it tries to find an *improving valid swap*. Here, a swap $(P, Q) \in \binom{\mathcal{F}}{\leq p} \times \binom{\mathcal{F}}{\leq p}$ is *valid* if $P \subseteq F$, $Q \subseteq \mathcal{F} \setminus F$, and $|P| = |Q|$, so that we close as many facilities as we open. A valid swap is *improving* if

$$\Phi((F \setminus P) \cup Q) < \Phi(F),$$

where Φ is as defined in (1.2). If the algorithm finds an improving valid swap (P, Q), it sets $F \leftarrow (F \setminus P) \cup Q$, and continues; if there are no such swaps it returns the local optimum F.

This algorithm can be made to run in polynomial time by only considering swaps that improve the potential by $(1 + \delta n^{-O(p)})$-factor; standard techniques (presented e.g. in Arya et al. [AGK$^+$01]) show that this changes the approximation factor by at most $(1 + \delta)$, since there are $n^{O(p)}$ many different swaps. Observe that checking whether we are at a (near)-local optimum, or finding an improving valid swap can be done in $n^{O(p)}$ time. In the rest of the paper we show the pseudo-approximation claimed in Theorem 1.1, i.e., the cost of a local optimum is comparable to the cost of any solution F^* with $k - r(\varepsilon)$ facilities, where $r(\varepsilon)$ is the number of extra local facilities.

Throughout the paper, we choose the swap size $p(\varepsilon)$ to be $M(\lceil 1/\varepsilon \rceil + 1)^{4\lceil 1/\varepsilon \rceil^{\lceil 1/\varepsilon \rceil}}$, and choose the number of extra local facilities to be $r(\varepsilon) = M(\lceil 1/\varepsilon \rceil + 1)^{1+16\lceil 1/\varepsilon \rceil^{\lceil 1/\varepsilon \rceil}}$ for a sufficiently large absolute constant M.

2.1 Proof Strategy

Let us fix some notation: fix a local optimum F of size k and a global optimum F^* of size $k - r(\varepsilon)$; we call the former the *local* and the latter the *optimal* facilities. For a client c, let

- $d^*(c) := d(c, F^*)$ be its cost and f^* its closest facility in the optimal solution F^*,
- $d_1(c)$ and $d_2(c)$ be its distances to the closest and second-closest facilities, and f_1 and f_2 be these facilities in F, and
- $\Phi^c := d_1(c) + \beta \min(d_2(c), \alpha d_1(c))$ be client c's contribution to the potential. From now on, we fix $\alpha = 3$ and $\beta = 1/5$.

Our proof of Theorem 1.1 is based on the fact that at the local optimum F, the potential change induced by a valid swap (P, Q) is non-negative, i.e., $\Phi((F \setminus P) \cup Q) - \Phi(F) \geq 0$. Defining the *potential change* of client c on swap (P, Q) to be

$$\delta_{(P,Q)}(c) := \Phi^c((F \setminus P) \cup Q) - \Phi^c(F), \tag{2.3}$$

we have

$$0 \leq \sum_{c \in \mathcal{C}} \delta_{(P,Q)}(c).$$

This inequality holds for all valid swaps (P, Q); it remains true even if we extend the definition of valid swaps to allow Q to intersect F and/or to have a size smaller than P, because doing so never decreases the potential change. We can thus take linear combinations of the inequality over all valid swaps (P, Q). In particular, for any random set $\mathcal{P}$ of valid swaps,

$$0 \leq \mathbb{E}_{\mathcal{P}}\Big[\sum_{(P,Q) \in \mathcal{P}} \sum_{c \in \mathcal{C}} \delta_{(P,Q)}(c)\Big] = \sum_{c \in \mathcal{C}} \mathbb{E}_{\mathcal{P}}\Big[\sum_{(P,Q) \in \mathcal{P}} \delta_{(P,Q)}(c)\Big].$$

Theorem 1.1 is thus implied by the following lemma (and observing that $\frac{2.5203}{0.8888} \leq 2.836$):

Lemma 2.1. *There is a distribution over sets $\mathcal{P}$ of valid swaps such that for all clients $c \in \mathcal{C}$,*

$$\mathbb{E}\Big[\sum_{(P,Q) \in \mathcal{P}} \delta_{(P,Q)}(c)\Big] \leq 2.5203\, d^*(c) - 0.8888\, d_1(c) + O(\varepsilon)\,(d^*(c) + d_1(c)).$$

In order to prove this lemma, we build a randomized procedure generating the set $\mathcal{P}$ of swaps (where we call elements of $\mathcal{P}$ *important swaps*), and divide our analysis into two cases: the *amenable* case and the *defiant* case. In particular, given a client c, we define a suitable amenable event $\mathcal{A}$ and its complement defiant event $\mathcal{D}$, and show the following two lemmas, which immediately imply Lemma 2.1.

Lemma 2.2 (Defiant Case)**.** *There is a distribution over sets $\mathcal{P}$ of valid swaps such that for all clients $c \in \mathcal{C}$,*

$$\mathbb{E}\Big[\mathbb{1}_{\mathcal{D}} \sum_{(P,Q) \in \mathcal{P}} \delta_{(P,Q)}(c)\Big] \leq O(\varepsilon)\,(d^*(c) + d_1(c)). \tag{2.4}$$

Lemma 2.3 (Amenable Case)**.** *For the distribution over valid swap sets from Lemma 2.2, for any $c \in \mathcal{C}$,*

$$\mathbb{E}\Big[\mathbb{1}_{\mathcal{A}} \sum_{(P,Q) \in \mathcal{P}} \delta_{(P,Q)}(c)\Big] \leq 2.5203\, d^*(c) - 0.8888\, d_1(c) + O(\varepsilon)\,(d^*(c) + d_1(c)). \tag{2.5}$$

In §3, we define the distribution over sets $\mathcal{P}$ of important swaps. In §4 we classify the clients into types. We define the amenable and defiant events for clients of each type and prove Lemma 2.2 in §5, and then prove Lemma 2.3 in §6.

3 Generating the Important Swaps

In this section, we describe our randomized procedure generating $\mathcal{P}$, the set of important swaps, that proves Lemmas 2.2 and 2.3. $\mathcal{P}$ contains valid swaps (P, Q), where $P \subseteq F$ has size at most $p(\varepsilon)$, and Q is an arbitrary set of facilities with size at most $|P|$. Every swap we generate has Q

being a subset of F^*, the set of optimal facilities. We say swap (P, Q) *closes* the local facilities in P, and *opens* the optimal facilities in Q. (By duplicating points in the metric space, we assume F and F^* are disjoint, and so are P, Q.) Sometimes we say the swap *contains* the local facilities in P and the optimal facilities in Q.

In order to prove Lemmas 2.2 and 2.3, we want to minimize the potential change of every client by always opening a "nearby" optimal facility whenever we close a local facility. Roughly, we generate both simple and tree swap sets by constructing a directed graph G over the vertex set $F \cup F^*$, where every edge connects "nearby" local and optimal facilities. We perform some surgery on this graph if needed: we remove vertices in F, duplicate vertices in both F and F^*, and remove some edges, so that every connected component of the resulting graph has a small size. Finally, we combine these connected components of G into small-sized groups so that the number of local facilities in each group is no smaller than the optimal facilities in it. The swap set $\mathcal{P}$ consists of the swap defined by each of these groups, closing/opening all the local/optimal facilities in it. In the following subsections, we describe in detail our procedures generating the simple and tree swap sets. (Again, recall this is all in the analysis, since the algorithm is just the p-swap local search that attempts to improve the potential.)

3.1 Generating the Important Simple Swaps

We start by constructing a random directed graph G_0 over vertices $F \cup F^*$. The graph is defined by a random function $\tau : F^* \to F$ that maps each optimal facility to a local facility: this gives a bipartite graph with F^* vertices have out-degree one, and F vertices having no out-degree. In previous analyses, $\tau(f^*)$ was defined as the closest local facility to f^*, but in our analysis, we choose $\tau(f^*)$ randomly from the two closest local facilities to f^* in order to cover a larger neighborhood with good balance. Indeed, independently for every optimal facility f^*, we choose $\tau(f^*)$ from η_1 and η_2, where $\eta_1 = \eta_1(f^*)$ and $\eta_2 = \eta_2(f^*) \in F$ are the first and second closest local facilities to f^*. The probability of choosing η_i depends on the value of $\rho = \rho(f^*) := \frac{d(f^*,\eta_1)}{d(f^*,\eta_2)} \in [0, 1]$. When $\rho(f^*) \leq 3/4$, we choose $\tau(f^*) = \eta_1$ with probability 1; when $\rho(f^*) > 3/4$, we choose $\tau(f^*) = \eta_1$ with probability $(5/2 - 2\rho)$ and $\tau(f^*) = \eta_2$ with the remaining probability $(2\rho - 3/2)$.

Intuitively, $\tau(f^*)$ is the facility used as a fallback to serve clients of f^*'s cluster when their closest local facility is swapped out. More precisely, we design the swaps such that either f^* or $\tau(f^*)$ is open. To bound the reassignment cost to $\tau(f^*)$, we therefore must ensure that $\tau(f^*)$ is as close as possible to f^*. When $\rho(f^*)$ is small, there is therefore a huge incentive in choosing $\tau(f^*) = \eta_1$. However, when $\rho(f^*)$ is close to 1, there is no difference between η_1 or η_2. Our probability distribution is chosen such as to implement that intuition. It has been tuned experimentally: using our LP formulation, we were able to look for a choice of of τ that gives a good approximation guarantee while being simple enough to prove that guarantee.

This defines the graph G_0. We wish to generate swaps according to the connected components of G_0, i.e., every swap closes all the local facilities in a connected component and opens all the optimal facilities in the same connected component. However, such swaps may not be valid because 1) the size of a connected component may be much larger than p, and 2) there may be more optimal facilities in a connected component than local facilities (since every connected component of G_0 contains exactly one local facility). We solve these issues by two procedures: *degree reduction* and *balancing*.

Degree reduction. The size of a connected component of G_0 being too large is caused by local facilities with high in-degree. We solve the problem by removing all local facilities that could potentially have high in-degree from the graph. We call these the *heavy* local facilities. To keep the number of local facilities in the graph unchanged, we duplicate other local facilities, which we call *local surrogates*. We formally define heavy local facilities and local surrogates as follows. We first define $N(f^*) \subseteq \{\eta_1, \eta_2\}$ and call it the set of local neighbors of f^*. If $\rho(f^*) \leq 2/3$, we define $N(f^*) = \{\eta_1\}$; otherwise, we define $N(f^*) = \{\eta_1, \eta_2\}$. We choose $t_\mathsf{d} = \lceil 1/\varepsilon \rceil$ as the *degree threshold*. Now the heavy local facilities are as follows:

Definition 3.1 (heavy local facility)**.** *A local facility $f \in F$ is heavy if it is a local neighbor of more than $t_\mathsf{d} + 1$ optimal facilities.*

Note that $\tau(f^*)$ must be a local neighbor of f^* because $3/4 > 2/3$. Therefore, only heavy local facilities can have in-degree more than $t_\mathsf{d} + 1$ in G_0. For every heavy local facility, we choose a local surrogate uniformly at random from the *local candidates* defined as follows:

Definition 3.2 (local candidate)**.** *A local facility $f \in F$ is a* local candidate *if it is not heavy and every optimal facility in $\tau^{-1}(f)$ has a heavy local neighbor.*

Note that, unlike our definition of heavy local facilities, the definition of local candidates depends on the random function τ. The following claim (proved in Appendix E.1) shows that there are enough local candidates from which the heavy local facilities can choose:

Claim 3.3. *The number of local candidates is at least $t_\mathsf{d}/2$ times the number of heavy local facilities.*

We are ready to describe our degree reduction procedure:

1. Remove all the edges incident to heavy local facilities;
2. Replace each heavy local facility f by its local surrogate s, chosen uniformly at random *without replacement* from the local candidates. Hence, in the graph the vertex labeled f (and now having no in-edges due to step 1) is replaced by one labeled s. So a local surrogate appears twice now: the original copy of s, and a single isolated vertex as a surrogate for f.

Let G_1 denote the graph after degree reduction. Clearly, every local facility has degree at most $t_\mathsf{d} + 1$ in G_1, and thus every connected component has size at most $t_\mathsf{d} + 2$. The next claim follows directly from Claim 3.3:

Claim 3.4. *The constructed graph G_1 satisfies following properties:*

i. Heavy local facilities do not appear in G_1.
ii. Local facilities chosen as local surrogates appear twice: once as the original copy and once as an isolated vertex.
iii. Other local facilities and all optimal facilities appear once.
vi. Every optimal facility f^* points to the original copy of $\tau(f^*)$ unless $\tau(f^*)$ is heavy.
v. Any local facility is chosen as a local surrogate with probability at most $2/t_\mathsf{d}$, and only when it is a local candidate.

Balancing. Since a connected component of G_1 may contain more optimal facilities than local ones we combine connected components together to form groups with at least as many local facilities as optimal ones, using the following claim (proved in Appendix E.2):

Claim 3.5 (Balancing Procedure)**.** *Consider a universe $U = R \cup G$ of red points R and green points G, with $|G| = |R| + r$. Let the collection of sets $S_1, \ldots, S_N$ partition U, and let $|S_i| \leq x$ for all i. Moreover, let H be a graph on the vertices $[N]$ with maximum degree at most $\theta \leq r$. Lastly, $r \geq \Omega\left(\frac{x^5\theta^3}{\varepsilon}\right)$ for some $0 \leq \varepsilon \leq 1$. Then we can merge these sets together into new sets $T_0, \ldots, T_M$ such that*

(i) each T_j has size $|T_j| \le O(x^2)$,
(ii) $|T_j \cap R| \le |T_j \cap G|$,
(iii) if there is an edge $\{i, j\}$ for $i, j \in [N]$, then S_i is not merged with S_j, and
(iv) for all $i \ne j$, S_i is merged with S_j with probability at most ε.

Recall that our degree reduction step did not change the total number of local and optimal facilities, so there are still $r(\varepsilon)$ more local facilities than optimal facilities. We identify F^*, F with R, G in Claim 3.5 respectively, and define every S_i as the set of facilities in every connected component of G_1. Note that $|S_i| \le t_{\mathsf{d}} + 2$. S_i and S_j are connected by an edge in H if and only if they contain two copies of the same local facility: one contains the original copy of a local facility and the other contains a new copy created as a local surrogate. The maximum degree of H is at most 1 due to Claim 3.4 and the fact that there is at most one local facility in each connected component. Since $r(\varepsilon) \ge \Omega((t_{\mathsf{d}} + 2)^5/\varepsilon)$, we use Claim 3.5 to combine components of G_1 into balanced groups, where every group contains at most $O((t_{\mathsf{d}} + 2)^2) \le p(\varepsilon)$ facilities. Every group thus defines a valid swap, and we define $\mathcal{P}$ as the set of these swaps. Figure 3.2 shows an example of the simple swap set $\mathcal{P}$ we generate.

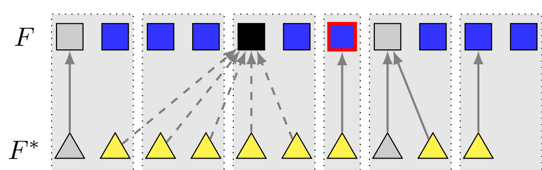


Figure 3.2: An example of a simple swap set $\mathcal{P}$. Edges correspond to $\tau(f^*)$'s. Dashed edges are removed. The black facility is a local surrogate replacing a heavy local facility. The original copy of the black surrogate is the facility with red boundary, chosen randomly from the local candidates (blue), assuming every yellow optimal facility has a heavy local neighbor. Gray boxes correspond to the swaps in $\mathcal{P}$.

3.2 Generating the Important Tree Swaps

Again, we start by constructing a directed graph G_0. Unlike simple swaps where only optimal facilities have out-edges, tree swaps require every local facility to also have an out-edge to an optimal facility in G_0. In particular, every local facility f has an out-edge to $\pi(f)$, the optimal facility closest to it. Every optimal facility still has an out-edge to $\tau(f^*) \in \{\eta_1, \eta_2\}$, but we pick $\tau(f^*)$ from a different distribution: if $\rho(f^*) \le 2/3$, then $\tau(f^*) = \eta_1$ with probability 1; else $\tau(f^*) = \eta_1$ with probability $1/2$ and $\tau(f^*) = \eta_2$ otherwise.

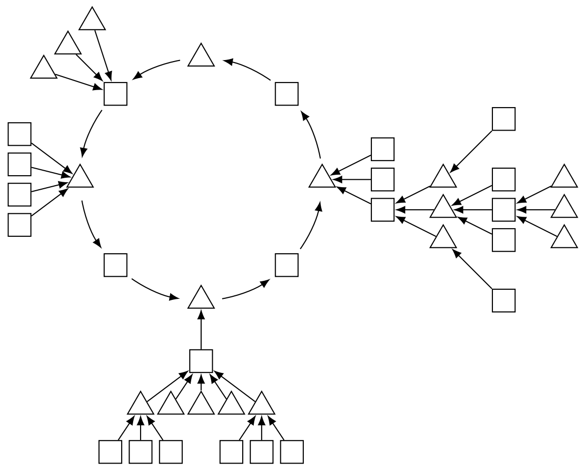


Figure 3.3: Every connected component of G_0 is a 1-tree. Local facilities are represented by squares, while optimal facilities are represented by triangles.

Since every vertex of G_0 has out-degree one, G_0 is a *1-forest*, with every connected component being a *1-tree*, i.e., a directed tree with a directed cycle as its root (see Figure 3.3), hence the name *tree swaps*. Having constructed G_0, we generate the tree swap set $\mathcal{P}$ by three procedures: degree reduction, edge deletion, and balancing. The balancing step remains essentially the same as in simple swaps, but the degree reduction step requires a new ingredient to deal with optimal facilities with high in-degree, which did not exist in the simple swaps case. The edge deletion step is also unique to tree swaps. Next, we describe these three steps in detail.

Degree reduction. We first modify G_0 so that every vertex has in-degree bounded by $t_\mathsf{d}+1$. In the same way as simple swaps, we can remove local facilities with high in-degree by removing heavy local facilities, but we need an additional procedure to deal with *heavy optimal facilities* with high in-degree. Specifically, we say f^* is a heavy optimal facility if it has in-degree more than t_d after heavy local facilities are removed, in other words, $|\pi^{-1}(f^*)\backslash\{\text{heavy local facilities}\}| > t_\mathsf{d}$. For such a heavy optimal facility f^* with in-degree s, we partition its children into $\lceil s/t_\mathsf{d}\rceil$ groups. Every group, except sometimes the last one, contains exactly t_d children. We make sure that the first group contains the t_d closest children to f^*. For each group other than the first one, we create a new copy of f^* and change the out-edges from the children in the group to point to the new copy of f^*. The new copy of f^* has an out-edge pointing to a new copy of a local facility f chosen uniformly at random from the previous group. We call the new copy of f an *optimal surrogate*. They are needed to keep the difference between the number of local and optimal facilities unchanged. We also add an out-edge from f pointing back to the new copy of f^*, as illustrated in Figure 3.4.

In summary, the degree reduction procedure for tree swaps consists of the following steps:

1. Remove edges incident to all heavy local facilities;
2. Replace every heavy local facility by its local surrogate, chosen uniformly at random without replacement from the local candidates;
3. Deal with heavy optimal facilities as above;
4. Add self-loops to vertices with no out-edge (due to step 1) to retain the 1-forest structure (this facilitates a cleaner presentation of our next procedure: edge deletion).

Let G_1 denote the graph after degree reduction. G_1 is still a 1-forest, and every vertex in G_1 now has in-degree at most $t_\mathsf{d}+1$. Moreover, the following claim is apparent (by observing that Claim 3.3

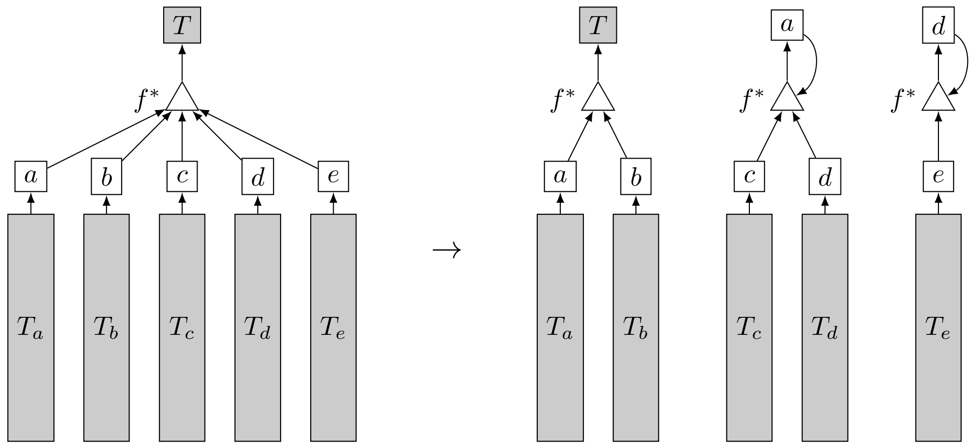


Figure 3.4: The figure shows the decomposition of high in-degree optimal facility f^* for $t_{\mathsf{d}} = 2$. Shaded rectangular boxes correspond to part of the original tree that does not change. Since the degree of f^* is 5 $\geq t_{\mathsf{d}}$, we create $\lceil 5/t_{\mathsf{d}} \rceil$ trees. The first tree stays in the original tree. Each remaining tree gets a f^*'s child chosen uniformly at random from the previous tree. f^* gets open 2 extra times, but we also close a and d to balance the number of opening and closure. In this example a and d are chosen as optimal surrogates. And a and b are two closest children to f^* among $\{a, b, c, d, e\}$.

still holds in the tree swaps case because its proof is completely independent of the distribution of $\tau(f^*)$):

Claim 3.6. *Constructed graph G_1 follows following properties:*

i. Every optimal facility appears in G_1 at least once.
ii. Every local facility appears in G_1 at most three times: once as the original copy, once as a local surrogate, and once as an optimal surrogate.
iii. Heavy local facilities do not appear in G_1.
iv. No two copies of the same facility appear in the same connected component.
v. The original copy of any optimal facility f^* points to the original copy of $\tau(f^*)$, unless $\tau(f^*)$ is heavy.
vi. The original copy of any local facility f points to $\pi(f)$, although it might be a new copy of $\pi(f)$.
vii. Any local facility is chosen as a local surrogate with probability at most $^2/_{t_{\mathsf{d}}}$, and as an optimal surrogate with probability at most $^1/_{t_{\mathsf{d}}}$.
viii. Every local surrogate is a local candidate.

The degree-reduction step ensures that vertices in G_1 have bounded in-degree, but a connected component of G_1 could still have large size (it could have large height or contain a long cycle). We deal with this problem in our next procedure: *edge deletion*.

Edge deletion. Next, we remove edges from G_1 to ensure that every connected component in the resulting graph is a tree of height at most $t_{\mathsf{h}} - 1$, where we choose the height threshold t_{h} uniformly at random from $2\lceil ^1/_\varepsilon \rceil, 2\lceil ^1/_\varepsilon \rceil^2, \cdots, 2\lceil ^1/_\varepsilon \rceil^{\lceil ^1/_\varepsilon \rceil}$. Specifically, for each connected component T of G_1, if the root cycle has length less than t_{h}, we insert dummy vertices into the cycle to make the length exactly t_{h}. Then we pick a vertex r in the root cycle uniformly at random, and delete the out-edge from r. This makes T a directed tree rooted at r. We then delete edges on the $a \cdot t_{\mathsf{h}}$-th levels for all $a \in \mathbb{N}$. See Figure 3.5 for an example.

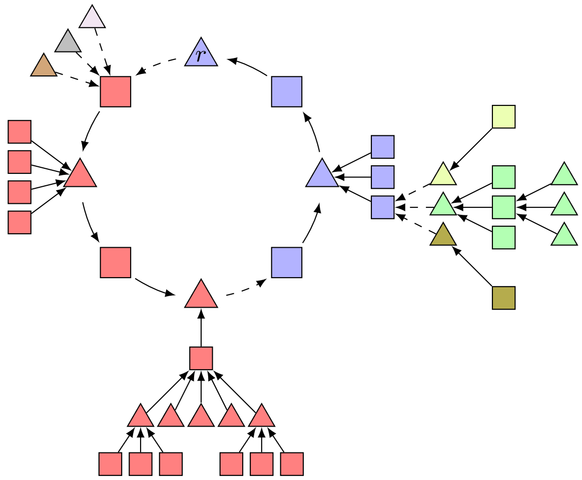


Figure 3.5: Example for $t_{\mathsf{h}} = 4$. Nodes with the same color correspond to nodes in the same connected component after edge deletion. We start from r (randomly chosen), and repeatedly cut edges on $a \cdot t_{\mathsf{h}}$ steps away from r (dashed edges).

Let G_2 be the graph after the edge deletion step. It is clear that every connected component of G_2 is a directed tree with height at most $t_{\mathsf{h}} - 1$, possibly containing some dummy vertices. Moreover, every vertex v has in-degree at most $t_{\mathsf{d}} + 1$ due to the degree reduction procedure. Therefore, the number of vertices in every connected component of G_2 is at most $(t_{\mathsf{d}}+1)^{t_{\mathsf{h}}}$. Moreover, we have the following claim for every connected component T of G_1, which is apparent from our edge deletion procedure:

Claim 3.7. *After dummy vertices are added into T, the edge out of vertex $v \in T$ is deleted if and only if the (unique) simple path from v to r has length divisible by t_{h}.*

If the cycle length of T is at most t_{h}, vertices on the cycle are still connected after edge deletion. Indeed, we delete only one edge in the cycle in this case. Therefore, after edge deletion, we ignore all the dummy vertices and still consider all the edges on the original cycle as not deleted by convention. This doesn't change the (non-dummy) vertices in every connected component of G_2, and thus doesn't change $\mathcal{P}$ we eventually generate. With this convention, we have the following corollary of Claim 3.7:

Corollary 3.8. *Any edge in G_1 is deleted with probability at most $^2/_{t_{\mathsf{h}}}$. Moreover, if the cycle length is at most t_{h}, edges on the cycle are never deleted.*

Proof. The second part is assumed by our convention. We thus assume henceforth that the edge is not on the cycle, or the cycle length is more than t_{h}. Suppose the edge is the out-edge of vertex v. By Claim 3.7, the edge is deleted if and only if the simple path p^* from v to r has length divisible by t_{h}. Suppose the cycle length after dummy vertices are added to it is $\ell \geq t_{\mathsf{h}}$, and let $\ell = ut_{\mathsf{h}} + w$ for $u, w \in \mathbb{Z}$ with $0 \leq w < t_{\mathsf{h}}$. There are at most $u+1$ choices of r such that p^* has length divisible by t_{h}. Therefore, the edge is deleted with probability at most $(u+1)/\ell = u/\ell + 1/\ell \leq 1/t_{\mathsf{h}} + 1/t_{\mathsf{h}} = 2/t_{\mathsf{h}}$. $\square$

After edge deletion, each connected component of G_2 contains at most $(t_{\mathsf{d}} + 1)^{t_{\mathsf{h}}} \leq p(\varepsilon)$ vertices. However, the number of local and optimal facilities in the component may not match (e.g., the blue tree containing r in Figure 3.5 has three extra local facilities, whereas the rightmost tree has one extra optimal facility). We fix this in the same way as in the simple swaps case using the balancing procedure.

Balancing. The balancing procedure is essentially the same as in the simple swaps case, based on Claim 3.5 again. The only difference is that the size of every connected component is now much larger ($(t_{\mathsf{d}}+1)^{t_{\mathsf{h}}}$), and the maximum degree of H is also much larger. Since optimal facilities may now have new copies, we may combine two connected components each containing a copy of the same optimal facility in the balancing step; this is fine because it only decreases the number of optimal facilities in a swap. However, we still need to make sure that no two copies of the same local facility are combined together, again by adding edges into H between connected components containing copies of the same local facility. Since a local facility can have at most 3 copies by Claim 3.6, the maximum degree of H is at most $2(t_{\mathsf{d}}+1)^{t_{\mathsf{h}}}$. Since we kept the number of extra local facilities unchanged, it's still $r(\varepsilon) \geq \Omega(((t_{\mathsf{d}}+1)^{t_{\mathsf{h}}})^5(2(t_{\mathsf{d}}+1)^{t_{\mathsf{h}}})^3/\varepsilon)$, so Claim 3.5 gives balanced groups each containing at most $O((t_{\mathsf{d}}+1)^{2t_{\mathsf{h}}}) \leq p(\varepsilon)$ facilities. Every group thus defines a valid swap, and we define $\mathcal{P}$ as the set of these tree swaps.

4 Client Types

We now classify the clients into a small number of types (based on how the client connects to facilities in the local and global solutions). The classification allows us to give a *client-by-client* analysis instead of a *swap-by-swap* analysis used in prior works. We make this change in perspective because the potential Φ depends on the two closest facilities, and so we need a better handle on the local neighborhood of a client to bound the reassignment costs when closing one of the close facilities.

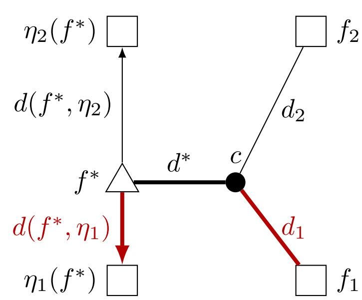
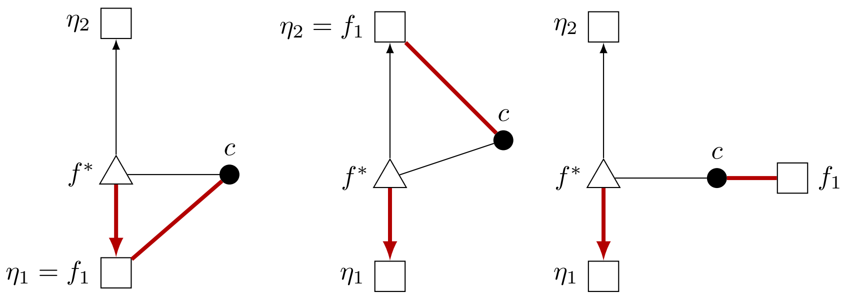


Figure 4.6: The squares are local facilities, triangles are optimal facilities, circle are clients. The thick red edge out of c goes to its closest local facility f_1; the thick red edge out of f^* goes to its closest local facility $\eta_1(f^*)$.

For a client c, recall that $f_1(c)$ and $f_2(c)$ are the closest and second-closest local facilities; we say f_1 and f_2 when there is no ambiguity. Figure 4.6 shows a picture of a generic client c and its related facilities.

We partition the set of clients into types based on the relationships between their local and optimal facilities, as follows. The *far* clients are those for which $d_2 \geq \alpha d_1$, and hence the potential just depends on the closest facility (f_1); the other kinds of clients are called *close*, for which both f_1 and f_2 are relevant.

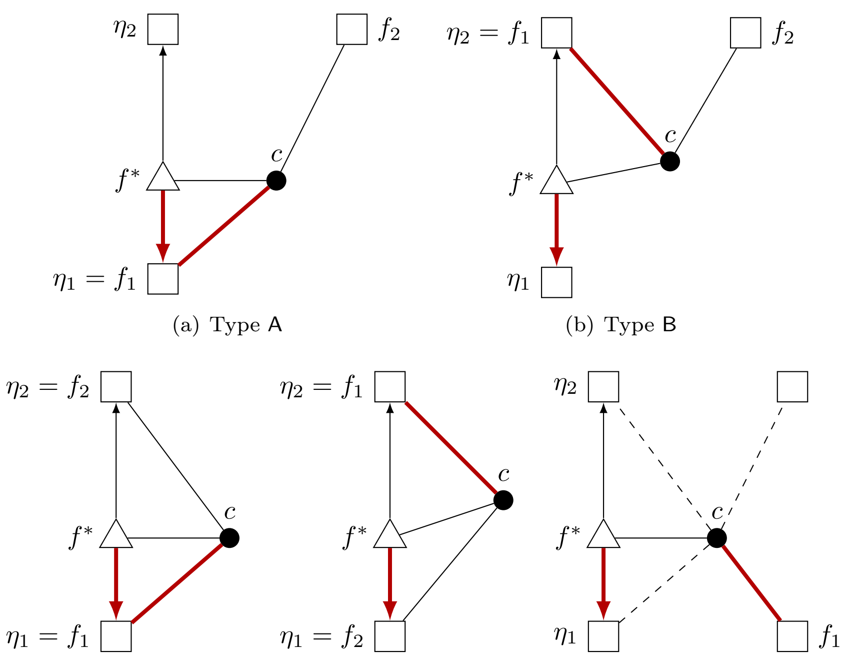
- **Far case** (where $d_2 \geq \alpha d_1$). Note that f_2 does not play any role in the far case, so the clients are classified according to how f_1 and f^* are related.
  - Type A: $\eta_1(f^*) = f_1$.
  - Type B: $\eta_2(f^*) = f_1$.
  - Type E: $f_1 \not\in \{\eta_1(f^*), \eta_2(f^*)\}$.



(a) Type A (b) Type B (c) Type E

Figure 4.7: Far Case

- *Close case* (where $d_2 \leq \alpha d_1$); now clients are classified according to how f_1, f_2 and f^* are related.
  - Type A: $\eta_1(f^*) = f_1$ and $\eta_2(f^*) \neq f_2$.
  - Type B: $\eta_1(f^*) \neq f_2$ and $\eta_2(f^*) = f_1$.
  - Type C: $\eta_1(f^*) = f_1$ and $\eta_2(f^*) = f_2$.
  - Type D: $\eta_1(f^*) = f_2$ and $\eta_2(f^*) = f_1$.
  - Type E: $f_1 \notin \{\eta_1(f^*), \eta_2(f^*)\}$.



(c) Type C (d) Type D (e) Type E

Figure 4.8: Close Case. For Type E*, the client's* f_2 *can be any one of dashed edges.*

5 Amenable and Defiant Events

Not all swaps are easy to argue about. Having fixed a client c, we define the *amenable event* and *defiant event* for this client—the former captures the case where the swaps in $\mathcal{P}$ are easy to reason about, and the latter the case where we throw up our hands and use a crude bound on the potential change. Thankfully, the latter happens very rarely, so the loss is small.

Recall that $f_1(c), f_2(c)$ are the two closest local facilities to c. Let $f^* = f^*(c)$ be the optimal facility that c is assigned to; then $\eta_1(f^*), \eta_2(f^*)$ are the two closest local facilties to f^*. We define the amenable and defiant events as follows:

Definition 5.1 (Amenable/Defiant)**.** *The* defiant event $\mathcal{D}$ *for a client c of type* A*,* B *or* E *is the union of the following events:*

(i) f_1, f_2, or $\tau(f^*)$ is chosen as a local or optimal surrogate in the degree reduction step;
(ii) $\mathcal{P}$ is a tree swaps set, and the out-edge from the original copy of f^*, f_1 or f_2 is deleted in the edge deletion step.
(iii) $\mathcal{P}$ is a simple swaps set, and two connected components each containing a facility in $\{f^*\} \cup \{f_1, f_2\} \cup \{\eta_1, \eta_2\}$ are grouped together in the balancing step.

The amenable event $\mathcal{A}$ *is the complement of $\mathcal{D}$.*

For type C and D clients, we enlarge the defiant event slightly to include $g^* := \pi(f_1)$ and

$$g := \mathrm{argmin}_{h \in F \setminus \{f_1, f_2\}} d(h, g^*)$$

as follows:

Definition 5.2 (Amenable/Defiant for type C and D)**.** *The* defiant event $\mathcal{D}$ *for a client c of type* C *or* D *is the union of the events (i), (ii), (iii) in Definition 5.1 and the following events:*

(i') $\tau(g^*)$ is chosen as a local or optimal surrogate in the degree reduction step;
(ii') $\mathcal{P}$ is a tree swaps set, and the out-edge from the original copy of g^* is deleted in the edge deletion step.
(iii') $\mathcal{P}$ is a simple swaps set, and two connected components each containing a facility in $\{f_1, g\}$ are grouped together in the balancing step.

The amenable event $\mathcal{A}$ *is the complement of $\mathcal{D}$.*

The events $\mathcal{A}$ and $\mathcal{D}$ depend on the client c, but we choose to omit c in our notation because we will always focus on a fixed client c in our proof. We now turn to proving Lemma 2.2 on the potential change due to defiant events. The approach is simple: we first show a crude upper bound that holds for all swap sets $\mathcal{P}$ that we generate, and then show that the probability of the defiant event is small enough so that we can afford to apply this crude upper bound.

Claim 5.3. *There is an absolute constant $\gamma > 0$ such that for any client c, and any swap set $\mathcal{P}$ that we generate, we have $\sum_{(P,Q) \in \mathcal{P}} \delta_{(P,Q)}(c) \le \gamma(d^*(c) + d_1(c))$.*

Claim 5.4. $\Pr[\mathcal{D}] \le O(\varepsilon)$ *for all clients c.*

The proof of Claim 5.4 follows from Claims 3.4 to 3.6, Corollary 3.8, and a trivial union bound. We defer the proof of Claim 5.3 to Appendix E.3. The two claims above imply Lemma 2.2, and hence control the effect of the defiant events. We focus next on the amenable events and the proof of Lemma 2.3.

6 The Potential Change due to Amenable Events

Having bounded the potential change due to defiant events, we now turn to bounding the potential change due to amenable events. Let us recall the claim we want to prove:

Lemma 2.3 (Amenable Case)**.** *For the distribution over valid swap sets from Lemma 2.2, for any* $c \in \mathcal{C}$*,*

$$\mathbb{E}\Big[\mathbb{1}_{\mathcal{A}} \sum_{(P,Q)\in\mathcal{P}} \delta_{(P,Q)}(c)\Big] \leq 2.5203\, d^*(c) - 0.8888\, d_1(c) + O(\varepsilon)\,(d^*(c) + d_1(c)). \tag{2.5}$$

This section gives an explicit proof that can be verified by hand. In §7 we show how to generate a much larger set of valid inequalities. Solving the resulting linear program gives improved bounds, but these are more tedious to verify manually.

6.1 Implications of Amenability

Claim 6.1 (Implications of amenability)**.** *For any client, swap sets* $\mathcal{P}$ *generated on the amenable event* $\mathcal{A}$ *have the following properties:*

- *(i) Any local facility* $f \in \{f_1, f_2\}$ *is closed in at most one swap in* $\mathcal{P}$*;*
- *(ii) Any swap in* $\mathcal{P}$ *closing* $\tau(f^*)$ *must open the original copy of* f^**;*
- *(Tii) If* $\mathcal{P}$ *is a tree swap set, any swap in* $\mathcal{P}$ *closing* $f \in \{f_1, f_2\}$ *must open* $\pi(f)$*;*
- *(Siii) If* $\mathcal{P}$ *is a simple swap set, no swap in* $\mathcal{P}$ *closes two local facilities in* $\{f_1, f_2\} \cup \{\eta_1, \eta_2\}$ *simultaneously;*
- *(Siv) If* $\mathcal{P}$ *is a simple swap set, any swap in* $\mathcal{P}$ *closing a local facility in* $\{f_1, f_2\}\backslash\{\tau(f^*)\}$ *does not open* f^**.*

For clients of type C *or* D*, we additionally have the following: (recall that we defined* g^* *as* $\pi(f_1)$*, and* g *as the local facility closest to* g^* *other than* f_1 *and* f_2*):*

- *(ii') Any swap in* $\mathcal{P}$ *closing* $\tau(g^*)$ *must open the original copy of* g^**;*
- *(Siii') If* $\mathcal{P}$ *is a simple swap set, no swap in* $\mathcal{P}$ *closes both* f_1 *and* g*.*

Proof of Claim 6.1. Recall that the amenable event $\mathcal{A}$ is the complement of the defiant event $\mathcal{D}$, defined in Definition 5.1.

Implication (i) follows from item (i) of Definition 5.1 directly.

Implication (ii) follows from items (i) and (ii) of Definition 5.1. Without loss of generality, we assume $\tau(f^*)$ is not heavy, since heavy local facilities are never closed. On the amenable event, $\tau(f^*)$ is closed only as its original copy, by item (i) of Definition 5.1. The edge to $\tau(f^*)$ from the original copy of f^* is never deleted by item (ii) of Definition 5.1, so the original copies of f^* and $\tau(f^*)$ must be in the same swap.

Implication (Tii) also follows from items (i) and (ii) of Definition 5.1, for a similar reason. Again, assume without loss of generality that neither f_1 nor f_2 is heavy. On the amenable event, f_1 and f_2 are closed only as their original copies by item (i) of Definition 5.1, and the edges $f_i \to \pi(f_i)$ are never deleted by item (ii).

Implications (Siii) and (Siv) both follow from item (iii) of Definition 5.1. When we generate the simple swap set, every connected component of the graph G_1 contains at most one local facility, and thus different facilities in $\{f_1, f_2\} \cup \{\eta_1, \eta_2\}$ must be in different connected components, which are not combined in the balancing step due to item (iii) of Definition 5.1. This proves implication

(Siii). Moreover, the connected component of f^* doesn't contain any local facility other than $\tau(f^*) \in \{\eta_1, \eta_2\}$. This proves implication (Siv).

(ii') and (Siii') can be proved in the same way as (ii) and (Siii) using Definition 5.2. $\square$

6.2 Notation and Useful Inequalities

Let $\Delta_{\mathcal{E}}(c)$ denote the expected potential change on client c restricted to some generic event $\mathcal{E}$:

$$\Delta_{\mathcal{E}}(c) := \mathbb{E}\Big[\mathbb{1}_{\mathcal{E}} \sum_{(P,Q)\in\mathcal{P}} \delta_{(P,Q)}(c)\Big].$$

Our goal in Lemma 2.3 is thus to upper bound $\Delta_{\mathcal{A}}(c)$ for the amenable event $\mathcal{A}$. In our proof, we consider sub-events $\mathcal{E}$ of $\mathcal{A}$, and prove worst-case upper-bounds for the potential change restricted to each sub-event $\mathcal{E}$. Formally, given a suitable partition $\mathcal{A} = \mathcal{E}_1 \cup \cdots \cup \mathcal{E}_t$, we define $\delta_{\mathcal{E}}(c) := \sum_{(P,Q)\in\mathcal{P}} \delta_{(P,Q)}(c)$ to be the worst-case (maximum) value for each event $\mathcal{E}$, and then use:

$$\Delta_{\mathcal{A}}(c) = \sum_{i=1}^{t} \Delta_{\mathcal{E}_i}(c) \le \sum_{i=1}^{t} \Pr[\mathcal{E}_i]\ \delta_{\mathcal{E}_i}(c). \tag{6.6}$$

For technical reasons, it is more convenient to assume $\delta_{\mathcal{E}}(c)$ is no smaller than, say, $-10d_1(c)$. We thus re-define $\delta_{\mathcal{E}}(c)$ as $-10d_1(c)$ when $\delta_{\mathcal{E}}(c) < -10d_1(c)$. This doesn't affect our analysis, as all our upper bounds for $\delta_{\mathcal{E}}(c)$ are larger than $-10d_1(c)$. Also, Claim 5.3 implies that $\delta_{\mathcal{E}}(c) \le O(d^* + d_1)$.

To apply (6.6), we need to understand $\Pr[\mathcal{E}]$ and $\delta_{\mathcal{E}}(c)$ for the following events (and their intersections): the *amenable event* $\mathcal{A}$ and its complement *defiant event* $\mathcal{D}$, the *simple event* $\mathcal{S}$ and its complement *tree event* $\mathcal{T}$. The simple event $\mathcal{S}$ is further partitioned into $\mathcal{S}_1$ and $\mathcal{S}_2$, and the tree event $\mathcal{T}$ is partitioned into $\mathcal{T}_1$ and $\mathcal{T}_2$, representing whether f^* points to η_1 or η_2. These events are defined for a fixed client c, and we omit c in our notations for brevity.

Recall that f^* is the optimal facility closest to c, and $\rho = \rho(f^*) := \frac{d(f^*,\eta_1(f^*))}{d(f^*,\eta_2(f^*))}$. To generate the set $\mathcal{P}$ of important swaps, we choose $\tau(f^*)$ from different distributions depending on the value of $\rho(f^*)$, and thus the probability of the events $\mathcal{S}_1, \mathcal{S}_2, \mathcal{T}_1, \mathcal{T}_2$ depends on $\rho(f^*)$ as follows:

Ratio-types	$\Pr[\mathcal{S}_1]$	$\Pr[\mathcal{S}_2]$	$\Pr[\mathcal{T}_1]$	$\Pr[\mathcal{T}_2]$
$0 \le \rho(f^*) \le 2/3$	$1/2$	$\cdot$	$1/2$	$\cdot$
$2/3 < \rho(f^*) \le 3/4$	$1/2$	$\cdot$	$1/4$	$1/4$
$3/4 < \rho(f^*) \le 1$	$5/4 - \rho$	$\rho - 3/4$	$1/4$	$1/4$

Table 1: Probability distribution for each ratio-type.

Since $\Pr[\mathcal{D}] = O(\varepsilon)$ due to Claim 5.4, the probability of any event $\mathcal{E} \cap \mathcal{A}$ is at least $\Pr[\mathcal{E}] - O(\varepsilon)$.

Bounding the worst-case change $\delta_{\mathcal{E}}(c)$. We fix an arbitrary swap set $\mathcal{P}$ generated under event $\mathcal{E}$, and analyze the effect of each swap in $\mathcal{P}$. Let $\langle\langle f^* \rangle\rangle$ denote the swap in $\mathcal{P}$ that opens f^*; such a swap always exists. There may be multiple such swaps in $\mathcal{P}$ when we perform tree swaps, in which case we let $\langle\langle f^* \rangle\rangle$ be the swap that opens the original copy of f^*. For a local facility $f \in \{f_1, f_2\}$, let $\langle\langle \neg f \rangle\rangle$ denote the swap in $\mathcal{P}$ that closes f. By implication (i) of amenability, there is at most one such swap as long as $\mathcal{E}$ is a sub-event of the amenable event $\mathcal{A}$. When there is no swap closing f (which happens when f is a heavy facility), we are often in a better situation because our bound for $\delta_{\langle\langle \neg f \rangle\rangle}(c)$ is often non-negative, so we will mostly focus on the case where $\langle\langle \neg f \rangle\rangle$ does exist.

Before we begin giving bounds for the various client types, let us record in Table 2 some inequalities we will frequently use. Recall that $\eta_1(f^*)$ and $\eta_2(f^*)$ are the closest and second-closest local facilities to f^*, and $\pi(f)$ is the closest optimal facility to f. These inequalities are proven in Appendix C.

Bound		Conditions (if any)
$d_2 \leq 2d^* + d_1$	(6.7)	$\eta_1(f^*) \neq f_1$
$d(c, \pi(f_1)) \leq 2d_1 + d^*$	(6.8)	
$\max\{d(c, \eta_1(f^*)), d(c, \eta_2(f^*))\} \leq 2d^* + d_1$	(6.9)	$\eta_1(f^*) \neq f_1$
$\min(d^*, d_1) + \beta \max(d^*, d_1) \leq (1-\beta)\, d^* + 2\beta\, d_1$	(6.10)	

Table 2: Useful Inequalities

6.3 Bounds for Clients of Type E

We now give an upper bound for the expected potential change $\Delta_{\mathcal{A}}(c)$ for any client c of type E. We give the entire proofs here; for clients of other types we will defer the proofs to the appendices.

Lemma 6.2. *For any client c of type E, we have*

$$\Delta_{\mathcal{A}}(c) \leq 2.5\, d^*(c) - 0.9\, d_1(c) + O(\varepsilon)(d^* + d_1).$$

In our proof, we partition the amenable event $\mathcal{A} = (\mathcal{S} \cap \mathcal{A}) \cup (\mathcal{T} \cap \mathcal{A})$ depending on whether we have a simple swap or a tree swap, and then bound $\Delta_{\mathcal{A}}(c)$ by

$$\begin{aligned} \Delta_{\mathcal{A}}(c) &\leq \Pr[\mathcal{S} \cap \mathcal{A}] \cdot \delta_{\mathcal{S}\cap\mathcal{A}}(c) + \Pr[\mathcal{T} \cap \mathcal{A}] \cdot \delta_{\mathcal{T}\cap\mathcal{A}}(c) \\ &\leq \Pr[\mathcal{S}] \cdot \delta_{\mathcal{S}\cap\mathcal{A}}(c) + \Pr[\mathcal{T}] \cdot \delta_{\mathcal{T}\cap\mathcal{A}}(c) + O(\varepsilon)(d^* + d_1). \end{aligned} \tag{6.11}$$

The second inequality is implied by Claim 5.4 and our assumption that $\delta_{\mathcal{E}}(c) \geq -10d_1$. To use (6.11) we give upper bounds for $\delta_{\mathcal{S}\cap\mathcal{A}}(c)$ and $\delta_{\mathcal{T}\cap\mathcal{A}}(c)$ for clients of both subtypes (close and far) in the next subsections. In other words, we pick an arbitrary swap set $\mathcal{P}$ generated under these events, and bound the potential change for client c due to the swaps in $\mathcal{P}$.

6.3.1 Far Clients of Type E: $d_2(c) \geq \alpha\, d_1(c)$

Simple Swaps. We fix a "far" client c and an arbitrary swap set $\mathcal{P}$ generated conditioned on the event $\mathcal{S} \cap \mathcal{A}$ for this client, and bound the sum $\sum_{(P,Q)\in\mathcal{P}} \delta_{(P,Q)}(c)$.

- Given the swap $\langle\langle f^* \rangle\rangle \in \mathcal{P}$ (which is not $\langle\langle \neg f_1 \rangle\rangle$ by implication (Siv) of amenability), c has an additional option of going to f^*, giving

$$\delta_{\langle\langle f^* \rangle\rangle}(c) \leq (d_1 + \beta\, d^*) - (1 + \alpha\beta)\, d_1.$$

- Next, by implication (i) of amenability, the set $\mathcal{P}$ contains at most one swap $\langle\langle \neg f_1 \rangle\rangle$. If $\langle\langle \neg f_1 \rangle\rangle$ does exist, both η_1 and η_2 are open (by implication (Siii) of amenability), and both at distance $\leq 2d^* + d_1$ from c. Therefore,

$$\delta_{\langle\langle \neg f_1 \rangle\rangle}(c) \leq (1+\beta)(2d^* + d_1) - (1 + \alpha\beta)\, d_1.$$

  This quantity is non-negative: since c has type E, $\eta_1 \neq f$ and also $d(c, \eta_1) \leq 2d^* + d_1$. But c is a far client, then $d(c, \eta_1) \geq \alpha d_1$. Putting the two together:

$$(1+\beta)(2d^* + d_1) - (1 + \alpha\beta)\, d_1 \geq d_1 + \beta(2d^* + d_1) - (1 + \alpha\beta)\, d_1 \geq 0.$$

- Finally, all other swaps in $\mathcal{P}$ leave f_1 open, and thus they cannot increase the potential for c.

Combining these, when the swap $\langle\langle \neg f_1 \rangle\rangle$ exists,

$$\sum_{(P,Q)\in\mathcal{P}} \delta_{(P,Q)}(c) \le \delta_{\langle\langle f^* \rangle\rangle}(c) + \delta_{\langle\langle \neg f_1 \rangle\rangle}(c) \le (2+3\beta)\, d^* - (2\alpha\beta - \beta)\, d_1. \tag{6.12}$$

In case $\langle\langle \neg f_1 \rangle\rangle$ does not exist, (6.12) still holds since our bound for $\delta_{\langle\langle \neg f_1 \rangle\rangle}$ is non-negative. Since $\mathcal{P}$ was a generic swap set conditioned on being amenable,

$$\delta_{\mathcal{S}\cap\mathcal{A}}(c) \le \boxed{(2+3\beta)\, d^* - (2\alpha\beta - \beta)\, d_1}.$$

Tree Swaps. We now turn to *tree swaps*, and fix an arbitrary swap set $\mathcal{P}$ generated on the event $\mathcal{T}\cap\mathcal{A}$. Again, $\mathcal{P}$ contains at most one swap $\langle\langle \neg f_1 \rangle\rangle$ that closes f_1, by amenability. We first consider the case where $\langle\langle \neg f_1 \rangle\rangle$ exists and is the same as $\langle\langle f^* \rangle\rangle$. In this case, all other swaps in $\mathcal{P}$ have non-positive potential changes, so

$$\sum_{(P,Q)\in\mathcal{P}} \delta_{(P,Q)}(c) \le \delta_{\langle\langle \neg f_1 \rangle\rangle}(c) \le (1+\alpha\beta)d^* - (1+\alpha\beta)d_1. \tag{6.13}$$

Next, consider the case where $\langle\langle \neg f_1 \rangle\rangle \ne \langle\langle f^* \rangle\rangle$. On swap $\langle\langle f^* \rangle\rangle$, client c can go to both d^* and d_1. On swap $\langle\langle \neg f_1 \rangle\rangle$, c can go to $\pi(f_1)$ at distance $\le 2d_1 + d^*$, and also to $\tau(f^*) \in \{\eta_1, \eta_2\}$ at distance $\le 2d^* + d_1$. Both these facilities $\pi(f_1)$ and $\tau(f^*)$ must be open after the swap $\langle\langle \neg f_1 \rangle\rangle$ due to implications (ii) and (Sii) of amenabilityx. All other swaps in $\mathcal{P}$ have non-positive potential changes, so

$$\begin{aligned}
\sum_{(P,Q)\in\mathcal{P}} \delta_{(P,Q)}(c) &\le \delta_{\langle\langle f^* \rangle\rangle}(c) + \delta_{\langle\langle \neg f_1 \rangle\rangle}(c) \\
(\delta_{\langle\langle f^* \rangle\rangle}) \qquad &\le d_1 + \beta\, d^* - (1+\alpha\beta)\, d_1 \\
(\delta_{\langle\langle \neg f_1 \rangle\rangle}) \qquad &\quad + (2d^* + d_1) + \beta(2d_1 + d^*) - (1+\alpha\beta)\, d_1 \\
(6.14) \qquad &= (2+2\beta)\, d^* - (2\alpha\beta - 2\beta)\, d_1.
\end{aligned}$$

In the case where $\langle\langle \neg f_1 \rangle\rangle$ doesn't exist, (6.14) still holds, because our bound for $\delta_{\langle\langle \neg f_1 \rangle\rangle}(c)$ is non-negative. By our choice of $\alpha = 3$ and $\beta = 1/5$, (6.13) is dominated by (6.14). Since $\mathcal{P}$ is a generic swap set,

$$\delta_{\mathcal{T}\cap\mathcal{A}}(c) \le \boxed{(2+2\beta)\, d^* - (2\alpha\beta - 2\beta)\, d_1}.$$

Summarizing the simple swaps case and the tree swaps case, we have

$$\begin{aligned}
\delta_{\mathcal{S}\cap\mathcal{A}}(c) &\le (2+3\beta)\, d^* - (2\alpha\beta - \beta)\, d_1 \le 2.6\, d^* - d_1, \\
\delta_{\mathcal{T}\cap\mathcal{A}}(c) &\le (2+2\beta)\, d^* - (2\alpha\beta - 2\beta)\, d_1 \le 2.4\, d^* - 0.8\, d_1.
\end{aligned}$$

Now substituting into (6.11), we get a bound for all type E far clients c:

$$\begin{aligned}
\Delta_{\mathcal{A}}(c) &\le 1/2 \cdot \delta_{\mathcal{S}\cap\mathcal{A}}(c) + 1/2 \cdot \delta_{\mathcal{T}\cap\mathcal{A}}(c) + O(\varepsilon)(d^* + d_1) \\
&\le \boxed{2.5\, d^* - 0.9\, d_1} + O(\varepsilon)(d^* + d_1).
\end{aligned} \tag{6.15}$$

This proves Lemma 6.2 for far clients of type E. The proof for all other types of clients will have a similar structure: we will identify which swaps affect client c, then we sum up the inequalities with the right probabilities. In some cases we will need to look at cases depending on ρ.

6.3.2 Close Clients of Type E: $d_2(c) \geq \alpha\, d_1(c)$

Simple swaps. Now we consider the case of close clients c. We fix an arbitrary swap set $\mathcal{P}$, and focus on $\langle\langle f^* \rangle\rangle$, $\langle\langle \neg f_1 \rangle\rangle$, and $\langle\langle \neg f_2 \rangle\rangle$ (All other swaps cause a non-positive potential change). Suppose these three swaps are different. When f^* opens, the client c can be served by both f^* and f_1. When f_1 closes, c can be served by f_2 and η_1, and when f_2 closes, c can be served by f_1 and η_1: in both these cases, we use implication (Siii) of amenability to ensure that both the corresponding facilities are open. We know that $d_2 \leq d(c, \eta_1)$ because c has type E; by (6.9) we get $d(c, \eta_1) \leq 2d^* + d_1$. Putting everything together, the three swaps yield:

$$
\begin{aligned}
&(\delta_{\langle\langle f^* \rangle\rangle}) && \delta_{\mathcal{S}\cap\mathcal{A}}(c) \leq d^* + \beta\, d_1 - d_1 - \beta\, d_2 \\
&(\delta_{\langle\langle \neg f_1 \rangle\rangle}) && \qquad + d_2 + \beta(2d^* + d_1) - d_1 - \beta\, d_2 \\
&(\delta_{\langle\langle \neg f_2 \rangle\rangle}) && \qquad + d_1 + \beta(2d^* + d_1) - d_1 - \beta\, d_2 \\
& && \qquad = \boxed{(1 + 4\beta)\, d^* - (2 - 3\beta)\, d_1 + (1 - 3\beta)\, d_2}.
\end{aligned}
$$

We address the assumption that the three swaps are different. As argued above, condition (Siv) of amenability for type E clients means that for simple swaps, $\langle\langle \neg f_1 \rangle\rangle \neq \langle\langle f^* \rangle\rangle$. However, f_2 could be $\tau(f^*)$, so it may happen that $\langle\langle \neg f_2 \rangle\rangle = \langle\langle f^* \rangle\rangle$, and hence that $\delta_{\mathcal{S}\cap\mathcal{A}}(c) \leq \delta_{\langle\langle f^* \rangle\rangle} + \delta_{\langle\langle \neg f_1 \rangle\rangle}$. Moreover, $\langle\langle \neg f_1 \rangle\rangle$ may not exist, in which case $\delta_{\mathcal{S}\cap\mathcal{A}}(c) \leq \delta_{\langle\langle f^* \rangle\rangle} + \delta_{\langle\langle \neg f_2 \rangle\rangle}$ or even $\delta_{\mathcal{S}\cap\mathcal{A}}(c) \leq \delta_{\langle\langle f^* \rangle\rangle}$. But since our bounds above for both $\delta_{\langle\langle \neg f_1 \rangle\rangle}$ and $\delta_{\langle\langle \neg f_2 \rangle\rangle}$ are non-negative, we infer that the boxed upper bound remains valid in all these cases.

Tree swaps. We now consider tree swaps. Fix an arbitrary swap set $\mathcal{P}$ generated on the event $\mathcal{T} \cap \mathcal{A}$. For a client c in the close case, there are three swaps that are relevant to c—those containing f^*, f_1, and f_2—although some of these swaps may coincide. (Also, no other swaps can increase the potential.)

When f_1 and f_2 belong to the same swap. First suppose that f_1 and f_2 belong to the same swap in $\mathcal{P}$. We start from the case where $\langle\langle f^* \rangle\rangle \neq \langle\langle \neg f_1, \neg f_2 \rangle\rangle$. For the swap $\langle\langle f^* \rangle\rangle$, the client c can be served by both f^* and f_1. And when f_1 and f_2 are both closed, c can be served by $\tau(f^*)$ (which is either η_1 or η_2) and $\pi(f_1)$. By (6.9) we get that $d(c, \tau(f^*))$ is at most $2d^* + d_1$, and by (6.8) we get $d(c, \pi(f_1)) \leq 2d_1 + d^*$. Hence,

$$
\begin{aligned}
& && \sum_{(P,Q)\in\mathcal{P}} \delta_{(P,Q)}(c) \leq \delta_{\langle\langle f^* \rangle\rangle} + \delta_{\langle\langle \neg f_1, \neg f_2 \rangle\rangle} \\
&(\delta_{\langle\langle f^* \rangle\rangle}) && \qquad \leq (d^* + \beta\, d_1) - (d_1 + \beta\, d_2) \\
&(\delta_{\langle\langle \neg f_1, \neg f_2 \rangle\rangle}) && \qquad\quad + (2d^* + d_1) + \beta(2d_1 + d^*) - (d_1 + \beta\, d_2) \\
& && \qquad = \boxed{(3 + \beta)\, d^* - (1 - 3\beta)\, d_1 - 2\beta\, d_2}.
\end{aligned}
$$

On the other hand, if f^*, f_1, and f_2 all belong to the same swap, we can assign c to f^*

$$
\begin{aligned}
& && \sum_{(P,Q)\in\mathcal{P}} \delta_{(P,Q)}(c) \leq \delta_{\langle\langle f^*, \neg f_1, \neg f_2 \rangle\rangle} \\
&(\delta_{\langle\langle f^*, \neg f_1, \neg f_2 \rangle\rangle}) && \qquad \leq (1 + \alpha\beta)\, d^* - d_1 - \beta\, d_2 \\
&(\text{since } 2d^* + d_1 \geq d_2) && \qquad\quad + \beta(2d^* + d_1 - d_2) \\
& && \qquad = (1 + \alpha\beta + 2\beta)\, d^* - (1 - \beta)\, d_1 - 2\beta\, d_2 \\
& && \qquad \leq 3.2\, d^* - 0.4\, d_1 - 0.4\, d_2.
\end{aligned}
$$

These two bounds are identical for our choices of $\alpha = 3$ and $\beta = 1/5$.

When f_1 and f_2 belong to different swaps. Next, consider the case when f_1 and f_2 belong to different swaps in $\mathcal{P}$. Let us first assume $\langle\!\langle f^* \rangle\!\rangle$ is neither $\langle\!\langle \neg f_1 \rangle\!\rangle$ nor $\langle\!\langle f_2 \rangle\!\rangle$. In the swap $\langle\!\langle f^* \rangle\!\rangle$ the client can served by f^* and f_1. When one of f_1 or f_2 is closed, the client c can be served by the other facility, and by $\tau(f^*)$, which is at distance at most $2d^* + d_1$ from c (by (6.8)). Hence,

$$\begin{aligned}
&(\delta_{\langle\!\langle f^* \rangle\!\rangle} \text{ with (6.10)}) && \sum_{(P,Q)\in\mathcal{P}} \delta_{(P,Q)}(c) \le (1-\beta)\, d^* + 2\beta\, d_1 - d_1 - \beta\, d_2 \\
&(\delta_{\langle\!\langle \neg f_1 \rangle\!\rangle}) && \qquad + d_2 + \beta(2d^* + d_1) - d_1 - \beta\, d_2 \\
&(\delta_{\langle\!\langle \neg f_2 \rangle\!\rangle}) && \qquad + d_1 + \beta(2d^* + d_1) - d_1 - \beta\, d_2 \\
& && \qquad = \boxed{(1+3\beta)\, d^* - (2-4\beta)\, d_1 + (1-3\beta)\, d_2}.
\end{aligned}$$

Our bound for $\delta_{\langle\!\langle f^* \rangle\!\rangle}$ does not require f_2 to remain open after the swap, and our bound for $\delta_{\langle\!\langle \neg f_2 \rangle\!\rangle}$ is non-negative. Therefore, the above bound also holds when $\langle\!\langle \neg f_2 \rangle\!\rangle = \langle\!\langle f^* \rangle\!\rangle$. When $\langle\!\langle \neg f_1 \rangle\!\rangle = \langle\!\langle f^* \rangle\!\rangle$, we still have the above bound:

$$\begin{aligned}
&(\delta_{\langle\!\langle f^*, \neg f_1 \rangle\!\rangle}) && \sum_{(P,Q)\in\mathcal{P}} \delta_{(P,Q)}(c) \le d^* + \beta\, d_2 - d_1 - \beta\, d_2 \\
&(\delta_{\langle\!\langle \neg f_2 \rangle\!\rangle}) && \qquad + d_1 + \beta(2d^* + d_1) - d_1 - \beta\, d_2 \\
&(\text{non-negative terms}) && \qquad + \beta\, d^* + \beta\, d_1 + (1-2\beta)(d_2 - d_1) \\
& && \qquad = \boxed{(1+3\beta)\, d^* - (2-4\beta)\, d_1 + (1-3\beta)\, d_2}.
\end{aligned}$$

Summarizing all these bounds (using that $\alpha = 3$ and $\beta = 0.2$),

$$\begin{aligned}
\delta_{\mathcal{S}\cap\mathcal{A}}(c) &\le (1+4\beta)\, d^* - (2-3\beta)\, d_1 + (1-3\beta)\, d_2 = 1.8\, d^* - 1.4\, d_1 + 0.4\, d_2 \\
\delta_{\mathcal{T}\cap\mathcal{A}}(c) &\le \max\{3.2\, d^* - 0.4\, d_1 - 0.4\, d_2, 1.6\, d^* - 1.2\, d_1 + 0.4\, d_2\}.
\end{aligned}$$

Combining and using (6.7) to get $d_2 \le 2d^* + d_1$ if the d_2 terms do not cancel out, we get for close clients c:

$$\begin{aligned}
\Delta_{\mathcal{A}}(c) &\le {}^1\!/\!_2 \cdot \delta_{\mathcal{S}\cap\mathcal{A}} + {}^1\!/\!_2 \cdot \delta_{\mathcal{T}\cap\mathcal{A}} + O(\varepsilon)(d^* + d_1) \\
&\le \boxed{2.5\, d^* - 0.9\, d_1} + O(\varepsilon)(d^* + d_1). \qquad (6.16)
\end{aligned}$$

Lemma 6.2 follows from the bound in (6.15) for the far clients and the one from (6.16) for the close clients.

6.4 All Other Client Types

Similarly, we can bound $\Delta_{\mathcal{A}}(c)$ for every other client type A–D. We summarize this in the following theorem: the calculations behind the expressions can be found in Appendix D.

Lemma 6.3. *For any far client c of type* A *or* B*, we have*

$$\Delta_{\mathcal{A}}(c) \le 2.47\, d^*(c) - 1.13\, d_1(c) + O(\varepsilon)(d^* + d_1) \qquad (6.17)$$

For any close client c_i of type $i \in \{\mathsf{A}, \mathsf{B}, \mathsf{C}, \mathsf{D}\}$, we have

$$\Delta_{\mathcal{A}}(c_{\mathsf{A}}) \le 2.375\, d^*(c_{\mathsf{A}}) - 0.9\, d_1(c_{\mathsf{A}}) + O(\varepsilon)(d^* + d_1) \qquad (6.18)$$

$$\Delta_{\mathcal{A}}(c_{\mathsf{B}}) \le 2.4\, d^*(c_{\mathsf{B}}) - 0.9\, d_1(c_{\mathsf{B}}) + O(\varepsilon)(d^* + d_1) \tag{6.19}$$

$$\Delta_{\mathcal{A}}(c_{\mathsf{C}}) \le 2.2\, d^*(c_{\mathsf{C}}) - 0.8888\, d_1(c_{\mathsf{C}}) + O(\varepsilon)(d^* + d_1) \tag{6.20}$$

$$\Delta_{\mathcal{A}}(c_{\mathsf{D}}) \le 2.5203\, d^*(c_{\mathsf{D}}) - 0.8888\, d_1(c_{\mathsf{D}}) + O(\varepsilon)(d^* + d_1) \tag{6.21}$$

Lemmas 6.2 and 6.3 imply that every client c satisfies

$$\Delta_{\mathcal{A}}(c) \le 2.5203\, d^*(c) - 0.8888\, d_1(c) + O(\varepsilon)(d^* + d_1).$$

This proves Lemma 2.3, and hence Lemma 2.1 and Theorem 1.1.

7 A Computer-Aided Analysis using Linear Programming

In this section we show how to generate a set of valid inequalities, then solve the resulting linear program to find an upper bound on our approximation ratio. We describe the ideas for the potential Φ_2 that only takes the second-closest facility into account, and indicate how to extend it to Φ_q for higher values of $q \ge 2$. Of course, the size of the LP increases exponentially as q increases.

To recall, our proof strategy in the previous section was to consider a local optimum, and then:

1. define a (randomized) collection of important swaps that are contained within our actual set of swaps;
2. for every client type, write constraints that apply to all clients of that type;
3. carefully combine those constraints to have only a few remaining constraints; and
4. manually check these remaining contraints.

An automated proof could avoid the last two steps by directly checking the entire set of constraints. Since every constraint we derive is a linear inequality on the distances, a linear program can be used for this automated proof. Put differently, our goal is to write a linear program that constructs a "worst-case example" for our potential function. Specifically, the program seeks values of the distances d_1, d_2, and d^* for each client type, so as to maximize the ratio between the costs of the optimum and local solutions, while respecting the set of constraints. [1]

Variables and constraints of the LP. Let us focus on simple swaps, the constraints for tree swaps are similar. We want to express the fact that simple swaps at a local optimum do not decrease the potential. We first classify facilities into types according to their ratio ρ; we consider only a fine net of values for ρ, and use continuity of the potential to control the loss due to this discretization. All facilities with a given ratio are treated the same way in the proof: our LP considers that all facilities of the same type are swapped at the same time. Specifically, we have a variable s_ρ corresponding to the difference in the potential function after applying simple swaps for all facilities with ratio ρ. The constraint saying that simple swaps do not decrease the potential is therefore $\sum_\rho s_\rho \ge 0$.

The value of the variable s_ρ is controlled by the clients connected to facilities having ratio ρ: each client type i has a contribution to it. In an S_j-swap (where $j \in \{1, 2\}$), let $\delta_{S_j}(i; x, y^1, y^2)$ be the potential change due to all clients of type-i connected to facilities with ratio ρ, in function of x, y^1 and y^2, respectively the total distance from those clients to the optimal solution, their closest and their second closest facility of the local solution. This difference of potential is described in

[1] In fact, it does not come up with a concrete example, since we do not maintain all the triangle inequalities between the clients, but only the triangle inequalities in some local neighborhood around each client. It is conceivable that using more triangle inequalities would lead to an even better result, but that increases the complexity even further.

Section 6: we illustrate it with clients of type A, in the far case. We denote AF those clients. As presented in D.1.1, the S_1 and S_2 swaps for those clients show

$$\begin{aligned}\delta_{\mathcal{S}_1\cap\mathcal{A}}(c) \leq&(1+\alpha\beta)\, d^* - (1+\alpha\beta)\, d_1\\ \delta_{\mathcal{S}_2\cap\mathcal{A}}(c) \leq&((1+1/\rho)(1+\alpha\beta)+\beta)\, d^* - ((1-1/\rho)(1+\alpha\beta)+\alpha\beta)\, d_1\end{aligned}$$

For bounding $\delta_{\mathcal{S}_2\cap\mathcal{A}}(c)$, we upper bounded the potential value of the swap by $d^*+\beta d_1$ when f^* is opened. However, we could be more precise: it could be the case that $d^* \leq d_1$ or $d_1 \leq \alpha d^*$. Therefore, this lead to 3 other possible upperbounds, namely $(1+\alpha\beta)d^*$, $d^*+\beta d_1$, and $(1+\alpha\beta)d_1$.

This translates to three other inequalities, one for each of those cases:

(when we choose $(1+\alpha\beta)d^*$) $\delta_{\mathcal{S}_2\cap\mathcal{A}}(c) \leq ((2+1/\rho)(1+\alpha\beta))\, d^* - ((2-1/\rho)(1+\alpha\beta))\, d_1$
(when we choose $d^*+\beta d_1$) $\delta_{\mathcal{S}_2\cap\mathcal{A}}(c) \leq ((1+1/\rho)(1+\alpha\beta)+1)\, d^* - ((2-1/\rho)(1+\alpha\beta)-\beta)\, d_1$
(when we choose $(1+\alpha\beta)d_1$) $\delta_{\mathcal{S}_2\cap\mathcal{A}}(c) \leq ((1+1/\rho)(1+\alpha\beta))\, d^* - ((1-1/\rho)(1+\alpha\beta))\, d_1$.

More generally, the LP encodes all possible combinations of variables giving valid bound on the potential after a swap. Note that then number of such inequalities grows exponentially with q, because each term $\min(\alpha_j d_1(c), d_j(c))$ doubles the number of valid inequalities.

Going back to type A, this gives rise to the constraints

$$\delta_{S_1}(\mathsf{AF}; x_{\mathsf{AF},\rho}, y^1_{\mathsf{AF},\rho}, y^2_{\mathsf{AF},\rho}) \leq (1+\alpha\beta)x_{\mathsf{AF},\rho} - (1+\alpha\beta)y^1_{\mathsf{AF},\rho} \tag{7.22}$$
$$\delta_{S_2}(\mathsf{AF}; x_{\mathsf{AF},\rho}, y^1_{\mathsf{AF},\rho}, y^2_{\mathsf{AF},\rho}) \leq ((1+1/\rho)(1+\alpha\beta)+\beta)\, x_{\mathsf{AF},\rho} - ((1-1/\rho)(1+\alpha\beta)+\alpha\beta)\, y^1_{\mathsf{AF},\rho} \tag{7.23}$$
$$\delta_{S_2}(\mathsf{AF}; x_{\mathsf{AF},\rho}, y^1_{\mathsf{AF},\rho}, y^2_{\mathsf{AF},\rho}) \leq ((2+1/\rho)(1+\alpha\beta))\, x_{\mathsf{AF},\rho} - ((2-1/\rho)(1+\alpha\beta))\, y^1_{\mathsf{AF},\rho} \tag{7.24}$$
$$\delta_{S_2}(\mathsf{AF}; x_{\mathsf{AF},\rho}, y^1_{\mathsf{AF},\rho}, y^2_{\mathsf{AF},\rho}) \leq ((1+1/\rho)(1+\alpha\beta)+1)\, x_{\mathsf{AF},\rho} - ((2-1/\rho)(1+\alpha\beta)-\beta)\, y^1_{\mathsf{AF},\rho} \tag{7.25}$$
$$\delta_{S_2}(\mathsf{AF}; x_{\mathsf{AF},\rho}, y^1_{\mathsf{AF},\rho}, y^2_{\mathsf{AF},\rho}) \leq ((1+1/\rho)(1+\alpha\beta))\, x_{\mathsf{AF},\rho} - ((1-1/\rho)(1+\alpha\beta))\, y^1_{\mathsf{AF},\rho}, \tag{7.26}$$

where the variables $x_{i,\rho}$ denote the total cost of clients of type i connected to facilities with ratio ρ in the optimal solution, and $y^j_{i,\rho}$ denote the total distance from those clients to their j-th closest facility in the local solution. This definition of $\delta_{S_j}(i)$ yields the following constraint on s_ρ : for $j=1,2$,

$$\sum_{i\in\mathcal{ST}} \delta_{S_j}(i; x_{i,\rho}, y^1_{i,\rho}, y^2_{i,\rho}) \geq s_\rho,$$

where $\mathcal{ST}$ is the set of client types.

Moreover, the triangle inequality gives constraints on the variables $x_{i,\rho}, y^1_{i,\rho}, y^2_{i,\rho}$: for instance, for type AF, we would have

$$y^2_{\mathsf{AF},\rho} \leq x_{\mathsf{AF},\rho} + (1/\rho)(\leq x_{\mathsf{AF},\rho} + y^1_{\mathsf{AF},\rho}). \tag{7.27}$$

The constraints due to tree swaps are defined analogously, with a variable t_ρ being the potential change after applying tree swaps for all facilities of type ρ, and $\delta_{T_j}(i)$ being the potential change due to all clients of type-i connected to facilities with ratio ρ. For $q>2$, we need to consider more than one ratio, so we let ρ be the vector of size $q-1$ that describes ratio of all two consecutive η_j and η_{j+1} for all $j \in \{1,...,q-1\}$. Let $\mathcal{R}$ be the set of values of ρ after discretization: we use $\mathcal{R} := \{\frac{i}{100} \mid i \in \{0,\dots,100\}\}^{q-1}$. In that case, all clients with ratio in $[i\cdot 10^{-2}, (i+1)\cdot 10^{-2})$ are considered to have $\rho = i\cdot 10^{-2}$ for each index. This means that our bounds for δ are slightly relaxed to cover an interval instead of a precise ρ. The j^{th} index of a ratio correspond to $\frac{d(f^*(c),\eta_j(f^*(c)))}{d(f^*(c),\eta_{j-1}(f^*(c)))}$.

Let $\mathcal{C}$ be the set of client types. For $q = 2$, each $i \in \mathcal{C}$ contains client type $(\mathsf{A}, \mathsf{B}, \mathsf{C}, \mathsf{D}, \mathsf{E})$, ratio ρ, underlying form of tree-graph (e.g. f_1 and f_2 belong to same tree in η_1-swap), and whether $d_2 \le \alpha d_1$ or not. For $j \in \{2, ..., q\}$, let $\mathcal{C}_j$ be the set of clients with $d_j \le \alpha_j d_1$. For $q = 2$, $\mathcal{C}_2$ is the set of close clients (i.e., $d_2 \le \alpha d_1$). Let C^ρ denote set of clients with ratio ρ. The general structure of the LP is the following:

$$
\begin{aligned}
&(7.28) && \max \sum_{i \in \mathcal{C}} y_i^1 && \\
&(7.29) && \text{s.t.} \sum_{i \in \mathcal{C}} x_i = 1 && \\
&(7.30) && y_i^j \le \alpha_j y_i^1 && \forall j \in \{2, \dots, q\}, i \in \mathcal{C}_j \\
&(7.31) && y_i^j \ge \alpha_j y_i^1 && \forall j \in \{2, \dots, q\}, i \notin \mathcal{C}_j \\
&(7.32) && \sum_{i \in \mathcal{C}^\rho} \delta_{S_j}(i; x_i, y_i^1, \dots, y_i^q) \ge s_\rho && \forall \rho \in \mathcal{R}, j \in \{1, \dots, q\} \\
&(7.33) && \sum_{i \in \mathcal{C}^\rho} \delta_{T_j}(i; x_i, y_i^1, \dots, y_i^q) \ge t_\rho && \forall \rho \in \mathcal{R}, j \in \{1, \dots, q\} \\
&(7.34) && \sum_{\rho \in \mathcal{R}} s_\rho \ge 0 && \\
&(7.35) && \sum_{\rho \in \mathcal{R}} t_\rho \ge 0 && \\
&(7.36) && \text{Triangle-inequalities} && \\
&(7.37) && \delta_{S_j}(i; x_i, y_i^1, \dots, y_i^q) \le \text{enumerated-upperbounds} && \forall i \in \mathcal{C}, j \in \{2, ..., q\} \\
&(7.38) && \delta_{T_j}(i; x_i, y_i^1, \dots, y_i^q) \le \text{enumerated-upperbounds} && \forall i \in \mathcal{C}, j \in \{2, ..., q\} \\
&(7.39) && y_i^j \ge 0, \quad x_i \ge 0 && \forall i \in \mathcal{C}, j \in \{1, \dots, q\}
\end{aligned}
$$

Note that $\sum_{i \in \mathcal{C}} y_i^1 / \sum_{i \in \mathcal{C}} x_i = \sum_{i \in \mathcal{C}} y_i^1$ is the locality gap. Constraints (7.30) and (7.31) restrict each distance based on whether they are 'far' client or 'close' client. Constraints (7.32) can be seen as the following: for each ratio ρ we pick $j \in \{1, ..., q\}$ that minimizes the sum of potential difference after performing S_j swap, then make $\tau(f^*) = \eta_j$ for all f^* with ratio ρ. Similarly, (7.33) chooses $\tau(\cdot)$ for tree swaps. Then (7.34) and (7.35) ensure the potential difference is non-negative after performing simple swap and tree swap respectively. We also add triangle inequalities (e.g., (7.27)). Lastly, we add upperbounds for each potential difference in (7.37) and (7.38) (e.g., (7.22) - (7.26)).

Implementing this approach, and then solving the resulting LP for for potential Φ_2 and Φ_3 gives us the following numbers:

Potential	Bound
Φ_2	2.7786
Φ_3	2.6861

For Φ_2, the LP finds that taking $\alpha = 3, \beta = 0.2$ yields the best result, whereas for Φ_3 we set manually $\alpha = 2.5, \beta = 0.3, \beta_2 = \beta \cdot 0.34$. As always, we get an additive ε term because of the defiant swaps. However, let us emphasize that these implementations should be considered preliminary, since they have not been formally verified. We hope that formal proofs of these results can be given in the near future.

A Locality Gap for Potential Φ_2

In this section, we give lower bounds on the locality gap, and prove Theorem 1.2. We show locality gap examples of $\max\{2, \alpha\}$, $3-2\beta$, and $1+4\beta$ for the potential function Φ_2. Putting these together, the locality gap is $\min_{\beta\in[0,1],\alpha\in[1,\infty)} \max\{3-2\beta, 1+4\beta\}, \max\{2, \alpha\}$. Note $\max\{3-2\beta, 1+4\beta\}$ is $2\frac{1}{3}$, when we set $\beta = 1/3$. Therefore we show a locality gap of 2.

In this section, we show a locality gap of 2 for Φ_2. We divide the cases into three main cases:

- When $\alpha \leq 2$
- When $\alpha > 2$ and $\beta \leq 1/3$
- When $\alpha > 2$ and $\beta > 1/3$

We mainly use two types of example that we call "bi-clique" and "double-bi-clique" described in Figure A.9 Figure A.10 respectively. In bi-clique we have $k + r$ local facilities on the right, where $r = O(1)$ is the number of extra local facilities, and k optimal facilities are on the left. There is a client between every (local, optimal) facility pair, at unit distance from the optimal facility, and at distance d from the local facility.

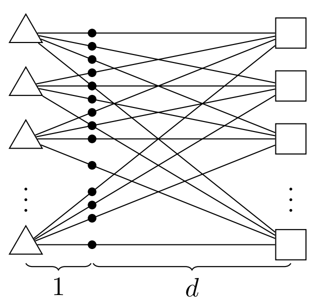


Figure A.9: An illustration of the bi-clique example with $r = 0$.

In double-bi-clique, we have two back-to-back bi-cliques as in Figure A.10. Each bi-clique is constructed the same way as Figure A.9 except the number of facilities are halved. Consider a client c with an edge going into f_i, create an edge at distance d between c and i^{th} local facility in the other bi-clique. Now every client has an optimal facility at distance 1, and two local facilities at distance d.

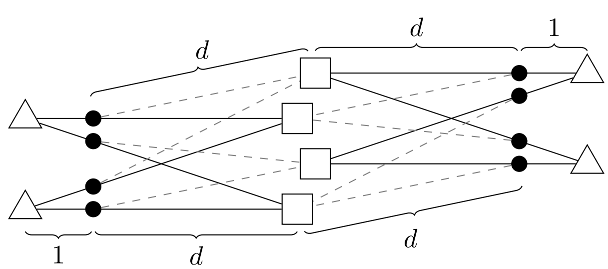


Figure A.10: The double-bi-clique example with $k = 4$ and $r = 0$.

For all cases we calculate the potential difference after performing a swap of size p. There are mainly 4 different types of clients.

- $\mathcal{C}_o$: the set of clients with their f^* opened.
- $\mathcal{C}_1$: Clients with their f_1 closed and f^* not opened
- $\mathcal{C}_2$: Clients with its f_2 closed and f^* not opened
- $\mathcal{C}_3$: Clients with its f^* closed, f_1 opened, and f_2 opened.

We use c_o, c_1, c_2, c_3 to denote generic client for sets $\mathcal{C}_o$, $\mathcal{C}_1$, $\mathcal{C}_2$, and $\mathcal{C}_3$ respectively. We first calculate potential difference for each client type, then sum them over.

We assume there is no client with their f_1 and f_2 both closed: those clients can only hurt the quality of the solution, and given a swap that closes f_1 and f_2 of some clients it is easy to construct a strictly better set of swaps with no such client.

A.1 When $\alpha \leq 2$

We first give lower bound examples when $\alpha \leq 2$. We divide the case further into two cases: when $\alpha \leq 4/3 + 1/(3\beta)$ and when $\alpha > 4/3 + 1/(3\beta)$.

Subcase I: $\alpha \leq 4/3 + 1/(3\beta)$**.** We create a bi-clique presented in Figure A.9 with $d = 2 - \varepsilon'$, where $\varepsilon' \approx O(1/k)$ is a small quantity to be specified later. Note that every client has $k + r - 1$ local facilities at distance $2 + d$, thus the second closest facility is never closed for any client. Then for each client we get the potential differences:

$$\begin{aligned}
\Delta\Phi^{c_o} &= 1 + \alpha\beta - (2 - \varepsilon') - (2 - \varepsilon')\alpha\beta = -1 - \alpha\beta + \varepsilon' + \varepsilon'\alpha\beta \\
\Delta\Phi^{c_1} &= (4 - \varepsilon') + (4 - \varepsilon')\beta - (2 - \varepsilon') - (2 - \varepsilon')\alpha\beta \geq 2 + 4\beta - 2\alpha\beta \\
\Delta\Phi^{c_2} &= 0 \\
\Delta\Phi^{c_3} &= 0
\end{aligned}$$

Note $|\mathcal{C}_o| = p(k + r)$ and $|\mathcal{C}_1| = p(k - p)$.

Summing up gives

$$\begin{aligned}
\sum_{c \in \mathcal{C}} \Delta\Phi^c &\geq p(k + r)(-1 - \alpha\beta + \varepsilon' + \varepsilon'\alpha\beta) + p(k - p)(2 + 4\beta - 2\alpha\beta) \\
&\geq pk(1 + 4\beta - 3\alpha\beta) + pk\varepsilon' + rp(-1 - \alpha\beta) - p^2(2 + 4\beta - 2\alpha\beta) \\
\text{(for } \alpha \leq 4/3 + 1/(3\beta)\text{)} \quad &\geq pk(1 + 4\beta - 3\alpha\beta) \geq 0
\end{aligned}$$

The second inequality holds for any $\varepsilon' \geq \frac{p^2(2+4\beta-2\alpha\beta)+rp(1+\alpha\beta)}{pk} = O(1/k)$. Hence, this example shows a locality gap of $2 - o(1)$ when $\alpha \leq \min(2, 4/3 + 1/(3\beta))$.

Subcase II: $\alpha > 4/3 + 1/(3\beta)$**.** Since we focus on $\alpha \leq 2$ and $\beta \leq 1$, this subcase implies that $\beta > 1/2$ and $\alpha > 5/3$. To deal with it, we create a double-bi-clique presented in Figure A.10 with $d = 2$. Note that every client has two local facilities at distance 2, and $k + r - 2$ facilities at distance 4. Then for each client, we get the following potential differences:

$$\begin{aligned}
\Delta\Phi^{c_o} &= 1 + \alpha\beta - 2 - 2\beta = -1 + \alpha\beta - 2\beta \\
\Delta\Phi^{c_1} &= 2 + 2\alpha\beta - 2 - 2\beta \geq 2\alpha\beta - 2\beta \\
\Delta\Phi^{c_2} &= 2 + 2\alpha\beta - 2 - 2\beta \geq 2\alpha\beta - 2\beta
\end{aligned}$$

$$\Delta\Phi^{c_3} = 0$$

Let p_1 and p_2 be the number of optimal facilities in the first clique that belong to the swap. Let p_2 be the number of optimal facilities in the second clique that belong to the swap. Then we have $|\mathcal{C}_o| = \frac{k+r}{2}p_1 + \frac{k+r}{2}p_2 = \frac{k+r}{2}p$. Similarly let p'_1 and p'_2 be the number of local facilities in the first and second clique that belong to the swap. Then we have $|\mathcal{C}_1| = p'_1(\frac{k}{2}-p_1)+p'_2(\frac{k}{2}-p_2) \geq (\frac{k}{2}-p)p$. Also note that $|\mathcal{C}_2| \geq (\frac{k}{2}-p)p$. Summing up gives

$$\begin{aligned}
\sum_{c\in\mathcal{C}} \Delta\Phi^c &\geq {}^{p}/_{2}(k+r)(-1+\alpha\beta-2\beta) + 2p({}^{k}/_{2}-p)(2\alpha\beta-2\beta) \\
&\geq {}^{pk}/_{2}(-1+5\alpha\beta-6\beta) - {}^{rp}/_{2}(1+2\beta-\alpha\beta) - 2p^2(2\alpha\beta-2\beta) \\
\text{(for } \alpha > 5/3 \text{ and } \beta > 1/2.) \quad &\geq pk(1/12) - {}^{r}/_{2}p(1+2\beta-\alpha\beta) - 2p^2(2\alpha\beta-2\beta) \geq 0
\end{aligned}$$

The last inequality holds for $r \leq \frac{k}{24(1+2\beta-\alpha\beta)}$ and $p \leq \frac{k}{48(2\alpha\beta-2\beta)}$. Since r and p are absolute constant (i.e, $o(k)$), the inequality is valid for big enough k. Hence, this example shows a locality gap of 2 when $4/3 + {}^{1}/_{3\beta} \leq \alpha \leq 2$, in particular when $\beta > {}^{1}/_{2}$ and $2 \geq \alpha \geq {}^{5}/_{3}$ This concludes therefore the case $\alpha \leq 2$.

A.2 When $\beta \leq 1/2$ and $\alpha > 2$

In this section we give a bi-clique example showing a locality gap when $\beta \leq 1/3$ and $\alpha > 2$. for constant-sized swap. Consider the bi-clique graph in Figure A.9 with distance $d = \min\{3-2\beta-\varepsilon', \alpha\}$. We divide the case into two subcases. In first case we consider when $3-2\beta \leq \alpha$. Then we consider when $3-2\beta > \alpha$.

Subcase I: $3-2\beta \leq \alpha$. We will first consider the case when $3-2\beta \leq \alpha$. In that case, the current potential value of a client in the local solution is $\Phi^c(F) = (3-2\beta-\varepsilon') + \beta(5-2\beta-\varepsilon')$ (since $\alpha > 2 \geq \frac{5-2\beta}{3-2\beta}$ for $\beta \leq {}^{1}/_{2}$).

For the $p(k+r)$ clients in $\mathcal{C}_o$, where $k+r$ is the number of local facilities, if $3-2\beta-\varepsilon' \leq \alpha$ we get the following potential difference:

$$\Delta\Phi^{c_o} \geq 1 + \beta(3-2\beta-\varepsilon') - (3-2\beta-\varepsilon') - \beta(5-2\beta-\varepsilon') = -2+\varepsilon'$$

Note that if a client's f^* is opened but its f_1 is closed, the client contributes $1+\beta(5-2\beta-\varepsilon') \geq 1+\beta(3-2\beta-\varepsilon')$, and hence the above inequality is still valid for those clients.

There are $p(k-p)$ clients in $\mathcal{C}_1$, and they induce the following potential difference:

$$\Delta\Phi^{c_1} = (5-2\beta-\varepsilon') + \beta(5-2\beta-\varepsilon') - (3-2\beta-\varepsilon') - \beta(5-2\beta-\varepsilon') = 2$$

Finally, clients in $\mathcal{C}_3$ and $\mathcal{C}_4$ do not induce a change in the potential value.

The sum over all clients yields

$$\begin{aligned}
\sum_{c\in\mathcal{C}} \Delta\Phi^c &\geq p(k+r)(\varepsilon'-2) + 2p(k-p) \\
&= pk(0) + p(k+r)(\varepsilon') - 2(pr) - 2p^2 \geq 0.
\end{aligned}$$

The last inequality holds for any $\varepsilon' \geq \frac{2pr+2p^2}{p(k+r)} = O(1/k)$.

Hence, in the case where $\alpha \geq 2$, $\beta \leq {}^{1}/_{2}$ and $3-2\beta \leq \alpha$, this example shows a locality gap of $3-2\beta-o(1) \geq 2$.

Case II: $2 \leq \alpha \leq 3-2\beta$**.** When $2 \leq \alpha \leq 3-2\beta$, client's closest distance is now α. Thus the potential value of a client before any swap is $\Phi^c(F) = \alpha + (2+\alpha)\beta$. Note that $2+\alpha \leq \alpha^2$ for $\alpha \geq 2$. We have the same number of clients in each set. Furthermore, we get the potential differences:

$$\begin{aligned}
\Delta\Phi^{c_o} &\geq 1 + \alpha\beta - \alpha - (2+\alpha)\beta \geq 1 - \alpha - 2\beta \\
\Delta\Phi^{c_1} &= (2+\alpha) + (2+\alpha)\beta - \alpha - (2+\alpha)\beta = 2 \\
\Delta\Phi^{c_2} &= 0 \\
\Delta\Phi^{c_3} &= 0
\end{aligned}$$

The sum over all clients yields

$$\begin{aligned}
\sum_{c\in\mathcal{C}} \Delta\Phi^c &\geq p(k+r)(1-\alpha-2\beta) + p(k-p)(2) \\
&\geq pk(3-2\beta-\alpha) + pr(1-\alpha-2\beta) - 2p^2 \\
(\alpha \leq 3-2\beta-\varepsilon') \qquad &\geq pk\varepsilon' - pr(\alpha+2\beta-1) - 2p^2 \geq 0.
\end{aligned}$$

The last inequality holds for any $\varepsilon' \geq \frac{pr(\alpha-1-2\beta)+2p^2}{pk} = O(1/k)$.

Hence, in the case where $\alpha \leq 3-2\beta$, $\alpha \geq 2$ and $\beta \leq 1/3$, this example shows a locality gap of $\alpha \geq 2$.

A.3 When $\beta > 1/2$ and $\alpha > 2$

Finally we give lower bound examples when $\beta > 1/2$ and $\alpha > 2$. We use double-bi-clique in described Figure A.10 with $d = \min\{1+4\beta-\varepsilon', \alpha\}$.

Subcase I: $1+4\beta \leq \alpha$**.** Here, the current potential function value for a client is $\Phi^c(F) = (1+4\beta-\varepsilon') + \beta(1+4\beta-\varepsilon')$.

There are $p(\frac{k+r}{2})$ clients in $\mathcal{C}_o$, and the potential difference for a client $c_o \in \mathcal{C}_o$ is

$$\Delta\Phi^{c_o} = 1 + \beta(1+4\beta-\varepsilon') - (1+\beta)(1+4\beta-\varepsilon') = -4\beta + \varepsilon'.$$

There are at least $p(\frac{k}{2}-p)$ clients in $\mathcal{C}_1$. Recall $\alpha > 2 \geq \frac{3+4\beta-\varepsilon'}{1+4\beta-\varepsilon'}$ for $\beta > 1/2$. The potential difference for $c_1 \in \mathcal{C}_1$ is

$$\Delta\Phi^{c_1} = (1+4\beta-\varepsilon') + \beta(3+4\beta-\varepsilon') - (1+\beta)(1+4\beta-\varepsilon') = 2\beta$$

There are $p(\frac{k}{2}-p)$ clients in $\mathcal{C}_2$, and they get the same swap value as clients in $\mathcal{C}_1$. Clients in $\mathcal{C}_3$ do not induce any change in the potential.

Then sum over all clients yields

$$\sum_{c\in\mathcal{C}} \Delta\Phi^c \geq -4\beta p\left(\frac{k+r}{2}\right) + \varepsilon' p\left(\frac{k+r}{2}\right) + 4\beta\left(\frac{pk}{2} - p^2\right) = \varepsilon'\left(\frac{p(k+r)}{2}\right) - 4\beta p\,(p + r/2) \geq 0$$

The difference in potential function is therefore positive for all $\varepsilon' > \frac{4\beta r + 8\beta p}{(k+r)} = O\left(\frac{1}{k}\right)$. Hence, this example shows a locality gap of $1+4\beta-o(1) \geq 2-o(1)$ when $\alpha \geq \max(2, 1+4\beta)$ and $\beta \geq 1/2$.

Case II: $2 \le \alpha \le 1 + 4\beta$**.** When $\alpha \le 1 + 4\beta$, clients' closest and the second closest local facilities are both at distance α. Thus the current potential value for a client is $\Phi^c(F) = \alpha + \alpha\beta$. We have the same number of clients in each set. We get the following potential differences:

$$\Delta\Phi^{c_o} = 1 + \alpha\beta - \alpha - \alpha\beta \ge 1 - \alpha$$

(Note $2 \le \alpha$ implies $(2 + \alpha) \le \alpha^2$.)

$$\Delta\Phi^{c_1} = \alpha + (2 + \alpha)\beta - \alpha - \alpha\beta = 2\beta$$

$$\Delta\Phi^{c_2} = \alpha + (2 + \alpha)\beta - \alpha - \alpha\beta = 2\beta$$

$$\Delta\Phi^{c_3} = 0$$

The sum over all clients yields

$$\sum_{c \in \mathcal{C}} \Delta\Phi^c \ge p\left(\frac{k+r}{2}\right)(1 - \alpha) + 2({}^{pk}/{}_2 - p^2)(2\beta)$$

$$\ge \frac{pk}{2}(1 - \alpha + 4\beta) - \frac{pr}{2}(\alpha - 1) - 2p^2(2\beta)$$

$(\alpha \le 1 + 4\beta - \varepsilon')$

$$\ge \frac{pk}{2}(\varepsilon') - \frac{pr}{2}(\alpha - 1) - 2p^2(2\beta) \ge 0$$

The last inequality holds for any $\varepsilon' \ge \frac{pr(\alpha-1)+8p^2\beta}{pk} = O(1/k)$. Hence, this example shows a locality gap of $\alpha - o(1) \ge 2 - o(1)$ when $2 \le \alpha \le 1 + 4\beta$ and $\beta \ge {}^1/{}_2$.

B Motivating our Swaps

In this section we present examples that motivate our choice of potential function and our swaps. In particular they show that the swap structures defined in previous works are not powerful enough to prove our results.

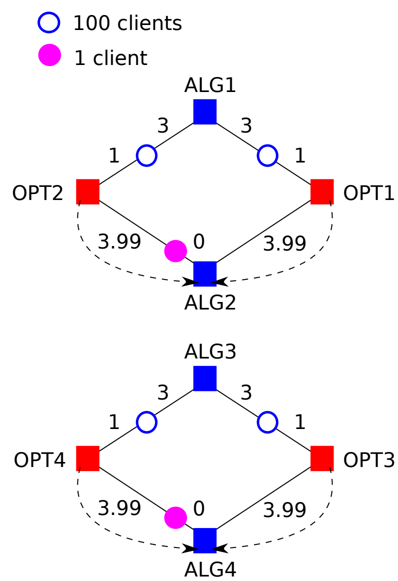


Figure B.11: A bad scenario for the swap structures defined by [GT08].

The analysis in [GT08] matches each optimal facility to its closest local facility. So it matches both OPT1, OPT2 to ALG2 and both OPT3, OPT4 to ALG4, leaving ALG1 and ALG3 with no facility of opt matched to them. Hence, two swaps are defined: (1) swapping in OPT1 and OPT2 and removing ALG2 and ALGj for some $j \in \{1, 3\}$ that remains unspecified in their analysis, and (2) swapping in OPT3 and OPT4 and removing ALG4 and ALG(4-j). Now, if we consider the swaps defined by choosing $j = 3$, the set of equations obtained does not allow us to deduce that the solution is not a local optimum, as long as $\alpha > 5/3$ and for any $\beta < 1$.

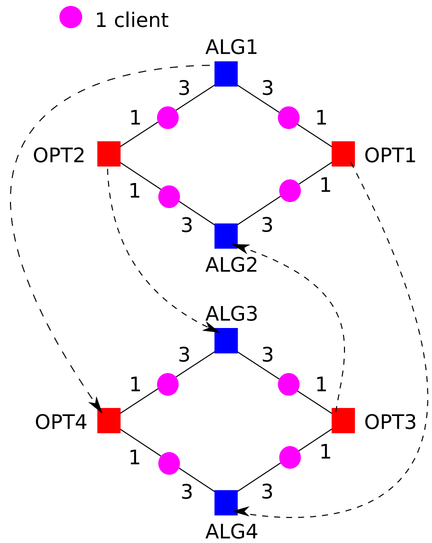


Figure B.12: A bad scenario for the swap structures defined by Arya et al. [AGK+01].

The definition of the swap structure in [AGK+01] does not uniquely identify which local facility is matched to which optimal facility. Hence, if the analysis matches ALG1 with OPT4, ALG2 with OPT3, ALG3 with OPT2 and ALG4 with OPT1, the set of linear equations obtained does not allow us to deduce that the instance is not a local optimum.

C Useful Inequalities

In this section, we prove the inequalities in Table 2. We also give some more inequalities in Table 3; these will be used in §D.

Bound		Coundition
$d_2 \geq (1/\rho)\, d_1 - (1 + 1/\rho)\, d^*$	(C.40)	$\eta_1(f^*) = f_1$
$d_2 \leq d^* + (1/\rho)(d_1 + d^*)$	(C.41)	$\eta_1(f^*) = f_1$
$d(c, \eta_2(f^*)) \leq d^* + 1/\rho(d^* + d_1)$	(C.42)	$\eta_1(f^*) = f_1$
$d_2 \leq d^* + \rho(d_1 + d^*)$	(C.43)	$\eta_1(f^*) \neq f_1$
$d(c, \eta_1(f^*)) \leq d^* + \rho(d^* + d_1) \leq 2d^* + d_1$	(C.44)	$\eta_1(f^*) \neq f_1$
$d(c, \eta_2(f^*)) \leq 2d^* + d_1$	(C.45)	$\eta_1(f^*) \neq f_1$
$d(c, \pi(f_2)) \leq 2d_2 + d^*$	(C.46)	$\eta_1(f^*) \neq f_1$

Table 3: More useful inequalities

For clients c with $\eta_1(f^*) = f_1$:

$$d_2 \geq d(f^*, \eta_2) - d^* = \tfrac{1}{\rho} d(f^*, \eta_1) - d^* \geq ({}^1\!/\!_\rho)\, d_1 - (1 + {}^1\!/\!_\rho)\, d^*, \qquad \text{(proving (C.40))}$$
$$d_2 \leq d^* + d(f^*, \eta_2) = d^* + \tfrac{1}{\rho} d(f^*, \eta_1) \leq d^* + \tfrac{1}{\rho}(d_1 + d^*). \qquad \text{(proving (C.41))}$$
$$d(c, \eta_2) \leq d(c, f^*) + d(f^*, \eta_2) = d^* + ({}^1\!/\!_\rho) d(f^*, f_1) \leq d^* + {}^1\!/\!_\rho (d^* + d_1). \qquad \text{(proving (C.42))}$$

Else when $\eta_1(f^*) \neq f_1$:

$$d(c, \eta_1) \leq d(c, f^*) + d(f^*, \eta_1) = d(c, f^*) + \rho\, d(f^*, \eta_2) \leq d^* + \rho(d^* + d_1) \quad \text{(proving (C.44))}$$
$$d(c, \eta_2) \leq d(c, f^*) + d(f^*, \eta_2) \leq d^* + (d^* + d_1). \qquad \text{(proving (C.45))}$$
$$d_2 \leq d(c, \eta_1) \overset{\text{(C.44)}}{\leq} d^* + \rho(d_1 + d^*), \qquad \text{(proving (C.43) and (6.7))}$$

Combining (C.44) and (C.45) gives (6.9).

Recalling that $\pi(f)$ is the closest optimal facility to f, we get for any client c,

$$d(c, \pi(f_1)) \leq d(c, f_1) + d(f_1, \pi(f_1)) \leq d_1 + d(f_1, f^*) \leq 2d_1 + d^* \qquad \text{(proving (6.8))}$$
$$d(c, \pi(f_2)) \leq d(c, f_2) + d(f_2, \pi(f_2)) \leq d_2 + d(f_2, f^*) \leq 2d_2 + d^*. \qquad \text{(proving (C.46))}$$

To prove (6.10), we use that for any $a, b \geq 0$ and $\beta \leq 1$, the expression $\min(a, b) + \beta \max(a, b) = \min(a+\beta b, b+\beta a)$ is smaller than any convex combination $(1-\lambda)(a+\beta b)+\lambda(b+\beta a)$ with $\lambda \in [0, 1]$. Setting $\lambda = \frac{\beta}{1-\beta}$ and simplifying gives $(1-\beta)a+2\beta b$. Using $a = d^*$ and $b = d_1$ completes the proof.

D Proof of Lemma 6.3

We now present the proof of Lemma 6.3, giving bounds for all the client types other than type E. The idea is the same for each one: First we fix a client c of some type. We partition the amenable event into some sub-events, and look on some sub-event $\mathcal{E}$. We consider a generic swap set $\mathcal{P}$ generated under that event, and give an upper bound for the maximum potential change for client c due to these swaps. Combining over all sub-events (with the correct probability values) gives the expected potential change. The largest such change for each client type is then shown to be the one recorded in Lemma 6.3.

When we prove upper bounds for the potential change caused by a swap set $\mathcal{P}$, we assume that both $\langle\!\langle \neg f_1 \rangle\!\rangle$ and $\langle\!\langle \neg f_2 \rangle\!\rangle$ exist in $\mathcal{P}$ (if f is heavy $\langle\!\langle \neg f \rangle\!\rangle$ does not exist). As we mentioned in Section 6.2, our bounds also hold in cases where either of them does not exist, because our upper bounds for $\delta_{\langle\!\langle \neg f \rangle\!\rangle}$ is non-negative as long as $\langle\!\langle f^* \rangle\!\rangle \neq \langle\!\langle \neg f \rangle\!\rangle$ for any $f \in \{f_1, f_2\}$.

In the rest of this section, we prove each inequality from Lemma 6.3.

Lemma 6.3. *For any far client c of type A or B, we have*

$$\Delta_{\mathcal{A}}(c) \leq 2.47\, d^*(c) - 1.13\, d_1(c) + O(\varepsilon)(d^* + d_1) \tag{6.17}$$

For any close client c_i of type $i \in \{\mathsf{A}, \mathsf{B}, \mathsf{C}, \mathsf{D}\}$, we have

$$\Delta_{\mathcal{A}}(c_{\mathsf{A}}) \leq 2.375\, d^*(c_{\mathsf{A}}) - 0.9\, d_1(c_{\mathsf{A}}) + O(\varepsilon)(d^* + d_1) \tag{6.18}$$
$$\Delta_{\mathcal{A}}(c_{\mathsf{B}}) \leq 2.4\, d^*(c_{\mathsf{B}}) - 0.9\, d_1(c_{\mathsf{B}}) + O(\varepsilon)(d^* + d_1) \tag{6.19}$$
$$\Delta_{\mathcal{A}}(c_{\mathsf{C}}) \leq 2.2\, d^*(c_{\mathsf{C}}) - 0.8888\, d_1(c_{\mathsf{C}}) + O(\varepsilon)(d^* + d_1) \tag{6.20}$$
$$\Delta_{\mathcal{A}}(c_{\mathsf{D}}) \leq 2.5203\, d^*(c_{\mathsf{D}}) - 0.8888\, d_1(c_{\mathsf{D}}) + O(\varepsilon)(d^* + d_1) \tag{6.21}$$

D.1 Proof of (6.17): Far Clients of Type A and B

In this section, we show that for any far case client c of type A or B, we have

$$\Delta_{\mathcal{A}}(c) \le 2.467\, d^*(c) - 1.13085\, d_1(c) + O(\varepsilon)(d^* + d_1).$$

We give different analysis depending on whether f^* points to η_1 or η_2; this is different from our type E analysis, where our bounds are the same in both cases. Formally, we partition the amenable event $\mathcal{A}$ as the union of $\mathcal{S}_1 \cap \mathcal{A}$, $\mathcal{S}_2 \cap \mathcal{A}$, $\mathcal{T}_1 \cap \mathcal{A}$, and $\mathcal{T}_2 \cap \mathcal{A}$. We upper-bound $\Delta_{\mathcal{A}}(c)$ by

$$\begin{aligned} \Delta_{\mathcal{A}}(c) &\le \Pr[\mathcal{S}_1 \cap \mathcal{A}]\delta_{\mathcal{S}_1\cap\mathcal{A}}(c) + \Pr[\mathcal{S}_2 \cap \mathcal{A}]\delta_{\mathcal{S}_2\cap\mathcal{A}}(c) + \Pr[\mathcal{T}_1 \cap \mathcal{A}]\delta_{\mathcal{T}_1\cap\mathcal{A}}(c) + \Pr[\mathcal{T}_2 \cap \mathcal{A}]\delta_{\mathcal{T}_2\cap\mathcal{A}}(c) \\ &\le \Pr[\mathcal{S}_1]\delta_{\mathcal{S}_1\cap\mathcal{A}}(c) + \Pr[\mathcal{S}_2]\delta_{\mathcal{S}_2\cap\mathcal{A}}(c) + \Pr[\mathcal{T}_1]\delta_{\mathcal{T}_1\cap\mathcal{A}}(c) + \Pr[\mathcal{T}_2]\delta_{\mathcal{T}_2\cap\mathcal{A}}(c) + O(\varepsilon)(d^* + d_1). \end{aligned} \tag{D.47}$$

The probabilities $\Pr[\mathcal{S}_1], \Pr[\mathcal{S}_2], \Pr[\mathcal{T}_1], \Pr[\mathcal{T}_2]$ are given in Table 1. We proceed by showing upper-bounds for the δ values, the potential changes of client c on the worst-case swap set $\mathcal{P}$, for far clients of type A and B in the following subsections.

D.1.1 Far clients of type A: $f_1 = \eta_1$

Simple swaps with $\tau(f^*) = \eta_1$ Type A clients have $f_1 = \eta_1$, which is the same as $\tau(f^*)$, so we have $\langle\langle f^*\rangle\rangle = \langle\langle \neg f_1\rangle\rangle$ in $\mathcal{P}$ by implication (ii) of amenability. On that swap, the client can be served by f^*. Therefore,

$$(\delta_{\langle\langle f^*, \neg f_1\rangle\rangle}) \qquad \delta_{\mathcal{S}_1\cap\mathcal{A}}(c) \le \boxed{(1+\alpha\beta)\, d^* - (1+\alpha\beta)\, d_1}.$$

Simple swaps with $\tau(f^*) = \eta_2$ Since $\tau(f^*) = \eta_2 \ne f_1$, we know $\langle\langle f^*\rangle\rangle \ne \langle\langle \neg f_1\rangle\rangle$ by implication (Siv) of amenability. On swap $\langle\langle f^*\rangle\rangle$, c can be served by both f^* and f_1. On swap $\langle\langle \neg f_1\rangle\rangle$, c can be served by η_2 (by implication (Siii) of amenability). Note that $d(c, \eta_2) \le d^* + {}^1\!/\!_\rho \cdot (d^* + d_1)$ by (C.42). Therefore,

$$\begin{aligned} (\delta_{\langle\langle f^*\rangle\rangle}) \qquad \delta_{\mathcal{S}_2\cap\mathcal{A}}(c) &\le d_1 + \beta\, d^* - (1+\alpha\beta)\, d_1 \\ (\delta_{\langle\langle \neg f_1\rangle\rangle}) \qquad &\quad + (1+\alpha\beta)(d^* + {}^1\!/\!_\rho \cdot (d^* + d_1)) - (1+\alpha\beta)\, d_1 \\ &= \boxed{((1 + {}^1\!/\!_\rho)(1+\alpha\beta) + \beta)\, d^* - ((1 - {}^1\!/\!_\rho)(1+\alpha\beta) + \alpha\beta)\, d_1}. \end{aligned}$$

Tree swaps with $\tau(f^*) = \eta_1$ We have $\langle\langle f^*\rangle\rangle = \langle\langle \neg f_1\rangle\rangle$ by implication (ii) of amenability. On that swap, the client can be served by $\langle\langle f^*\rangle\rangle$. Therefore,

$$(\delta_{\langle\langle f^*, \neg f_1\rangle\rangle}) \qquad \delta_{\mathcal{T}_1\cap\mathcal{A}}(c) \le \boxed{(1+\alpha\beta)\, d^* - (1+\alpha\beta)\, d_1}.$$

Tree swaps with $\tau(f^*) = \eta_2$ If $\langle\langle f^*\rangle\rangle = \langle\langle \neg f_1\rangle\rangle$, then we have the same bound as above:

$$\sum_{(P,Q)\in\mathcal{P}} \delta_{(P,Q)}(c) \le (1+\alpha\beta)\, d^* - (1+\alpha\beta)\, d_1. \tag{D.48}$$

If $\langle\langle f^*\rangle\rangle \ne \langle\langle \neg f_1\rangle\rangle$, then on swap $\langle\langle \neg f_1\rangle\rangle$, c can be served by η_2 and $\pi(f_1)$, by implications (ii) and (Tii) of amenability. We already showed $d(c, \eta_2) \le d^* + {}^1\!/\!_\rho \cdot (d^* + d_1)$. We also have $d(c, \pi(f_1)) \le 2d_1 + d^*$ by (6.8). Therefore,

$$(\delta_{\langle\langle f^*\rangle\rangle}) \qquad \sum_{(P,Q)\in\mathcal{P}} \delta_{(P,Q)}(c) \le d_1 + \beta\, d^* - (1+\alpha\beta)\, d_1$$

$(\delta_{\langle\langle \neg f_1 \rangle\rangle})$ $\qquad + (d^* + 1/\rho \cdot (d^* + d_1)) + \beta(2d_1 + d^*) - (1 + \alpha\beta)\, d_1$

(D.49) $\qquad = (1 + 1/\rho + 2\beta)\, d^* - (1 + 2\alpha\beta - 2\beta - 1/\rho)\, d_1.$

For our choice of α, β, (D.49) is larger than (D.48), so we have

$$\delta_{\mathcal{T}_2 \cap \mathcal{A}}(c) \leq \boxed{(1 + 1/\rho + 2\beta)\, d^* - (1 + 2\alpha\beta - 2\beta - 1/\rho)\, d_1}.$$

Summarizing, we have

$$\begin{aligned}
\delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) &\leq (1 + \alpha\beta)\, d^* - (1 + \alpha\beta)\, d_1 && = 1.6\, d^* - 1.6\, d_1 \\
\delta_{\mathcal{S}_2 \cap \mathcal{A}}(c) &\leq ((1 + 1/\rho)(1 + \alpha\beta) + \beta)\, d^* - ((1 - 1/\rho)(1 + \alpha\beta) + \alpha\beta)\, d_1 && \\
& && = (1.8 + 1.6/\rho)\, d^* - (2.2 - 1.6/\rho)\, d_1 \\
\delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) &\leq (1 + \alpha\beta)\, d^* - (1 + \alpha\beta)\, d_1 && = 1.6\, d^* - 1.6\, d_1 \\
\delta_{\mathcal{T}_2 \cap \mathcal{A}}(c) &\leq (1 + 1/\rho + 2\beta)\, d^* - (1 + 2\alpha\beta - 2\beta - 1/\rho)\, d_1 && = (1.4 + 1/\rho)\, d^* - (1.8 - 1/\rho)\, d_1
\end{aligned}$$

We now combine these inequalities using (D.47). If $\rho(f^*) \leq 2/3$, we have $\Pr[\mathcal{S}_1] = \Pr[\mathcal{T}_1] = 1/2$ and $\Pr[\mathcal{S}_2] = \Pr[\mathcal{T}_2] = 0$. Therefore,

$$\begin{aligned}
\Delta_{\mathcal{A}}(c) &\leq 1/2 \cdot \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) + 1/2 \cdot \delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) + O(\varepsilon)(d^* + d_1) \\
&\leq \boxed{1.6\, d^* - 1.6\, d_1} + O(\varepsilon)(d^* + d_1).
\end{aligned}$$

If $2/3 < \rho(f^*) \leq 3/4$, we have $\Pr[\mathcal{S}_1] = 1/2, \Pr[\mathcal{S}_2] = 0, \Pr[\mathcal{T}_1] = 1/4, \Pr[\mathcal{T}_2] = 1/4$. Therefore,

$$\begin{aligned}
\Delta_{\mathcal{A}}(c) &\leq 1/2 \cdot \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_2 \cap \mathcal{A}}(c) + O(\varepsilon)(d^* + d_1) \\
&\leq 3/4 \cdot (1.6\, d^* - 1.6\, d_1) + 1/4 \cdot ((1.4 + 3/2)d^* - (1.8 - 3/2)d_1) + O(\varepsilon)(d^* + d_1) \\
&= \boxed{1.925\, d^* - 1.275\, d_1} + O(\varepsilon)(d^* + d_1).
\end{aligned}$$

If $\rho(f^*) > 3/4$, we have $\Pr[\mathcal{S}_1] = 5/4 - \rho, \Pr[\mathcal{S}_2] = \rho - 3/4, \Pr[\mathcal{T}_1] = \Pr[\mathcal{T}_2] = 1/4$. Therefore,

$$\begin{aligned}
\Delta_{\mathcal{A}}(c) &\leq (5/4 - \rho) \cdot \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) + (\rho - 3/4) \cdot \delta_{\mathcal{S}_2 \cap \mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_2 \cap \mathcal{A}}(c) + O(\varepsilon)(d^* + d_1) \\
&\leq (3 + 0.2\rho - 0.95/\rho)\, d^* - (0.6\rho + 0.95/\rho - 0.4)\, d_1 + O(\varepsilon)(d^* + d_1) \\
&\leq (3 + 0.2 - 0.95)\, d^* - (0.6 + 0.95 - 0.4)\, d_1 + O(\varepsilon)(d^* + d_1) \\
&= \boxed{2.25\, d^* - 1.15\, d_1} + O(\varepsilon)(d^* + d_1).
\end{aligned}$$

D.1.2 Far clients of type B: $f_1 = \eta_2$

When c is a far client of type B, we have $\langle\langle f^* \rangle\rangle = \langle\langle \neg f_1 \rangle\rangle$ on $\mathcal{S}_2 \cap \mathcal{A}$ and $\mathcal{T}_2 \cap \mathcal{A}$. This is exactly the situation for type A clients on $\mathcal{S}_1 \cap \mathcal{A}$ and $\mathcal{T}_1 \cap \mathcal{A}$. Therefore, we have the same bound for all of these cases:

$(\delta_{\langle\langle f^*, \neg f_1 \rangle\rangle})$ $\qquad \delta_{\mathcal{S}_2 \cap \mathcal{A}}(c) \leq \boxed{(1 + \alpha\beta)\, d^* - (1 + \alpha\beta)\, d_1},$

$(\delta_{\langle\langle f^*, \neg f_1 \rangle\rangle})$ $\qquad \delta_{\mathcal{T}_2 \cap \mathcal{A}}(c) \leq \boxed{(1 + \alpha\beta)\, d^* - (1 + \alpha\beta)\, d_1}.$

We continue to bound $\delta_{\mathcal{S}_1 \cap \mathcal{A}}(c)$ and $\delta_{\mathcal{T}_1 \cap \mathcal{A}}(c)$.

Simple swaps with $\tau(f^*) = \eta_1$ By implication (Siv) of amenability, we have $\langle\langle f^* \rangle\rangle \neq \langle\langle \neg f_1 \rangle\rangle$. On swap $\langle\langle f^* \rangle\rangle$, the client can be served by f^* and f_1. On swap $\langle\langle \neg f_1 \rangle\rangle$, the client can be served by η_1 (by implication (Siii) of amenability). Also, $d(c, \eta_1) \le d^* + \rho(d^* + d_1)$ by (C.44). Therefore,

$$
\begin{aligned}
&(\delta_{\langle\langle f^* \rangle\rangle}) && \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) \le d_1 + \beta\, d^* - (1 + \alpha\beta)\, d_1 \\
&(\delta_{\langle\langle \neg f_1 \rangle\rangle}) && \qquad + (1+\alpha\beta)(d^* + \rho(d^* + d_1)) - (1+\alpha\beta)\, d_1 \\
& && \quad = \boxed{((1+\rho)(1+\alpha\beta) + \beta)\, d^* - ((1-\rho)(1+\alpha\beta) + \alpha\beta)\, d_1}.
\end{aligned}
$$

Tree swaps with $\tau(f^*) = \eta_1$ If $\langle\langle f^* \rangle\rangle = \langle\langle \neg f_1 \rangle\rangle$, then we have

$$
\sum_{(P,Q) \in \mathcal{P}} \delta_{(P,Q)}(c) \le (1+\alpha\beta)\, d^* - (1+\alpha\beta)\, d_1. \tag{D.50}
$$

If $\langle\langle f^* \rangle\rangle \neq \langle\langle \neg f_1 \rangle\rangle$, then on swap $\langle\langle \neg f_1 \rangle\rangle$, c can be served by η_1 and $\pi(f_1)$ by implications (ii) and (Tii) of amenability. We showed $d(c, \eta_1) \le d^* + \rho(d^* + d_1)$. We also have $d(c, \pi(f_1)) \le 2d_1 + d^*$ by (6.8). Therefore,

$$
\begin{aligned}
&(\delta_{\langle\langle f^* \rangle\rangle}) && \sum_{(P,Q) \in \mathcal{P}} \delta_{(P,Q)}(c) \le d_1 + \beta\, d^* - (1+\alpha\beta)\, d_1 \\
&(\delta_{\langle\langle \neg f_1 \rangle\rangle}) && \qquad + (d^* + \rho(d^* + d_1)) + \beta(2d_1 + d^*) - (1+\alpha\beta)\, d_1 \\
&(\text{D.51}) && \quad = (1 + \rho + 2\beta)\, d^* - (1 + 2\alpha\beta - 2\beta - \rho)\, d_1.
\end{aligned}
$$

Taking the maximum of (D.50) and (D.51) using $\alpha = 3, \beta = 0.2$, we have

$$
\delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) \le \boxed{(1.4 + \max\{\rho, 0.2\})\, d^* - (1.8 - \max\{\rho, 0.2\})\, d_1}.
$$

Summarizing, we have

$$
\begin{aligned}
\delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) &\le ((1+\rho)(1+\alpha\beta) + \beta)\, d^* - ((1-\rho)(1+\alpha\beta) + \alpha\beta)\, d_1 \\
&= (1.8 + 1.6\rho)\, d^* - (2.2 - 1.6\rho)\, d_1 \\
\delta_{\mathcal{S}_2 \cap \mathcal{A}}(c) &\le (1+\alpha\beta)\, d^* - (1+\alpha\beta)\, d_1 = 1.6\, d^* - 1.6\, d_1 \\
\delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) &\le (1.4 + \max\{\rho, 0.2\})\, d^* - (1.8 - \max\{\rho, 0.2\})\, d_1 \\
\delta_{\mathcal{T}_2 \cap \mathcal{A}}(c) &\le (1+\alpha\beta)\, d^* - (1+\alpha\beta)\, d_1 = 1.6\, d^* - 1.6\, d_1
\end{aligned}
$$

We now combine these inequalities using (D.47). If $\rho(f^*) \le 2/3$, we have $\Pr[\mathcal{S}_1] = \Pr[\mathcal{T}_1] = 1/2$ and $\Pr[\mathcal{S}_2] = \Pr[\mathcal{T}_2] = 0$. Therefore,

$$
\begin{aligned}
\Delta_{\mathcal{A}}(c) &\le 1/2 \cdot \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) + 1/2 \cdot \delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) + O(\varepsilon)(d^* + d_1) \\
&\le 1/2 \cdot ((1.8 + 1.6 \times 2/3)d^* - (2.2 - 1.6 \times 2/3)d_1) \\
&\quad + 1/2 \cdot ((1.4 + 2/3)d^* - (1.8 - 2/3)d_1) \\
&\quad + O(\varepsilon)(d^* + d_1) \\
&\le \boxed{2.46667d^* - 1.13333d_1} + O(\varepsilon)(d^* + d_1).
\end{aligned}
$$

If $2/3 < \rho(f^*) \le 3/4$, we have $\Pr[\mathcal{S}_1] = 1/2, \Pr[\mathcal{S}_2] = 0, \Pr[\mathcal{T}_1] = 1/4, \Pr[\mathcal{T}_2] = 1/4$. Therefore,

$$
\Delta_{\mathcal{A}}(c) \le 1/2 \cdot \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_2 \cap \mathcal{A}}(c) + O(\varepsilon)(d^* + d_1)
$$

$$
\begin{aligned}
&\le {}^1\!/\!_2 \cdot ((1.8 + 1.6 \times {}^3\!/\!_4)d^* - (2.2 - 1.6 \times {}^3\!/\!_4)d_1) \\
&\quad + {}^1\!/\!_4 \cdot ((1.4 + {}^3\!/\!_4)d^* - (1.8 - {}^3\!/\!_4)d_1) \\
&\quad + {}^1\!/\!_4 \cdot (1.6d^* - 1.6d_1) \\
&\quad + O(\varepsilon)(d^* + d_1) \\
&= \boxed{2.4375d^* - 1.1625d_1} + O(\varepsilon)(d^* + d_1).
\end{aligned}
$$

If $\rho(f^*) > {}^3\!/\!_4$, we have $\Pr[\mathcal{S}_1] = {}^5\!/\!_4 - \rho, \Pr[\mathcal{S}_2] = \rho - {}^3\!/\!_4, \Pr[\mathcal{T}_1] = \Pr[\mathcal{T}_2] = {}^1\!/\!_4$. Therefore,

$$
\begin{aligned}
\Delta_{\mathcal{A}}(c) &\le ({}^5\!/\!_4 - \rho) \cdot \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) + (\rho - {}^3\!/\!_4) \cdot \delta_{\mathcal{S}_2 \cap \mathcal{A}}(c) + {}^1\!/\!_4 \cdot \delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) + {}^1\!/\!_4 \cdot \delta_{\mathcal{T}_2 \cap \mathcal{A}}(c) + O(\varepsilon)(d^* + d_1) \\
&\le (1.8 + 2.05\rho - 1.6\rho^2)d^* - (2.4 - 2.85\rho + 1.6\rho^2)d_1 + O(\varepsilon)(d^* + d_1) \\
&\le (1.8 + 2.05 \cdot {}^3\!/\!_4 - 1.6 \cdot ({}^3\!/\!_4)^2)d^* - (2.4 - 2.85 \cdot {}^{2.85}\!/\!_{3.2} + 1.6 \cdot ({}^{2.85}\!/\!_{3.2})^2)d_1 + O(\varepsilon)(d^* + d_1) \\
&\le \boxed{2.4375d^* - 1.13085d_1} + O(\varepsilon)(d^* + d_1).
\end{aligned}
$$

D.2 Proof of (6.18): Close Clients of Type A

In this section, we show that for any close case client c with type A, we have

$$\Delta_{\mathcal{A}}(c) \le 2.375\, d^*(c) - 0.9\, d_1(c).$$

D.2.1 Clients with $\rho(f^*) \le {}^2\!/\!_3$

We first consider the case where $\rho(f^*) \le {}^2\!/\!_3$. Our analysis for this case is very simple: we directly use $\Delta_{\mathcal{A}}(c) \le \Pr[\mathcal{A}]\delta_{\mathcal{A}}(c) \le \delta_{\mathcal{A}}(c) + O(\varepsilon)(d^* + d_1)$ without considering sub-events of $\mathcal{A}$.

Let us fix a generic swap set $\mathcal{P}$ generated on the amenable event $\mathcal{A}$. $\rho(f^*) \le {}^2\!/\!_3$ implies that $\tau(f^*)$ always equals to $\eta_1 = f_1$. By implication (ii) of amenability, we have $\langle\!\langle f^* \rangle\!\rangle = \langle\!\langle \neg f_1 \rangle\!\rangle$ in $\mathcal{P}$. Therefore,

$$
\begin{aligned}
&(\delta_{\langle\!\langle f^*, \neg f_1 \rangle\!\rangle}) && \delta_{\mathcal{A}}(c) \le (1 + \alpha\beta)\, d^* - d_1 - \beta\, d_2 \\
&(\delta_{\langle\!\langle \neg f_2 \rangle\!\rangle}) && \qquad + (1 + \alpha\beta)\, d_1 - d_1 - \beta\, d_2 \\
& && \quad = (1 + \alpha\beta)\, d^* - (1 - \alpha\beta)\, d_1 - 2\beta\, d_2.
\end{aligned}
$$

Note that this bound also holds when $\langle\!\langle \neg f_2 \rangle\!\rangle = \langle\!\langle f^*, \neg f_1 \rangle\!\rangle$, because our bound for $\langle\!\langle f^*, \neg f_1 \rangle\!\rangle$ does not require f_2 to remain open after the swap and $\delta_{\langle\!\langle \neg f_2 \rangle\!\rangle}$ is non-negative.

If $d_1 \le d^*$, then we have

$$
\begin{aligned}
&(d_1 \le d^*) && \delta_{\mathcal{A}}(c) \le (1 + \alpha\beta + \beta)\, d^* - (1 - \alpha\beta + \beta)\, d_1 - 2\beta\, d_2 \\
&(d_2 \ge d_1) && \qquad \le (1 + \alpha\beta + \beta)\, d^* - (1 - \alpha\beta + 3\beta)\, d_1 \\
& && \qquad = \boxed{1.8\, d^* - d_1}.
\end{aligned}
$$

If $d_1 \ge d^*$, we have

$$
\begin{aligned}
& && \delta_{\mathcal{A}}(c) \le (1 + \alpha\beta)\, d^* - (1 - \alpha\beta)\, d_1 - 2\beta\, d_2 \\
&(\text{averaging (C.40) with } d_2 \ge d_1) && \qquad \le (1 + \alpha\beta)\, d^* - (1 - \alpha\beta)\, d_1 - 2\beta \left(\frac{1 + {}^1\!/\!_\rho}{2} d_1 - \frac{1 + {}^1\!/\!_\rho}{2} d^* \right)
\end{aligned}
$$

$$(\rho \le 2/3 \text{ and } d^* \le d_1) \qquad \le (1+\alpha\beta)\,d^* - (1-\alpha\beta)\,d_1 - 2\beta\left(\frac{5}{4}d_1 - \frac{5}{4}d^*\right)$$
$$= (1+\alpha\beta+2.5\beta)\,d^* - (1-\alpha\beta+2.5\beta)\,d_1$$
$$= \boxed{2.1\,d^* - 0.9\,d_1}.$$

D.2.2 Clients with $\rho(f^*) > 2/3$

Now we turn to close clients of type A with $\rho(f^*) > 2/3$. Our analysis for simple swaps adopts the usual strategy:

$$\begin{aligned}\Delta_{\mathcal{S}\cap\mathcal{A}}(c) &\le \Pr[\mathcal{S}_1 \cap \mathcal{A}]\delta_{\mathcal{S}_1\cap\mathcal{A}}(c) + \Pr[\mathcal{S}_2 \cap \mathcal{A}]\delta_{\mathcal{S}_2\cap\mathcal{A}}(c)\\ &\le \Pr[\mathcal{S}_1]\delta_{\mathcal{S}_1\cap\mathcal{A}}(c) + \Pr[\mathcal{S}_2]\delta_{\mathcal{S}_2\cap\mathcal{A}}(c) + O(\varepsilon)(d^* + d_1).\end{aligned}$$

However, we will be a little more careful in our tree swaps analysis. We further partition the tree events $\mathcal{T}_1$ and $\mathcal{T}_2$ as $\mathcal{T}_1 = \mathcal{T}_{11} \cup \mathcal{T}_{12}$ and $\mathcal{T}_2 = \mathcal{T}_{21} \cup \mathcal{T}_{22}$ in the following way. $\mathcal{T}_{11}$ is defined as the intersection of $\mathcal{T}_1$ and the event that $\langle\langle f^*\rangle\rangle$ is the only swap closing any facility in $\{f_1, f_2\}$. $\mathcal{T}_{21}$ is defined as the intersection of $\mathcal{T}_2$ and the event that there is a swap which closes both f_1 and f_2 but does not open the original copy of f^*. $\mathcal{T}_{12}$ and $\mathcal{T}_{22}$ are defined accordingly: $\mathcal{T}_{12} = \mathcal{T}_1\backslash\mathcal{T}_{11}$ and $\mathcal{T}_{22} = \mathcal{T}_2\backslash\mathcal{T}_{21}$.

Recall that $\rho(f^*) > 2/3$ implies $\Pr[\mathcal{T}_1] = \Pr[\mathcal{T}_2] = 1/4$. The naive way to bound $\Delta_{\mathcal{T}\cap\mathcal{A}}(c)$ is by the following:

$$\begin{aligned}\Delta_{\mathcal{T}\cap\mathcal{A}}(c) &\le \Pr[\mathcal{T}_1]\delta_{\mathcal{T}_1\cap\mathcal{A}}(c) + \Pr[\mathcal{T}_2]\delta_{\mathcal{T}_2\cap\mathcal{A}}(c) + O(\varepsilon)(d^* + d_1)\\ &= 1/4 \cdot \max\{\delta_{\mathcal{T}_{11}\cap\mathcal{A}}(c), \delta_{\mathcal{T}_{12}\cap\mathcal{A}}(c)\} + 1/4 \cdot \max\{\delta_{\mathcal{T}_{21}\cap\mathcal{A}}(c), \delta_{\mathcal{T}_{22}\cap\mathcal{A}}(c)\} + O(\varepsilon)(d^* + d_1).\end{aligned}$$

If we ignore the $O(\varepsilon)(d^* + d_1)$ term, the above bound is equal to $1/4$ times the maximum of all four sums: $\delta_{\mathcal{T}_{11}\cap\mathcal{A}}(c) + \delta_{\mathcal{T}_{21}\cap\mathcal{A}}(c), \delta_{\mathcal{T}_{11}\cap\mathcal{A}}(c) + \delta_{\mathcal{T}_{22}\cap\mathcal{A}}(c), \delta_{\mathcal{T}_{12}\cap\mathcal{A}}(c) + \delta_{\mathcal{T}_{21}\cap\mathcal{A}}(c), \delta_{\mathcal{T}_{12}\cap\mathcal{A}}(c) + \delta_{\mathcal{T}_{22}\cap\mathcal{A}}(c)$. However, by relating the probabilities of $\mathcal{T}_{11}$ and $\mathcal{T}_{21}$, we have the following lemma (proved in Appendix E.4), which gives an improved bound by not taking $\delta_{\mathcal{T}_{12}\cap\mathcal{A}}(c)+\delta_{\mathcal{T}_{21}\cap\mathcal{A}}(c)$ into the maximum.

Lemma D.1 (Type A averaging)**.** *For a close client of type* A *with $\rho(f^*) > 2/3$, we have*

$$\Delta_{\mathcal{T}\cap\mathcal{A}}(c) \le 1/4 \cdot \max\{\delta_{\mathcal{T}_{11}\cap\mathcal{A}} + \delta_{\mathcal{T}_{21}\cap\mathcal{A}}, \delta_{\mathcal{T}_{11}\cap\mathcal{A}} + \delta_{\mathcal{T}_{22}\cap\mathcal{A}}, \delta_{\mathcal{T}_{12}\cap\mathcal{A}} + \delta_{\mathcal{T}_{22}\cap\mathcal{A}}\} + O(\varepsilon)(d^* + d_1).$$

We now proceed to show upper bounds for the worst-case potential change on each event.

Simple swaps with $\tau(f^*) = \eta_1$ We have $\langle\langle f^*\rangle\rangle = \langle\langle \neg f_1\rangle\rangle \ne \langle\langle \neg f_2\rangle\rangle$ by implications (ii) and (Siv) of amenability. On swap $\langle\langle f^*\rangle\rangle$, the client can be served by f^*, and on $\langle\langle \neg f_2\rangle\rangle$, the client can be served by f_1 (at distance d_1) and η_2 (at distance $\le d^* + 1/\rho \cdot (d^* + d_1)$), by implication (Siii) of amenability. Therefore,

$$\begin{aligned}(\delta_{\langle\langle f^*,\neg f_1\rangle\rangle}) \qquad \delta_{\mathcal{S}_1\cap\mathcal{A}}(c) &\le (1+\alpha\beta)\,d^* - d_1 - \beta\,d_2\\ (\delta_{\langle\langle \neg f_2\rangle\rangle}) \qquad &\quad + d_1 + \beta(d^* + 1/\rho \cdot (d^* + d_1)) - d_1 - \beta\,d_2\\ &= \boxed{(1+\alpha\beta+\beta+\beta/\rho)\,d^* - (1-\beta/\rho)\,d_1 - 2\beta\,d_2}.\end{aligned}$$

Simple swaps with $\tau(f^*) = \eta_2$ By implications (Siii) and (Siv) of amenability, the three swaps $\langle\langle f^*\rangle\rangle$,$\langle\langle \neg f_1\rangle\rangle$, $\langle\langle \neg f_2\rangle\rangle$ are all different. On swap $\langle\langle f^*\rangle\rangle$, the client can be served by f^* and f_1. On swap $\langle\langle \neg f_1 \rangle\rangle$, the client can be served by f_2 and η_2. On swap $\langle\langle \neg f_2 \rangle\rangle$, the client can be served by f_1 and η_2. Therefore,

$$\begin{aligned}
&(\delta_{\langle\langle f^* \rangle\rangle}) && \delta_{\mathcal{S}_2 \cap \mathcal{A}}(c) \le d^* + \beta\, d_1 - d_1 - \beta\, d_2 \\
&(\delta_{\langle\langle \neg f_1 \rangle\rangle}) && \qquad + d_2 + \beta(d^* + {}^1\!/\!_\rho \cdot (d^* + d_1)) - d_1 - \beta\, d_2 \\
&(\delta_{\langle\langle \neg f_2 \rangle\rangle}) && \qquad + d_1 + \beta(d^* + {}^1\!/\!_\rho \cdot (d^* + d_1)) - d_1 - \beta\, d_2 \\
& && \qquad = \boxed{(1 + 2\beta + {}^{2\beta}\!/\!_\rho)\, d^* - (2 - \beta - {}^{2\beta}\!/\!_\rho)\, d_1 + (1 - 3\beta)\, d_2}.
\end{aligned}$$

Tree swaps with $\tau(f^*) = \eta_1$ On $\mathcal{T}_{11} \cap \mathcal{A}$, $\langle\langle f^* \rangle\rangle$ is the only swap closing any facility in $\{f_1, f_2\}$ by the definition of $\mathcal{T}_{11}$. In other words, both $\langle\langle \neg f_1 \rangle\rangle$ and $\langle\langle \neg f_2 \rangle\rangle$ coincide with $\langle\langle f^* \rangle\rangle$ as long as they exist. Therefore,

$$(\delta_{\langle\langle f^*, \neg f_1, \neg f_2 \rangle\rangle}) \qquad \delta_{\mathcal{T}_{11} \cap \mathcal{A}}(c) \le \boxed{(1 + \alpha\beta)\, d^* - d_1 - \beta\, d_2}.$$

On $\mathcal{T}_{12} \cap \mathcal{A}$, we have $\langle\langle f^* \rangle\rangle = \langle\langle \neg f_1 \rangle\rangle \ne \langle\langle \neg f_2 \rangle\rangle$ by implication (ii) of amenability. On swap $\langle\langle f^* \rangle\rangle$, the client can be served by f^*. On swap $\langle\langle \neg f_2 \rangle\rangle$, the client can be served by f_1 and $\pi(f_2)$ by implication (Tii) of amenability. We have $d(c, \pi(f_2)) \le 2d_2 + d^*$ by (C.46). Therefore,

$$\begin{aligned}
&(\delta_{\langle\langle f^*, \neg f_1 \rangle\rangle}) && \delta_{\mathcal{T}_{12} \cap \mathcal{A}}(c) \le (1 + \alpha\beta)\, d^* - d_1 - \beta\, d_2 \\
&(\delta_{\langle\langle \neg f_2 \rangle\rangle}) && \qquad + d_1 + \beta(2\, d_2 + d^*) - d_1 - \beta\, d_2 \\
& && \qquad = \boxed{(1 + \alpha\beta + \beta)\, d^* - d_1}.
\end{aligned}$$

Tree swaps with $\tau(f^*) = \eta_2$ On $\mathcal{T}_{21} \cap \mathcal{A}$, we have $\langle\langle \neg f_1 \rangle\rangle = \langle\langle \neg f_2 \rangle\rangle \ne \langle\langle f^* \rangle\rangle$. On swap $\langle\langle f^* \rangle\rangle$, the client can be served by f^* and f_1. On swap $\langle\langle \neg f_1, \neg f_2 \rangle\rangle$, the client can be served by η_2 and $\pi(f_1)$ by implications (ii) and (Tii) of amenability. We have $d(c, \eta_2) \le d^* + {}^1\!/\!_\rho \cdot (d^* + d_1)$ and $d(c, \pi(f_1)) \le 2d_1 + d^*$. Therefore,

$$\begin{aligned}
&(\delta_{\langle\langle f^* \rangle\rangle}) && \delta_{\mathcal{T}_{21} \cap \mathcal{A}}(c) \le d^* + \beta\, d_1 - d_1 - \beta\, d_2 \\
&(\delta_{\langle\langle \neg f_1, \neg f_2 \rangle\rangle}) && \qquad + (d^* + {}^1\!/\!_\rho(d^* + d_1)) + \beta(2d_1 + d^*) - d_1 - \beta\, d_2 \\
& && \qquad = \boxed{(2 + \beta + {}^1\!/\!_\rho)\, d^* - (2 - 3\beta - {}^1\!/\!_\rho)\, d_1 - 2\beta\, d_2}.
\end{aligned}$$

On $\mathcal{T}_{22} \cap \mathcal{A}$, we first consider the case where all three swaps $\langle\langle f^* \rangle\rangle, \langle\langle \neg f_1 \rangle\rangle, \langle\langle \neg f_2 \rangle\rangle$ are different. On swap $\langle\langle f^* \rangle\rangle$, the client can be served by f^*. On swap $\langle\langle \neg f_1 \rangle\rangle$, the client can be served by f_2 and $\pi(f_1)$ (at distance $\le 2d_1 + d^*$), by implication (Tii) of amenability. On swap $\langle\langle \neg f_2 \rangle\rangle$, the client can be served by f_1 and $\pi(f_2)$ (at distance $\le 2d_2 + d^*$), again by implication (Tii) of amenability. Therefore,

$$\begin{aligned}
&(\delta_{\langle\langle f^* \rangle\rangle}) && \delta_{\mathcal{T}_{22} \cap \mathcal{A}}(c) \le (1 + \alpha\beta)\, d^* - d_1 - \beta\, d_2 \\
&(\delta_{\langle\langle \neg f_1 \rangle\rangle}) && \qquad + d_2 + \beta(2d_1 + d^*) - d_1 - \beta\, d_2 \\
&(\delta_{\langle\langle \neg f_2 \rangle\rangle}) && \qquad + d_1 + \beta(2d_2 + d^*) - d_1 - \beta\, d_2 \\
& && \qquad \le \boxed{(1 + \alpha\beta + 2\beta)\, d^* - (2 - 2\beta)\, d_1 + (1 - \beta)\, d_2}.
\end{aligned}$$

Since our bound for $\delta_{\langle\langle f^* \rangle\rangle}(c)$ doesn't require either f_1 or f_2 to remain open after the swap, and both $\delta_{\langle\langle \neg f_1 \rangle\rangle}$ and $\delta_{\langle\langle \neg f_2 \rangle\rangle}$ are non-negative, the above bound also holds when $\langle\langle \neg f_1 \rangle\rangle$ and/or $\langle\langle \neg f_2 \rangle\rangle$ coincides with $\langle\langle f^* \rangle\rangle$.

Summarizing, we have

$$\begin{aligned}
\delta_{\mathcal{S}_1\cap\mathcal{A}}(c) \leq &(1+\alpha\beta+\beta+\beta/\rho)\, d^* - (1-\beta/\rho)\, d_1 - 2\beta\, d_2 \\
&= (1.8+0.2/\rho)\, d^* - (1-0.2/\rho)\, d_1 - 0.4\, d_2 \\
\delta_{\mathcal{S}_2\cap\mathcal{A}}(c) \leq &(1+2\beta+2\beta/\rho)\, d^* - (2-\beta-2\beta/\rho)\, d_1 + (1-3\beta)\, d_2 \\
&= (1.4+0.4/\rho)\, d^* - (1.8-0.4/\rho)\, d_1 + 0.4\, d_2 \\
\delta_{\mathcal{T}_{11}\cap\mathcal{A}}(c) \leq &(1+\alpha\beta)\, d^* - d_1 - \beta\, d_2 \quad = 1.6\, d^* - d_1 - 0.2\, d_2 \\
\delta_{\mathcal{T}_{12}\cap\mathcal{A}}(c) \leq &(1+\alpha\beta+\beta)\, d^* - d_1 \quad = 1.8\, d^* - d_1 \\
\delta_{\mathcal{T}_{21}\cap\mathcal{A}}(c) \leq &(2+\beta+1/\rho)\, d^* - (2-3\beta-1/\rho)\, d_1 - 2\beta\, d_2 \\
&= (2.2+1/\rho)\, d^* - (1.4-1/\rho)\, d_1 - 0.4\, d_2 \\
\delta_{\mathcal{T}_{22}\cap\mathcal{A}}(c) \leq &(1+\alpha\beta+2\beta)\, d^* - (2-2\beta)\, d_1 + (1-\beta)\, d_2 \\
&= 2\, d^* - 1.6\, d_1 + 0.8\, d_2
\end{aligned}$$

We now combine these bounds to show an upper bound for $\Delta_{\mathcal{A}}(c)$ using Lemma D.1. Note that our bound for $\delta_{\mathcal{T}_{11}\cap\mathcal{A}}(c)$ is smaller than our bound for $\delta_{\mathcal{T}_{12}\cap\mathcal{A}}(c)$, so we only need to consider cases where the maximum in Lemma D.1 is attained at either $\delta_{\mathcal{T}_{11}\cap\mathcal{A}}(c) + \delta_{\mathcal{T}_{21}\cap\mathcal{A}}(c)$ or $\delta_{\mathcal{T}_{12}\cap\mathcal{A}}(c) + \delta_{\mathcal{T}_{22}\cap\mathcal{A}}(c)$.

When $2/3 < \rho(f^*) \leq 3/4$, we have $\Pr[\mathcal{S}_1] = 1/2$ and $\Pr[\mathcal{S}_2] = 0$. Therefore, if the maximum in Lemma D.1 is attained at $\delta_{\mathcal{T}_{11}\cap\mathcal{A}}(c) + \delta_{\mathcal{T}_{21}\cap\mathcal{A}}(c)$, we have

$$\begin{aligned}
\Delta_{\mathcal{A}}(c) &\leq 1/2 \cdot \delta_{\mathcal{S}_1\cap\mathcal{A}}(c) + 1/4 \cdot (\delta_{\mathcal{T}_{11}\cap\mathcal{A}}(c) + \delta_{\mathcal{T}_{21}\cap\mathcal{A}}(c)) + O(\varepsilon)(d^* + d_1) \\
&\leq (1.85 + 0.35/\rho)d^* - (1.1 - 0.35/\rho)d_1 - 0.35d_2 + O(\varepsilon)(d^* + d_1) \\
(d_2 \geq d_1) \quad &\leq (1.85 + 0.35 \times 3/2)d^* - (1.45 - 0.35 \times 3/2)d_1 + O(\varepsilon)(d^* + d_1) \\
&= \boxed{2.375d^* - 0.925d_1} + O(\varepsilon)(d^* + d_1).
\end{aligned}$$

If the maximum in Lemma D.1 is attained at $\delta_{\mathcal{T}_{12}\cap\mathcal{A}}(c) + \delta_{\mathcal{T}_{22}\cap\mathcal{A}}(c)$, we have

$$\begin{aligned}
\Delta_{\mathcal{A}}(c) &\leq 1/2 \cdot \delta_{\mathcal{S}_1\cap\mathcal{A}}(c) + 1/4 \cdot (\delta_{\mathcal{T}_{12}\cap\mathcal{A}}(c) + \delta_{\mathcal{T}_{22}\cap\mathcal{A}}(c)) + O(\varepsilon)(d^* + d_1) \\
&\leq (1.85 + 0.1/\rho)d^* - (1.15 - 0.1/\rho)d_1 + O(\varepsilon)(d^* + d_1) \\
&\leq (1.85 + 0.1 \times 3/2)d^* - (1.15 - 0.1 \times 3/2)d_1 + O(\varepsilon)(d^* + d_1) \\
&= \boxed{2\, d^* - d_1} + O(\varepsilon)(d^* + d_1).
\end{aligned}$$

When $\rho(f^*) > 3/4$, we have $\Pr[\mathcal{S}_1] = 5/4 - \rho$ and $\Pr[\mathcal{S}_2] = \rho - 3/4$. Therefore, if the maximum in Lemma D.1 is attained at $\delta_{\mathcal{T}_{11}\cap\mathcal{A}}(c) + \delta_{\mathcal{T}_{21}\cap\mathcal{A}}(c)$, we have

$$\begin{aligned}
\Delta_{\mathcal{A}}(c) &\leq (5/4 - \rho) \cdot \delta_{\mathcal{S}_1\cap\mathcal{A}}(c) + (\rho - 3/4) \cdot \delta_{\mathcal{S}_2\cap\mathcal{A}}(c) + 1/4 \cdot (\delta_{\mathcal{T}_{11}\cap\mathcal{A}}(c) + \delta_{\mathcal{T}_{21}\cap\mathcal{A}}(c)) + O(\varepsilon)(d^* + d_1) \\
&\leq (2.35 + 0.2/\rho - 0.4\rho)d^* - (0.3 - 0.2/\rho + 0.8\rho)d_1 - (0.95 - 0.8\rho)d_2 + O(\varepsilon)(d^* + d_1) \\
(d_2 \geq d_1) \quad & \\
&\leq (2.35 + 0.2/\rho - 0.4\rho)d^* - (1.25 - 0.2/\rho)d_1 + O(\varepsilon)(d^* + d_1) \\
&\leq (2.35 + 0.2 \times 4/3 - 0.4 \times 3/4)d^* - (1.25 - 0.2 \times 4/3)d_1 + O(\varepsilon)(d^* + d_1) \\
&\leq \boxed{2.31667\, d^* - 0.98333\, d_1} + O(\varepsilon)(d^* + d_1).
\end{aligned}$$

If the maximum in Lemma D.1 is attained at $\delta_{\mathcal{T}_{12}\cap\mathcal{A}}(c) + \delta_{\mathcal{T}_{22}\cap\mathcal{A}}(c)$, we have

$$\Delta_{\mathcal{A}}(c) \leq (5/4 - \rho) \cdot \delta_{\mathcal{S}_1\cap\mathcal{A}}(c) + (\rho - 3/4) \cdot \delta_{\mathcal{S}_2\cap\mathcal{A}}(c) + 1/4 \cdot (\delta_{\mathcal{T}_{12}\cap\mathcal{A}}(c) + \delta_{\mathcal{T}_{22}\cap\mathcal{A}}(c)) + O(\varepsilon)(d^* + d_1)$$

$$\leq (2.35 - 0.05/\rho - 0.4\rho)d^* - (0.35 + 0.05/\rho + 0.8\rho)d_1 + (0.8\rho - 0.6)d_2 + O(\varepsilon)(d^* + d_1)$$

$(d_2 \leq d^* + {}^1\!/\!_\rho \cdot (d^* + d_1))$

$$\leq (2.55 - 0.65/\rho + 0.4\rho)d^* - (0.65/\rho + 0.8\rho - 0.45)d_1 + O(\varepsilon)(d^* + d_1)$$
$$\leq (2.55 - 0.65 + 0.4)d^* - (0.65 \cdot 4/\sqrt{13} + 0.8 \cdot \sqrt{13}/4 - 0.45)d_1 + O(\varepsilon)(d^* + d_1)$$
$$\leq \boxed{2.3\,d^* - 0.99222\,d_1} + O(\varepsilon)(d^* + d_1).$$

D.3 Proof of (6.19): Close Clients of Type B

In this section, we show that for any close case client c with type B, we have

$$\Delta_{\mathcal{A}}(c) \leq 2.4\,d^*(c) - 0.9\,d_1(c).$$

In our type A analysis, we further partitioned the tree events $\mathcal{T}_1$ and $\mathcal{T}_2$ as $\mathcal{T}_1 = \mathcal{T}_{11} \cup \mathcal{T}_{12}$ and $\mathcal{T}_2 = \mathcal{T}_{21} \cup \mathcal{T}_{22}$. We require this partitioning also in our type B analysis, with the roles of η_1 and η_2 flipped. Specifically, we define $\mathcal{T}_{11}$ as the intersection of $\mathcal{T}_1$ and the event that there is a swap which closes both f_1 and f_2 but does not open the original copy of f^*. We define $\mathcal{T}_{21}$ as the intersection of $\mathcal{T}_2$ and the event that $\langle\langle f^* \rangle\rangle$ is the only swap closing any facility in $\{f_1, f_2\}$. We define $\mathcal{T}_{12}$ and $\mathcal{T}_{22}$ accordingly as $\mathcal{T}_{12} = \mathcal{T}_1 \backslash \mathcal{T}_{11}$ and $\mathcal{T}_{22} = \mathcal{T}_2 \backslash \mathcal{T}_{21}$. Similar to Lemma D.1, we have the following lemma for type B:

Lemma D.2 (Type B averaging)**.** *For a close client of type* B *with $\rho(f^*) > 2/3$, we have*

$$\Delta_{\mathcal{T} \cap \mathcal{A}}(c) \leq 1/4 \cdot \max\{\delta_{\mathcal{T}_{11} \cap \mathcal{A}} + \delta_{\mathcal{T}_{21} \cap \mathcal{A}}, \delta_{\mathcal{T}_{12} \cap \mathcal{A}} + \delta_{\mathcal{T}_{21} \cap \mathcal{A}}, \delta_{\mathcal{T}_{12} \cap \mathcal{A}} + \delta_{\mathcal{T}_{22} \cap \mathcal{A}}\} + O(\varepsilon)(d^* + d_1).$$

We now proceed to bound the worst-case potential changes in different events.

Simple swaps with $\tau(f^*) = \eta_1$ By implications (Siii) and (Siv), all three swaps $\langle\langle f^* \rangle\rangle, \langle\langle \neg f_1 \rangle\rangle, \langle\langle \neg f_2 \rangle\rangle$ are different. On swap $\langle\langle f^* \rangle\rangle$, the client can be served by f^* and f_1. On swap $\langle\langle \neg f_1 \rangle\rangle$, the client can be served by f_2 and η_1 by implication (Siii) of amenability. On swap $\langle\langle \neg f_2 \rangle\rangle$, the client can be served by f_1 and η_1, again by implication (Siii) of amenability. Note that $d(c, \eta_1) \leq d^* + \rho(d^* + d_1)$ by (C.44). Therefore,

$$
\begin{aligned}
&(\delta_{\langle\langle f^* \rangle\rangle}) && \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) \leq d^* + \beta\,d_1 - d_1 - \beta\,d_2 \\
&(\delta_{\langle\langle \neg f_1 \rangle\rangle}) && \qquad + d_2 + \beta(d^* + \rho(d^* + d_1)) - d_1 - \beta\,d_2 \\
&(\delta_{\langle\langle \neg f_2 \rangle\rangle}) && \qquad + d_1 + \beta(d^* + \rho(d^* + d_1)) - d_1 - \beta\,d_2 \\
& && \qquad = \boxed{(1 + 2\beta + 2\rho\beta)\,d^* - (2 - \beta - 2\rho\beta)\,d_1 + (1 - 3\beta)\,d_2}.
\end{aligned}
$$

Simple swaps with $\tau(f^*) = \eta_2$ By implications (ii) and (Siv), we have $\langle\langle f^* \rangle\rangle = \langle\langle \neg f_1 \rangle\rangle \neq \langle\langle \neg f_2 \rangle\rangle$. On swap $\langle\langle f^* \rangle\rangle$, the client can be served by f^*. On swap $\langle\langle \neg f_2 \rangle\rangle$, the client can be served by f_1 and η_1, by implication (Siii) of amenability. Therefore,

$$
\begin{aligned}
&(\delta_{\langle\langle f^*, \neg f_1 \rangle\rangle}) && \delta_{\mathcal{S}_2 \cap \mathcal{A}}(c) \leq (1 + \alpha\beta)\,d^* - d_1 - \beta\,d_2 \\
&(\delta_{\langle\langle \neg f_2 \rangle\rangle}) && \qquad + d_1 + \beta(d^* + \rho(d^* + d_1)) - d_1 - \beta\,d_2 \\
& && \qquad = \boxed{(1 + \alpha\beta + \beta + \rho\beta)\,d^* - (1 - \rho\beta)\,d_1 - 2\beta\,d_2}.
\end{aligned}
$$

Tree swaps with $\tau(f^*) = \eta_1$ On $\mathcal{T}_{11} \cap \mathcal{A}$, we have $\langle\langle \neg f_1 \rangle\rangle = \langle\langle \neg f_2 \rangle\rangle \neq \langle\langle f^* \rangle\rangle$. On swap $\langle\langle f^* \rangle\rangle$, the client can be served by f^* and f_1. On swap $\langle\langle \neg f_1, \neg f_2 \rangle\rangle$, the client can be served by η_1 and $\pi(f_1)$ by implication (ii) and (Tii) of amenability. We have $d(c, \eta_1) \leq d^* + \rho(d^* + d_1)$ and $d(c, \pi(f_1)) \leq 2d_1 + d^*$ by (6.8). Therefore,

$$\begin{aligned}
&(\delta_{\langle\langle f^* \rangle\rangle}) & \delta_{\mathcal{T}_{11} \cap \mathcal{A}}(c) &\leq d^* + \beta\, d_1 - d_1 - \beta\, d_2 \\
&(\delta_{\langle\langle \neg f_1, \neg f_2 \rangle\rangle}) & &\quad + (d^* + \rho(d^* + d_1)) + \beta(2d_1 + d^*) - d_1 - \beta\, d_2 \\
& & &\leq \boxed{(2 + \beta + \rho)\, d^* - (2 - 3\beta - \rho)\, d_1 - 2\beta\, d_2}.
\end{aligned}$$

On $\mathcal{T}_{12} \cap \mathcal{A}$, we first consider the case where all three swaps $\langle\langle f^* \rangle\rangle, \langle\langle \neg f_1 \rangle\rangle, \langle\langle \neg f_2 \rangle\rangle$ are different. On swap $\langle\langle f^* \rangle\rangle$, the client can be served by f^*. On swap $\langle\langle \neg f_1 \rangle\rangle$, the client can be served by f_2 and η_1. On swap $\langle\langle \neg f_2 \rangle\rangle$, the client can be served by f_1 and η_1. After both $\langle\langle \neg f_1 \rangle\rangle$ and $\langle\langle \neg f_2 \rangle\rangle$, η_1 is open by implication (ii) of amenability. Therefore,

$$\begin{aligned}
&(\delta_{\langle\langle f^* \rangle\rangle}) & \delta_{\mathcal{T}_{12} \cap \mathcal{A}}(c) &\leq (1 + \alpha\beta)\, d^* - d_1 - \beta\, d_2 \\
&(\delta_{\langle\langle \neg f_1 \rangle\rangle}) & &\quad + d_2 + \beta(d^* + \rho(d^* + d_1)) - d_1 - \beta\, d_2 \\
&(\delta_{\langle\langle \neg f_2 \rangle\rangle}) & &\quad + d_1 + \beta(d^* + \rho(d^* + d_1)) - d_1 - \beta\, d_2 \\
& & &\leq \boxed{(1 + \alpha\beta + 2\beta + 2\rho\beta)\, d^* - (2 - 2\rho\beta)\, d_1 + (1 - 3\beta)\, d_2}.
\end{aligned}$$

Since our bound for $\delta_{\langle\langle f^* \rangle\rangle}(c)$ doesn't require either f_1 or f_2 to remain open after the swap, the above bound also holds when $\langle\langle \neg f_1 \rangle\rangle$ and/or $\langle\langle \neg f_2 \rangle\rangle$ coincides with $\langle\langle f^* \rangle\rangle$.

Tree swaps with $\tau(f^*) = \eta_2$ On $\mathcal{T}_{21} \cap \mathcal{A}$, $\langle\langle f^* \rangle\rangle$ is the only swap closing any facility in $\{f_1, f_2\}$ by the definition of $\mathcal{T}_{21}$. In other words, both $\langle\langle \neg f_1 \rangle\rangle$ and $\langle\langle \neg f_2 \rangle\rangle$ coincide with $\langle\langle f^* \rangle\rangle$ as long as they exist. Therefore,

$$(\delta_{\langle\langle f^*, \neg f_1, \neg f_2 \rangle\rangle}) \qquad \delta_{\mathcal{T}_{21} \cap \mathcal{A}}(c) \leq \boxed{(1 + \alpha\beta)\, d^* - d_1 - \beta\, d_2}.$$

On $\mathcal{T}_{22} \cap \mathcal{A}$, we have $\langle\langle f^* \rangle\rangle = \langle\langle \neg f_1 \rangle\rangle \neq \langle\langle \neg f_2 \rangle\rangle$ by implication (ii) of amenability. On swap $\langle\langle f^* \rangle\rangle$, the client can be served by f^*. On swap $\langle\langle \neg f_2 \rangle\rangle$, the client can be served by f_1 and $\pi(f_2)$. Therefore,

$$\begin{aligned}
&(\delta_{\langle\langle f^*, \neg f_1 \rangle\rangle}) & \delta_{\mathcal{T}_{22} \cap \mathcal{A}}(c) &\leq (1 + \alpha\beta)\, d^* - d_1 - \beta\, d_2 \\
&(\delta_{\langle\langle \neg f_2 \rangle\rangle}) & &\quad + d_1 + \beta(2d_2 + d^*) - d_1 - \beta\, d_2 \\
& & &= \boxed{(1 + \alpha\beta + \beta)\, d^* - d_1}.
\end{aligned}$$

Summarizing, we have

$$\begin{aligned}
\delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) &\leq (1 + 2\beta + 2\rho\beta)\, d^* - (2 - \beta - 2\rho\beta)\, d_1 + (1 - 3\beta)\, d_2 \\
&\qquad = (1.4 + 0.4\rho)\, d^* - (1.8 - 0.4\rho)\, d_1 + 0.4\, d_2 \\
\delta_{\mathcal{S}_2 \cap \mathcal{A}}(c) &\leq (1 + \alpha\beta + \beta + \rho\beta)\, d^* - (1 - \rho\beta)\, d_1 - 2\beta\, d_2 \qquad = (1.8 + 0.2\rho)\, d^* - (1 - 0.2\rho)\, d_1 - 0.4\, d_2 \\
\delta_{\mathcal{T}_{11} \cap \mathcal{A}}(c) &\leq (2 + \beta + \rho)\, d^* - (2 - 3\beta - \rho)\, d_1 - 2\beta\, d_2 \qquad = (2.2 + \rho)\, d^* - (1.4 - \rho)\, d_1 - 0.4\, d_2 \\
\delta_{\mathcal{T}_{12} \cap \mathcal{A}}(c) &\leq (1 + \alpha\beta + 2\beta + 2\rho\beta)\, d^* - (2 - 2\rho\beta)\, d_1 + (1 - 3\beta)\, d_2 \\
&\qquad = (2 + 0.4\rho)\, d^* - (2 - 0.4\rho)\, d_1 + 0.4\, d_2 \\
\delta_{\mathcal{T}_{21} \cap \mathcal{A}}(c) &\leq (1 + \alpha\beta)\, d^* - d_1 - \beta\, d_2 \qquad = 1.6\, d^* - d_1 - 0.2\, d_2 \\
\delta_{\mathcal{T}_{22} \cap \mathcal{A}}(c) &\leq (1 + \alpha\beta + \beta)\, d^* - d_1 \qquad = 1.8\, d^* - d_1
\end{aligned}$$

Now we combine these inequalities to get an upper bound for $\Delta_{\mathcal{A}}(c)$. When $\rho(f^*) \le 2/3$, we have $\Pr[\mathcal{S}_1] = \Pr[\mathcal{T}_1] = 1/2$ and $\Pr[\mathcal{S}_2] = \Pr[\mathcal{T}_2] = 0$. Therefore,

$$\Delta_{\mathcal{A}}(c) \le 1/2 \cdot \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) + 1/2 \cdot \max\{\delta_{\mathcal{T}_{11} \cap \mathcal{A}}(c), \delta_{\mathcal{T}_{12} \cap \mathcal{A}}(c)\} + O(\varepsilon)(d^* + d_1).$$

If the maximum is attained at $\delta_{\mathcal{T}_{11} \cap \mathcal{A}}(c)$, we have

$$\begin{aligned}
\Delta_{\mathcal{A}}(c) &\le 1/2 \cdot \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) + 1/2 \cdot \delta_{\mathcal{T}_{11} \cap \mathcal{A}}(c) + O(\varepsilon)(d^* + d_1) \\
&\le (1.8 + 0.7\rho)d^* - (1.6 - 0.7\rho)d_1 + O(\varepsilon)(d^* + d_1) \\
&\le (1.8 + 0.7 \times 2/3)d^* - (1.6 - 0.7 \times 2/3)d_1 + O(\varepsilon)(d^* + d_1) \\
&= \boxed{2.26667d^* - 1.13333d_1} + O(\varepsilon)(d^* + d_1).
\end{aligned}$$

If the maximum is attained at $\delta_{\mathcal{T}_{12} \cap \mathcal{A}}(c)$, we have

$$\begin{aligned}
\Delta_{\mathcal{A}}(c) &\le 1/2 \cdot \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) + 1/2 \cdot \delta_{\mathcal{T}_{12} \cap \mathcal{A}}(c) + O(\varepsilon)(d^* + d_1) \\
&\le (1.7 + 0.4\rho)d^* - (1.9 - 0.4\rho)d_1 + 0.4d_2 + O(\varepsilon)(d^* + d_1) \\
&\text{($d_2 \le 1/2 \cdot (d^* + \rho(d^* + d_1)) + 1/2 \cdot \alpha d_1$)} \\
&\le (1.9 + 0.6\rho)d^* - (1.3 - 0.6\rho)d_1 + O(\varepsilon)(d^* + d_1) \\
&\le (1.9 + 0.6 \times 2/3)d^* - (1.3 - 0.6 \times 2/3)d_1 + O(\varepsilon)(d^* + d_1) \\
&= \boxed{2.3\, d^* - 0.9\, d_1} + O(\varepsilon)(d^* + d_1).
\end{aligned}$$

When $\rho(f^*) > 2/3$, we apply Lemma D.2 to combine the inequalities. Note that our bound for $\delta_{\mathcal{T}_{21} \cap \mathcal{A}}(c)$ is smaller than our bound for $\delta_{\mathcal{T}_{22} \cap \mathcal{A}}(c)$, so we only need to consider cases where the maximum in Lemma D.2 is attained at either $\delta_{\mathcal{T}_{11} \cap \mathcal{A}}(c) + \delta_{\mathcal{T}_{21} \cap \mathcal{A}}(c)$ or $\delta_{\mathcal{T}_{12} \cap \mathcal{A}}(c) + \delta_{\mathcal{T}_{22} \cap \mathcal{A}}(c)$.

When $2/3 < \rho(f^*) \le 3/4$, we have $\Pr[\mathcal{S}_1] = 1/2, \Pr[\mathcal{S}_2] = 0, \Pr[\mathcal{T}_1] = \Pr[\mathcal{T}_2] = 1/4$. If the maximum in Lemma D.2 is attained at $\delta_{\mathcal{T}_{11} \cap \mathcal{A}}(c) + \delta_{\mathcal{T}_{21} \cap \mathcal{A}}(c)$, we have

$$\begin{aligned}
\Delta_{\mathcal{A}}(c) &\le 1/2 \cdot \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) + 1/4 \cdot (\delta_{\mathcal{T}_{11} \cap \mathcal{A}}(c) + \delta_{\mathcal{T}_{21} \cap \mathcal{A}}(c)) + O(\varepsilon)(d^* + d_1) \\
&\le (1.65 + 0.45\rho)\, d^* - (1.5 - 0.45\rho)\, d_1 + 0.05\, d_2 + O(\varepsilon)(d^* + d_1) \\
\text{($d_2 \le \alpha d_1$)} \quad &\le (1.65 + 0.45\rho)\, d^* - (1.35 - 0.45\rho)\, d_1 + O(\varepsilon)(d^* + d_1) \\
&\le (1.65 + 0.45 \times 3/4)d^* - (1.35 - 0.45 \times 3/4)d_1 + O(\varepsilon)(d^* + d_1) \\
&= \boxed{1.9875d^* - 1.0125d_1} + O(\varepsilon)(d^* + d_1).
\end{aligned}$$

If the maximum in Lemma D.2 is attained at $\delta_{\mathcal{T}_{12} \cap \mathcal{A}}(c) + \delta_{\mathcal{T}_{22} \cap \mathcal{A}}(c)$, we have

$$\begin{aligned}
\Delta_{\mathcal{A}}(c) &\le 1/2 \cdot \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) + 1/4 \cdot (\delta_{\mathcal{T}_{12} \cap \mathcal{A}}(c) + \delta_{\mathcal{T}_{22} \cap \mathcal{A}}(c)) + O(\varepsilon)(d^* + d_1) \\
&\le (1.65 + 0.3\rho)\, d^* - (1.65 - 0.3\rho)\, d_1 + 0.3\, d_2 + O(\varepsilon)(d^* + d_1) \\
\text{($d_2 \le d^* + \rho(d^* + d_1)$)} \quad &\le (1.95 + 0.6\rho)\, d^* - (1.65 - 0.6\rho)\, d_1 + O(\varepsilon)(d^* + d_1) \\
&\le (1.95 + 0.6 \times 3/4)d^* - (1.65 - 0.6 \times 3/4)d_1 + O(\varepsilon)(d^* + d_1) \\
&= \boxed{2.4d^* - 1.2d_1} + O(\varepsilon)(d^* + d_1).
\end{aligned}$$

When $\rho(f^*) > 3/4$, we have $\Pr[\mathcal{S}_1] = 5/4 - \rho, \Pr[\mathcal{S}_2] = \rho - 3/4, \Pr[\mathcal{T}_1] = \Pr[\mathcal{T}_2] = 1/4$. If the maximum in Lemma D.2 is attained at $\delta_{\mathcal{T}_{11} \cap \mathcal{A}}(c) + \delta_{\mathcal{T}_{21} \cap \mathcal{A}}(c)$, we have

$$\Delta_{\mathcal{A}}(c) \le (5/4 - \rho) \cdot \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) + (\rho - 3/4) \cdot \delta_{\mathcal{S}_2 \cap \mathcal{A}}(c) + 1/4 \cdot (\delta_{\mathcal{T}_{11} \cap \mathcal{A}}(c) + \delta_{\mathcal{T}_{21} \cap \mathcal{A}}(c)) + O(\varepsilon)(d^* + d_1)$$

$$\leq (1.35 + \rho - 0.2\rho^2)\, d^* - (2.1 - 1.4\rho + 0.2\rho^2)\, d_1 + (0.65 - 0.8\rho)\, d_2 + O(\varepsilon)(d^* + d_1).$$

When $\rho < 0.8125 = {}^{0.65}/_{0.8}$, we use $d_2 \leq \alpha d_1$:

$$\begin{aligned}\Delta_{\mathcal{A}}(c) &\leq (1.35 + \rho - 0.2\rho^2)\, d^* - (0.15 + \rho + 0.2\rho^2)\, d_1 + O(\varepsilon)(d^* + d_1)\\ &\leq (1.35 + 0.8125 - 0.2 \times 0.8125^2)d^* - (0.15 + {}^3/_4 + 0.2 \times ({}^3/_4)^2)d_1 + O(\varepsilon)(d^* + d_1)\\ &= \boxed{2.03047d^* - 1.0125d_1} + O(\varepsilon)(d^* + d_1).\end{aligned}$$

When $\rho \geq 0.8125$, we use $d_2 \geq d_1$:

$$\begin{aligned}\Delta_{\mathcal{A}}(c) &\leq (1.35 + \rho - 0.2\rho^2)\, d^* - (1.45 - 0.6\rho + 0.2\rho^2)\, d_1 + O(\varepsilon)(d^* + d_1)\\ &\leq (1.35 + 1 - 0.2)d^* - (1.45 - 0.6 + 0.2)d_1 + O(\varepsilon)(d^* + d_1)\\ &= \boxed{2.15\, d^* - 1.05\, d_1} + O(\varepsilon)(d^* + d_1).\end{aligned}$$

If the maximum in Lemma D.2 is attained at $\delta_{\mathcal{T}_{12}\cap\mathcal{A}}(c) + \delta_{\mathcal{T}_{22}\cap\mathcal{A}}(c)$, we have

$$\begin{aligned}\Delta_{\mathcal{A}}(c) &\leq ({}^5/_4 - \rho) \cdot \delta_{\mathcal{S}_1\cap\mathcal{A}}(c) + (\rho - {}^3/_4) \cdot \delta_{\mathcal{S}_2\cap\mathcal{A}}(c) + {}^1/_4 \cdot (\delta_{\mathcal{T}_{12}\cap\mathcal{A}}(c) + \delta_{\mathcal{T}_{22}\cap\mathcal{A}}(c)) + O(\varepsilon)(d^* + d_1)\\ &\leq (1.35 + 0.85\rho - 0.2\rho^2)\, d^* - (2.25 - 1.25\rho + 0.2\rho^2)\, d_1 + (0.9 - 0.8\rho)\, d_2 + O(\varepsilon)(d^* + d_1)\end{aligned}$$

$(d_2 \leq d^* + \rho(d^* + d_1))$

$$\begin{aligned}&\leq (2.25 + 0.95\rho - \rho^2)\, d^* - (2.25 - 2.15\rho + \rho^2)\, d_1 + O(\varepsilon)(d^* + d_1)\\ &\leq (2.25 + 0.95 \times {}^3/_4 - ({}^3/_4)^2)d^* - (2.25 - 2.15 + 1)d_1 + O(\varepsilon)(d^* + d_1)\\ &= \boxed{2.4d^* - 1.1d_1} + O(\varepsilon)(d^* + d_1).\end{aligned}$$

D.4 Proof of (6.20): Clients of Type C

In this section, we show that for any client c with type C, we have

$$\Delta_{\mathcal{A}}(c) \leq 2.2\, d^*(c) - 0.8888\, d_1(c).$$

If the client c satisfies $\rho(f^*) \leq {}^2/_3$, we have the same bound as in the type A case in Appendix D.2.1, where our analysis was independent of whether $f_2 = \eta_2$ or not. That is

$$\Delta_{\mathcal{A}}(c) \leq \boxed{2.1\, d^*(c) - 0.9\, d_1(c)} + O(\varepsilon)(d^* + d_1).$$

We thus focus on clients with $\rho(f^*) > {}^2/_3$. Compared to our analysis for other client types, our analysis for type C involves a larger neighborhood of the client. In particular, the optimal facility $g^* := \pi(f_1)$ and the local facilities close to it play a crucial role in our analysis. This makes it important to consider finer-grained events. Recall that we used $\mathcal{S}_1, \mathcal{S}_2, \mathcal{T}_1, \mathcal{T}_2$ to denote simple/tree events restricted to f^* pointing to η_1 or η_2. We now also define events $\mathcal{S}'_1, \mathcal{S}'_2, \mathcal{T}'_1, \mathcal{T}'_2$ similarly, except that they depend on where g^* points to, rather than f^*. We classify clients into subtypes according to the characteristics of the swap sets generated on these events:

Claim D.3 (Subtypes within type C). *For a client c of type C, one of the following is true:*

(a) f_1 is heavy.

(b) f_2 is heavy.

(c) A facility h is open near c after the simple swap closing f_1. Formally, a facility $h \neq f_2$ is open after swap $\langle\!\langle \neg f_1 \rangle\!\rangle$ at distance $d(c,h) \leq 3d_1 + 2d^*$ on $\mathcal{S} \cap \mathcal{A}$.

(d) $g^* \neq f^*$, $\rho(g^*) > 3/4$, and for all $b = 1, 2$, any swap set $\mathcal{P}$ generated on $\mathcal{S}'_b \cap \mathcal{A}$, a facility $h \neq f_2$ is open after swap $\langle\!\langle \neg f_1 \rangle\!\rangle$ at distance $d(c,h) \leq \begin{cases} 2d_1 + d^*, & \text{if } b = 1 \\ 2d_1 + d^* + 4/3(d_1 + d^*), & \text{if } b = 2 \end{cases}$.

(e) For any swap set $\mathcal{P}$ generated on $\mathcal{T}_2 \cap \mathcal{A}$, $\langle\!\langle f^* \rangle\!\rangle$ closes both f_1 and f_2.

(f) $g^* \neq f^*$, $\rho(g^*) > 2/3$, and there exists $b \in \{1,2\}$ such that for any swap set $\mathcal{P}$ generated on $\mathcal{T}'_b \cap \mathcal{A}$, $\langle\!\langle f^* \rangle\!\rangle$ closes both f_1 and f_2.

We prove this claim in Appendix E.5. Below we present our bounds for each of these subtypes.

D.4.1 When f_1 is a heavy facility

f_1 being heavy implies that the swap $\langle\!\langle \neg f_1 \rangle\!\rangle$ doesn't exist. We thus focus on $\langle\!\langle f^* \rangle\!\rangle$ and $\langle\!\langle \neg f_2 \rangle\!\rangle$.

Simple swaps with $\tau(f^*) = \eta_1$ By implication (Siv) of amenability, we have $\langle\!\langle f^* \rangle\!\rangle \neq \langle\!\langle \neg f_2 \rangle\!\rangle$. On swap $\langle\!\langle f^* \rangle\!\rangle$, the client can be served by f^*, and on swap $\langle\!\langle \neg f_2 \rangle\!\rangle$, the client can be served by f_1. Therefore,

$$\begin{aligned}
&(\delta_{\langle\!\langle f^* \rangle\!\rangle}) && \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) \leq (1+\alpha\beta)d^* - d_1 - \beta\, d_2 \\
&(\delta_{\langle\!\langle \neg f_2 \rangle\!\rangle}) && \qquad + (1+\alpha\beta)d_1 - d_1 - \beta\, d_2 \\
& && = \boxed{(1+\alpha\beta)d^* - (1-\alpha\beta)d_1 - 2\beta d_2}.
\end{aligned}$$

Tree swaps with $\tau(f^*) = \eta_1$ Let us first assume $\langle\!\langle f^* \rangle\!\rangle \neq \langle\!\langle \neg f_2 \rangle\!\rangle$. On swap $\langle\!\langle f^* \rangle\!\rangle$, the client can be served by f^*, and on swap $\langle\!\langle \neg f_2 \rangle\!\rangle$, the client can be served by f_1 and $\pi(f_2)$ by implication (Tii) of amenability. Note that $d(c, \pi(f_2)) \leq 2d_2 + d^*$ by (C.46). Therefore,

$$\begin{aligned}
&(\delta_{\langle\!\langle f^* \rangle\!\rangle}) && \delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) \leq (1+\alpha\beta)d^* - d_1 - \beta\, d_2 \\
&(\delta_{\langle\!\langle \neg f_2 \rangle\!\rangle}) && \qquad + d_1 + \beta(2d_2 + d^*) - d_1 - \beta\, d_2 \\
& && = \boxed{(1+\alpha\beta+\beta)d^* - d_1}.
\end{aligned}$$

The inequality also holds when $\langle\!\langle f^* \rangle\!\rangle = \langle\!\langle \neg f_2 \rangle\!\rangle$ since our bound for $\delta_{\langle\!\langle f^* \rangle\!\rangle}$ does not require f_2 to remain open after the swap.

Simple & tree swaps with $\tau(f^*) = \eta_2$ We have $\langle\!\langle f^* \rangle\!\rangle = \langle\!\langle \neg f_2 \rangle\!\rangle$ by implication (ii) of amenability. On that swap, the client can be served by f^*. Therefore,

$$(\delta_{\langle\!\langle f^*, \neg f_2 \rangle\!\rangle}) \qquad \delta_{\mathcal{S}_2 \cap \mathcal{A}}(c) \leq \boxed{(1+\alpha\beta)d^* - d_1 - \beta\, d_2},$$

$$(\delta_{\langle\!\langle f^*, \neg f_2 \rangle\!\rangle}) \qquad \delta_{\mathcal{T}_2 \cap \mathcal{A}}(c) \leq \boxed{(1+\alpha\beta)d^* - d_1 - \beta\, d_2}.$$

Summarizing, we have

$$\begin{aligned}
\delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) &\leq (1+\alpha\beta)\, d^* - (1-\alpha\beta)d_1 - 2\beta\, d_2 &&= 1.6\, d^* - 0.4 d_1 - 0.4\, d_2 \\
\delta_{\mathcal{S}_2 \cap \mathcal{A}}(c) &\leq (1+\alpha\beta)\, d^* - d_1 - \beta\, d_2 &&= 1.6\, d^* - d_1 - 0.2\, d_2 \\
\delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) &\leq (1+\alpha\beta+\beta)\, d^* - d_1 &&= 1.8\, d^* - d_1 \\
\delta_{\mathcal{T}_2 \cap \mathcal{A}}(c) &\leq (1+\alpha\beta)\, d^* - d_1 - \beta\, d_2 &&= 1.6\, d^* - d_1 - 0.2\, d_2
\end{aligned}$$

We now combine these inequalities to get an upper bound for $\Delta_{\mathcal{A}}(c)$.

When $2/3 < \rho(f^*) \le 3/4$, we have $\Pr[\mathcal{S}_1] = 1/2, \Pr[\mathcal{S}_2] = 0, \Pr[\mathcal{T}_1] = \Pr[\mathcal{T}_2] = 1/4$. Therefore,

$$\begin{aligned}
\Delta_{\mathcal{A}}(c) &\le 1/2 \cdot \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_2 \cap \mathcal{A}}(c) + O(\varepsilon)(d^* + d_1) \\
&\le 1.65\, d^* - 0.7\, d_1 - 0.25\, d_2 + O(\varepsilon)(d^* + d_1) \\
(d_2 \ge d_1) \qquad &\le \boxed{1.65\, d^* - 0.95\, d_1} + O(\varepsilon)(d^* + d_1).
\end{aligned}$$

When $\rho(f^*) > 3/4$, we have $\Pr[\mathcal{S}_1] = 5/4 - \rho, \Pr[\mathcal{S}_2] = \rho - 3/4, \Pr[\mathcal{T}_1] = \Pr[\mathcal{T}_2] = 1/4$. Therefore,

$$\begin{aligned}
\Delta_{\mathcal{A}}(c) &\le (5/4 - \rho) \cdot \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) + (\rho - 3/4) \cdot \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_2 \cap \mathcal{A}}(c) + O(\varepsilon)(d^* + d_1) \\
&\le 1.65\, d^* - (0.6\rho + 0.25)\, d_1 - (0.4 - 0.2\rho)\, d_2 + O(\varepsilon)(d^* + d_1) \\
(d_2 \ge d_1) \\
&\le 1.65\, d^* - (0.4\rho + 0.65)\, d_1 + O(\varepsilon)(d^* + d_1) \\
&\le 1.65\, d^* - (0.4 \times 3/4 + 0.65)\, d_1 + O(\varepsilon)(d^* + d_1) \\
&= \boxed{1.65\, d^* - 0.95\, d_1} + O(\varepsilon)(d^* + d_1).
\end{aligned}$$

D.4.2 When f_2 is a heavy facility

f_2 being heavy implies that $\langle\!\langle \neg f_2 \rangle\!\rangle$ does not exist. We thus focus on $\langle\!\langle f^* \rangle\!\rangle$ and $\langle\!\langle \neg f_1 \rangle\!\rangle$.

Simple & tree swaps with $\tau(f^*) = \eta_1$ We have $\langle\!\langle f^* \rangle\!\rangle = \langle\!\langle \neg f_1 \rangle\!\rangle$ by implication (ii) of amenability. On that swap, the client can be served by f^*. Therefore,

$$\begin{aligned}
(\delta_{\langle\!\langle f^*, \neg f_1 \rangle\!\rangle}) \qquad \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) &\le \boxed{(1 + \alpha\beta)d^* - d_1 - \beta\, d_2}, \\
(\delta_{\langle\!\langle f^*, \neg f_1 \rangle\!\rangle}) \qquad \delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) &\le \boxed{(1 + \alpha\beta)d^* - d_1 - \beta\, d_2}.
\end{aligned}$$

Simple swaps with $\tau(f^*) = \eta_2$ Implication (Siv) of amenability implies that $\langle\!\langle f^* \rangle\!\rangle \ne \langle\!\langle \neg f_1 \rangle\!\rangle$. On swap $\langle\!\langle f^* \rangle\!\rangle$, the client can be served by f^*. On swap $\langle\!\langle \neg f_1 \rangle\!\rangle$, the client can be served by f_2. Therefore,

$$\begin{aligned}
(\delta_{\langle\!\langle f^* \rangle\!\rangle}) \qquad \delta_{\mathcal{S}_2 \cap \mathcal{A}}(c) &\le (1 + \alpha\beta)d^* - d_1 - \beta\, d_2 \\
(\delta_{\langle\!\langle \neg f_1 \rangle\!\rangle}) \qquad &\quad + (1 + \alpha\beta)d_2 - d_1 - \beta\, d_2 \\
\text{(D.52)} \qquad &= \boxed{(1 + \alpha\beta)d^* - 2\, d_1 + (1 + \alpha\beta - 2\beta)d_2}.
\end{aligned}$$

Tree swaps with $\tau(f^*) = \eta_2$ We first assume that $\langle\!\langle f^* \rangle\!\rangle \ne \langle\!\langle \neg f_1 \rangle\!\rangle$. On swap $\langle\!\langle f^* \rangle\!\rangle$, the client can be served by f^*. On swap $\langle\!\langle \neg f_1 \rangle\!\rangle$, the client can be served by f_2 and $\pi(f_1)$, by implication (Tii) of amenability. Note that $d(c, \pi(f_1)) \le d_1 + d(f_1, \pi(f_1)) \le 2d_1 + d^*$. Therefore,

$$\begin{aligned}
(\delta_{\langle\!\langle f^* \rangle\!\rangle}) \qquad \delta_{\mathcal{T}_2 \cap \mathcal{A}}(c) &\le (1 + \alpha\beta)d^* - d_1 - \beta\, d_2 \\
(\delta_{\langle\!\langle \neg f_1 \rangle\!\rangle}) \qquad &\quad + d_2 + \beta(2d_1 + d^*) - d_1 - \beta\, d_2 \\
&= \boxed{(1 + \alpha\beta + \beta)d^* - (2 - 2\beta)\, d_1 + (1 - 2\beta)d_2}.
\end{aligned}$$

The above inequality also holds when $\langle\!\langle f^* \rangle\!\rangle = \langle\!\langle \neg f_1 \rangle\!\rangle$ because our bound for $\delta_{\langle\!\langle f^* \rangle\!\rangle}$ does not require f_1 to remain open after the swap. and $\delta_{\langle\!\langle \neg f_1 \rangle\!\rangle}$ is non-negative.

Summarizing, we have

$$\begin{aligned}
\delta_{\mathcal{S}_1\cap\mathcal{A}}(c) &\le (1+\alpha\beta)\,d^* - d_1 - \beta\,d_2 && = 1.6\,d^* - d_1 - 0.2\,d_2\\
\delta_{\mathcal{S}_2\cap\mathcal{A}}(c) &\le (1+\alpha\beta)\,d^* - 2\,d_1 + (1+\alpha\beta-2\beta)\,d_2 && = 1.6\,d^* - 2\,d_1 + 1.2\,d_2\\
\delta_{\mathcal{T}_1\cap\mathcal{A}}(c) &\le (1+\alpha\beta)\,d^* - d_1 - \beta\,d_2 && = 1.6\,d^* - d_1 - 0.2\,d_2\\
\delta_{\mathcal{T}_2\cap\mathcal{A}}(c) &\le (1+\alpha\beta+\beta)\,d^* - (2-2\beta)\,d_1 + (1-2\beta)\,d_2 && = 1.8\,d^* - 1.6\,d_1 + 0.6\,d_2
\end{aligned}$$

We now combine these inequalities to get an upper bound for $\Delta_{\mathcal{A}}(c)$.

When $2/3 < \rho(f^*) \le 3/4$, we have $\Pr[\mathcal{S}_1] = 1/2, \Pr[\mathcal{S}_2] = 0, \Pr[\mathcal{T}_1] = \Pr[\mathcal{T}_2] = 1/4$. Therefore,

$$\begin{aligned}
\Delta_{\mathcal{A}}(c) &\le 1/2 \cdot \delta_{\mathcal{S}_1\cap\mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_1\cap\mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_2\cap\mathcal{A}}(c) + O(\varepsilon)(d^* + d_1)\\
&\le \boxed{1.65\,d^* - 1.15\,d_1} + O(\varepsilon)(d^* + d_1).
\end{aligned}$$

When $\rho(f^*) > 3/4$, we have $\Pr[\mathcal{S}_1] = 5/4 - \rho, \Pr[\mathcal{S}_2] = \rho - 3/4, \Pr[\mathcal{T}_1] = \Pr[\mathcal{T}_2] = 1/4$. Therefore,

$$\begin{aligned}
\Delta_{\mathcal{A}}(c) &\le (5/4 - \rho) \cdot \delta_{\mathcal{S}_1\cap\mathcal{A}}(c) + (\rho - 3/4) \cdot \delta_{\mathcal{S}_2\cap\mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_1\cap\mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_2\cap\mathcal{A}}(c) + O(\varepsilon)(d^* + d_1)\\
&\le 1.65\,d^* - (\rho + 0.4)\,d_1 + (1.4\rho - 1.05)\,d_2 + O(\varepsilon)(d^* + d_1)
\end{aligned}$$

$(\rho > 3/4$ and $d_2 \le d^* + 1/\rho(d^* + d_1))$

$$\begin{aligned}
&\le (2 + 1.4\rho - 1.05/\rho)\,d^* - (\rho + 1.05/\rho - 1)\,d_1 + O(\varepsilon)(d^* + d_1)\\
&\le (2 + 1.4 - 1.05)\,d^* - (1 + 1.05 - 1)\,d_1 + O(\varepsilon)(d^* + d_1)\\
&= \boxed{2.35\,d^* - 1.05\,d_1} + O(\varepsilon)(d^* + d_1).
\end{aligned}$$

D.4.3 There exists h such that $d(c,h) \le 3d_1 + 2d^*$ in simple swaps

Simple swaps with $\tau(f^*) = \eta_1$ By implications (ii) and (Siv) of amenablity, we know $\langle\langle f^*\rangle\rangle = \langle\langle \neg f_1\rangle\rangle \ne \langle\langle \neg f_2\rangle\rangle$. On swap $\langle\langle f^*, \neg f_1\rangle\rangle$, the client can be served by f^*. On swap $\langle\langle \neg f_2\rangle\rangle$, the client can be served by f_1. Therefore,

$$\begin{aligned}
&(\delta_{\langle\langle f^*,\neg f_1\rangle\rangle}) && \delta_{\mathcal{S}_1\cap\mathcal{A}}(c) \le (1+\alpha\beta)\,d^* - d_1 - \beta\,d_2\\
&(\delta_{\langle\langle \neg f_2\rangle\rangle}) && \qquad + (1+\alpha\beta)\,d_1 - d_1 - \beta\,d_2\\
& && \qquad = \boxed{(1+\alpha\beta)\,d^* - (1-\alpha\beta)\,d_1 - 2\beta\,d_2}.
\end{aligned}$$

Simple swaps with $\tau(f^*) = \eta_2$ By implications (ii) and (Siv) of amenablity, we know $\langle\langle f^*\rangle\rangle = \langle\langle \neg f_2\rangle\rangle \ne \langle\langle \neg f_1\rangle\rangle$. On swap $\langle\langle f^*, \neg f_2\rangle\rangle$, the client can be served by f^*. On swap $\langle\langle \neg f_1\rangle\rangle$, the client can be served by f_2 and h. Therefore,

$$\begin{aligned}
&(\delta_{\langle\langle f^*,\neg f_2\rangle\rangle}) && \delta_{\mathcal{S}_2\cap\mathcal{A}}(c) \le (1+\alpha\beta)\,d^* - d_1 - \beta\,d_2\\
&(\delta_{\langle\langle \neg f_1\rangle\rangle}) && \qquad + d_2 + \beta(3d_1 + 2d^*) - d_1 - \beta\,d_2\\
& && \qquad = \boxed{(1+\alpha\beta+2\beta)\,d^* - (2-3\beta)\,d_1 + (1-2\beta)\,d_2}.
\end{aligned}$$

Tree swaps with $\tau(f^*) = \eta_1$ By implication (ii) of amenability, we know $\langle\langle f^*\rangle\rangle = \langle\langle \neg f_1\rangle\rangle$. Let us first assume that $\langle\langle \neg f_2\rangle\rangle \ne \langle\langle f^*, \neg f_1\rangle\rangle$. On swap $\langle\langle f^*, \neg f_1\rangle\rangle$, the client can be served by f^*. On swap $\langle\!\langle \neg f_2 \rangle\!\rangle$, the client can be served by f_1 and $\pi(f_2)$ by implication (Tii) of amenability. We have $d(c,\pi(f_2)) \le 2d_2 + d^*$ by (C.46). Therefore,

$$\begin{aligned}
&(\delta_{\langle\!\langle f^*,\neg f_1\rangle\!\rangle}) \qquad && \delta_{\mathcal{T}_1\cap\mathcal{A}}(c) \le (1+\alpha\beta)\, d^* - d_1 - \beta\, d_2 \\
&(\delta_{\langle\!\langle \neg f_2\rangle\!\rangle}) && \qquad\qquad + d_1 + \beta(2d_2 + d^*) - d_1 - \beta\, d_2 \\
& && \qquad = \boxed{(1+\alpha\beta+\beta)\, d^* - d_1}.
\end{aligned}$$

This inequality also holds when $\langle\!\langle \neg f_2 \rangle\!\rangle = \langle\!\langle f^*, \neg f_1 \rangle\!\rangle$, because our bound for $\delta_{\langle\!\langle f^*,\neg f_1\rangle\!\rangle}$ does not require f_2 to remain open after the swap and $\delta_{\langle\!\langle \neg f_2\rangle\!\rangle}$ is non-negative.

Tree swaps with $\tau(f^*) = \eta_2$ By implication (ii) of amenability, we know $\langle\!\langle f^* \rangle\!\rangle = \langle\!\langle \neg f_2 \rangle\!\rangle$. Again, let us first assume that $\langle\!\langle \neg f_1 \rangle\!\rangle \neq \langle\!\langle f^*, \neg f_2 \rangle\!\rangle$. On swap $\langle\!\langle f^*, \neg f_2 \rangle\!\rangle$, the client can be served by f^*. On swap $\langle\!\langle \neg f_1 \rangle\!\rangle$, the client can be served by f_2 and $\pi(f_1)$ by implication (ii) of amenability. We have $d(c,\pi(f_1)) \le 2d_1 + d^*$ by (6.8). Therefore,

$$\begin{aligned}
&(\delta_{\langle\!\langle f^*,\neg f_2\rangle\!\rangle}) \qquad && \delta_{\mathcal{T}_2\cap\mathcal{A}}(c) \le (1+\alpha\beta)\, d^* - d_1 - \beta\, d_2 \\
&(\delta_{\langle\!\langle \neg f_1\rangle\!\rangle}) && \qquad\qquad + d_2 + \beta(2d_1 + d^*) - d_1 - \beta\, d_2 \\
& && \qquad = \boxed{(1+\alpha\beta+\beta)\, d^* - (2-2\beta)\, d_1 + (1-2\beta)\, d_2}.
\end{aligned}$$

This inequality also holds when $\langle\!\langle \neg f_1 \rangle\!\rangle = \langle\!\langle f^*, \neg f_2 \rangle\!\rangle$, because our bound for $\delta_{\langle\!\langle f^*,\neg f_2\rangle\!\rangle}$ does not require f_1 to remain open after the swap and $\delta_{\langle\!\langle \neg f_1\rangle\!\rangle}$ is non-negative.

Summarizing, we have

$$\begin{aligned}
\delta_{\mathcal{S}_1\cap\mathcal{A}}(c) &\le (1+\alpha\beta)\, d^* - (1-\alpha\beta)\, d_1 - 2\beta\, d_2 && = 1.6\, d^* - 0.4\, d_1 - 0.4\, d_2 \\
\delta_{\mathcal{S}_2\cap\mathcal{A}}(c) &\le (1+\alpha\beta+2\beta)\, d^* - (2-3\beta)\, d_1 + (1-2\beta)\, d_2 && = 2d^* - 1.4\, d_1 + 0.6\, d_2 \\
\delta_{\mathcal{T}_1\cap\mathcal{A}}(c) &\le (1+\alpha\beta+\beta)\, d^* - d_1 && = 1.8\, d^* - d_1 \\
\delta_{\mathcal{T}_2\cap\mathcal{A}}(c) &\le (1+\alpha\beta+\beta)\, d^* - (2-2\beta)\, d_1 + (1-2\beta)\, d_2 && = 1.8\, d^* - 1.6\, d_1 + 0.6\, d_2
\end{aligned}$$

We now combine these inequalities to get an upper bound for $\Delta_{\mathcal{A}}(c)$.

When $2/3 < \rho(f^*) \le 3/4$, we have $\Pr[\mathcal{S}_1] = 1/2, \Pr[\mathcal{S}_2] = 0, \Pr[\mathcal{T}_1] = \Pr[\mathcal{T}_2] = 1/4$. Therefore,

$$\begin{aligned}
& && \Delta_{\mathcal{A}}(c) \le 1/2 \cdot \delta_{\mathcal{S}_1\cap\mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_1\cap\mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_2\cap\mathcal{A}}(c) + O(\varepsilon)(d^* + d_1) \\
& && \qquad\quad \le 1.7\, d^* - 0.85\, d_1 - 0.05\, d_2 + O(\varepsilon)(d^* + d_1) \\
&(d_2 \ge d_1) && \qquad\quad \le \boxed{1.7\, d^* - 0.9\, d_1} + O(\varepsilon)(d^* + d_1).
\end{aligned}$$

When $\rho(f^*) > 3/4$, we have $\Pr[\mathcal{S}_1] = 5/4 - \rho, \Pr[\mathcal{S}_2] = \rho - 3/4, \Pr[\mathcal{T}_1] = \Pr[\mathcal{T}_2] = 1/4$. Therefore,

$$\begin{aligned}
\Delta_{\mathcal{A}}(c) &\le (5/4 - \rho) \cdot \delta_{\mathcal{S}_1\cap\mathcal{A}}(c) + (\rho - 3/4) \cdot \delta_{\mathcal{S}_2\cap\mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_1\cap\mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_2\cap\mathcal{A}}(c) + O(\varepsilon)(d^* + d_1) \\
&\le (1.4 + 0.4\rho)\, d^* - (0.1 + \rho)\, d_1 + (\rho - 0.8)\, d_2 + O(\varepsilon)(d^* + d_1).
\end{aligned}$$

When $\rho \le 0.8$, we use $d_2 \ge d_1$:

$$\begin{aligned}
\Delta_{\mathcal{A}}(c) &\le (1.4 + 0.4\rho)\, d^* - 0.9\, d_1 + O(\varepsilon)(d^* + d_1) \\
&\le \boxed{1.72\, d^* - 0.9\, d_1} + O(\varepsilon)(d^* + d_1).
\end{aligned}$$

When $\rho > 0.8$, we use $d_2 \le d^* + d(f^*, d_2) \le d^* + {}^1\!/\!_\rho \cdot (d^* + d_1)$:

$$\begin{aligned}\Delta_{\mathcal{A}}(c) &\le (1.6 + 1.4\rho - 0.8/\rho)\, d^* - (\rho + 0.8/\rho - 0.9)\, d_1 + O(\varepsilon)(d^* + d_1)\\ &\le (1.6 + 1.4 - 0.8)\, d^* - (2\sqrt{0.8} - 0.9)\, d_1 + O(\varepsilon)(d^* + d_1)\\ &\le \boxed{2.2\, d^* - 0.88885\, d_1} + O(\varepsilon)(d^* + d_1).\end{aligned}$$

D.4.4 $d(c, h) \le 2d_1 + d^*$ or $d(c, h) \le 2d_1 + d^* + {}^4\!/\!_3(d^* + d_1)$ in simple swaps

We have the same bound for $\Delta_{\mathcal{A}}(c)$ in this case as the previous case. Indeed, our previous bounds for $\delta_{\mathcal{S}_1\cap\mathcal{A}}(c), \delta_{\mathcal{T}_1\cap\mathcal{A}}(c)$ and $\delta_{\mathcal{T}_2\cap\mathcal{A}}(c)$ remain valid. We replace our bound for $\delta_{\mathcal{S}_2\cap\mathcal{A}}(c)$ by a bound for

$$\delta'_{\mathcal{S}_2\cap\mathcal{A}}(c) := \Pr[\mathcal{S}'_1|\mathcal{S}_2]\delta_{\mathcal{S}'_1\cap\mathcal{S}_2\cap\mathcal{A}}(c) + \Pr[\mathcal{S}'_2|\mathcal{S}_2]\delta_{\mathcal{S}'_2\cap\mathcal{S}_2\cap\mathcal{A}}(c). \tag{D.53}$$

We show that we can upper-bound $\delta'_{\mathcal{S}_2\cap\mathcal{A}}(c)$ by the same expression as in (D.52). Our previous bound for $\delta_{\mathcal{S}_2\cap\mathcal{A}}(c)$ is linear in $d(c, h)$ with a non-negative coefficient: $\delta_{\mathcal{S}_2\cap\mathcal{A}}(c) \le A \cdot d(c, h) + B$ with $A \ge 0$, so

$$\begin{aligned}\delta_{\mathcal{S}'_1\cap\mathcal{S}_2\cap\mathcal{A}}(c) &\le A \cdot (2d_1 + d^*) + B\\ \delta_{\mathcal{S}'_2\cap\mathcal{S}_2\cap\mathcal{A}}(c) &\le A \cdot (2d_1 + d^* + {}^4\!/\!_3(d^* + d_1)) + B.\end{aligned}$$

Plugging them into (D.53), we have

$$\begin{aligned}\delta'_{\mathcal{S}_2\cap\mathcal{A}}(c) &\le A \cdot (\Pr[\mathcal{S}'_1|\mathcal{S}_2] \cdot (2d_1 + d^*) + \Pr[\mathcal{S}'_2|\mathcal{S}_2] \cdot (2d_1 + d^* + {}^4\!/\!_3(d^* + d_1))) + B\\ &\le A \cdot ({}^1\!/\!_2 \cdot (2d_1 + d^*) + {}^1\!/\!_2 \cdot (2d_1 + d^* + {}^4\!/\!_3(d^* + d_1))) + B\\ &\le A \cdot (3d_1 + 2d^*) + B.\end{aligned}$$

D.4.5 $\langle\!\langle f^* \rangle\!\rangle$ closes f_1 and f_2 on $\mathcal{T}_2 \cap \mathcal{A}$

If $\tau(f^*) = \eta_1$, we get the same bounds as before:

$$\begin{aligned}\delta_{\mathcal{S}_1\cap\mathcal{A}}(c) &\le \boxed{(1 + \alpha\beta)\, d^* - (1 - \alpha\beta)\, d_1 - 2\beta\, d_2},\\ \delta_{\mathcal{T}_1\cap\mathcal{A}}(c) &\le \boxed{(1 + \alpha\beta + \beta)\, d^* - d_1}.\end{aligned}$$

We continue to bound $\delta_{\mathcal{S}_2\cap\mathcal{A}}(c)$ and $\delta_{\mathcal{T}_2\cap\mathcal{A}}(c)$.

Simple swaps with $\tau(f^*) = \eta_2$ By implications (ii) and (Siv), we have $\langle\!\langle f^* \rangle\!\rangle = \langle\!\langle \neg f_2 \rangle\!\rangle \ne \langle\!\langle \neg f_1 \rangle\!\rangle$. On swap $\langle\!\langle f^*, \neg f_2 \rangle\!\rangle$, the client can be served by f^*. On swap $\langle\!\langle \neg f_1 \rangle\!\rangle$, the client can be served by f_2. Therefore,

$$\begin{aligned}(\delta_{\langle\!\langle f^*,\neg f_2\rangle\!\rangle}) \qquad \delta_{\mathcal{S}_2\cap\mathcal{A}}(c) \le&(1 + \alpha\beta)\, d^* - d_1 - \beta\, d_2\\ (\delta_{\langle\!\langle \neg f_1\rangle\!\rangle}) \qquad &+ (1 + \alpha\beta)\, d_2 - d_1 - \beta\, d_2\\ =& \boxed{(1 + \alpha\beta)\, d^* - 2\, d_1 + (1 + \alpha\beta - 2\beta)\, d_2}.\end{aligned}$$

Tree swaps with $\tau(f^*) = \eta_2$ On $\mathcal{T}_2 \cap \mathcal{A}$, we know $\langle\!\langle f^* \rangle\!\rangle$ closes both f_1 and f_2. Therefore,

$$\delta_{\mathcal{T}_2\cap\mathcal{A}}(c) \le \boxed{(1 + \alpha\beta)\, d^* - d_1 - \beta\, d_2}. \tag{$\delta_{\langle\!\langle f^*,\neg f_1,\neg f_2\rangle\!\rangle}$}$$

Summarizing, we have

$$
\begin{aligned}
\delta_{\mathcal{S}_1\cap\mathcal{A}}(c) &\le (1+\alpha\beta)\,d^* - (1-\alpha\beta)\,d_1 - 2\beta\,d_2 && = 1.6\,d^* - 0.4\,d_1 - 0.4\,d_2 \\
\delta_{\mathcal{S}_2\cap\mathcal{A}}(c) &\le (1+\alpha\beta)\,d^* - 2\,d_1 + (1+\alpha\beta-2\beta)\,d_2 && = 1.6\,d^* - 2\,d_1 + 1.2\,d_2 \\
\delta_{\mathcal{T}_1\cap\mathcal{A}}(c) &\le (1+\alpha\beta+\beta)\,d^* - d_1 && = 1.8\,d^* - d_1 \\
\delta_{\mathcal{T}_2\cap\mathcal{A}}(c) &\le (1+\alpha\beta)\,d^* - d_1 - \beta\,d_2 && = 1.6\,d^* - d_1 - 0.2\,d_2
\end{aligned}
$$

We now combine these inequalities to get an upper bound for $\Delta_{\mathcal{A}}(c)$.

When $2/3 < \rho(f^*) \le 3/4$, we have $\Pr[\mathcal{S}_1] = 1/2, \Pr[\mathcal{S}_2] = 0, \Pr[\mathcal{T}_1] = \Pr[\mathcal{T}_2] = 1/4$. Therefore,

$$
\begin{aligned}
\Delta_{\mathcal{A}}(c) &\le 1/2\cdot\delta_{\mathcal{S}_1\cap\mathcal{A}}(c) + 1/4\cdot\delta_{\mathcal{T}_1\cap\mathcal{A}}(c) + 1/4\cdot\delta_{\mathcal{T}_2\cap\mathcal{A}}(c) + O(\varepsilon)(d^*+d_1) \\
&\le 1.65\,d^* - 0.7\,d_1 - 0.25\,d_2 + O(\varepsilon)(d^*+d_1) \\
(d_2 \ge d_1) \qquad &\le \boxed{1.65\,d^* - 0.95\,d_1} + O(\varepsilon)(d^*+d_1).
\end{aligned}
$$

When $\rho(f^*) > 3/4$, we have $\Pr[\mathcal{S}_1] = 5/4-\rho, \Pr[\mathcal{S}_2] = \rho - 3/4, \Pr[\mathcal{T}_1] = \Pr[\mathcal{T}_2] = 1/4$. Therefore,

$$
\begin{aligned}
\Delta_{\mathcal{A}}(c) &\le (5/4-\rho)\cdot\delta_{\mathcal{S}_1\cap\mathcal{A}}(c) + (\rho-3/4)\cdot\delta_{\mathcal{S}_2\cap\mathcal{A}}(c) + 1/4\cdot\delta_{\mathcal{T}_1\cap\mathcal{A}}(c) + 1/4\cdot\delta_{\mathcal{T}_2\cap\mathcal{A}}(c) + O(\varepsilon)(d^*+d_1) \\
&\le 1.65\,d^* - (1.6\rho - 0.5)\,d_1 + (1.6\rho - 1.45)\,d_2 + O(\varepsilon)(d^*+d_1).
\end{aligned}
$$

When $\rho \le 1.45/1.6$, we use $d_2 \ge d_1$:

$$
\Delta_{\mathcal{A}}(c) \le \boxed{1.65\,d^* - 0.95\,d_1} + O(\varepsilon)(d^*+d_1).
$$

When $\rho > 1.45/1.6$, we use $d_2 \le d^* + 1/\rho\cdot(d^*+d_1)$:

$$
\begin{aligned}
\Delta_{\mathcal{A}}(c) &\le (1.8 + 1.6\rho - 1.45/\rho)\,d^* - (1.6\rho + 1.45/\rho - 2.1)\,d_1 + O(\varepsilon)(d^*+d_1) \\
&\le (1.8 + 1.6 - 1.45)\,d^* - (2\sqrt{1.6\times 1.45} - 2.1)\,d_1 + O(\varepsilon)(d^*+d_1) \\
&\le \boxed{1.95\,d^* - 0.94630\,d_1} + O(\varepsilon)(d^*+d_1).
\end{aligned}
$$

D.4.6 $\langle\langle f^*\rangle\rangle$ closes f_1 and f_2 on $\mathcal{T}'_b \cap \mathcal{A}$ for some $b \in \{1,2\}$

Bounds for simple swaps remain the same as before:

$$
\begin{aligned}
\delta_{\mathcal{S}_1\cap\mathcal{A}}(c) &\le \boxed{(1+\alpha\beta)\,d^* - (1-\alpha\beta)\,d_1 - 2\beta\,d_2}, \\
\delta_{\mathcal{S}_2\cap\mathcal{A}}(c) &\le \boxed{(1+\alpha\beta)\,d^* - 2\,d_1 + (1+\alpha\beta-2\beta)\,d_2}.
\end{aligned}
$$

For tree swaps, we partition $\mathcal{T}\cap\mathcal{A}$ as the union of $\mathcal{T}_1 \cap \mathcal{T}'_{3-b} \cap \mathcal{A}$, $\mathcal{T}_2 \cap \mathcal{T}'_{3-b} \cap \mathcal{A}$ and $\mathcal{T}'_b \cap \mathcal{A}$. On the first two events, our bounds are the same as in Appendix D.4.3:

$$
\begin{aligned}
\delta_{\mathcal{T}_1\cap\mathcal{T}'_{3-b}\cap\mathcal{A}}(c) &\le \boxed{(1+\alpha\beta+\beta)\,d^* - d_1}, \\
\delta_{\mathcal{T}_2\cap\mathcal{T}'_{3-b}\cap\mathcal{A}}(c) &\le \boxed{(1+\alpha\beta+\beta)\,d^* - (2-2\beta)\,d_1 + (1-2\beta)\,d_2}.
\end{aligned}
$$

On $\mathcal{T}'_b \cap \mathcal{A}$, we have $\langle\langle f^*\rangle\rangle$ closes both f_1 and f_2. Therefore,

$$
(\delta_{\langle\langle f^*, \neg f_1, \neg f_2\rangle\rangle}) \qquad \delta_{\mathcal{T}'_b\cap\mathcal{A}}(c) \le \boxed{(1+\alpha\beta)\,d^* - d_1 - \beta\,d_2}.
$$

Summarizing, we have

$$\begin{aligned}
\delta_{\mathcal{S}_1\cap\mathcal{A}}(c) &\le (1+\alpha\beta)\, d^* - (1-\alpha\beta)\, d_1 - 2\beta\, d_2 && = 1.6\, d^* - 0.4\, d_1 - 0.4\, d_2 \\
\delta_{\mathcal{S}_2\cap\mathcal{A}}(c) &\le (1+\alpha\beta)\, d^* - 2\, d_1 + (1+\alpha\beta-2\beta)\, d_2 && = 1.6\, d^* - 2\, d_1 + 1.2\, d_2 \\
\delta_{\mathcal{T}_1\cap\mathcal{T}'_{3-b}\cap\mathcal{A}}(c) &\le (1+\alpha\beta+\beta)\, d^* - d_1 && = 1.8\, d^* - d_1 \\
\delta_{\mathcal{T}_2\cap\mathcal{T}'_{3-b}\cap\mathcal{A}}(c) &\le (1+\alpha\beta+\beta)\, d^* - (2-2\beta)\, d_1 + (1-2\beta)\, d_2 && = 1.8\, d^* - 1.6\, d_1 + 0.6\, d_2 \\
\delta_{\mathcal{T}'_b\cap\mathcal{A}}(c) &\le (1+\alpha\beta)\, d^* - d_1 - \beta\, d_2 && = 1.6\, d^* - d_1 - 0.2\, d_2
\end{aligned}$$

We now combine these inequalities to get an upper bound for $\Delta_{\mathcal{A}}(c)$.

When $2/3 < \rho(f^*) \le 3/4$, we have $\Pr[\mathcal{S}_1] = 1/2, \Pr[\mathcal{S}_2] = 0, \Pr[\mathcal{T}_1 \cap \mathcal{T}'_{3-b}] = \Pr[\mathcal{T}_2 \cap \mathcal{T}'_{3-b}] = 1/8, \Pr[\mathcal{T}'_b] = 1/4$. Therefore,

$$\begin{aligned}
\Delta_{\mathcal{A}}(c) &\le 1/2 \cdot \delta_{\mathcal{S}_1\cap\mathcal{A}}(c) + 1/8 \cdot \delta_{\mathcal{T}_1\cap\mathcal{T}'_{3-b}\cap\mathcal{A}}(c) + 1/8 \cdot \delta_{\mathcal{T}_2\cap\mathcal{T}'_{3-b}\cap\mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}'_b\cap\mathcal{A}}(c) + O(\varepsilon)(d^* + d_1) \\
&\le 1.65\, d^* - 0.775\, d_1 - 0.175\, d_2 + O(\varepsilon)(d^* + d_1) \\
(d_2 \ge d_1) \quad &\le \boxed{1.65\, d^* - 0.95\, d_1} + O(\varepsilon)(d^* + d_1).
\end{aligned}$$

When $\rho(f^*) > 3/4$, we have $\Pr[\mathcal{S}_1] = 5/4 - \rho, \Pr[\mathcal{S}_2] = \rho - 3/4, \Pr[\mathcal{T}_1 \cap \mathcal{T}'_{3-b}] = \Pr[\mathcal{T}_2 \cap \mathcal{T}'_{3-b}] = 1/8, \Pr[\mathcal{T}'_b] = 1/4$. Therefore,

$$\begin{aligned}
\Delta_{\mathcal{A}}(c) \le{}& (5/4 - \rho) \cdot \delta_{\mathcal{S}_1\cap\mathcal{A}}(c) + (\rho - 3/4) \cdot \delta_{\mathcal{S}_2\cap\mathcal{A}}(c) + 1/8 \cdot \delta_{\mathcal{T}_1\cap\mathcal{T}'_{3-b}\cap\mathcal{A}}(c) + 1/8 \cdot \delta_{\mathcal{T}_2\cap\mathcal{T}'_{3-b}\cap\mathcal{A}}(c) \\
&+ 1/4 \cdot \delta_{\mathcal{T}'_b\cap\mathcal{A}}(c) + O(\varepsilon)(d^* + d_1) \\
\le{}& 1.65\, d^* - (1.6\rho - 0.425)\, d_1 + (1.6\rho - 1.375)\, d_2 + O(\varepsilon)(d^* + d_1).
\end{aligned}$$

When $\rho \le 1.375/1.6$, we use $d_2 \ge d_1$:

$$\Delta_{\mathcal{A}}(c) \le \boxed{1.65\, d^* - 0.95\, d_1} + O(\varepsilon)(d^* + d_1).$$

When $\rho > 1.375/1.6$, we use $d_2 \le d^* + 1/\rho(d^* + d_1)$:

$$\begin{aligned}
\Delta_{\mathcal{A}}(c) &\le (1.875 + 1.6\rho - 1.375/\rho)\, d^* - (1.6\rho + 1.375/\rho - 2.025)\, d_1 + O(\varepsilon)(d^* + d_1) \\
&\le (1.875 + 1.6 - 1.375)\, d^* - (2\sqrt{1.6 \times 1.375} - 2.025)\, d_1 + O(\varepsilon)(d^* + d_1) \\
&\le \boxed{2.1\, d^* - 0.94147\, d_1} + O(\varepsilon)(d^* + d_1).
\end{aligned}$$

D.5 Proof of (6.21): Clients of Type D

In this section, we show that for any client c with type D, we have

$$\Delta_{\mathcal{A}}(c) \le 2.5203\, d^*(c) - 0.8888\, d_1(c).$$

Similar to Claim D.3 for type C clients, we also have the following claim classifying type D clients into subtypes. The only change is in item (e), where we replace $\mathcal{T}_2$ by $\mathcal{T}_1$ because the roles of η_1 and η_2 are now swapped.

Claim D.4 (Type D subcases)**.** *For a client c of type D, one of the following is true:*

(a) f_1 *is heavy.*

(b) f_2 *is heavy.*

(c) *A facility* h *is open near* c *after the simple swap closing* f_1*. Formally, a facility* $h \neq f_2$ *is open after swap* $\langle\langle \neg f_1 \rangle\rangle$ *at distance* $d(c,h) \leq 3d_1 + 2d^*$ *on* $\mathcal{S} \cap \mathcal{A}$*.*

(d) $g^* \neq f^*$*,* $\rho(g^*) > 3/4$*, and for all* $b = 1, 2$*, any swap set* $\mathcal{P}$ *generated on* $\mathcal{S}'_b \cap \mathcal{A}$*, a facility* $h \neq f_2$ *is open after swap* $\langle\langle \neg f_1 \rangle\rangle$ *at distance* $d(c,h) \leq \begin{cases} 2d_1 + d^*, & \text{if } b = 1 \\ 2d_1 + d^* + 4/3(d_1 + d^*), & \text{if } b = 2 \end{cases}$.

(e) *For any swap set* $\mathcal{P}$ *generated on* $\mathcal{T}_1 \cap \mathcal{A}$*,* $\langle\langle f^* \rangle\rangle$ *closes both* f_1 *and* f_2*;*

(f) $g^* \neq f^*$*,* $\rho(g^*) > 2/3$*, and there exists* $b \in \{1, 2\}$ *such that for any swap set* $\mathcal{P}$ *generated on* $\mathcal{T}'_b \cap \mathcal{A}$*,* $\langle\langle f^* \rangle\rangle$ *closes both* f_1 *and* f_2*.*

D.5.1 When f_1 is a heavy facility

f_1 being heavy implies that $\langle\langle \neg f_1 \rangle\rangle$ doesn't exist. We thus focus on $\langle\langle f^* \rangle\rangle$ and $\langle\langle \neg f_2 \rangle\rangle$.

Simple & tree swaps with $\tau(f^*) = \eta_1$ We have $\langle\langle f^* \rangle\rangle = \langle\langle \neg f_2 \rangle\rangle$ by implication (ii) of amenability. Therefore,

$$(\delta_{\langle\langle f^*, \neg f_2 \rangle\rangle}) \qquad \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) \leq \boxed{(1 + \alpha\beta)d^* - d_1 - \beta\, d_2},$$

$$(\delta_{\langle\langle f^*, \neg f_2 \rangle\rangle}) \qquad \delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) \leq \boxed{(1 + \alpha\beta)d^* - d_1 - \beta\, d_2}.$$

Simple swaps with $\tau(f^*) = \eta_2$ By implication (Siv) of amenability, we have $\langle\langle f^* \rangle\rangle \neq \langle\langle \neg f_2 \rangle\rangle$. On swap $\langle\langle f^* \rangle\rangle$ the client can be served by f^*. On swap $\langle\langle \neg f_2 \rangle\rangle$, the client can be served by f_1. Therefore,

$$\begin{aligned}
&(\delta_{\langle\langle f^* \rangle\rangle}) & \delta_{\mathcal{S}_2 \cap \mathcal{A}}(c) &\leq (1 + \alpha\beta)d^* - d_1 - \beta\, d_2 \\
&(\delta_{\langle\langle \neg f_2 \rangle\rangle}) & &\quad + (1 + \alpha\beta)d_1 - d_1 - \beta\, d_2 \\
& & &= \boxed{(1 + \alpha\beta)d^* - (1 - \alpha\beta)\, d_1 - 2\beta\, d_2}
\end{aligned}$$

Tree swaps with $\tau(f^*) = \eta_2$ Let us first assume that $\langle\langle f^* \rangle\rangle \neq \langle\langle \neg f_2 \rangle\rangle$. On swap $\langle\langle f^* \rangle\rangle$, the client can be served by f^*. On swap $\langle\langle \neg f_2 \rangle\rangle$, the client can be served by f_1 and $\pi(f_2)$, by implication (Tii) of amenability. We have $d(c, \pi(f_2)) \leq\leq 2d_2 + d^*$ by (C.46). Therefore,

$$\begin{aligned}
&(\delta_{\langle\langle f^* \rangle\rangle}) & \delta_{\mathcal{T}_2 \cap \mathcal{A}}(c) &\leq (1 + \alpha\beta)d^* - d_1 - \beta\, d_2 \\
&(\delta_{\langle\langle \neg f_2 \rangle\rangle}) & &\quad + d_1 + \beta(2d_2 + d^*) - d_1 - \beta\, d_2 \\
& & &= \boxed{(1 + \alpha\beta + \beta)d^* - d_1}
\end{aligned}$$

This inequality also holds when $\langle\langle \neg f_2 \rangle\rangle = \langle\langle f^* \rangle\rangle$, because our bound for $\delta_{\langle\langle f^* \rangle\rangle}$ does not require f_2 to remain open after the swap and $\delta_{\langle\langle \neg f_2 \rangle\rangle}$ is non-negative.

$$\begin{aligned}
\delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) &\leq (1 + \alpha\beta)\, d^* - d_1 - \beta\, d_2 &&= 1.6\, d^* - d_1 - 0.2\, d_2 \\
\delta_{\mathcal{S}_2 \cap \mathcal{A}}(c) &\leq (1 + \alpha\beta)\, d^* - (1 - \alpha\beta)\, d_1 - 2\beta\, d_2 &&= 1.6\, d^* - 0.4\, d_1 - 0.4\, d_2 \\
\delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) &\leq (1 + \alpha\beta)\, d^* - d_1 - \beta\, d_2 &&= 1.6\, d^* - d_1 - 0.2\, d_2 \\
\delta_{\mathcal{T}_2 \cap \mathcal{A}}(c) &\leq (1 + \alpha\beta + \beta)\, d^* - d_1 &&= 1.8\, d^* - d_1
\end{aligned}$$

We now combine these inequalities to get an upper bound for $\Delta_{\mathcal{A}}(c)$.

When $\rho(f^*) \leq 2/3$, we have $\Pr[\mathcal{S}_1] = \Pr[\mathcal{T}_1] = 1/2, \Pr[\mathcal{S}_2] = \Pr[\mathcal{T}_2] = 0$. Therefore,

$$\begin{aligned} \Delta_{\mathcal{A}}(c) &\leq 1/2 \cdot \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) + 1/2 \cdot \delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) + O(\varepsilon)(d^* + d_1) \\ (d_1 \leq d_2) \qquad &\leq \boxed{1.6\, d^* - 1.2\, d_1} + O(\varepsilon)(d^* + d_1). \end{aligned}$$

When $2/3 < \rho(f^*) \leq 3/4$, we have $\Pr[\mathcal{S}_1] = 1/2, \Pr[\mathcal{S}_2] = 0, \Pr[\mathcal{T}_1] = \Pr[\mathcal{T}_2] = 1/4$. Therefore,

$$\begin{aligned} \Delta_{\mathcal{A}}(c) &\leq 1/2 \cdot \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_2 \cap \mathcal{A}}(c) + O(\varepsilon)(d^* + d_1) \\ &\leq 1.65\, d^* - d_1 - 0.15\, d_2 + O(\varepsilon)(d^* + d_1) \\ (d_1 \leq d_2) \qquad &\leq \boxed{1.65\, d^* - 1.15\, d_1} + O(\varepsilon)(d^* + d_1). \end{aligned}$$

When $\rho(f^*) > 3/4$, we have $\Pr[\mathcal{S}_1] = 5/4 - \rho, \Pr[\mathcal{S}_2] = \rho - 3/4, \Pr[\mathcal{T}_1] = \Pr[\mathcal{T}_2] = 1/4$. Therefore,

$$\begin{aligned} \Delta_{\mathcal{A}}(c) &\leq (5/4 - \rho) \cdot \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) + (\rho - 3/4) \cdot \delta_{\mathcal{S}_2 \cap \mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_2 \cap \mathcal{A}}(c) + O(\varepsilon)(d^* + d_1) \\ &\leq 1.65\, d^* - (1.45 - 0.6\rho)\, d_1 - (0.2\rho)\, d_2 + O(\varepsilon)(d^* + d_1) \\ (d_1 \leq d_2) \qquad & \\ &\leq 1.65\, d^* - (1.45 - 0.4\rho)\, d_1 + O(\varepsilon)(d^* + d_1) \\ &\leq \boxed{1.65\, d^* - 1.05\, d_1} + O(\varepsilon)(d^* + d_1) \end{aligned}$$

D.5.2 When f_2 is a heavy facility

f_2 being heavy implies that the swap $\langle\langle \neg f_2 \rangle\rangle$ doesn't exist. We thus focus on $\langle\langle f^* \rangle\rangle$ and $\langle\langle \neg f_1 \rangle\rangle$.

Simple swaps with $\tau(f^*) = \eta_1$ We have $\langle\langle f^* \rangle\rangle \neq \langle\langle \neg f_1 \rangle\rangle$ by implication (Siv) of amenability. On swap $\langle\langle f^* \rangle\rangle$, the client can be served by f^* and f_1. On swap $\langle\langle \neg f_1 \rangle\rangle$, the client can be served by f_2. Therefore,

$$\begin{aligned} (\delta_{\langle\langle f^* \rangle\rangle}) \qquad \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) &\leq d^* + \beta d_1 - d_1 - \beta\, d_2 \\ (\delta_{\langle\langle \neg f_1 \rangle\rangle}) \qquad &\quad + (1 + \alpha\beta) d_2 - d_1 - \beta\, d_2 \\ &= \boxed{d^* - (2 - \beta)\, d_1 + (1 + \alpha\beta - 2\beta) d_2}. \end{aligned}$$

We can also use $(1 - \beta) d^* + 2\beta\, d_1$ to upper-bound $\delta_{\langle\langle f^* \rangle\rangle}$ (by (6.10)) and get

$$\delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) \leq \boxed{(1 - \beta)\, d^* - (2 - 2\beta)\, d_1 + (1 + \alpha\beta - 2\beta) d_2}.$$

Tree swaps with $\tau(f^*) = \eta_1$ Let us first assume that $\langle\langle f^* \rangle\rangle \neq \langle\langle \neg f_1 \rangle\rangle$. On swap $\langle\langle f^* \rangle\rangle$, the client can be served by f^* and f_1. On swap $\langle\langle \neg f_1 \rangle\rangle$, the client can be served by f_2 and $\pi(f_1)$ by implication (Tii) of amenability. We have $d(c, \pi(f_1)) \leq 2d_1 + d^*$ by (6.8). Therefore,

$$\begin{aligned} (\delta_{\langle\langle f^* \rangle\rangle}) \qquad \delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) &\leq d^* + \beta\, d_1 - d_1 - \beta\, d_2 \\ (\delta_{\langle\langle \neg f_1 \rangle\rangle}) \qquad &\quad + d_2 + \beta(2d_1 + d^*) - d_1 - \beta\, d_2 \\ &= \boxed{(1 + \beta) d^* - (2 - 3\beta) d_1 + (1 - 2\beta)\, d_2}. \end{aligned}$$

If $\langle\langle \neg f_1 \rangle\rangle = \langle\langle f^* \rangle\rangle$, we still have the same bound:

$$\begin{aligned}
&(\delta_{\langle\langle f^*, \neg f_1 \rangle\rangle}) && \delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) \le d^* + \beta\, d_2 - d_1 - \beta\, d_2 \\
&(\text{non-negative terms}) && \qquad + (1-2\beta)(d_2 - d_1) + \beta\, d^* + \beta\, d_1 \\
& && \quad = (1+\beta) d^* - (2 - 3\beta) d_1 + (1-2\beta)\, d_2.
\end{aligned}$$

Simple & tree swaps with $\tau(f^*) = \eta_2$ We have $\langle\langle f^* \rangle\rangle = \langle\langle \neg f_1 \rangle\rangle$ by implication (ii) of amenability. Therefore,

$$\begin{aligned}
&(\delta_{\langle\langle f^*, \neg f_1 \rangle\rangle}) && \delta_{\mathcal{S}_2 \cap \mathcal{A}}(c) \le \boxed{(1+\alpha\beta) d^* - d_1 - \beta\, d_2}, \\
&(\delta_{\langle\langle f^*, \neg f_1 \rangle\rangle}) && \delta_{\mathcal{T}_2 \cap \mathcal{A}}(c) \le d^* + \beta d_2 - d_1 - \beta\, d_2 \\
& && \qquad = \boxed{d^* - d_1}.
\end{aligned}$$

Summarizing, we have

$$\begin{aligned}
\delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) &\le d^* - (2-\beta)\, d_1 + (1 + \alpha\beta - 2\beta)\, d_2 && = d^* - 1.8\, d_1 + 1.2\, d_2 \\
\delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) &\le (1-\beta)\, d^* - (2 - 2\beta)\, d_1 + (1 + \alpha\beta - 2\beta) d_2 && = 0.8\, d^* - 1.6\, d_1 + 1.2\, d_2 \\
\\
\delta_{\mathcal{S}_2 \cap \mathcal{A}}(c) &\le (1+\alpha\beta) d^* - d_1 - \beta d_2 && = 1.6\, d^* - d_1 - 0.2\, d_2 \\
\delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) &\le (1+\beta) d^* - (2-3\beta) d_1 + (1-2\beta) d_2 && = 1.2\, d^* - 1.4\, d_1 + 0.6\, d_2 \\
\delta_{\mathcal{T}_2 \cap \mathcal{A}}(c) &\le d^* - d_1 && = d^* - d_1
\end{aligned}$$

We now combine these inequalities to get an upper bound for $\Delta_{\mathcal{A}}(c)$.

When $\rho(f^*) \le 2/3$, we have $\Pr[\mathcal{S}_1] = \Pr[\mathcal{T}_1] = 1/2, \Pr[\mathcal{S}_2] = \Pr[\mathcal{T}_2] = 0$. In this case we use the second inequality for $\delta_{\mathcal{S}_1 \cap \mathcal{A}}(c)$. Therefore,

$$\begin{aligned}
\Delta_{\mathcal{A}}(c) &\le 1/2 \cdot \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) + 1/2 \cdot \delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) + O(\varepsilon)(d^* + d_1) \\
&\le d^* - 1.5\, d_1 + 0.9\, d_2 + O(\varepsilon)(d^* + d_1) \\
(d_2 \le d^* + \rho(d^* + d_1)) \qquad &\le (1.9 + 0.9\rho)\, d^* - (1.5 - 0.9\rho)\, d_1 + O(\varepsilon)(d^* + d_1) \\
&\le (1.9 + 0.9 \cdot 2/3) d^* - (1.5 - 0.9 \cdot 2/3)\, d_1 + O(\varepsilon)(d^* + d_1) \\
&\le \boxed{2.5\, d^* - 0.9\, d_1} + O(\varepsilon)(d^* + d_1).
\end{aligned}$$

When $2/3 < \rho(f^*) \le 3/4$, we have $\Pr[\mathcal{S}_1] = 1/2, \Pr[\mathcal{S}_2] = 0, \Pr[\mathcal{T}_1] = \Pr[\mathcal{T}_2] = 1/4$. In this case we use the first inequality for $\delta_{\mathcal{S}_1 \cap \mathcal{A}}(c)$. Therefore,

$$\begin{aligned}
\Delta_{\mathcal{A}}(c) &\le 1/2 \cdot \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_2 \cap \mathcal{A}}(c) + O(\varepsilon)(d^* + d_1) \\
&\le 1.05\, d^* - 1.5\, d_1 + 0.75\, d_2 + O(\varepsilon)(d^* + d_1) \\
(d_2 \le d^* + \rho(d^* + d_1)) \qquad &\le (1.8 + 0.75\rho)\, d^* - (1.5 - 0.75\rho)\, d_1 + O(\varepsilon)(d^* + d_1) \\
&\le (1.8 + 0.75 \cdot 3/4) - (1.5 - 0.75 \cdot 3/4)\, d_1 + O(\varepsilon)(d^* + d_1) \\
&\le \boxed{2.3625\, d^* - 0.9375\, d_1} + O(\varepsilon)(d^* + d_1).
\end{aligned}$$

When $\rho(f^*) > 3/4$, we have $\Pr[\mathcal{S}_1] = 5/4 - \rho, \Pr[\mathcal{S}_2] = \rho - 3/4, \Pr[\mathcal{T}_1] = \Pr[\mathcal{T}_2] = 1/4$. In this case we use the first inequality for $\delta_{\mathcal{S}_1 \cap \mathcal{A}}(c)$. Therefore,

$$\Delta_{\mathcal{A}}(c) \le (5/4 - \rho) \cdot \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) + (\rho - 3/4) \cdot \delta_{\mathcal{S}_2 \cap \mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_2 \cap \mathcal{A}}(c) + O(\varepsilon)(d^* + d_1)$$

$$\leq (0.6 + 0.6\rho)\, d^* - (2.1 - 0.8\rho)\, d_1 + (1.8 - 1.4\rho)\, d_2 + O(\varepsilon)(d^* + d_1)$$

$(d_2 \leq d^* + \rho(d^* + d_1))$

$$\leq (2.4 + \rho - 1.4\rho^2)\, d^* - (2.1 - 2.6\rho + 1.4\rho^2)\, d_1 + O(\varepsilon)(d^* + d_1)$$

$$\leq (2.4 + 3/4 - 1.4(3/4)^2)\, d^* - (2.1 - 2.6 \cdot 13/14 + 1.4(13/14)^2)\, d_1 + O(\varepsilon)(d^* + d_1)$$

$$\leq \boxed{2.3625\, d^* - 0.8928\, d_1} + O(\varepsilon)(d^* + d_1).$$

D.5.3 There exists a facility h such that $d(c, h) \leq 3d_1 + 2d^*$ in simple swaps

Simple swaps with $\tau(f^*) = \eta_1$ Implications (ii) and (Siv) of amenability imply $\langle\langle f^* \rangle\rangle = \langle\langle \neg f_2 \rangle\rangle \neq \langle\langle \neg f_1 \rangle\rangle$. On swap $\langle\langle f^*, \neg f_2 \rangle\rangle$, the client can be served by f^* and f_1. On swap $\langle\langle \neg f_1 \rangle\rangle$, the client can be served by f_2 and h. Therefore,

$$\begin{aligned}
&(\delta_{\langle\langle f^*, \neg f_2 \rangle\rangle}) && \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) \leq d^* + \beta\, d_1 - d_1 - \beta\, d_2 \\
&(\delta_{\langle\langle \neg f_1 \rangle\rangle}) && \qquad + d_2 + \beta(3d_1 + 2d^*) - d_1 - \beta\, d_2 \\
& && \qquad = \boxed{(1 + 2\beta)\, d^* - (2 - 4\beta)\, d_1 + (1 - 2\beta)\, d_2}.
\end{aligned}$$

In $\delta_{\langle\langle f^*, \neg f_2 \rangle\rangle}$, we can use $0.776(1 + \alpha\beta)d^* + 0.224(d^* + \beta\, d_1)$ instead of $d^* + \beta\, d_1$. This gives

$$\delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) \leq \boxed{(1 + 2\beta + 0.776\alpha\beta)\, d^* - (2 - 3.224\beta)\, d_1 + (1 - 2\beta)\, d_2}.$$

Simple swaps with $\tau(f^*) = \eta_2$ Implications (ii) and (Siv) of amenability imply $\langle\langle f^* \rangle\rangle = \langle\langle \neg f_1 \rangle\rangle \neq \langle\langle \neg f_2 \rangle\rangle$. On swap $\langle\langle f^*, \neg f_1 \rangle\rangle$, the client can be served by f^*. On swap $\langle\langle \neg f_2 \rangle\rangle$, the client can be served by f_1. Therefore,

$$\begin{aligned}
&(\delta_{\langle\langle f^*, \neg f_1 \rangle\rangle}) && \delta_{\mathcal{S}_2 \cap \mathcal{A}}(c) \leq (1 + \alpha\beta)\, d^* - d_1 - \beta\, d_2 \\
&(\delta_{\langle\langle \neg f_2 \rangle\rangle}) && \qquad + (1 + \alpha\beta)\, d_1 - d_1 - \beta\, d_2 \\
& && \qquad = \boxed{(1 + \alpha\beta)\, d^* - (1 - \alpha\beta)\, d_1 - 2\beta\, d_2}.
\end{aligned}$$

Tree swaps with $\tau(f^*) = \eta_1$ We have $\langle\langle f^* \rangle\rangle = \langle\langle \neg f_2 \rangle\rangle$ by implication (ii) of amenability. Let us first assume that $\langle\langle \neg f_1 \rangle\rangle \neq \langle\langle f^*, \neg f_2 \rangle\rangle$. On swap $\langle\langle f^*, \neg f_2 \rangle\rangle$, the client can be served by f^*. On swap $\langle\langle \neg f_1 \rangle\rangle$, the client can be served by f_2 and $\pi(f_1)$ by implication (Tii) of amenability. Note that $d(c, \pi(f_1)) \leq 2d_1 + d^*$ by (6.8). Therefore,

$$\begin{aligned}
&(\delta_{\langle\langle f^*, \neg f_2 \rangle\rangle}) && \delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) \leq (1 + \alpha\beta)\, d^* - d_1 - \beta\, d_2 \\
&(\delta_{\langle\langle \neg f_1 \rangle\rangle}) && \qquad + d_2 + \beta(2d_1 + d^*) - d_1 - \beta\, d_2 \\
& && \qquad = \boxed{(1 + \alpha\beta + \beta)\, d^* - (2 - 2\beta)\, d_1 + (1 - 2\beta)\, d_2}.
\end{aligned}$$

This inequality also holds when $\langle\langle \neg f_1 \rangle\rangle = \langle\langle f^*, \neg f_2 \rangle\rangle$, because our bound for $\delta_{\langle\langle f^*, \neg f_2 \rangle\rangle}$ does not require f_1 to remain open after the swap and $\delta_{\langle\langle \neg f_1 \rangle\rangle}$ is non-negative.

In $\delta_{\langle\langle f^*, \neg f_2 \rangle\rangle}$, we can use $d^* + \beta\, d_1$ instead of $(1 + \alpha\beta)\, d^*$. This gives

$$\delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) \leq \boxed{(1 + \beta)\, d^* - (2 - 3\beta)\, d_1 + (1 - 2\beta)\, d_2}.$$

This bound also holds when $\langle\langle \neg f_1 \rangle\rangle = \langle\langle f^*, \neg f_2 \rangle\rangle$ because in this case we have

$$(\delta_{\langle\langle f^*, \neg f_2, \neg f_1 \rangle\rangle}) \qquad \delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) \leq d^* + \beta\, d_1 - d_1 - \beta\, d_2$$

$$\text{(non-negative terms)} \qquad + (1-\beta)(d_2 - d_1) + \beta d_1 + \beta d^*$$
$$= (1+\beta)\, d^* - (2-3\beta)\, d_1 + (1-2\beta)\, d_2.$$

Tree swaps with $\tau(f^*) = \eta_2$ We have $\langle\langle f^* \rangle\rangle = \langle\langle \neg f_1 \rangle\rangle$ by implication (ii) of amenability. Let us first assume that $\langle\langle \neg f_2 \rangle\rangle \neq \langle\langle f^*, \neg f_1 \rangle\rangle$. On swap $\langle\langle f^*, \neg f_1 \rangle\rangle$, the client can be served by f^*. On swap $\langle\langle \neg f_2 \rangle\rangle$, the client can be served by f_1 and $\pi(f_2)$ by implication (Tii) of amenability. Note that $d(c, \pi(f_2)) \le d_2 + d(f_2, \pi(f_2)) \le d_2 + d(f_2, f^*) \le d_2 + \rho d(f_1, f^*) \le d_2 + \rho(d_1 + d^*)$. Therefore,

$$\begin{aligned}
&(\delta_{\langle\langle f^*, \neg f_1 \rangle\rangle}) && \delta_{\mathcal{T}_2 \cap \mathcal{A}}(c) \le (1+\alpha\beta)\, d^* - d_1 - \beta\, d_2 \\
&(\delta_{\langle\langle \neg f_2 \rangle\rangle}) && \qquad + d_1 + \beta(d_2 + \rho(d^* + d_1)) - d_1 - \beta\, d_2 \\
& && = \boxed{(1+\alpha\beta+\rho\beta)\, d^* - (1-\rho\beta)\, d_1 - \beta\, d_2}.
\end{aligned}$$

This inequality also holds when $\langle\langle \neg f_2 \rangle\rangle = \langle\langle f^*, \neg f_1 \rangle\rangle$, because our bound for $\delta_{\langle\langle f^*, \neg f_1 \rangle\rangle}$ does not require f_2 to remain open after the swap and $\delta_{\langle\langle \neg f_2 \rangle\rangle}$ is non-negative.

Summarizing, we have

$$\begin{aligned}
\delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) &\le (1+2\beta)\, d^* - (2-4\beta)\, d_1 + (1-2\beta)\, d_2 && = 1.4\, d^* - 1.2\, d_1 + 0.6\, d_2 \\
\delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) &\le (1+2\beta+0.776\alpha\beta)\, d^* - (2-3.224\beta)\, d_1 + (1-2\beta)\, d_2 && = 1.8656\, d^* - 1.3552\, d_1 + 0.6\, d_2 \\
\delta_{\mathcal{S}_2 \cap \mathcal{A}}(c) &\le (1+\alpha\beta)\, d^* - (1-\alpha\beta)\, d_1 - 2\beta\, d_2 && = 1.6\, d^* - 0.4\, d_1 - 0.4\, d_2 \\
\delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) &\le (1+\alpha\beta+\beta)\, d^* - (2-2\beta)\, d_1 + (1-2\beta)\, d_2 && = 1.8\, d^* - 1.6\, d_1 + 0.6\, d_2 \\
\delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) &\le (1+\beta)\, d^* - (2-3\beta)\, d_1 + (1-2\beta)\, d_2 && = 1.2\, d^* - 1.4\, d_1 + 0.6\, d_2 \\
\delta_{\mathcal{T}_2 \cap \mathcal{A}}(c) &\le (1+\alpha\beta+\rho\beta)\, d^* - (1-\rho\beta)\, d_1 - \beta\, d_2 && = (1.6+0.2\rho)\, d^* - (1-0.2\rho)\, d_1 - 0.2\, d_2
\end{aligned}$$

We now combine these inequalities to get an upper bound for $\Delta_{\mathcal{A}}(c)$.

When $\rho(f^*) \le 2/3$, we have $\Pr[\mathcal{S}_1] = \Pr[\mathcal{T}_1] = 1/2, \Pr[\mathcal{S}_2] = \Pr[\mathcal{T}_2] = 0$. We use the first bound for $\delta_{\mathcal{S}_1 \cap \mathcal{A}}(c)$ and the second bound for $\delta_{\mathcal{T}_1 \cap \mathcal{A}}(c)$. Therefore,

$$\begin{aligned}
& && \Delta_{\mathcal{A}}(c) \le 1/2 \cdot \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) + 1/2 \cdot \delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) + O(\varepsilon)(d^* + d_1) \\
& && \le 1.3\, d^* - 1.3\, d_1 + 0.6\, d_2 + O(\varepsilon)(d^* + d_1) \\
&(d_2 \le d^* + \rho(d^* + d_1)) && \le (1.9 + 0.6\rho)\, d^* - (1.3 - 0.6\rho)\, d_1 + O(\varepsilon)(d^* + d_1) \\
& && \le (1.9 + 0.6 \cdot 2/3)\, d^* - (1.3 - 0.6 \cdot 2/3)\, d_1 + O(\varepsilon)(d^* + d_1) \\
& && \le \boxed{2.3\, d^* - 0.9\, d_1} + O(\varepsilon)(d^* + d_1)
\end{aligned}$$

When $2/3 < \rho(f^*) \le 3/4$, we have $\Pr[\mathcal{S}_1] = 1/2, \Pr[\mathcal{S}_2] = 0, \Pr[\mathcal{T}_1] = \Pr[\mathcal{T}_2] = 1/4$. We use the first bound for both $\delta_{\mathcal{S}_1 \cap \mathcal{A}}(c)$ and $\delta_{\mathcal{T}_1 \cap \mathcal{A}}(c)$. Therefore,

$$\begin{aligned}
& && \Delta_{\mathcal{A}}(c) \le 1/2 \cdot \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_2 \cap \mathcal{A}}(c) + O(\varepsilon)(d^* + d_1) \\
& && \le (1.55 + 0.05\rho)\, d^* - (1.25 - 0.05\rho)\, d_1 + 0.4\, d_2 + O(\varepsilon)(d^* + d_1) \\
&(d_2 \le d^* + \rho(d^* + d_1)) && \le (1.95 + 0.45\rho)\, d^* - (1.25 - 0.45\rho)\, d_1 + O(\varepsilon)(d^* + d_1)
\end{aligned}$$

$$\leq (1.95 + 0.45 \cdot 3/4)\, d^* - (1.35 - 0.45 \cdot 3/4)\, d_1 + O(\varepsilon)(d^* + d_1)$$
$$= \boxed{2.2875\, d^* - 0.9125\, d_1} + O(\varepsilon)(d^* + d_1).$$

When $\rho(f^*) > 3/4$, we have $\Pr[\mathcal{S}_1] = 5/4 - \rho, \Pr[\mathcal{S}_2] = \rho - 3/4, \Pr[\mathcal{T}_1] = \Pr[\mathcal{T}_2] = 1/4$. We use the second bound for $\delta_{\mathcal{S}_1 \cap \mathcal{A}}(c)$ and the first bound for $\delta_{\mathcal{T}_1 \cap \mathcal{A}}(c)$. Therefore,

$$\begin{aligned}
\Delta_{\mathcal{A}}(c) &\leq (5/4 - \rho) \cdot \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) + (\rho - 3/4) \cdot \delta_{\mathcal{S}_2 \cap \mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_2 \cap \mathcal{A}}(c) + O(\varepsilon)(d^* + d_1) \\
&\leq (1.982 - 0.2156\rho)\, d^* - (2.044 - 1.0052\rho)\, d_1 + (1.15 - \rho)\, d_2 + O(\varepsilon)(d^* + d_1) \\
(d_2 \leq d^* + \rho(d^* + d_1)) \quad &\leq (3.132 - 0.0656\rho - \rho^2)\, d^* - (2.044 - 2.1552\rho + \rho^2)\, d_1 + O(\varepsilon)(d^* + d_1) \\
&\leq (3.132 - 0.0656 \cdot 3/4 - 3/4^2)\, d^* - (2.044 - 2.1552 + 1^2)\, d_1 + O(\varepsilon)(d^* + d_1) \\
&\leq \boxed{2.5203\, d^* - 0.8888\, d_1} + O(\varepsilon)(d^* + d_1).
\end{aligned}$$

D.5.4 $d(c,h) \leq 2d_1 + d^*$ or $d(c,h) \leq 2d_1 + d^* + 4/3(d^* + d_1)$ in simple swaps

Similarly to Appendix D.4.4, our bound for $\Delta_{\mathcal{A}}(c)$ in the previous case remains valid in this case.

D.5.5 $\langle\!\langle f^* \rangle\!\rangle$ closes f_1 and f_2 on $\mathcal{T}_1 \cap \mathcal{A}$

If $\tau(f^*) = \eta_2$, we get the same bounds as before:

$$\delta_{\mathcal{S}_2 \cap \mathcal{A}}(c) \leq \boxed{(1 + \alpha\beta)\, d^* - (1 - \alpha\beta)\, d_1 - 2\beta\, d_2},$$
$$\delta_{\mathcal{T}_2 \cap \mathcal{A}}(c) \leq \boxed{(1 + \alpha\beta + \rho\beta)\, d^* - (1 - \rho\beta)\, d_1 - \beta\, d_2}.$$

We proceed to bound $\delta_{\mathcal{S}_1 \cap \mathcal{A}}(c)$ and $\delta_{\mathcal{T}_1 \cap \mathcal{A}}(c)$.

Simple swaps with $\tau(f^*) = \eta_1$ Implications (ii) and (Siv) of amenability implies $\langle\!\langle f^* \rangle\!\rangle = \langle\!\langle \neg f_2 \rangle\!\rangle \neq \langle\!\langle \neg f_1 \rangle\!\rangle$. On swap $\langle\!\langle f^*, \neg f_2 \rangle\!\rangle$, the client can be served by f^* and f_1. On swap $\langle\!\langle \neg f_1 \rangle\!\rangle$, the client can be served by f_2. Therefore,

$$\begin{aligned}
(\delta_{\langle\!\langle f^*, \neg f_2 \rangle\!\rangle}) \quad \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) &\leq d^* + \beta\, d_1 - d_1 - \beta\, d_2 \\
(\delta_{\langle\!\langle \neg f_1 \rangle\!\rangle}) \quad &\quad + (1 + \alpha\beta)\, d_2 - d_1 - \beta\, d_2 \\
&= \boxed{d^* - (2 - \beta)\, d_1 + (1 + \alpha\beta - 2\beta)\, d_2}.
\end{aligned}$$

Tree swaps with $\tau(f^*) = \eta_1$ On $\mathcal{T}_1 \cap \mathcal{A}$, we assumed that $\langle\!\langle f^* \rangle\!\rangle$ closes f_1 and f_2. Therefore,

$$(\delta_{\langle\!\langle f^*, \neg f_1, \neg f_2 \rangle\!\rangle}) \quad \delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) \leq \boxed{(1 + \alpha\beta)\, d^* - d_1 - \beta\, d_2}.$$

Summarizing, we have

$$\begin{aligned}
\delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) &\leq d^* - (2 - \beta)\, d_1 + (1 + \alpha\beta - 2\beta)\, d_2 && = d^* - 1.8\, d_1 + 1.2\, d_2 \\
\delta_{\mathcal{S}_2 \cap \mathcal{A}}(c) &\leq (1 + \alpha\beta)\, d^* - (1 - \alpha\beta)\, d_1 - 2\beta\, d_2 && = 1.6\, d^* - 0.4\, d_1 - 0.4\, d_2 \\
\delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) &\leq (1 + \alpha\beta)\, d^* - d_1 - \beta\, d_2 && = 1.6\, d^* - d_1 - 0.2\, d_2 \\
\delta_{\mathcal{T}_2 \cap \mathcal{A}}(c) &\leq (1 + \alpha\beta + \rho\beta)\, d^* - (1 - \rho\beta)\, d_1 - \beta\, d_2 && = (1.6 + 0.2\rho)\, d^* - (1 - 0.2\rho)\, d_1 - 0.2\, d_2
\end{aligned}$$

We now combine these inequalities to get an upper bound for $\Delta_{\mathcal{A}}(c)$.

When $\rho(f^*) \leq 2/3$, we have $\Pr[\mathcal{S}_1] = \Pr[\mathcal{T}_1] = 1/2, \Pr[\mathcal{S}_2] = \Pr[\mathcal{T}_2] = 0$. Therefore,

$$\begin{aligned}
\Delta_{\mathcal{A}}(c) &\leq 1/2 \cdot \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) + 1/2 \cdot \delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) + O(\varepsilon)(d^* + d_1) \\
&\leq 1.3\, d^* - 1.4\, d_1 + 0.5\, d_2 + O(\varepsilon)(d^* + d_1) \\
(d_2 \leq d^* + \rho(d^* + d_1)) \quad &\leq (1.8 + 0.5\rho)\, d^* - (1.4 - 0.5\rho)\, d_1 + O(\varepsilon)(d^* + d_1) \\
&\leq (1.8 + 0.5 \times 2/3)\, d^* - (1.4 - 0.5 \times 2/3)\, d_1 + O(\varepsilon)(d^* + d_1) \\
&\leq \boxed{2.13334\, d^* - 1.06666\, d_1} + O(\varepsilon)(d^* + d_1).
\end{aligned}$$

When $2/3 < \rho(f^*) \leq 3/4$, we have $\Pr[\mathcal{S}_1] = 1/2, \Pr[\mathcal{S}_2] = 0, \Pr[\mathcal{T}_1] = \Pr[\mathcal{T}_2] = 1/4$. Therefore,

$$\begin{aligned}
\Delta_{\mathcal{A}}(c) &\leq 1/2 \cdot \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_2 \cap \mathcal{A}}(c) + O(\varepsilon)(d^* + d_1) \\
&\leq (1.3 + 0.05\rho)\, d^* - (1.4 - 0.05\rho)\, d_1 + 0.5\, d_2 + O(\varepsilon)(d^* + d_1) \\
(d_2 \leq d^* + \rho(d^* + d_1)) \quad &\leq (1.8 + 0.55\rho)\, d^* - (1.4 - 0.55\rho)\, d_1 + O(\varepsilon)(d^* + d_1) \\
&\leq (1.8 + 0.55 \times 3/4)\, d^* - (1.4 - 0.55 \times 3/4)\, d_1 + O(\varepsilon)(d^* + d_1) \\
&= \boxed{2.2125\, d^* - 0.9875\, d_1} + O(\varepsilon)(d^* + d_1).
\end{aligned}$$

When $\rho(f^*) > 3/4$, we have $\Pr[\mathcal{S}_1] = 5/4 - \rho, \Pr[\mathcal{S}_2] = \rho - 3/4, \Pr[\mathcal{T}_1] = \Pr[\mathcal{T}_2] = 1/4$. Therefore,

$$\begin{aligned}
\Delta_{\mathcal{A}}(c) &\leq (5/4 - \rho) \cdot \delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) + (\rho - 3/4) \cdot \delta_{\mathcal{S}_2 \cap \mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_1 \cap \mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_2 \cap \mathcal{A}}(c) + O(\varepsilon)(d^* + d_1) \\
&\leq (0.85 + 0.65\rho)\, d^* - (2.45 - 1.45\rho)\, d_1 + (1.7 - 1.6\rho)\, d_2 + O(\varepsilon)(d^* + d_1) \\
&(d_2 \leq d^* + \rho(d^* + d_1)) \\
&\leq (2.55 + 0.75\rho - 1.6\rho^2)\, d^* - (2.45 - 3.15\rho + 1.6\rho^2)\, d_1 + O(\varepsilon)(d^* + d_1) \\
&\leq (2.55 + 0.75 \times 3/4 - 1.6 \times (3/4)^2)\, d^* - (2.45 - 3.15 \times 3.15/3.2 + 1.6 \times (3.15/3.2)^2)\, d_1 \\
&\quad + O(\varepsilon)(d^* + d_1) \\
&\leq \boxed{2.2125\, d^* - 0.89960\, d_1} + O(\varepsilon)(d^* + d_1).
\end{aligned}$$

D.5.6 $\langle\langle f^* \rangle\rangle$ closes f_1 and f_2 on $\mathcal{T}'_b \cap \mathcal{A}$

Bounds for simple swaps remain the same as before:

$$\begin{aligned}
\delta_{\mathcal{S}_1 \cap \mathcal{A}}(c) &\leq \boxed{d^* - (2 - \beta)\, d_1 + (1 + \alpha\beta - 2\beta)\, d_2}, \\
\delta_{\mathcal{S}_2 \cap \mathcal{A}}(c) &\leq \boxed{(1 + \alpha\beta)\, d^* - (1 - \alpha\beta)\, d_1 - 2\beta\, d_2}.
\end{aligned}$$

For tree swaps, we partition $\mathcal{T} \cap \mathcal{A}$ as the union of $\mathcal{T}_1 \cap \mathcal{T}'_{3-b} \cap \mathcal{A}$, $\mathcal{T}_2 \cap \mathcal{T}'_{3-b} \cap \mathcal{A}$ and $\mathcal{T}'_b \cap \mathcal{A}$. On the first two events, our bounds are the same as in Appendix D.5.3:

$$\begin{aligned}
\delta_{\mathcal{T}_1 \cap \mathcal{T}'_{3-b} \cap \mathcal{A}}(c) &\leq \boxed{(1 + \alpha\beta + \beta)\, d^* - (2 - 2\beta)\, d_1 + (1 - 2\beta)\, d_2}, \\
\delta_{\mathcal{T}_2 \cap \mathcal{T}'_{3-b} \cap \mathcal{A}}(c) &\leq \boxed{(1 + \alpha\beta + \rho\beta)\, d^* - (1 - \rho\beta)\, d_1 - \beta\, d_2}.
\end{aligned}$$

On $\mathcal{T}'_b \cap \mathcal{A}$, we assumed that $\langle\langle f^* \rangle\rangle$ closes f_1 and f_2. Therefore,

$$(\delta_{\langle\langle f^*, \neg f_1, \neg f_2 \rangle\rangle}) \qquad \delta_{\mathcal{T}'_b \cap \mathcal{A}} \leq \boxed{(1 + \alpha\beta)\, d^* - d_1 - \beta\, d_2}.$$

Summarizing, we have

$$
\begin{aligned}
\delta_{\mathcal{S}_1\cap\mathcal{A}}(c) &\le d^* - (2-\beta)\,d_1 + (1+\alpha\beta-2\beta)\,d_2 && = d^* - 1.8\,d_1 + 1.2\,d_2\\
\delta_{\mathcal{S}_2\cap\mathcal{A}}(c) &\le (1+\alpha\beta)\,d^* - (1-\alpha\beta)\,d_1 - 2\beta\,d_2 && = 1.6\,d^* - 0.4\,d_1 - 0.4\,d_2\\
\delta_{\mathcal{T}_1\cap\mathcal{T}'_{3-b}\cap\mathcal{A}}(c) &\le (1+\alpha\beta+\beta)\,d^* - (2-2\beta)\,d_1 + (1-2\beta)\,d_2 &&\\
& && = 1.8\,d^* - 1.6\,d_1 + 0.6\,d_2\\
\delta_{\mathcal{T}_2\cap\mathcal{T}'_{3-b}\cap\mathcal{A}}(c) &\le (1+\alpha\beta+\rho\beta)\,d^* - (1-\rho\beta)\,d_1 - \beta\,d_2 && = (1.6+0.2\rho)\,d^* - (1-0.2\rho)\,d_1 - 0.2\,d_2\\
\delta_{\mathcal{T}'_b\cap\mathcal{A}}(c) &\le (1+\alpha\beta)\,d^* - d_1 - \beta\,d_2 && = 1.6\,d^* - d_1 - 0.2\,d_2
\end{aligned}
$$

We now combine these inequalities to get an upper bound for $\Delta_{\mathcal{A}}(c)$.

When $\rho(f^*) \le 2/3$, we have $\Pr[\mathcal{S}_1] = 1/2, \Pr[\mathcal{T}_1 \cap \mathcal{T}'_{3-b}] = \Pr[\mathcal{T}'_b] = 1/4, \Pr[\mathcal{S}_2] = \Pr[\mathcal{T}_2 \cap \mathcal{T}'_{3-b}] = 0$. Therefore,

$$
\begin{aligned}
\Delta_{\mathcal{A}}(c) &\le 1/2 \cdot \delta_{\mathcal{S}_1\cap\mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}_1\cap\mathcal{T}'_{3-b}\cap\mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}'_b\cap\mathcal{A}}(c) + O(\varepsilon)(d^*+d_1)\\
&\le 1.35\,d^* - 1.55\,d_1 + 0.7\,d_2 + O(\varepsilon)(d^*+d_1)\\
(d_2 \le d^* + \rho(d^*+d_1)) \qquad &\le (2.05+0.7\rho)\,d^* - (1.55-0.7\rho)\,d_1 + O(\varepsilon)(d^*+d_1)\\
&\le (2.05+0.7\times 2/3)\,d^* - (1.55-0.7\times 2/3)\,d_1 + O(\varepsilon)(d^*+d_1)\\
&\le \boxed{2.51667\,d^* - 1.08333\,d_1} + O(\varepsilon)(d^*+d_1).
\end{aligned}
$$

When $2/3 < \rho \le 3/4$, we have $\Pr[\mathcal{S}_1] = 1/2, \Pr[\mathcal{S}_2] = 0, \Pr[\mathcal{T}_1 \cap \mathcal{T}'_{3-b}] = \Pr[\mathcal{T}_2 \cap \mathcal{T}'_{3-b}] = 1/8, \Pr[\mathcal{T}'_b] = 1/4$. Therefore,

$$
\begin{aligned}
\Delta_{\mathcal{A}}(c) &\le 1/2 \cdot \delta_{\mathcal{S}_1\cap\mathcal{A}}(c) + 1/8 \cdot \delta_{\mathcal{T}_2\cap\mathcal{T}'_{3-b}\cap\mathcal{A}}(c) + 1/8 \cdot \delta_{\mathcal{T}_2\cap\mathcal{T}'_{3-b}\cap\mathcal{A}}(c) + 1/4 \cdot \delta_{\mathcal{T}'_b\cap\mathcal{A}}(c) + O(\varepsilon)(d^*+d_1)\\
&\le (1.325+0.025\rho)\,d^* - (1.475-0.025\rho)\,d_1 + 0.6\,d_2 + O(\varepsilon)(d^*+d_1)\\
&(d_2 \le d^* + \rho(d^*+d_1))\\
&\le (1.925+0.625\rho)\,d^* - (1.475-0.625\rho)\,d_1 + O(\varepsilon)(d^*+d_1)\\
&\le (1.925+0.625\times 3/4)\,d^* - (1.475-0.625\times 3/4)\,d_1 + O(\varepsilon)(d^*+d_1)\\
&= \boxed{2.39375\,d^* - 1.00625\,d_1} + O(\varepsilon)(d^*+d_1).
\end{aligned}
$$

When $\rho > 3/4$, we have $\Pr[\mathcal{S}_1] = 5/4-\rho, \Pr[\mathcal{S}_2] = \rho-3/4, \Pr[\mathcal{T}_1\cap\mathcal{T}'_{3-b}] = \Pr[\mathcal{T}_2\cap\mathcal{T}'_{3-b}] = 1/8, \Pr[\mathcal{T}'_b] = 1/4$. Therefore,

$$
\begin{aligned}
\Delta_{\mathcal{A}}(c) &\le (5/4-\rho)\cdot \delta_{\mathcal{S}_1\cap\mathcal{A}}(c) + (\rho-3/4)\cdot \delta_{\mathcal{S}_2\cap\mathcal{A}}(c) + 1/8 \cdot \delta_{\mathcal{T}_2\cap\mathcal{T}'_{3-b}\cap\mathcal{A}}(c) + 1/8 \cdot \delta_{\mathcal{T}_2\cap\mathcal{T}'_{3-b}\cap\mathcal{A}}(c)\\
&\quad + 1/4 \cdot \delta_{\mathcal{T}'_b\cap\mathcal{A}}(c) + O(\varepsilon)(d^*+d_1)\\
&\le (0.875+0.625\rho)\,d^* - (2.525-1.425\rho)\,d_1 + (1.8-1.6\rho)\,d_2 + O(\varepsilon)(d^*+d_1)\\
&(d_2 \le d^* + \rho(d^*+d_1))\\
&\le (2.675+0.825\rho-1.6\rho^2)\,d^* - (2.525-3.225\rho+1.6\rho^2)\,d_1 + O(\varepsilon)(d^*+d_1)\\
&\le (2.675+0.825\times 3/4 - 1.6\times(3/4)^2)\,d^* - (2.525-3.225+1.6)\,d_1 + O(\varepsilon)(d^*+d_1)\\
&= \boxed{2.39375\,d^* - 0.9\,d_1} + O(\varepsilon)(d^*+d_1).
\end{aligned}
$$

E Omitted Proofs

E.1 Proof of Claim 3.3: There are enough local candidates

Claim 3.3. *The number of local candidates is at least $t_d/2$ times the number of heavy local facilities.*

Proof. Let F_h be the set of heavy local facilities, $F_p \subseteq F \setminus F_h$ be the set of local facilities pointed to by at least one optimal facility with no heavy local neighbor, and F_c be the remaining local facilities, which are exactly the local candidates. $|F_h| + |F_p| + |F_c| =$ the number of local facilities, which in turn is at least the number of optimal facilities. There are at least $(t_d + 2)|F_h|/2$ many optimal facilities having a heavy local neighbor because 1) a heavy local facility is a neighbor of at least $t_d + 2$ optimal facilities, and 2) each optimal facility has at most 2 local neighbors. Finally, each local facility in F_p is pointed to by an optimal facility with no heavy local neighbor, so the total number of optimal facilities is at least $(t_d + 2)|F_h|/2 + |F_p|$. In other words, $|F_c| \geq \frac{t_d}{2}|F_h|$.

□

E.2 Proof of Claim 3.5: Balancing Procedure

Claim 3.5 (Balancing Procedure)**.** *Consider a universe $U = R \cup G$ of red points R and green points G, with $|G| = |R| + r$. Let the collection of sets $S_1, \ldots, S_N$ partition U, and let $|S_i| \leq x$ for all i. Moreover, let H be a graph on the vertices $[N]$ with maximum degree at most $\theta \leq r$. Lastly, $r \geq \Omega(\frac{x^5\theta^3}{\varepsilon})$ for some $0 \leq \varepsilon \leq 1$. Then we can merge these sets together into new sets $T_0, \ldots, T_M$ such that*

- *(i) each T_j has size $|T_j| \leq O(x^2)$,*
- *(ii) $|T_j \cap R| \leq |T_j \cap G|$,*
- *(iii) if there is an edge $\{i, j\}$ for $i, j \in [N]$, then S_i is not merged with S_j, and*
- *(iv) for all $i \neq j$, S_i is merged with S_j with probability at most ε.*

Proof of Claim 3.5. Recall $|G| = |R| + r$, where $r \geq \frac{16x^5\theta^2(\theta+1)}{\varepsilon}$ suffices. For each integer $s \in \{-x, \ldots, x\}$ let D_s be the sets S with discrepancy $|S \cap G| - |S \cap R|$. Each set in D_0 can be output immediately. If for some i, j we have $|D_i| \geq j/\varepsilon$ and $|D_{-j}| \geq i/\varepsilon$, and there is no edge in H, then we can choose some j sets uniformly at random from D_i, and i sets from D_{-j}, and merge these together.

However, since there are forbidden sets (a set S_1 and S_2 are forbidden if there is an edge between them in H), we need one more ingredient. We claim that if some D_i, D_{-j} have $\geq 8x^2\theta$ sets, then we can find j sets from D_i and i sets from D_{-j} that are not forbidden for each other. Indeed, pick a random collection of j sets from D_i and i sets from D_{-j}. The probability that any one set has an edge to any of the other $i + j - 1$ sets is $\leq \frac{(i+j-1)\theta}{8x^2\theta} < \frac{1}{4x}$. Hence, a union bound over all the $i + j$ sets says that with probability at least a half, this collection does not have any edges of H within it, and hence we can merge this collection together.

However, above procedure does not ensure two sets are combined with probability at most ε. To do so, if we find some pair D_i, D_{-j} with $\geq \frac{8x^2\theta}{\varepsilon}$ sets, then we can randomly partition each of D_i and D_{-j} into $1/\varepsilon$ equal-sized subgroups with $8x^2\theta$ sets each. Now we can merge some j sets from any subgroup from D_i with some i sets from a randomly chosen subgroup of D_{-j} to form a set with equal number of greens and reds, exactly as above. Henceforth, we assume that for each D_i, D_{-j}, at least one has fewer than $\frac{8x^2\theta}{\varepsilon}$ sets.

Finally, since the greens outnumber the reds by r, we know there exists a value $j > 0$ such that

$|D_j| \geq r/x = 16x^4\theta^2(\theta+1)/\varepsilon$. Thus, we know each D_s with $s < 0$ has at most $\frac{8x^2\theta}{\varepsilon}$ sets each. We randomly divide D_j into $\frac{16x^3\theta^2}{\varepsilon}$ parts of of size $x(\theta+1)$ sets each. Note any two sets S_a and S_b fall in the same part with probability at most $\frac{\varepsilon}{16x^3\theta^2} \leq \varepsilon$. From each part pick x sets that have no edge in H between themselves and call them a positive group; this can be done because the maximum degree of H is at most θ. Each such positive group has at least x extra green points. On the other hand, there are at most $x \cdot \frac{8x^2\theta}{\varepsilon} = \frac{8x^3\theta}{\varepsilon}$ negative sets, i.e., in $\{D_i\}_{i<0}$. Each negative set has edges to at most θ sets, so there are at most $\frac{8x^3\theta^2}{\varepsilon}$ sets with an edge to some negative set. Since there are $\frac{16x^3\theta^2}{\varepsilon}$ positive groups, there are at least $\frac{8x^3\theta^2}{\varepsilon}$ positive groups with no edge to any negative set, so we can merge each negative set with a randomly-chosen such positive group. This ensures that each new set has more green points than red, and two sets are combined with probability at most $\frac{\varepsilon}{8x^3\theta^2} \leq \varepsilon$. The newly-created sets have of size at most $O(x^2)$. Finally, each remaining set can form a group by itself, because they have more green points. □

E.3 Proof of Claim 5.3: Crude Upper Bound of Potential Change

Claim 5.3. *There is an absolute constant $\gamma > 0$ such that for any client c, and any swap set $\mathcal{P}$ that we generate, we have $\sum_{(P,Q)\in\mathcal{P}} \delta_{(P,Q)}(c) \leq \gamma(d^*(c) + d_1(c))$.*

Proof. Since every local facility is closed by at most 3 swaps in $\mathcal{P}$, there are at most 6 swaps in $\mathcal{P}$ that closes any facility in $\{f_1, f_2\}$. Thus, it suffices to show that $\delta_{(P,Q)}(c) \leq O(d^* + d_1)$ for these 6 swaps (P, Q).

If f^* has a heavy local neighbor h, the client can be served by h at distance $\leq d^* + {}^3\!/_2(d^* + d_1)$. We assume henceforth that f^* has no heavy local neighbor, which means $\tau(f^*)$ is not heavy and never closed as a local surrogate.

When $\mathcal{P}$ is a simple swap set, the client can be served by either f^* (at distance $\leq d^*$) or $\tau(f^*)$ (at distance $\leq d^* + {}^4\!/_3(d^* + d_1)$). When $\mathcal{P}$ is a tree swap set, we show that one of the following facilities must be open after every swap in $\mathcal{P}$:

$$\begin{aligned} f^* \text{ at distance } &\leq d^*, \\ \tau(f^*) \text{ at distance } &\leq d^* + {}^3\!/_2(d^* + d_1), \\ \pi(\tau(f^*)) \text{ at distance } &\leq d^* + 2 \cdot {}^3\!/_2(d^* + d_1). \end{aligned}$$

It suffices to show that any swap closing $\tau(f^*)$ must open either f^* or $\pi(\tau(f^*))$. If $\tau(f^*)$ is closed as an optimal surrogate, $\pi(\tau(f^*))$ must be open because edges on short cycles are not deleted in the edge deletion step (Corollary 3.8). We thus focus on the swap closing the original copy of $\tau(f^*)$ henceforth.

Consider the 1-forest G_1 before edge deletion. The edges in G_1 from f^* to $\tau(f^*)$ and from $\tau(f^*)$ to $\pi(\tau(f^*))$ cannot both be deleted in the edge deletion step, because we always choose t_{h} as an even number and G_1 is bipartite (when self-loops are ignored). Therefore, either f^* or $\pi(\tau(f^*))$ must be in the same swap with $\tau(f^*)$, as desired. □

E.4 Proof of Lemma D.1: Combining Type A Inequalities

Lemma D.1 (Type A averaging). *For a close client of type A with $\rho(f^*) > {}^2\!/_3$, we have*

$$\Delta_{\mathcal{T}\cap\mathcal{A}}(c) \leq {}^1\!/_4 \cdot \max\{\delta_{\mathcal{T}_{11}\cap\mathcal{A}} + \delta_{\mathcal{T}_{21}\cap\mathcal{A}}, \delta_{\mathcal{T}_{11}\cap\mathcal{A}} + \delta_{\mathcal{T}_{22}\cap\mathcal{A}}, \delta_{\mathcal{T}_{12}\cap\mathcal{A}} + \delta_{\mathcal{T}_{22}\cap\mathcal{A}}\} + O(\varepsilon)(d^* + d_1).$$

We first prove Lemma D.1 assuming the following lemma, which we prove later.

Lemma E.1. *For a close client of type* A *with* $\rho(f^*) > 2/3$*, we have*

$$\Pr[\mathcal{T}_{21}] \le \Pr[\mathcal{T}_{11}] + O(\varepsilon).$$

Proof of Lemma D.1. Define $p_{ij} := \Pr[\mathcal{T}_{ij}]$. Note that $\rho(f^*) > 2/3$ implies that $p_{11} + p_{12} = p_{21} + p_{22} = 1/4$. Define $p_\Delta := p_{11} - p_{21} = p_{22} - p_{12}$. Lemma E.1 implies $p_\Delta \ge -O(\varepsilon)$. Define $\delta_{\max} := \max\{\delta_{\mathcal{T}_{11}\cap\mathcal{A}} + \delta_{\mathcal{T}_{21}\cap\mathcal{A}}, \delta_{\mathcal{T}_{11}\cap\mathcal{A}} + \delta_{\mathcal{T}_{22}\cap\mathcal{A}}, \delta_{\mathcal{T}_{12}\cap\mathcal{A}} + \delta_{\mathcal{T}_{22}\cap\mathcal{A}}\}$. Claim 5.3 implies $\delta_{\max} \le O(d^* + d_1)$. Lemma D.1 is proved by the following chain of inequalities:

$$\begin{aligned}
\Delta_{\mathcal{T}\cap\mathcal{A}}(c) &\le \Pr[\mathcal{T}_{11}\cap\mathcal{A}]\delta_{\mathcal{T}_{11}\cap\mathcal{A}}(c) + \Pr[\mathcal{T}_{12}\cap\mathcal{A}]\delta_{\mathcal{T}_{12}\cap\mathcal{A}}(c) \\
&\quad + \Pr[\mathcal{T}_{21}\cap\mathcal{A}]\delta_{\mathcal{T}_{21}\cap\mathcal{A}}(c) + \Pr[\mathcal{T}_{22}\cap\mathcal{A}]\delta_{\mathcal{T}_{22}\cap\mathcal{A}}(c) \\
&\le p_{11}\delta_{\mathcal{T}_{11}\cap\mathcal{A}}(c) + p_{12}\delta_{\mathcal{T}_{12}\cap\mathcal{A}}(c) \\
&\quad + p_{21}\delta_{\mathcal{T}_{21}\cap\mathcal{A}}(c) + p_{22}\delta_{\mathcal{T}_{22}\cap\mathcal{A}}(c) \\
\text{(Claim 5.4 and } \delta_{\mathcal{E}}(c) \ge -10d_1) &\quad + O(\varepsilon)(d^* + d_1) \\
&= p_{21}\delta_{\mathcal{T}_{11}\cap\mathcal{A}}(c) + p_\Delta\delta_{\mathcal{T}_{11}\cap\mathcal{A}}(c) + p_{12}\delta_{\mathcal{T}_{12}\cap\mathcal{A}}(c) \\
&\quad + p_{21}\delta_{\mathcal{T}_{21}\cap\mathcal{A}}(c) + p_\Delta\delta_{\mathcal{T}_{22}\cap\mathcal{A}}(c) + p_{12}\delta_{\mathcal{T}_{22}\cap\mathcal{A}}(c) \\
&\quad + O(\varepsilon)(d^* + d_1) \\
&\le p_{21}\delta_{\max} + p_\Delta(\delta_{\mathcal{T}_{11}\cap\mathcal{A}}(c) + \delta_{\mathcal{T}_{22}\cap\mathcal{A}}(c)) + p_{12}\delta_{\max} + O(\varepsilon)(d^* + d_1) \\
(p_\Delta \ge -O(\varepsilon),\ \delta_{\mathcal{E}}(c) \ge -10d_1 \text{ and } \delta_{\max} \le O(d^* + d_1)) & \\
&\le p_{21}\delta_{\max} + p_\Delta\delta_{\max} + p_{12}\delta_{\max} + O(\varepsilon)(d^* + d_1) \\
&= (p_{21} + p_\Delta + p_{12})\delta_{\max} + O(\varepsilon)(d^* + d_1) \\
&= 1/4 \cdot \delta_{\max} + O(\varepsilon)(d^* + d_1).
\end{aligned}$$

□

We now turn to proving Lemma E.1. Before doing so, we need some deeper understandings of the edge deletion procedure, which we establish in Appendix E.4.1. The proof of Lemma E.1 is presented in Appendix E.4.2.

E.4.1 Probability of Surviving Edge Deletion

Let T be a 1-tree in the 1-forest G_1 before the edge deletion procedure. The edge deletion procedure splits T into several connected components by deleting some edges from T. In this section, we prove upper and lower bounds on the probabilities that paths in T remain connected after edge deletion.

Let $\ell > 0$ denote the cycle length of T. Condition on the height threshold t_{h} being fixed. We prove the following two lemmas:

Lemma E.2 (Upper bound)**.** *Suppose* p *is a directed simple path in* T *of length* s*. If* $\ell \ge t_{\mathsf{h}}$*, then the probability that no edge in* p *is deleted is at most* $\max\{\frac{t_{\mathsf{h}}-s}{t_{\mathsf{h}}}, 0\}(1 + t_{\mathsf{h}}/\ell)$*. If* $\ell \le t_{\mathsf{h}}$*, and we further assume that* p *doesn't contain any cycle edge, then the probability is exactly* $\max\{\frac{t_{\mathsf{h}}-s}{t_{\mathsf{h}}}, 0\}$*.*

Proof. If $s \ge t_{\mathsf{h}}$, the lemma is trivial because any path after edge deletion has length at most $t_{\mathsf{h}} - 1$. We assume $s < t_{\mathsf{h}}$ henceforth.

Suppose vertices on p are $v_0 \leftarrow v_1 \leftarrow \cdots \leftarrow v_s$. We first consider the case where $\ell \ge t_{\mathsf{h}}$. We prove that as long as the (unique) simple path p^* from v_0 to r has length equal to $-1, -2, \cdots, -s$ modulo t_{h}, some edge on path p is deleted. Indeed, suppose p^* has length $-i$ modulo t_{h}. If p^* doesn't contain any vertex in $\{v_1, \cdots, v_s\}$, then the edge out of v_i is deleted by Claim 3.7. Otherwise, r must be one of $v_1, v_2, \cdots, v_s$, in which case the edge out of r is deleted.

Suppose $\ell = ut_{\mathsf{h}} + w$ for $u, w \in \mathbb{Z}$ where $0 \le w < t_{\mathsf{h}}$. There are at most $(t_{\mathsf{h}} - s)(u+1)$ choices of r such that p^* has length not in $\{-1, \cdots, -s\}$ modulo t_{h}. Therefore, when $\ell \ge t_{\mathsf{h}}$, the probability that no edge in p is deleted is at most $(t_{\mathsf{h}} - s)(u+1)/\ell = \frac{t_{\mathsf{h}}-s}{t_{\mathsf{h}}} \cdot (\frac{ut_{\mathsf{h}}}{\ell} + \frac{t_{\mathsf{h}}}{\ell}) \le \frac{t_{\mathsf{h}}-s}{t_{\mathsf{h}}} \cdot (1 + t_{\mathsf{h}}/\ell)$.

When $\ell \le t_h$ and p doesn't contain a cycle edge, an edge on the path p is deleted if and only if p^* has length $-1, -2, \cdots, -s$ modulo t_{h} by Claim 3.7. Since the cycle length is exactly t_{h} after dummy vertices are inserted on it, the probability that no edge on p is deleted is exactly $\frac{t_{\mathsf{h}}-s}{t_{\mathsf{h}}}$. □

Lemma E.3 (Lower bound). *Let v_1, v_2, v^* be vertices in T and p_1, p_2 be directed simple paths in T from v_1 and v_2 to v^*, respectively. Suppose both p_1 and p_2 have lengths no greater than s. If $\ell \ge t_{\mathsf{h}}$, then the probability that no edge on either path p_1, p_2 is deleted is at least* $\max\{\frac{t_{\mathsf{h}}-s}{t_{\mathsf{h}}}, 0\}(1 - 2t_{\mathsf{h}}/\ell)$. *If $\ell \le t_{\mathsf{h}}$, and we further assume that v_1 is on the cycle, then the probability is at least* $\max\{\frac{t_{\mathsf{h}}-s}{t_{\mathsf{h}}}, 0\}$.

Proof. Again, the lemma is trivial if $s \ge t_{\mathsf{h}}$. Assume $s < t_{\mathsf{h}}$ henceforth.

Let us first consider the case where $\ell \ge t_{\mathsf{h}}$. Consider the vertices on the cycle that are different from v^* but have paths to v^* with length at most s. There are at most s such vertices, and they form a contiguous part of the cycle. If r is not among these vertices, then the simple path p^* from v^* to r contains no vertex on p_1 or p_2 except v^* itself. If we further assume that p^* has length not in $-1, -2, \cdots, -s$ modulo t_{h}, then by Claim 3.7 no edge on either path p_1, p_2 is deleted. Therefore, assuming $\ell - s = ut_{\mathsf{h}} + w$ for $u, w \in \mathbb{Z}$ where $0 \le w < t_{\mathsf{h}}$, the probability that no edge on either path is deleted is at least $u(t_{\mathsf{h}} - s)/\ell = \frac{t_{\mathsf{h}}-s}{t_{\mathsf{h}}} \cdot \frac{ut_{\mathsf{h}}}{\ell} = \frac{t_{\mathsf{h}}-s}{t_{\mathsf{h}}} \cdot (1 - \frac{s+w}{\ell}) \ge \frac{t_{\mathsf{h}}-s}{t_{\mathsf{h}}} \cdot (1 - 2t_{\mathsf{h}}/\ell)$.

When $\ell \le t_{\mathsf{h}}$ and v_1 is on the cycle, every edge on p_1 must be on the cycle. Since no edge on the cycle is deleted by our convention, the probability that no edge on either path is deleted is lower bounded by the probability that no edge on the shortest path p' from v_2 to the cycle is deleted. p' is a part of p_2, so p' has length at most s. By the second part of the previous lemma, the probability that no edge on p' is deleted is at least $\max\{\frac{t_{\mathsf{h}}-s}{t_{\mathsf{h}}}, 0\}$. □

E.4.2 Proof of Lemma E.1

We are now ready to prove Lemma E.1. Define $\mathcal{D}'$ as the union of the defiant event $\mathcal{D}$ (Definition 5.1) and the following events:

- (i) $\mathcal{P}$ is a tree swap set, and, before edge deletion, the cycle in the 1-tree containing the original copy of f^* has length ℓ in the range $(t_{\mathsf{h}}, \lceil 1/\varepsilon \rceil \cdot t_{\mathsf{h}})$;
- (ii) $\mathcal{P}$ is a tree swap set, and two connected components each containing a facility in $\{f_1, f_2\}$ are combined in the balancing procedure.

Event (i) happens with probability $O(\varepsilon)$ because our height threshold t_{h} is chosen uniformly at random from $2\lceil 1/\varepsilon \rceil, 2\lceil 1/\varepsilon \rceil^2, \cdots, 2\lceil 1/\varepsilon \rceil^{\lceil 1/\varepsilon \rceil}$. Event (ii) happens with probability $O(\varepsilon)$ as well due to Claims 3.5 and 3.6. By a union bound with Claim 5.4, we have

Claim E.4. *The event $\mathcal{D}'$ happens with probability $O(\varepsilon)$.*

Proof of Lemma E.1. If either f_1 or f_2 is heavy, then $\mathcal{T}_{21}$ never happens. Indeed, $\mathcal{T}_{21}$ assumes the existence of a swap closing both f_1 and f_2, but heavy local facilities are never closed. Hence, we assume neither f_1 nor f_2 is heavy.

By Claim E.4 and the union bound, it suffices to prove $\Pr[\mathcal{T}_{21} \backslash \mathcal{D}'] \le (1 + O(\varepsilon)) \Pr[\mathcal{T}_{11} \cup \mathcal{D}']$. By law of total probability, it suffices to prove

$$\Pr[\mathcal{T}_{21} \backslash \mathcal{D}' | \mathcal{E}_i] \le (1 + O(\varepsilon)) \Pr[\mathcal{T}_{11} \cup \mathcal{D}' | \mathcal{E}_i] \tag{E.54}$$

for a partition $\mathcal{E}_1, \mathcal{E}_2, \cdots, \mathcal{E}_t$ of the entire probability space.

If $\mathcal{E}_i = \mathcal{S}$, then both sides of (E.54) become zero. Let us condition on the tree event $\mathcal{T}$ henceforth. Conditioned on $\mathcal{T}$, the probabilities of $\tau(f^*) = \eta_1$ and $\tau(f^*) = \eta_2$ are both $1/2$ since $\rho(f^*) > 2/3$. Note that the set of heavy local/optimal facilities doesn't depend on the random function τ. Therefore, if we condition on the τ's of all optimal facilities except f^*, the out-edges of the original copies of all facilities in G_1 except f^* are determined, where G_1 is the 1-forest after degree reduction but before edge deletion. Let G_1^* be G_1 with the out-edge of the original copy of f^* removed. If we ignore the identity of the local and optimal surrogates, everything else in G_1^* is determined. Moreover, the conditioning we did is independent of $\tau(f^*)$, so the conditional probabilities of $\tau(f^*) = \eta_1$ and $\tau(f^*) = \eta_2$ are both still $1/2$.

Note that f^* may be a heavy optimal facility, in which case f^* has new copies in G_1. We use f^* to refer to only the original copy. $\tau(f^*)$ may also be a heavy local facility when $\tau(f^*) = \eta_2$ (note that we assumed $\eta_1 = f_1$ is not heavy), in which case f^* points to itself in G_1. If either f_1 or f_2 is chosen as a surrogate, then $\mathcal{T}_{21} \backslash \mathcal{D}'$ cannot happen because $\mathcal{D}'$ happens. We thus assume f_1 and f_2 only appear as their original copies in G_1. Since f^* is the only vertex in G_1^* that doesn't have an out-edge, f^* is the root of a tree, and all other connected components of G_1^* are 1-trees.

We divide our proof into five cases depending on the structure of G_1^*:

1. f^*, f_1, f_2 are all in the different connected components;
2. f^*, f_1 are in the same tree, different from f_2;
3. f^*, f_2 are in the same tree, different from f_1;
4. f_1, f_2 are in the same 1-tree (denoted by T), different from f^*;
5. all three are in the same tree (denoted by T^*).

Let $\mathcal{E}_1$ denote the event that f_1 and f_2 are in the same connected component in G_2, where G_2 is the graph after the edge deletion procedure. Since $\mathcal{D}'$ includes the case where the edge from f^* to $\eta_1 = f_1$ is deleted in the edge deletion step, we have $\mathcal{E}_1 \cap \mathcal{T}_1 \subseteq \mathcal{T}_{11} \cup \mathcal{D}'$. Let $\mathcal{E}_0$ denote the event that f_1 and f_2 are in the same connected component in G_2 but different from f^*. Since subtracting $\mathcal{D}'$ rules out the possibility of f_1 and f_2 being combined in the balancing step, we have $\mathcal{T}_{21} \backslash \mathcal{D}' \subseteq \mathcal{E}_0 \subseteq \mathcal{E}_1$.

In case 1, $\mathcal{T}_{21} \backslash \mathcal{D}'$ never happens because $\mathcal{E}_1$ never happens. Indeed, f_1, f_2 must be in different connected components in G_1 and thus must be in different connected components in G_2.

In cases 2&3, $\mathcal{T}_{21} \backslash \mathcal{D}'$ never happens either because $\mathcal{E}_0$ never happens. Indeed, the only way f_1 can connect to f_2 (by an undirected path in G_1) is through f^*, and in the edge deletion procedure, there is no way to put f_1, f_2 in the same connected component of G_2 without also putting f^* in it.

In case 4, f^* is not on the cycle part of T, so the height threshold t_{h} and the choice of $r \in T$ in the edge deletion step are both independent of $\tau(f^*)$. Once conditioned on t_{h}, r, whether or not f_1 and f_2 are in the same connected component in G_2 is determined. We assume that f_1 and f_2 are in the same connected component of G_2 because otherwise $\mathcal{T}_{21} \backslash \mathcal{D}'$ never happens. If $\tau(f^*) = \eta_1 (= f_1)$, then we know $\mathcal{T}_{11} \cup \mathcal{D}'$ must happen, because $\mathcal{E}_1 \cap \mathcal{T}_1$ happens. Moreover, $\mathcal{T}_{21} \backslash \mathcal{D}'$ happens only when $\tau(f^*) = \eta_2$ simply because $\mathcal{T}_{21} \subseteq \mathcal{T}_2$. Therefore, if we let $\mathcal{E}$ be the event summarizing all the conditioning we did so far, we have

$$\begin{aligned} \Pr[\mathcal{T}_{11} \cup \mathcal{D}' | \mathcal{E}] &= \Pr[\tau(f^*) = \eta_1 | \mathcal{E}] = 1/2, \\ \Pr[\mathcal{T}_{21} \backslash \mathcal{D}' | \mathcal{E}] &\leq \Pr[\tau(f^*) = \eta_2 | \mathcal{E}] = 1/2, \end{aligned}$$

and thus (E.54) holds for $\mathcal{E}_i = \mathcal{E}$.

Case 5 is a little tricky since the cycle structure of T, the 1-tree in G_1 containing all of f^*, f_1, f_2, may depend on where f^* points to. Condition on the height threshold t_{h} being fixed, and let $\mathcal{E}$ be the event summarizing all the conditioning we did so far. Let f_a be the least common ancestor of f_1 and f_2 in T^*, and let s denote the path length from f_i to f_a maximized over $i = 1, 2$.

Conditioned on $\tau(f^*) = \eta_1$, or equivalently $\mathcal{T}_1$, the probability of $\mathcal{T}_{11} \cup \mathcal{D}'$ is 1 if the cycle length ℓ of T is in the range $(t_{\mathsf{h}}, \lceil 1/\varepsilon \rceil \cdot t_{\mathsf{h}})$, and if ℓ is not in the range, the conditional probability of $\mathcal{T}_{11} \cup \mathcal{D}'$ is at least the conditional probability of $\mathcal{E}_1$, which is at least $\max\{\frac{t_{\mathsf{h}} - s}{t_{\mathsf{h}}}, 0\}(1 - O(\varepsilon))$ by Lemma E.3 (Observe that $f_1 = \eta_1$ is on the cycle of T because f^* points to it on event $\mathcal{T}_1$). Therefore,

$$\begin{aligned} \Pr[\mathcal{T}_{11} \cup \mathcal{D}' | \mathcal{E}] &\geq \Pr[\tau(f^*) = \eta_1 | \mathcal{E}] \cdot \max\left\{\frac{t_{\mathsf{h}} - s}{t_{\mathsf{h}}}, 0\right\}(1 - O(\varepsilon)) \\ &= 1/2 \cdot \max\left\{\frac{t_{\mathsf{h}} - s}{t_{\mathsf{h}}}, 0\right\}(1 - O(\varepsilon)). \end{aligned} \tag{E.55}$$

On the other hand, $\mathcal{T}_{21} \backslash \mathcal{D}'$ happens only when $\tau(f^*) = \eta_2$. Condition on $\tau(f^*) = \eta_2$. If the cycle length ℓ is in the range $(t_{\mathsf{h}}, \lceil 1/\varepsilon \rceil \cdot t_{\mathsf{h}})$, then $\mathcal{T}_{21} \backslash \mathcal{D}'$ never happens. If $\ell \leq t_{\mathsf{h}}$, and f_a is on the cycle, then $\mathcal{T}_{21} \backslash \mathcal{D}'$ never happens either because $\mathcal{E}_0$ never happens. Indeed, the only possible undirected path in T connecting f_1 with f_2 without passing through f^* intersects the cycle, so f_1, f_2 have to connect to the cycle after edge deletion to make $\mathcal{E}_0$ happen, but the cycle contains f^* and remains connected after edge deletion (because $\ell \leq t_{\mathsf{h}}$). Therefore, we assume either $\ell \geq \lceil 1/\varepsilon \rceil \cdot t_{\mathsf{h}}$, or $\ell \leq t_{\mathsf{h}}$ and f_a is not on the cycle. In this case, the conditional probability of $\mathcal{T}_{21} \backslash \mathcal{D}'$ is at most the conditional probability of $\mathcal{E}_0$, which is at most $\max\left\{\frac{t_{\mathsf{h}} - s}{t_{\mathsf{h}}}, 0\right\}(1 + O(\varepsilon))$ by Lemma E.2. Therefore,

$$\begin{aligned} \Pr[\mathcal{T}_{21} \backslash \mathcal{D}' | \mathcal{E}] &\leq \Pr[\tau(f^*) = \eta_2 | \mathcal{E}] \cdot \max\left\{\frac{t_{\mathsf{h}} - s}{t_{\mathsf{h}}}, 0\right\}(1 + O(\varepsilon)) \\ &= 1/2 \cdot \max\left\{\frac{t_{\mathsf{h}} - s}{t_{\mathsf{h}}}, 0\right\}(1 + O(\varepsilon)). \end{aligned} \tag{E.56}$$

Combining (E.55) and (E.56), we know (E.54) holds for $\mathcal{E}_i = \mathcal{E}$. □

E.5 Proof of Claim D.3: Subtypes within Type C

Claim D.3 (Subtypes within type C). *For a client c of type C, one of the following is true:*

(a) f_1 is heavy.

(b) f_2 is heavy.

(c) A facility h is open near c after the simple swap closing f_1. Formally, a facility $h \neq f_2$ is open after swap $\langle\!\langle \neg f_1 \rangle\!\rangle$ at distance $d(c, h) \leq 3d_1 + 2d^*$ on $\mathcal{S} \cap \mathcal{A}$.

(d) $g^* \neq f^*$, $\rho(g^*) > 3/4$, and for all $b = 1, 2$, any swap set $\mathcal{P}$ generated on $\mathcal{S}'_b \cap \mathcal{A}$, a facility $h \neq f_2$ is open after swap $\langle\!\langle \neg f_1 \rangle\!\rangle$ at distance $d(c,h) \leq \begin{cases} 2d_1 + d^*, & \text{if } b = 1 \\ 2d_1 + d^* + 4/3(d_1 + d^*), & \text{if } b = 2 \end{cases}$.

(e) For any swap set $\mathcal{P}$ generated on $\mathcal{T}_2 \cap \mathcal{A}$, $\langle\!\langle f^* \rangle\!\rangle$ closes both f_1 and f_2.

(f) $g^* \neq f^*$, $\rho(g^*) > 2/3$, and there exists $b \in \{1, 2\}$ such that for any swap set $\mathcal{P}$ generated on $\mathcal{T}'_b \cap \mathcal{A}$, $\langle\!\langle f^* \rangle\!\rangle$ closes both f_1 and f_2.

Proof. Recall that g^* is $\pi(f_1)$ and $\mathcal{S}'_b$ is the event that $\mathcal{P}$ is a simple swap and g^* points to $\eta_b(g^*)$. Similarly $\mathcal{T}'_b$ is the event that $\mathcal{P}$ is a tree swap and g^* points to $\eta_b(g^*)$.

If either f_1 or f_2 is heavy, then condition (a) or (b) holds. We assume neither f_1 nor f_2 is heavy henceforth. In other words, the swaps $\langle\langle \neg f_1 \rangle\rangle$ and $\langle\langle \neg f_2 \rangle\rangle$ both exist.

Let g be the closest local facility to g^* that is different from f_1 and f_2. Intuitively, we show that either a client is close to g or there is a tree that contains all f_1, f_2, and f^*.

If $d(g, g^*) \le d(f_1, g^*)$, then we have $d(c, g) \le d_1 + d(f_1, g^*) + d(g, g^*) \le d_1 + 2d(f_1, g^*) \le 3d_1 + 2d^*$. Furthermore, when we generate tree swaps, f_1 points to $g^* = \pi(f_1)$ in the 1-forest G_1 after degree reduction. If f_1 points to a new copy of g^*, we know that f_1 is not among the t_{d} closest local facilities to g^* in $\pi^{-1}(g^*)$. Therefore, we know $d(g, g^*) \le d(f_1, g^*)$. Note that g and f_1 are not closed in the same simple swap by implication (Siii') of amenability, so condition (c) holds in this case.

We can now assume that f_1 points to the original copy of g^* and $d(g, g^*) > d(f_1, g^*)$. If $g^* = f^*$, we know condition (e) holds, because both edges $f_1 \to f^*, f^* \to f_2$ remain after the edge deletion step by amenability. We assume $g^* \ne f^*$ henceforth.

If $\rho(g^*) \le 2/3$, we know $\tau(g^*) = \eta_1(g^*)$ deterministically. Moreover, $d(g, g^*) > d(f_1, g^*)$ implies that $\tau(g^*)$ is either f_1 or f_2. If $\eta_1(g^*) = f_1$, then $\langle\langle \neg f_1 \rangle\rangle$ must open g^* by implication (ii') of amenability, so condition (c) holds in this case since $d(c, g^*) \le d_1 + d(f_1, g^*) \le 2d_1 + d^*$. Otherwise, $\eta_1(g^*) = f_2$, and then condition (e) holds, because the edges $f_1 \to g^*, g^* \to f_2, f^* \to f_2$ all survive edge deletion by amenability, so f_1, f_2, f^* must all be in the same swap.

It remains to consider the case where $\rho(g^*) > 2/3$. If $f_2 = \eta_b(g^*) \in \{\eta_1(g^*), \eta_2(g^*)\}$, then condition (f) holds because the edges $f_1 \to g^*, g^* \to f_2, f^* \to \tau(f^*) \in \{f_1, f_2\}$ all survive edge deletion on $\mathcal{T}'_b \cap \mathcal{A}$ (see the left graph in Figure E.13). Otherwise, $f_2 \notin \{\eta_1(g^*), \eta_2(g^*)\}$, and in this case we know $\eta_1(g^*) = f_1$ and $\eta_2(g^*) = g$ because $d(g, g^*) > d(f_1, g^*)$. We show that condition (c) or (d) holds, depending on whether $\rho(g^*) \le 3/4$. Indeed, on $\mathcal{S}'_1 \cap \mathcal{A}$, we know $\langle\langle \neg f_1 \rangle\rangle$ opens g^* at distance $\le 2d_1 + d^*$ by implication (ii') of amenability, and on $\mathcal{S}'_2 \cap \mathcal{A}$, we know either g^* or g is open after swap $\langle\langle \neg f_1 \rangle\rangle$, again by implication (ii') of amenability, and $d(c, g) \le 2d_1 + d^* + 1/\rho(g^*) \cdot (d^* + d_1)$ (see the right graph in Figure E.13). $\square$

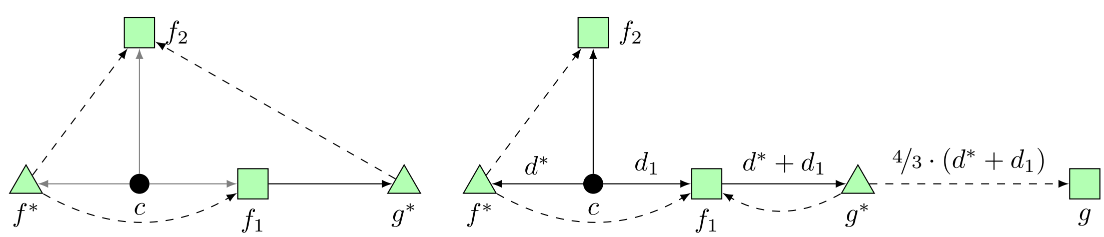


Figure E.13: In the figure, dashed edges represent the random function τ. In the left graph, whenever g^* points to f_2, f_1, f_2, f^* are all in the same swap, so condition (f) holds. In the right graph, condition (d) holds.

Acknowledgments

We thank Amit Kumar, Ola Svensson, and Justin Ward for fruitful discussions. Special thanks to Guru Guruganesh, with whom we obtained some early results on this problem. LH is supported by NSF Award IIS-1908774 and a VMware fellowship. This work was conducted in part while LH was an undergraduate at Tsinghua University visiting CMU and TTI-Chicago.

Supported in part by NSF awards CCF-1907820, CCF1955785, and CCF-2006953.

This work was [partially] funded by the grant ANR-19-CE48-0016 from the French National Research Agency (ANR).

References

[ABS10] Pranjal Awasthi, Avrim Blum, and Or Sheffet. Stability yields a PTAS for k-median and k-means clustering. In *51th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2010, October 23-26, 2010, Las Vegas, Nevada, USA*, pages 309–318, 2010.

[AFS13] Sara Ahmadian, Zachary Friggstad, and Chaitanya Swamy. Local-search based approximation algorithms for mobile facility location problems. In Sanjeev Khanna, editor, *Proceedings of the Twenty-Fourth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2013, New Orleans, Louisiana, USA, January 6-8, 2013*, pages 1607–1621. SIAM, 2013.

[AGK+01] Vijay Arya, Naveen Garg, Rohit Khandekar, Adam Meyerson, Kamesh Munagala, and Vinayaka Pandit. Local search heuristic for k-median and facility location problems. In *Proceedings of the Thirty-Third Annual ACM Symposium on Theory of Computing*, STOC '01, page 21–29, New York, NY, USA, 2001. Association for Computing Machinery.

[Ali94] Paola Alimonti. New local search approximation techniques for maximum generalized satisfiability problems. In *Algorithms and Complexity, Second Italian Conference, CIAC '94, Rome, Italy, February 23-25, 1994, Proceedings*, pages 40–53, 1994.

[ANSW17] Sara Ahmadian, Ashkan Norouzi-Fard, Ola Svensson, and Justin Ward. Better guarantees for k-means and Euclidean k-median by primal-dual algorithms. In *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*, pages 61–72, 2017.

[BPR+15] Jarosław Byrka, Thomas Pensyl, Bartosz Rybicki, Aravind Srinivasan, and Khoa Trinh. An improved approximation for k-median, and positive correlation in budgeted optimization. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '15, page 737–756, USA, 2015. Society for Industrial and Applied Mathematics.

[CG99] Moses Charikar and Sudipto Guha. Improved combinatorial algorithms for the facility location and k-median problems. In *40th Annual Symposium on Foundations of Computer Science, FOCS '99, 17-18 October, 1999, New York, NY, USA*, pages 378–388, 1999.

[CGTS99] Moses Charikar, Sudipto Guha, Éva Tardos, and David B. Shmoys. A constant-factor approximation algorithm for the k-median problem (extended abstract). In *Proceedings of the Thirty-First Annual ACM Symposium on Theory of Computing*, STOC '99, page 1–10, New York, NY, USA, 1999. Association for Computing Machinery.

[CL12] Moses Charikar and Shi Li. A dependent LP-rounding approach for the k-median problem. In *Automata, Languages, and Programming - 39th International Colloquium, ICALP 2012, Warwick, UK, July 9-13, 2012, Proceedings, Part I*, pages 194–205, 2012.

[FFSW17] Moran Feldman, Yuval Filmus, Roy Schwartz, and Justin Ward. A tight combinatorial algorithm for submodular maximization subject to a matroid constraint. `https://www.cs.toronto.edu/~yuvalf/NOLS.pdf`, January 2017.

[FW12] Yuval Filmus and Justin Ward. The power of local search: Maximum coverage over a matroid. In *29th Symposium on Theoretical Aspects of Computer Science (STACS 2012)*, pages 601–612, 2012.

[FW14] Yuval Filmus and Justin Ward. Monotone submodular maximization over a matroid via non-oblivious local search. *SIAM J. Comput.*, 43(2):514–542, 2014.

[GGK+18] Martin Groß, Anupam Gupta, Amit Kumar, Jannik Matuschke, Daniel R. Schmidt, Melanie Schmidt, and José Verschae. A local-search algorithm for steiner forest. In *9th Innovations in Theoretical Computer Science Conference, ITCS 2018, January 11-14, 2018, Cambridge, MA, USA*, pages 31:1–31:17, 2018.

[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.

[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 of the Thiry-Fourth Annual ACM Symposium on Theory of Computing*, STOC '02, page 731–740, New York, NY, USA, 2002. Association for Computing Machinery.

[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.

[KMN+02] Tapas Kanungo, David M. Mount, Nathan S. Netanyahu, Christine D. Piatko, Ruth Silverman, and Angela Y. Wu. A local search approximation algorithm for k-means clustering. In *Proceedings of the Eighteenth Annual Symposium on Computational Geometry*, SCG '02, page 10–18, New York, NY, USA, 2002. Association for Computing Machinery.

[KMSV98] Sanjeev Khanna, Rajeev Motwani, Madhu Sudan, and Umesh V. Vazirani. On syntactic versus computational views of approximability. *SIAM J. Comput.*, 28(1):164–191, 1998.

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