

# Strategyproof Facility Location in Perturbation Stable Instances<sup>\*</sup>

Dimitris Fotakis and Panagiotis Patsilinakos

School of Electrical and Computer Engineering  
National Technical University of Athens, 15780 Athens, Greece  
fotakis@cs.ntua.gr, patsilinak@corelab.ntua.gr

**Abstract.** We consider  $k$ -Facility Location games, where  $n$  strategic agents report their locations on the real line, and a mechanism maps them to  $k \geq 2$  facilities. Each agent seeks to minimize her distance to the nearest facility. We are interested in (deterministic or randomized) strategyproof mechanisms without payments that achieve a reasonable approximation ratio to the optimal social cost of the agents. To circumvent the inapproximability of  $k$ -Facility Location by deterministic strategyproof mechanisms, we restrict our attention to perturbation stable instances. An instance of  $k$ -Facility Location on the line is  $\gamma$ -perturbation stable (or simply,  $\gamma$ -stable), for some  $\gamma \geq 1$ , if the optimal agent clustering is not affected by moving any subset of consecutive agent locations closer to each other by a factor at most  $\gamma$ . We show that the optimal solution is strategyproof in  $(2 + \sqrt{3})$ -stable instances whose optimal solution does not include any singleton clusters, and that allocating the facility to the agent next to the rightmost one in each optimal cluster (or to the unique agent, for singleton clusters) is strategyproof and  $(n - 2)/2$ -approximate for 5-stable instances (even if their optimal solution includes singleton clusters). On the negative side, we show that for any  $k \geq 3$  and any  $\delta > 0$ , there is no deterministic anonymous mechanism that achieves a bounded approximation ratio and is strategyproof in  $(\sqrt{2} - \delta)$ -stable instances. We also prove that allocating the facility to a random agent of each optimal cluster is strategyproof and 2-approximate in 5-stable instances. To the best of our knowledge, this is the first time that the existence of deterministic (resp. randomized) strategyproof mechanisms with a bounded (resp. constant) approximation ratio is shown for a large and natural class of  $k$ -Facility Location instances.

## 1 Introduction

We consider  $k$ -Facility Location games, where  $k \geq 2$  facilities are placed on the real line based on the preferences of  $n$  strategic agents. Such problems are motivated by natural scenarios in Social Choice, where a local authority plans to build a fixed number of public facilities in an area (see e.g., [40]). The choice of the locations is based on the preferences of local people, or *agents*. Each agent reports her ideal location, and the local authority applies a (deterministic or randomized) *mechanism* that maps the agents' preferences to  $k$  facility locations.

Each agent evaluates the mechanism's outcome according to her *connection cost*, i.e., the distance of her ideal location to the nearest facility. The agents seek to minimize their connection cost and may misreport their ideal locations in an attempt of manipulating the mechanism. Therefore, the mechanism should be *strategyproof*, i.e., it should ensure that no agent can benefit from misreporting her location, or even *group strategyproof*, i.e., resistant to coalitional manipulations. The local authority's objective is to minimize the *social cost*, namely the sum of agent connections costs. In addition to allocating the facilities in an incentive compatible way, which is formalized by (group) strategyproofness, the mechanism should result in a socially desirable outcome, which is quantified by the mechanism's approximation ratio to the optimal social cost.

Since Procaccia and Tennenholtz [42] initiated the research agenda of *approximate mechanism design without money*,  $k$ -Facility Location has served as the benchmark problem in the area and

<sup>\*</sup> This work was supported by the Hellenic Foundation for Research and Innovation (H.F.R.I.) under the "First Call for H.F.R.I. Research Projects to support Faculty members and Researchers' and the procurement of high-cost research equipment grant", project BALSAM, HFRI-FM17-1424

its approximability by deterministic or randomized strategyproof mechanisms has been studied extensively in virtually all possible variants and generalizations. For instance, previous work has considered multiple facilities on the line (see e.g., [27,28,32,37,41]) and in general metric spaces [26,36]), different objectives (e.g., social cost, maximum cost, the  $L_2$  norm of agent connection costs [23,42,28]), restricted metric spaces more general than the line (cycle, plane, trees, see e.g., [2,17,25,31,39]), facilities that serve different purposes (see e.g., [34,35,48]), and different notions of private information about the agent preferences that should be declared to the mechanism (see e.g., [16,21,38] and the references therein).

Due to the significant research interest in the topic, the fundamental and most basic question of approximating the optimal social cost by strategyproof mechanisms for  $k$ -Facility Location on the line has been relatively well-understood. For a single facility ( $k = 1$ ), placing the facility at the median location is group strategyproof and optimizes the social cost. For two facilities ( $k = 2$ ), the best possible approximation ratio is  $n - 2$  and is achieved by a natural group strategyproof mechanism that places the facilities at the leftmost and the rightmost location [27,42]. However, for three or more facilities ( $k \geq 3$ ), there do not exist any deterministic anonymous<sup>1</sup> strategyproof mechanisms for  $k$ -Facility Location with a bounded (in terms of  $n$  and  $k$ ) approximation ratio [27]. On the positive side, there is a randomized anonymous group strategyproof mechanism<sup>2</sup> with an approximation ratio of  $n$  [28] (see also Section 1.1 for a selective list of additional references).

**Perturbation Stability in  $k$ -Facility Location Games.** Our work aims to circumvent the strong impossibility result of [27] and is motivated by the recent success on the design of polynomial-time exact algorithms for perturbation stable clustering instances (see e.g., [3,9,10,11,43,44]). An instance of a clustering problem, like  $k$ -Facility Location (a.k.a.  $k$ -median in the optimization and approximation algorithms literature), is  $\gamma$ -*perturbation stable* (or simply,  $\gamma$ -*stable*), for some  $\gamma \geq 1$ , if the optimal clustering is not affected by scaling down any subset of the entries of the distance matrix by a factor at most  $\gamma$ . Perturbation stability was introduced by Bilu and Linial [12] and Awasthi, Blum and Sheffet [7] (and has motivated a significant volume of followup work since then, see e.g., [3,9,11,44] and the references therein) in an attempt to obtain a theoretical understanding of the superior practical performance of relatively simple clustering algorithms for well known NP-hard clustering problems (such as  $k$ -Facility Location in general metric spaces). Intuitively, the optimal clusters of a  $\gamma$ -stable instance are somehow well separated, and thus, relatively easy to identify (see also the main properties of stable instances in Section 3). As a result, natural extensions of simple algorithms, like single-linkage (a.k.a. Kruskal’s algorithm), can recover the optimal clustering in polynomial time, provided that  $\gamma \geq 2$  [3], and standard approaches, like dynamic programming (resp. local search), work in almost linear time for  $\gamma > 2 + \sqrt{3}$  (resp.  $\gamma > 5$ ) [1].

In this work, we investigate whether restricting our attention to stable instances allows for improved strategyproof mechanisms with bounded (and ideally, constant) approximation guarantees for  $k$ -Facility Location on the line, with  $k \geq 2$ . We note that the impossibility results of [27] crucially depend on the fact that the clustering (and the subsequent facility placement) produced by any deterministic mechanism with a bounded approximation ratio must be sensitive to location misreports by certain agents (see also Section 6). Hence, it is very natural to investigate whether the restriction to  $\gamma$ -stable instances allows for some nontrivial approximation guarantees by deterministic or randomized strategyproof mechanisms for  $k$ -Facility Location on the line.

<sup>1</sup> A mechanism is *anonymous* if its outcome depends only on the agent locations, not on their identities.

<sup>2</sup> The result of [28] applies to the more general setting where the agent connection cost is a nondecreasing concave function of the distance to the nearest facility.

To study the question above, we adapt to the real line the stricter<sup>3</sup> notion of  $\gamma$ -metric stability [3], where the definition also requires that the distances form a metric after the  $\gamma$ -perturbation. In our notion of *linear  $\gamma$ -stability*, the instances should retain their linear structure after a  $\gamma$ -perturbation. Hence, a  $\gamma$ -perturbation of a linear  $k$ -Facility Location instance is obtained by moving any subset of pairs of consecutive agent locations closer to each other by a factor at most  $\gamma \geq 1$ . We say that a  $k$ -Facility Location instance is  $\gamma$ -stable, if the original instance and any  $\gamma$ -perturbation of it admit the same unique optimal clustering<sup>4</sup> (see also Definition 1).

Interestingly, for  $\gamma$  sufficiently large,  $\gamma$ -stable instances of  $k$ -Facility Location have additional structure that one could exploit towards the design of strategyproof mechanisms with good approximation guarantees (see also Section 3). E.g., each agent location is  $\gamma - 1$  times closer to the nearest facility than to any location in a different cluster (Proposition 1). Moreover, for  $\gamma \geq 2 + \sqrt{3}$ , the distance between any two consecutive clusters is larger than their diameter (Lemma 1).

From a conceptual viewpoint, our work is motivated by a reasoning very similar to that discussed by Bilu, Daniely, Linial and Saks [13] and summarized in “*clustering is hard only when it doesn’t matter*” by Roughgarden [46]. In a nutshell, we expect that when  $k$  public facilities (such as schools, libraries, hospitals, representatives) are to be allocated to some communities (e.g., cities, villages or neighborhoods, as represented by the locations of agents on the real line) the communities are already well formed, relatively easy to identify and difficult to radically reshape by small distance perturbations or agent location misreports. Moreover, in natural practical applications of  $k$ -Facility Location games, agents tend to misreport “locally” (i.e., they tend to declare a different ideal location in their neighborhood, trying to manipulate the location of the local facility), which usually does not affect the cluster formation. In practice, this happens because the agents do not have enough knowledge about locations in other neighborhoods, and because “large non-local” misreports are usually easy to identify by combining publicly available information about the agents (e.g., occupation, address, habits, lifestyle). Hence, we believe that the class of  $\gamma$ -stable instances, especially for relatively small values of  $\gamma$ , provides a reasonably accurate abstraction of the instances of  $k$ -Facility Location games that a mechanism designer is more likely to deal with in practice. Thus, we feel that our work takes a small first step towards justifying that (not only clustering but also) strategyproof facility location is hard only when it doesn’t matter.

**Contributions and Techniques.** Our conceptual contribution is that we initiate the study of efficient (wrt. their approximation ratio for the social cost) strategyproof mechanisms for the large and natural class of  $\gamma$ -stable instances of  $k$ -Facility Location on the line. Our technical contribution is that we show the existence of deterministic (resp. randomized) strategyproof mechanisms with a bounded (resp. constant) approximation ratio for  $\gamma$ -stable instances and any number of facilities  $k \geq 2$ . Moreover, we show that the optimal solution is strategyproof for  $(2 + \sqrt{3})$ -stable instances, if the optimal clustering does not include any singleton clusters (which is likely to be the case in virtually all practical applications). To provide evidence that restriction to stable instances does not make the problem trivial, we strengthen the impossibility result of Fotakis and Tzamos [27], so that it applies to  $\gamma$ -stable instances, with  $\gamma < \sqrt{2}$ . Specifically, we show that that for any  $k \geq 3$  and

<sup>3</sup> The notion of  $\gamma$ -metric stability is “stricter” than standard  $\gamma$ -stability in the sense that the former excludes some perturbations allowed by the latter. Hence, the class of  $\gamma$ -metric stable instances includes the class of  $\gamma$ -stable instances. More generally, the stricter a notion of stability is, the larger the class of instances qualified as stable, and the more general the positive results that one gets. Similarly, for any  $\gamma' > \gamma \geq 1$ , the class of  $\gamma$ -metric stable instances includes the class of  $\gamma'$ -metric instances. Hence, a smaller value of  $\gamma$  makes a positive result stronger and more general.

<sup>4</sup> As for the optimal centers, in case of ties, the center of an optimal cluster is determined by a fixed deterministic tie-breaking rule, e.g., the center is always the left median point of the cluster.

any  $\delta > 0$ , there do not exist any deterministic anonymous strategyproof mechanisms for  $k$ -Facility Location on  $(\sqrt{2} - \delta)$ -stable instances with bounded (in terms of  $n$  and  $k$ ) approximation ratio.

At the conceptual level, we interpret the stability assumption as a prior on the class of true instances that the mechanism should be able to deal with. Namely, we assume that the mechanism has only to deal with  $\gamma$ -stable true instances, a restriction motivated by (and fully consistent with) how the stability assumption is used in the literature on efficient algorithms for stable clustering (see e.g., [3,9,11,12], where the algorithms are analyzed for stable instances only). More specifically, our mechanisms expect as input a declared instance such that in the optimal clustering, the distance between any two consecutive clusters is at least  $\frac{(\gamma-1)^2}{2\gamma}$  times larger than the diameters of the two clusters (a.k.a. *cluster-separation* property, see Lemma 1). This condition is necessary (but not sufficient) for  $\gamma$ -stability and can be easily checked. If the declared instance does not satisfy the cluster-separation property, our mechanisms do not allocate any facilities. Otherwise, our mechanisms allocate  $k$  facilities (even if the instance is not stable). We prove that for all  $\gamma$ -stable true instances (with the exact stability factor  $\gamma$  depending on the mechanism), if agents can only deviate so that the declared instance satisfies the cluster-separation property (and does not have singleton clusters, for the optimal mechanism), our mechanisms are strategyproof and achieve the desired approximation guarantee. Hence, if we restrict ourselves to  $\gamma$ -stable true instances and to agent deviations that do not obviously violate  $\gamma$ -stability, our mechanisms should only deal with  $\gamma$ -stable declared instances, due to strategyproofness. On the other hand, if non-stable true instances may occur, the mechanisms cannot distinguish between a stable true instance and a declared instance, which appears to be stable, but is obtained from a non-stable instance through location misreports.

The restriction that the agents of a  $\gamma$ -stable instance are only allowed to deviate so that the declared instance satisfies the cluster-separation property (and does not have any singleton clusters, for the optimal mechanism) bears a strong conceptual resemblance to the notion of strategyproof mechanisms with *local verification* (see e.g., [6,4,14,15,29,30,33]), where the set of each agent's allowable deviations is restricted to a so-called *correspondence set*, which typically depends on the agent's true type, but not on the types of the other agents. Instead of restricting the correspondence set of each individual agent independently, we impose a structural condition on the entire declared instance, which restricts the set of the agents' allowable deviations, but in a global and observable sense. As a result, we can actually implement our notion of verification, by checking some simple properties of the declared instance, instead of just assuming that any deviation outside an agent's correspondence set will be caught and penalized (which is the standard approach in mechanisms with local verification [4,15,14], but see e.g., [6,26] for noticeable exceptions).

On the technical side, we start, in Section 3, with some useful properties of stables instances of  $k$ -Facility Location on the line. Among others, we show (i) the *cluster-separation* property (Lemma 1), which states that in any  $\gamma$ -stable instance, the distance between any two consecutive clusters is at least  $\frac{(\gamma-1)^2}{2\gamma}$  times larger than their diameters; and (ii) the so-called *no direct improvement from singleton deviations* property (Lemma 2), i.e., that in any 3-stable instance, no agent who deviates to a location, which becomes a singleton cluster in the optimal clustering of the resulting instance, can improve her connection cost through the facility of that singleton cluster.

In Section 4, we show that for  $(2 + \sqrt{3})$ -stable instances whose optimal clustering does not include any singleton clusters, the optimal solution is strategyproof (Theorem 1). For the analysis, we observe that since placing the facility at the median location of any fixed cluster is strategyproof, a misreport cannot be profitable for an agent, unless it results in a different optimal clustering. The key step is to show that for  $(2 + \sqrt{3})$ -stable instances without singleton clusters, a profitable misreport cannot change the optimal clustering, unless the instance obtained from the misreport violates the cluster-separation property. To the best of our knowledge, the idea of penalizing (and thus, es-

essentially forbidding) a whole class of potentially profitable misreports by identifying how they affect a key structural property of the original instance, which becomes possible due to our restriction to stable instances, has not been used before in the design of strategyproof mechanisms for  $k$ -Facility Location (see also the discussion above about resemblance to mechanisms with verification).

We should also motivate our restriction to stable instances without singleton clusters in their optimal clustering. So, let us consider the rightmost agent  $x_j$  of an optimal cluster  $C_i$  in a  $\gamma$ -stable instance  $\vec{x}$ . No matter the stability factor  $\gamma$ , it is possible that  $x_j$  performs a so-called *singleton deviation*. Namely,  $x_j$  deviates to a remote location  $x'$  (potentially very far away from any location in  $\vec{x}$ ), which becomes a singleton cluster in the optimal clustering of the resulting instance. Such a singleton deviation might cause cluster  $C_i$  to merge with (possibly part of the next) cluster  $C_{i+1}$ , which in turn, might bring the median of the new cluster much closer to  $x_j$  (see also Fig. 1 in Section 3). It is not hard to see that if we stick to the optimal solution, where the facilities are located at the median of each optimal cluster, there are  $\gamma$ -stable instances<sup>5</sup>, with arbitrarily large  $\gamma \geq 1$ , where some agents can deviate to a remote location and gain, by becoming singleton clusters, while maintaining the desirable stability factor of the declared instance (see also Fig. 1).

To deal with singleton deviations<sup>6</sup>, we should place the facility either at a location close to an extreme one, as we do in Section 5 with the ALMOSTRIGHTMOST mechanism, or at a random location, as we do in Section 7 with the RANDOM mechanism. More specifically, in Section 5, we show that the ALMOSTRIGHTMOST mechanism, which places the facility of any non-singleton optimal cluster at the location of the second rightmost agent, is strategyproof for 5-stable instances of  $k$ -Facility Location (even if their optimal clustering includes singleton clusters) and achieves an approximation ratio at most  $(n - 2)/2$  (Theorem 2). Moreover, in Section 7, we show that the RANDOM mechanism, which places the facility of any optimal cluster at a location chosen uniformly at random, is strategyproof for 5-stable instances (again even if their optimal clustering includes singleton clusters) and achieves an approximation ratio of 2 (Theorem 4).

To obtain a deeper understanding of the challenges behind the design of strategyproof mechanisms for stable instances of  $k$ -Facility Location on the line, we strengthen the impossibility result of [27, Theorem 3.7] so that it applies to  $\gamma$ -stable instances with  $\gamma < \sqrt{2}$  (Section 6). Through a careful analysis of the image sets of deterministic strategyproof mechanisms, we show that for any  $k \geq 3$ , any  $\delta > 0$ , and any  $\rho \geq 1$ , there do not exist any  $\rho$ -approximate deterministic anonymous strategyproof mechanisms for  $(\sqrt{2} - \delta)$ -stable instances of  $k$ -Facility Location on the line (Theorem 3). The proof of Theorem 3 requires additional ideas and extreme care (and some novelty) in the agent deviations, so as to only consider stable instances, compared against the proof of [27, Theorem 3.7]. Interestingly, singleton deviations play a crucial role in the proof of Theorem 3.

## 1.1 Other Related Work

Approximate mechanism design without money for variants and generalizations of Facility Location games on the line has been a very active and productive area of research in the last decade.

Previous work has shown that deterministic strategyproof mechanisms can only achieve a bounded approximation ratio for  $k$ -Facility Location on the line, only if we have at most 2 facilities [27,42]. Notably, stable (called *well-separated* in [27]) instances with  $n = k + 1$  agents play a

<sup>5</sup> E.g., let  $k = 2$  and consider the  $\Theta(\gamma)$ -stable instance  $(0, 1 - \varepsilon, 1, 6\gamma, 6\gamma + \varepsilon, 6\gamma + 1, 6\gamma + 1 + \varepsilon, 6\gamma + 2)$ , for any  $\gamma \geq 1$ . Then, the agent at location  $6\gamma$  can decrease its connection cost (from 1) to  $\varepsilon$  by deviating to location  $(6\gamma)^2$ .

<sup>6</sup> Another natural way to deal efficiently with singleton deviations is through some means of *location verification*, such as winner-imposing verification [26] or  $\varepsilon$ -symmetric verification [30,29]. Adding e.g., winner-imposing verification to the optimal mechanism, discussed in Section 4, results in a strategyproof mechanism for  $(2 + \sqrt{3})$ -stable instances whose optimal clustering may include singleton clusters.

key role in the proof of inapproximability of  $k$ -Facility Location by deterministic anonymous strategyproof mechanisms [27, Theorem 3.7]. On the other hand, randomized mechanisms are known to achieve a better approximation ratio for  $k = 2$  facilities [37], a constant approximation ratio if we have  $k \geq 2$  facilities and only  $n = k + 1$  agents [19,28], and an approximation ratio of  $n$  for any  $k \geq 3$  [28]. Fotakis and Tzamos [26] considered winner-imposing randomized mechanisms that achieve an approximation ratio of  $4k$  for  $k$ -Facility Location in general metric spaces. In fact, the approximation ratio can be improved to  $\Theta(\ln k)$ , using the analysis of [5].

For the objective of maximum agent cost, Alon et al. [2] almost completely characterized the approximation ratios achievable by randomized and deterministic strategyproof mechanisms for 1-Facility Location in general metrics and rings. Fotakis and Tzamos [28] presented a 2-approximate randomized group strategyproof mechanism for  $k$ -Facility Location on the line and the maximum cost objective. For 1-Facility Location on the line and the objective of minimizing the sum of squares of the agent connection costs, Feldman and Wilf [23] proved that the best approximation ratio is 1.5 for randomized and 2 for deterministic mechanisms. Golomb and Tzamos [32] presented tight (resp. almost tight) additive approximation guarantees for locating a single (resp. multiple) facilities on the line and the objectives of the maximum cost and the social cost.

Regarding the application of perturbation stability, we follow the approach of *beyond worst-case analysis* (see e.g., [43,44]), where researchers seek a theoretical understanding of the superior practical performance of certain algorithms by formally analyzing them on practically relevant instances. The beyond worst-case approach is not anything new for Algorithmic Mechanism Design. *Bayesian* analysis, where the bidder valuations are drawn as independent samples from a distribution known to the mechanism, is standard in revenue maximization when we allocate private goods (see e.g., [45]) and has led to many strong and elegant results for social welfare maximization in combinatorial auctions by truthful posted price mechanisms (see e.g., [18,22]). However, in this work, instead of assuming (similar to Bayesian analysis) that the mechanism designer has a relatively accurate knowledge of the distribution of agent locations on the line (and use e.g., an appropriately optimized percentile mechanism [49]), we employ a deterministic restriction on the class of instances (namely, perturbation stability), and investigate if deterministic (resp. randomized) strategyproof mechanisms with a bounded (resp. constant) approximation ratio are possible for locating any number  $k \geq 2$  facilities on such instances. To the best of our knowledge, the only previous work where the notion of perturbation stability is applied to Algorithmic Mechanism Design (to combinatorial auctions, in particular) is [24] (but see also [8,20] where the similar in spirit assumption of endowed valuations was applied to combinatorial markets).

## 2 Notation, Definitions and Preliminaries

We let  $[n] = \{1, \dots, n\}$ . For any  $x, y \in \mathbb{R}$ , we let  $d(x, y) = |x - y|$  be the distance of locations  $x$  and  $y$  on the real line. For a tuple  $\vec{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ , we let  $\vec{x}_{-i}$  denote the tuple  $\vec{x}$  without coordinate  $x_i$ . For a non-empty set  $S$  of indices, we let  $\vec{x}_S = (x_i)_{i \in S}$  and  $\vec{x}_{-S} = (x_i)_{i \notin S}$ . We write  $(\vec{x}_{-i}, a)$  to denote the tuple  $\vec{x}$  with  $a$  in place of  $x_i$ ,  $(\vec{x}_{-\{i,j\}}, a, b)$  to denote the tuple  $\vec{x}$  with  $a$  in place of  $x_i$  and  $b$  in place of  $x_j$ , and so on. For a random variable  $X$ ,  $\mathbb{E}(X)$  denotes the expectation of  $X$ . For an event  $E$  in a sample space,  $\mathbb{P}(E)$  denotes the probability that  $E$  occurs.

**Instances.** We consider  $k$ -Facility Location with  $k \geq 2$  facilities and  $n \geq k + 1$  agents on the real line. We let  $N = \{1, \dots, n\}$  be the set of agents. Each agent  $i \in N$  resides at a location  $x_i \in \mathbb{R}$ , which is  $i$ 's private information. We usually refer to a locations profile  $\vec{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ ,  $x_1 \leq \dots \leq x_n$ , as an *instance*. By slightly abusing the notation, we use  $x_i$  to refer both to the agent  $i$ 's location and sometimes to the agent  $i$  (i.e., the strategic entity) herself.

**Mechanisms.** A *deterministic mechanism*  $M$  for  $k$ -Facility Location maps an instance  $\vec{x}$  to a  $k$ -tuple  $(c_1, \dots, c_k) \in \mathbb{R}^k$ ,  $c_1 \leq \dots \leq c_k$ , of facility locations. We let  $M(\vec{x})$  denote the outcome of  $M$  in instance  $\vec{x}$ , and let  $M_j(\vec{x})$  denote  $c_j$ , i.e., the  $j$ -th smallest coordinate in  $M(\vec{x})$ . We write  $c \in M(\vec{x})$  to denote that  $M(\vec{x})$  places a facility at location  $c$ . A *randomized mechanism*  $M$  maps an instance  $\vec{x}$  to a probability distribution over  $k$ -tuples  $(c_1, \dots, c_k) \in \mathbb{R}^k$ .

**Connection Cost and Social Cost.** Given a  $k$ -tuple  $\vec{c} = (c_1, \dots, c_k)$ ,  $c_1 \leq \dots \leq c_k$ , of facility locations, the connection cost of agent  $i$  wrt.  $\vec{c}$ , denoted  $d(x_i, \vec{c})$ , is  $d(x_i, \vec{c}) = \min_{1 \leq j \leq k} |x_i - y_j|$ . Given a deterministic mechanism  $M$  and an instance  $\vec{x}$ ,  $d(x_i, M(\vec{x}))$  denotes the connection cost of agent  $i$  wrt. the outcome of  $M(\vec{x})$ . If  $M$  is a randomized mechanism, the expected connection cost of agent  $i$  is  $\mathbb{E}_{\vec{c} \sim M(\vec{x})}(d(x_i, \vec{c}))$ . The *social cost* of a deterministic mechanism  $M$  for an instance  $\vec{x}$  is  $\text{cost}(\vec{x}, M(\vec{x})) = \sum_{i=1}^n d(x_i, M(\vec{x}))$ . The social cost of a facility locations profile  $\vec{c} \in \mathbb{R}^k$  is  $\text{cost}(\vec{x}, \vec{c}) = \sum_{i=1}^n d(x_i, \vec{c})$ . The *expected social cost* of a randomized mechanism  $M$  on instance  $\vec{x}$  is

$$\text{cost}(\vec{x}, M(\vec{x})) = \sum_{i=1}^n \mathbb{E}_{\vec{c} \sim M(\vec{x})}(d(x_i, \vec{c})).$$

The *optimal social cost* for an instance  $\vec{x}$  is  $\text{cost}^*(\vec{x}) = \min_{\vec{c} \in \mathbb{R}^k} \sum_{i=1}^n d(x_i, \vec{c})$ . For  $k$ -Facility Location, the optimal social cost (and the corresponding optimal facility locations profile) can be computed in  $O(kn \log n)$  time by standard dynamic programming.

**Approximation Ratio.** A mechanism  $M$  has an approximation ratio of  $\rho \geq 1$ , if for any instance  $\vec{x}$ ,  $\text{cost}(\vec{x}, M(\vec{x})) \leq \rho \text{cost}^*(\vec{x})$ . We say that the approximation ratio  $\rho$  of  $M$  is *bounded*, if  $\rho$  is bounded from above either by a constant or by a (computable) function of  $n$  and  $k$ .

**Strategyproofness.** A deterministic mechanism  $M$  is *strategyproof*, if no agent can benefit from misreporting her location. Formally,  $M$  is strategyproof, if for all location profiles  $\vec{x}$ , any agent  $i$ , and all locations  $y$ ,  $d(x_i, M(\vec{x})) \leq d(x_i, M((\vec{x}_{-i}, y)))$ . Similarly, a randomized mechanism  $M$  is strategyproof (in expectation), if for all location profiles  $\vec{x}$ , any agent  $i$ , and all locations  $y$ ,  $\mathbb{E}_{\vec{c} \sim M(\vec{x})}(d(x_i, \vec{c})) \leq \mathbb{E}_{\vec{c} \sim M((\vec{x}_{-i}, y))}(d(x_i, \vec{c}))$ .

**Clusterings.** A *clustering* (or  $k$ -clustering, if  $k$  is not clear from the context) of an instance  $\vec{x}$  is any partitioning  $\vec{C} = (C_1, \dots, C_k)$  of  $\vec{x}$  into  $k$  sets of consecutive agent locations. We index clusters from left to right. I.e.,  $C_1 = \{x_1, \dots, x_{|C_1|}\}$ ,  $C_2 = \{x_{|C_1|+1}, \dots, x_{|C_1|+|C_2|}\}$ , and so on. We refer to a cluster  $C_i$  that includes only one agent (i.e., with  $|C_i| = 1$ ) as a *singleton* cluster. We sometimes use  $(\vec{x}, \vec{C})$  to highlight that we consider  $\vec{C}$  as a clustering of instance  $\vec{x}$ .

Two clusters  $C$  and  $C'$  are identical, denoted  $C = C'$ , if they include the exact same locations. Two clusterings  $\vec{C} = (C_1, \dots, C_k)$  and  $\vec{Y} = (Y_1, \dots, Y_k)$  of an instance  $\vec{x}$  are the same, if  $C_i = Y_i$ , for all  $i \in [k]$ . Abusing the notation, we say that a clustering  $\vec{C}$  of an instance  $\vec{x}$  is identical to a clustering  $\vec{Y}$  of a  $\gamma$ -perturbation  $\vec{x}'$  of  $\vec{x}$  (see also Definition 1), if  $|C_i| = |Y_i|$ , for all  $i \in [k]$ .

We let  $x_{i,l}$  and  $x_{i,r}$  denote the leftmost and the rightmost agent of each cluster  $C_i$ . Under this notation,  $x_{i-1,r} < x_{i,l} \leq x_{i,r} < x_{i+1,l}$ , for all  $i \in \{2, \dots, k-1\}$ . Exploiting the linearity of instances, we extend this notation to refer to other agents by their relative location in each cluster. Namely,  $x_{i,l+1}$  (resp.  $x_{i,r-1}$ ) is the second agent from the left (resp. right) of cluster  $C_i$ . The *diameter* of a cluster  $C_i$  is  $D(C_i) = d(x_{i,l}, x_{i,r})$ . The distance of clusters  $C_i$  and  $C_j$  is  $d(C_i, C_j) = \min_{x \in C_i, y \in C_j} \{d(x, y)\}$ , i.e., the minimum distance between a location  $x \in C_i$  and a location  $y \in C_j$ .

A  $k$ -facility locations (or  $k$ -centers) profile  $\vec{c} = (c_1, \dots, c_k)$  induces a clustering  $\vec{C} = (C_1, \dots, C_k)$  of an instance  $\vec{x}$  by assigning each agent / location  $x_j$  to the cluster  $C_i$  with facility  $c_i$  closest to  $x_j$ . Formally, for each  $i \in [k]$ ,  $C_i = \{x_j \in \vec{x} : d(x_j, c_i) = d(x_j, \vec{c})\}$ . The *optimal clustering* of an instance  $\vec{x}$  is the clustering of  $\vec{x}$  induced by the facility locations profile with minimum social cost.

The social cost of a clustering  $\vec{C}$  induced by a  $k$ -facility locations profile  $\vec{c}$  on an instance  $\vec{x}$  is simply  $\text{cost}(\vec{x}, \vec{c})$ , i.e., the social cost of  $\vec{c}$  for  $\vec{x}$ . We sometimes refer to the social cost  $\text{cost}(\vec{x}, \vec{C})$  of a clustering  $\vec{C}$  for an instance  $\vec{x}$ , without any explicit reference to the corresponding facility locations profile. Then, we refer to the social cost  $\text{cost}(\vec{x}, \vec{c})$ , where each facility  $c_i$  is located at the median location of  $C_i$  (the left median location of  $C_i$ , if  $|C_i|$  is even).

We often consider certain structural changes in a clustering due to agent deviations. Let  $\vec{C}$  be a clustering of an instance  $\vec{x}$ , which due to an agent deviation, changes to a different clustering  $\vec{C}'$ . We say that cluster  $C_i$  is *split* when  $\vec{C}$  changes to  $\vec{C}'$ , if not all agents in  $C_i$  are served by the same facility in  $\vec{C}'$ . We say that  $C_i$  is *merged* in  $\vec{C}'$ , if all agents in  $C_i$  are served by the same facility, but this facility also serves in  $\vec{C}'$  some agents not in  $C_i$ .

### 3 Perturbation Stability on the Line: Definition and Properties

Next, we introduce the notion of  $\gamma$ -(linear) stability and prove some useful properties of  $\gamma$ -stable instances of  $k$ -Facility Location, which are repeatedly used in the analysis of our mechanisms.

**Definition 1 ( $\gamma$ -Perturbation and  $\gamma$ -Stability).** Let  $\vec{x} = (x_1, \dots, x_n)$  be a locations profile. A locations profile  $\vec{x}' = (x'_1, \dots, x'_n)$  is a  $\gamma$ -perturbation of  $\vec{x}$ , for some  $\gamma \geq 1$ , if  $x'_1 = x_1$  and for every  $i \in [n-1]$ ,  $d(x_i, x_{i+1})/\gamma \leq d(x'_i, x'_{i+1}) \leq d(x_i, x_{i+1})$ . A  $k$ -Facility Location instance  $\vec{x}$  is  $\gamma$ -perturbation stable (or simply,  $\gamma$ -stable), if  $\vec{x}$  has a unique optimal clustering  $(C_1, \dots, C_k)$  and every  $\gamma$ -perturbation  $\vec{x}'$  of  $\vec{x}$  has the same unique optimal clustering  $(C_1, \dots, C_k)$ .

Namely, a  $\gamma$ -perturbation  $\vec{x}'$  of an instance  $\vec{x}$  is obtained by moving a subset of pairs of consecutive locations closer by a factor at most  $\gamma \geq 1$ . A  $k$ -Facility Location instance  $\vec{x}$  is  $\gamma$ -stable, if  $\vec{x}$  and any  $\gamma$ -perturbation  $\vec{x}'$  of  $\vec{x}$  admit the same unique optimal clustering (where clustering identity for  $\vec{x}$  and  $\vec{x}'$  is understood as explained in Section 2). We consistently select the optimal center  $c_i$  of each optimal cluster  $C_i$  with an even number of points as the left median point of  $C_i$ .

Our notion of linear perturbation stability naturally adapts the notion of metric perturbation stability [3, Definition 2.5] to the line. We note, the class of  $\gamma$ -stable linear instances, according to Definition 1, is at least as large as the class of metric  $\gamma$ -stable linear instances, according to [3, Definition 2.5]. Similarly to [3, Theorem 3.1] (see also [46, Lemma 7.1] and [7, Corollary 2.3]), we can show that for all  $\gamma \geq 1$ , every  $\gamma$ -stable instance  $\vec{x}$ , which admits an optimal clustering  $C_1, \dots, C_k$  with optimal centers  $c_1, \dots, c_k$ , satisfies the following  $\gamma$ -center proximity property: For all cluster pairs  $C_i$  and  $C_j$ , with  $i \neq j$ , and all locations  $x \in C_i$ ,  $d(x, c_j) > \gamma d(x, c_i)$ .

We repeatedly use the following immediate consequence of  $\gamma$ -center proximity (see also [46, Lemma 7.2]). The proof generalizes the proof of [46, Lemma 7.2] to any  $\gamma \geq 2$ .

**Proposition 1.** Let  $\gamma \geq 2$  and let  $\vec{x}$  be any  $\gamma$ -stable instance, with unique optimal clustering  $C_1, \dots, C_k$  and optimal centers  $c_1, \dots, c_k$ . Then, for all clusters  $C_i$  and  $C_j$ , with  $i \neq j$ , and all locations  $x \in C_i$  and  $y \in C_j$ ,  $d(x, y) > (\gamma - 1)d(x, c_i)$ .

The following observation, which allows us to treat stability factors multiplicatively, is an immediate consequence of Definition 1.

**Observation 1** Every  $\alpha$ -perturbation followed by a  $\beta$ -perturbation of a locations profile can be implemented by a  $(\alpha\beta)$ -perturbation and vice versa. Hence, a  $\gamma$ -stable instance remains  $(\gamma/\gamma')$ -stable after a  $\gamma'$ -perturbation, with  $\gamma' < \gamma$ , is applied to it.

We next show that for  $\gamma$  large enough, the optimal clusters of a  $\gamma$ -stable instance are well-separated, in the sense that the distance of two consecutive clusters is larger than their diameters.



**Lemma 1 (Cluster-Separation Property).** *For any  $\gamma$ -stable instance on the line with optimal clustering  $(C_1, \dots, C_k)$  and all clusters  $C_i$  and  $C_j$ , with  $i \neq j$ ,  $d(C_i, C_j) > \frac{(\gamma-1)^2}{2\gamma} \max\{D(C_i), D(C_j)\}$ .*

The cluster-separation property of Lemma 1 was first obtained in [1] as a consequence of  $\gamma$ -cluster proximity. For completeness, in Section A, we present a different proof that exploits the linear structure of the instance. Setting  $\gamma \geq 2 + \sqrt{3}$ , we get that:

**Corollary 1.** *Let  $\gamma \geq 2 + \sqrt{3}$  and let  $\vec{x}$  be any  $\gamma$ -stable instance with unique optimal clustering  $(C_1, \dots, C_k)$ . Then, for all clusters  $C_i$  and  $C_j$ , with  $i \neq j$ ,  $d(C_i, C_j) > \max\{D(C_i), D(C_j)\}$ .*

The following is an immediate consequence of the cluster-separation property in Lemma 1.

**Observation 2** *Let  $\vec{x}$  be a  $k$ -Facility Location with a clustering  $\vec{C} = (C_1, \dots, C_k)$  such that for any two clusters  $C_i$  and  $C_j$ ,  $\max\{D(C_i), D(C_j)\} < d(C_i, C_j)$ . Then, if in the optimal clustering of  $\vec{x}$ , there is a facility at the location of some  $x \in C_i$ , no agent in  $C_i$  is served by a facility at  $x_j \notin C_i$ .*

Next, we establish the so-called *no direct improvement from singleton deviations* property, used to show the strategyproofness of the ALMOSTRIGHTMOST and RANDOM mechanisms. Namely, we show that in any 3-stable instance, no agent deviating to a singleton cluster in the optimal clustering of the resulting instance can improve her connection cost through the facility of that singleton cluster. The proof is deferred to Appendix B.

**Lemma 2.** *Let  $\vec{x}$  be a  $\gamma$ -stable instance with  $\gamma \geq 3$  and optimal clustering  $\vec{C} = (C_1, \dots, C_k)$  and cluster centers  $(c_1, \dots, c_k)$ , and let an agent  $x_i \in C_i \setminus \{c_i\}$  and a location  $x'$  such that  $x'$  is a singleton cluster in the optimal clustering of the resulting instance  $(\vec{x}_{-i}, x')$ . Then,  $d(x_i, x') > d(x_i, c_i)$ .*

The following shows that for 5-stable instances  $\vec{x}$ , an agent cannot form a singleton cluster, unless she deviates by a distance larger than the diameter of her cluster in  $\vec{x}$ 's optimal clustering.

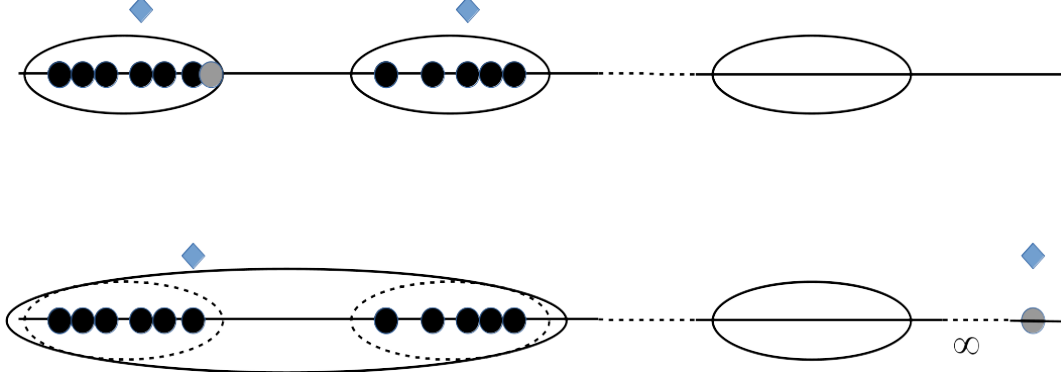
**Lemma 3.** *Let  $\vec{x}$  be any  $\gamma$ -stable instance with  $\gamma \geq 5$  and optimal clustering  $\vec{C} = (C_1, \dots, C_k)$ . Let  $x_i \in C_i \setminus \{c_i\}$  be any agent and  $x'$  any location such that  $x'$  is a singleton cluster in the optimal clustering of instance  $\vec{x}' = (\vec{x}_{-i}, x')$ , where  $x_i$  has deviated to  $x'$ . Then,  $d(x', x_i) > D(C_i)$ .*

*Proof (Sketch.).* Initially, we show that a clustering  $\vec{C}'$  of instance  $\vec{x}' = (\vec{x}_{-i}, x')$ , with  $d(x', x_i) \leq D(C_i)$ , cannot be optimal and contain  $x'$  as a singleton cluster, unless some agent  $\vec{x} \setminus C_i$  is clustered together with some agent in  $C_i$ . To this end, we use the lower bound on the distance between difference clusters for 5-stable instances show in Lemma 1. Then, using stability arguments, i.e. that the optimal clustering should not change for instance  $\vec{x}$ , even when we decrease, by a factor of 4, the distances between consecutive agents in  $\vec{x} \setminus C_i$ , we show that in  $\vec{C}'$  agents in  $\vec{x} \setminus C_i$  experience an increase in cost of at least  $2D(C_i)$  (notice that  $\vec{x} \setminus C_i = \vec{x}' \setminus (C_i \cup \{x'\})$ ). But the additional cost of serving  $x'$  from  $c_i$  in clustering  $\vec{C}$  is at most  $2D(C_i)$ , since  $d(x', x_i) \leq D(C_i)$  and  $d(x_i, c_i) \leq D(C_i)$ . Hence retaining clustering  $\vec{C}$  and serving location  $x'$  from  $c_i$  would have a smaller cost than the supposedly optimal clustering  $\vec{C}'$ . The complete proof follows by a careful case analysis and can be found in Appendix C.  $\square$

#### 4 The Optimal Solution is Strategyproof for $(2 + \sqrt{3})$ -Stable Instances

We next show that the OPTIMAL mechanism, which allocates the facilities optimally, is strategyproof for  $(2 + \sqrt{3})$ -stable instances of  $k$ -Facility Location whose optimal clustering does not include any singleton clusters. More specifically, in this section, we analyze Mechanism 1.

In general, due to the incentive compatibility of the median location in a single cluster, a deviation can be profitable only if it results in a  $k$ -clustering different from the optimal clustering



**Fig. 1.** An example of a so-called *singleton deviation*. The deviating agent (grey) declares a remote location, becomes a singleton cluster, and essentially turns the remaining agents into a  $(k - 1)$ -Facility Location instance. Thus, the deviating agent can benefit from her singleton deviation, due to the subsequent cluster merge.

---

**Mechanism 1: OPTIMAL**

---

**Result:** An allocation of  $k$ -facilities

**Input:** A  $k$ -Facility Location instance  $\vec{x}$ .

- 1 Compute the optimal clustering  $(C_1, \dots, C_k)$ . Let  $c_i$  be the left median point of each cluster  $C_i$ .
  - 2 **if**  $(\exists i \in [k] \text{ with } |C_i| = 1) \text{ or } (\exists i \in [k - 1] \text{ with } \max\{D(C_i), D(C_{i+1})\} \geq d(C_i, C_{i+1}))$  **then**
    - | **Output:** “FACILITIES ARE NOT ALLOCATED”.
  - 3 **else**
    - | **Output:** The  $k$ -facility allocation  $(c_1, \dots, c_k)$
- 

$(C_1, \dots, C_k)$  of  $\vec{x}$ . For  $\gamma$  is sufficiently large,  $\gamma$ -stability implies that the optimal clusters are well identified so that any attempt to alter the optimal clustering (without introducing singleton clusters and without violating the cluster separation property, which is necessary of stability) results in an increased cost for the deviating agent. We should highlight that Mechanism 1 may also “serve” non-stable instances that satisfy the cluster separation property. We next prove that the mechanism is strategyproof if the true instance is  $(2 + \sqrt{3})$ -stable and its optimal clustering does not include any singleton clusters, when the agent deviations do not introduce any singleton clusters and not result in instances that violate the cluster separation property (i.e. are served by the mechanism) .

**Theorem 1.** *The OPTIMAL mechanism applied to  $(2 + \sqrt{3})$ -stable instances of  $k$ -Facility Location without singleton clusters in their optimal clustering is strategyproof and minimizes the social cost.*

*Proof.* We first recall some of the notation about clusterings, introduced in Section 2. Specifically, for a clustering  $\vec{C} = (C_1, \dots, C_k)$  of an instance  $\vec{x}$  with centers  $\vec{c} = (c_1, \dots, c_k)$ , the cost of an agent (or a location)  $x$  is  $d(x, \vec{C}) = \min_{j \in [k]} \{d(x, c_j)\}$ . The cost of a set of agents  $X$  in a clustering  $\vec{C}$  is  $\text{cost}(X, \vec{C}) = \sum_{x \in X} d(x, \vec{C})$ . Finally, the cost of an instance  $\vec{x}$  in a clustering  $\vec{C}$  is  $\text{cost}(\vec{x}, \vec{C}) = \sum_{x_j \in \vec{x}} d(x_j, \vec{C})$ . This general notation allows us to refer to the cost of the same clustering for different instances. I.e, if  $\vec{C}$  is the optimal clustering of  $\vec{x}$ , then  $\text{cost}(\vec{y}, \vec{C})$  denotes the cost of instance  $\vec{y}$  in clustering  $\vec{C}$  (where we select the same centers as in clustering  $\vec{C}$  for  $\vec{x}$ ).

The fact that if OPTIMAL outputs  $k$  facilities, they optimize the social cost is straightforward. So, we only need to establish strategyproofness. To this end, we show the following: Let  $\vec{x}$  be any  $(2 + \sqrt{3})$ -perturbation stable  $k$ -Facility Location instance with optimal clustering  $\vec{C} = (C_1, \dots, C_k)$ . For any agent  $i$  and any location  $y$ , let  $\vec{Y}$  be the optimal clustering of the instance  $\vec{y} = (\vec{x}_{-i}, y)$  resulting from the deviation of  $i$  from  $x_i$  to  $y$ . Then, if  $y$  does not form a singleton cluster in  $(\vec{y}, \vec{Y})$ , either  $d(x_i, \vec{C}) < d(x_i, \vec{Y})$ , or there is an  $i \in [k - 1]$  for which  $\max\{D(Y_i), D(Y_{i+1})\} \geq d(Y_i, Y_{i+1})$ .

So, we let  $x_i \in C_i$  deviate to a location  $y$ , resulting in  $\vec{y} = (\vec{x}_{-i}, y)$  with optimal clustering  $\vec{Y}$ . Since  $y$  is not a singleton cluster, it is clustered with agents belonging in one or two clusters of  $\vec{C}$ , say either in cluster  $C_j$  or in clusters  $C_{j-1}$  and  $C_j$ . By optimality of  $\vec{C}$  and  $\vec{Y}$ , the number of facilities serving  $C_{j-1} \cup C_j \cup \{y\}$  in  $(\vec{y}, \vec{Y})$  is no less than the number of facilities serving  $C_{j-1} \cup C_j$  in  $(\vec{x}, \vec{C})$ . Hence, there is at least one facility in either  $C_{j-1}$  or  $C_j$ .

Wlog., suppose that a facility is allocated to an agent in  $C_j$  in  $(\vec{y}, \vec{Y})$ . By Corollary 1 and Observation 2, no agent in  $C_j$  is served by a facility in  $\vec{x} \setminus C_j$  in  $\vec{Y}$ . Thus we get the following cases about what happens with the optimal clustering  $\vec{Y}$  of instance  $\vec{y} = (\vec{x}_{-i}, y)$ :

**Case 1:**  $y$  is not allocated a facility in  $\vec{Y}$ : This can happen in one of two ways:

**Case 1a:**  $y$  is clustered together with some agents from cluster  $C_j$  and no facility placed in  $C_j$  serves agents in  $\vec{x} \setminus C_j$  in  $\vec{Y}$ .

**Case 1b:**  $y$  is clustered together with some agents from a cluster  $C_j$  and at least one of the facilities placed in  $C_j$  serve agents in  $\vec{x} \setminus C_j$  in  $\vec{Y}$ .

**Case 2:**  $y$  is allocated a facility in  $\vec{Y}$ . This can happen in one of two ways:

**Case 2a:**  $y$  only serves agents that belong in  $C_j$  (by optimality,  $y$  must be the median location of the new cluster, which implies that either  $y < x_{i,l}$  and  $y$  only serves  $x_{i,l}$  or  $x_{j,l} \leq y \leq x_{j,r}$ ).

**Case 2b:** In  $\vec{Y}$ ,  $y$  serves agents that belong in both  $C_{j-1}$  and  $C_j$ .

We next show that the cost of the original clustering  $\vec{C}$  is less than the cost of clustering  $\vec{Y}$  in  $\vec{y}$ . Hence, mechanism OPTIMAL would also select clustering  $\vec{C}$  for  $\vec{y}$ , which would make  $x_i$ 's deviation to  $y$  non-profitable. In particular, it suffices to show that:

$$\begin{aligned} \text{cost}(\vec{y}, \vec{C}) &< \text{cost}(\vec{y}, \vec{Y}) \Leftrightarrow \\ \text{cost}(\vec{x}, \vec{C}) + d(y, \vec{C}) - d(x_i, \vec{C}) &< \text{cost}(\vec{x}, \vec{Y}) + d(y, \vec{Y}) - d(x_i, \vec{Y}) \Leftrightarrow \\ d(y, \vec{C}) - d(y, \vec{Y}) &< \text{cost}(\vec{x}, \vec{Y}) - \text{cost}(\vec{x}, \vec{C}) + d(x_i, \vec{C}) - d(x_i, \vec{Y}) \end{aligned}$$

Since  $x_i$ 's deviation to  $y$  is profitable,  $d(x_i, \vec{C}) - d(x_i, \vec{Y}) > 0$ . Hence, it suffices to show that:

$$\begin{aligned} d(y, \vec{C}) - d(y, \vec{Y}) &\leq \text{cost}(\vec{x}, \vec{Y}) - \text{cost}(\vec{x}, \vec{C}) \\ &= \text{cost}(C_j, \vec{Y}) - \text{cost}(C_j, \vec{C}) + \text{cost}(\vec{x} \setminus C_j, \vec{Y}) - \text{cost}(\vec{x} \setminus C_j, \vec{C}) \end{aligned} \quad (1)$$

We first consider Case 1a and Case 2a, i.e., the cases where  $\vec{Y}$  allocates facilities to agents of  $C_j$  (between  $x_{j,l}$  and  $x_{j,r}$ ) that serve only agents in  $C_j$ . Note that in case 2a,  $y$  can also be located outside of  $C_j$  and serve only  $x_{i,l}$ . We can treat this case as Case 1a, since it is equivalent to placing the facility on  $x_{i,l}$  and serving  $y$  from there.

In Case 1a and Case 2a, we note that (1) holds if the clustering  $\vec{Y}$  allocates a single facility to agents in  $C_j \cup \{y\}$ , because the facility is allocated to the median of  $C_j \cup \{y\}$ , hence  $d(y, \vec{C}) - d(y, \vec{Y}) = \text{cost}(C_j, \vec{Y}) - \text{cost}(C_j, \vec{C})$ , while  $\text{cost}(\vec{x} \setminus C_j, \vec{Y}) - \text{cost}(\vec{x} \setminus C_j, \vec{C}) \geq 0$ , since  $\vec{C}$  is optimal for  $\vec{x}$ . So, we focus on the most interesting case where the agents in  $C_j \cup \{y\}$  are allocated at least two facilities. We observe that (1) follows from:

$$d(y, \vec{C}) - d(y, \vec{Y}) \leq \frac{1}{\gamma} \left( \text{cost}(\vec{x} \setminus C_j, \vec{Y}) - \text{cost}(\vec{x} \setminus C_j, \vec{C}) \right) \quad (2)$$

$$\text{cost}(C_j, \vec{C}) - \text{cost}(C_j, \vec{Y}) \leq \left( 1 - \frac{1}{\gamma} \right) \left( \text{cost}(\vec{x} \setminus C_j, \vec{Y}) - \text{cost}(\vec{x} \setminus C_j, \vec{C}) \right) \quad (3)$$

To establish (2) and (3), we first consider the valid  $\gamma$ -perturbation of the original instance  $\vec{x}$  where all distances between consecutive agent pairs to the left of  $C_j$  (i.e. agents  $\{x_1, x_2, \dots, x_{j-1,r}\}$ )

and between consecutive agent pairs to the right of  $C_j$  (i.e. agents  $\{x_{j+1,l}, \dots, x_{k,r}\}$ ) are scaled down by  $\gamma$ . By stability, the clustering  $\vec{C}$  remains the unique optimal clustering for the perturbed instance  $\vec{x}'$ . Moreover, since agents in  $\vec{x} \setminus C_j$  are not served by a facility in  $C_j$  in  $\vec{C}$  and  $\vec{Y}$ , and since all distances outside  $C_j$  are scaled down by  $\gamma$ , while all distances within  $C_j$  remain the same, the cost of the clusterings  $\vec{C}$  and  $\vec{Y}$  for the perturbed instance  $\vec{x}'$  is  $\text{cost}(C_j, \vec{C}) + \text{cost}(\vec{x} \setminus C_j, \vec{C})/\gamma$  and  $\text{cost}(C_j, \vec{Y}) + \text{cost}(\vec{x} \setminus C_j, \vec{Y})/\gamma$ , respectively. Using  $\text{cost}(\vec{x}', \vec{C}) < \text{cost}(\vec{x}', \vec{Y})$  and  $\gamma \geq 2$ , we obtain:

$$\text{cost}(C_j, \vec{C}) - \text{cost}(C_j, \vec{Y}) < \frac{1}{\gamma} \left( \text{cost}(\vec{x} \setminus C_j, \vec{Y}) - \text{cost}(\vec{x} \setminus C_j, \vec{C}) \right) \quad (4)$$

$$\leq \left(1 - \frac{1}{\gamma}\right) \left( \text{cost}(\vec{x} \setminus C_j, \vec{Y}) - \text{cost}(\vec{x} \setminus C_j, \vec{C}) \right) \quad (5)$$

Moreover, if  $C_j \cup \{y\}$  is served by at least two facilities in  $\vec{Y}$ , the facility serving  $y$  (and some agents of  $C_j$ ) is placed at the median location of  $\vec{Y}$ 's cluster that contains  $y$ . Wlog., we assume that  $y$  lies on the left of the median of  $C_j$ . Then, the decrease in the cost of  $y$  due to the additional facility in  $\vec{Y}$  is equal to the decrease in the cost of  $x_{i,l}$  in  $\vec{Y}$ , which bounds from below the total decrease in the cost of  $C_j$  due to the additional facility in  $\vec{Y}$ . Hence,

$$d(y, \vec{C}) - d(y, \vec{Y}) \leq \text{cost}(C_j, \vec{C}) - \text{cost}(C_j, \vec{Y}) \quad (6)$$

We conclude Case 1a and Case 2a, by observing that (2) follows directly from (6) and (4).

Finally, we study Case 1b and Case 2b, i.e. the cases where some agents of  $C_j$  are clustered with agents of  $\vec{x} \setminus C_j$  in  $\vec{Y}$ . Let  $C'_{j1}$  and  $C'_{j2}$  denote the clusters of  $(\vec{y}, \vec{Y})$  including all agents of  $C_j$  (i.e.,  $C_j \subseteq C'_{j1} \cup C'_{j2}$ ). By hypothesis, at least one of  $C'_{j1}$  and  $C'_{j2}$  contains an agent  $z \in \vec{x} \setminus C_j$ . Suppose this is true for the cluster  $C'_{j1}$ . Then,  $D(C'_{j1}) > D(C_j)$ , since by Corollary 1, for any  $\gamma \geq (2 + \sqrt{3})$ , the distance of any agent  $z$  outside  $C_j$  to the nearest agent in  $C_j$  is larger than  $C_j$ 's diameter. But since both  $C'_{j1}$  and  $C'_{j2}$  contain agents of  $C_j$ , we have that  $d(C'_{j1}, C'_{j2}) < D(C_j)$ . Therefore,  $D(C'_{j1}) > d(C'_{j1}, C'_{j2})$  and the cluster-separation property is violated. Hence the resulting instance  $\vec{y}$  is not  $\gamma$ -stable and Mechanism 1 does not allocated any facilities for it.  $\square$

## 5 A Deterministic Mechanism Resistant to Singleton Deviations

Next, we present a deterministic strategyproof mechanism for 5-stable instances of  $k$ -Facility Location whose optimal clustering may include singleton clusters. To make singleton cluster deviations non profitable, cluster merging has to be discouraged by the facility allocation rule. So, we allocate facilities near the edge of each optimal cluster, ending up with a significantly larger approximation ratio and a requirement for larger stability, in order to achieve strategyproofness. Specifically, we now need to ensure that no agent can become a singleton cluster close enough to her original location. Moreover, since agents can now gain by splitting their (true) optimal cluster, we need to ensure that such deviations are either non profitable or violate the cluster-separation property.

**Theorem 2.** *ALMOSTRIGHTMOST (Mechanism 2) is strategyproof for 5-stable instances of  $k$ -Facility Location and achieves an approximation ratio of  $(n - 2)/2$ .*

*Proof.* The approximation ratio of  $(n - 2)/2$  follows directly from the fact that the mechanism allocates the facility to the second rightmost agent of each non-singleton optimal cluster.

As for strategyproofness, let  $\vec{x}$  denote the true instance and  $\vec{C} = (C_1, \dots, C_k)$  its optimal clustering. We consider an agent  $x_i \in C_j$  deviating to location  $y$ , resulting in an instance  $\vec{y} = (\vec{x}_{-i}, y)$  with optimal clustering  $\vec{Y}$ . Agent  $x_i$ 's cost is at most  $D(C_j)$ . Agent  $x_i$  could profitably declare false location  $y$  in the following ways:

**Case 1:** The agents in  $C_j$  are clustered together in  $\vec{Y}$  and  $y$  is allocated a facility with  $d(y, x_i) < d(x_i, x_{i,r-1}) \leq D(C_j)$  ( $x_{i,r-1}$  is the location of  $x_i$ 's facility, when she is truthful).

**Case 1a:**  $y$  is a singleton cluster and  $d(y, x_i) < D(C_j)$ . For 5-stable instances, Lemma 3 implies that  $x_i \in C_j$  has to move by at least  $D(C_j)$  to become a singleton cluster, a contradiction.

**Case 1b:**  $y$  is the second rightmost agent of a cluster  $C'_j$  in  $(\vec{y}, \vec{Y})$ . Then, the agent  $x_i$  can gain only if  $d(y, x_i) < D(C_j)$ . In Case 1, the agents in  $C_j$  are clustered together in  $\vec{Y}$ . If  $y < x_i$ ,  $y$  must be the second rightmost agent of a cluster on the left of  $x_{j,l}$  and by Lemma 1,  $d(x_i, y) \geq d(x_{j,l}, x_{j-1,r}) > D(C_j)$ . Hence, such a deviation cannot be profitable for  $x_i$  (note how this case crucially uses the facility allocation to the second rightmost agent of a cluster). If  $y > x_i$ ,  $x_i$  can only gain if  $y$  is the second rightmost agent of a cluster including  $C_j \cup \{y, x_{j+1,l}\}$  and possibly some agents on the left of  $C_j$ , which is treated below.

**Case 2:** The agents in  $C_j$  are clustered together in  $\vec{Y}$  and  $C_j$  is merged with some agents from  $C_{j+1}$  and possibly some other agents to the left of  $x_{j,l}$  (note that merging  $C_j$  only with agents to the left of  $x_{j,l}$  does not change the facility of  $x_i$ ). Then, we only need to consider the case where the deviating agent  $x_i$  is  $x_{j,r}$ , since any other agent to the left of  $x_{j,r-1}$  cannot gain, because cluster merging can only move their serving facility further to the right. As for  $x_{j,r}$ , we note that by optimality and the hypothesis that agents in  $C_j$  belong in the same cluster of  $\vec{Y}$ ,  $x_{i,r}$  cannot cause the clusters  $C_j$  and  $C_{j+1}$  to merge in  $\vec{Y}$  by deviating in the range  $[x_{j,r}, x_{j+1,l}]$ . The reason is that the set of agents  $(C_i \setminus \{x_{j,r}\}) \cup \{y\} \cup C_{j+1}$  cannot be served optimally by a single facility, when the set of agents  $C_j \cup C_{j+1}$  requires two facilities in the optimal clustering  $\vec{C}$ . Hence, unless  $C_{j+1}$  is split in  $\vec{Y}$  (which is treated similarly to Case 3a),  $x_{j,r}$  can only move her facility to  $C_{j+1}$ , which is not profitable for her, due to Lemma 1.

**Case 3:**  $C_j$  is split into two clusters in  $\vec{Y}$ . Hence, the leftmost agents, originally in  $C_j$ , are served by a different facility than the rest of the agents originally in  $C_j$ . We next show that in any profitable deviation of  $x_i$  where  $C_j$  is split, either the deviation is not feasible or the cluster-separation property is violated. The case analysis below is similar to the proof of Theorem 1.

**Case 3a:** Agents in  $C_j$  are clustered together with some agents of  $\vec{x} \setminus C_j$  in  $\vec{Y}$ . By hypothesis, there are agents  $z, w \in C_j$  placed in different clusters of  $\vec{Y}$ , and at least one of them, say  $z$ , is clustered together with an agent  $p \in C_\ell$ , with  $\ell \neq j$ , in  $\vec{Y}$ . For brevity, we refer to the (different) clusters in which  $z$  and  $w$  are placed in clustering  $\vec{Y}$  as  $C'_z$  and  $C'_w$ , respectively. Then,  $D(C'_z) \geq d(p, z) > D(C_j)$ , by Lemma 1. But also  $d(C'_z, C'_w) < d(z, w) \leq D(C_j)$ , and consequently,  $D(C'_z) > d(C'_z, C'_w)$ , which implies that the cluster-separation property is violated and Mechanism 2 does not allocate any facilities in this case.

---

**Mechanism 2: ALMOSTRIGHTMOST**

---

**Result:** An allocation of  $k$ -facilities

**Input:** A  $k$ -Facility Location instance  $\vec{x}$ .

```

1 Find the optimal clustering  $\vec{C} = (C_1, \dots, C_k)$  of  $\vec{x}$ .
2 if there are two consecutive clusters  $C_i$  and  $C_{i+1}$  with  $\max\{D(C_i), D(C_{i+1})\} \geq d(C_i, C_{i+1})$  then
    | Output: "FACILITIES ARE NOT ALLOCATED".
3 for  $i \in \{1, \dots, k\}$  do
4     | if  $|C_i| > 1$  then
5         | Allocate a facility to the location of the second rightmost agent of  $C_i$ , i.e.,  $c_i \leftarrow x_{i,r-1}$ .
6     | else
7         | Allocate a facility to the single agent location of  $C_i$ :  $c_i \leftarrow x_{i,l}$ 
8     | end
9 end
Output: The  $k$ -facility allocation  $\vec{c} = (c_1, \dots, c_k)$ .

```

---

**Case 3b:** Agents in  $C_j$  are split and are not clustered together with any agents of  $\vec{x} \setminus C_j$  in  $\vec{Y}$ . Hence,  $y$  is not clustered with any agents in  $\vec{x} \setminus C_j$  in  $\vec{Y}$ . Otherwise, i.e., if  $y$  is not clustered with agents of  $C_j$  in  $\vec{Y}$ , it would be suboptimal for clustering  $\vec{Y}$  to allocate more than one facility to agents of  $C_j \setminus \{x_i\}$  and at most  $k - 2$  facilities to  $(\vec{x} \cup \{y\}) \setminus C_j$ , while the optimal clustering  $\vec{C}$  allocates a single facility to  $C_j$  and  $k - 1$  facilities to  $\vec{x} \setminus C_j$ . But again if  $y$  is only clustered with agents of  $C_j$ , it is suboptimal for clustering  $\vec{Y}$  to allocate more than one facility to agents of  $(C_j \cup \{y\}) \setminus \{x_i\}$  and at most  $k - 2$  facilities to  $\vec{x} \setminus C_j$ , while the optimal clustering  $\vec{C}$  allocates a single facility to  $C_j$  and  $k - 1$  facilities to  $\vec{x} \setminus C_j$ , as shown in the proof of Theorem 1.  $\square$

## 6 Low Stability and Inapproximability by Deterministic Mechanisms

We next extend the impossibility result of [27, Theorem 3.7] to  $\sqrt{2}$ -stable instances of  $k$ -Facility Location on the line, with  $k \geq 3$ . Thus, we provide strong evidence that restricting our attention to stable instances does not make strategyproof mechanism design trivial.

### 6.1 Image Sets, Holes and Well-Separated Instances

We start with some basic facts about strategyproof mechanisms and by adapting the technical machinery of well-separating instances from [27, Section 2.2] to stable instances.

**Image Sets and Holes.** Given a mechanism  $M$ , the *image set*  $I_i(\vec{x}_{-i})$  of an agent  $i$  with respect to an instance  $\vec{x}_{-i}$  is the set of facility locations the agent  $i$  can obtain by varying her reported location. Formally,  $I_i(\vec{x}_{-i}) = \{a \in \mathbb{R} : \exists y \in \mathbb{R} \text{ with } M(\vec{x}_{-i}, y) = a\}$ .

If  $M$  is strategyproof, any image set  $I_i(\vec{x}_{-i})$  is a collection of closed intervals (see e.g., [47, p. 249]). Moreover, a strategyproof mechanism  $M$  places a facility at the location in  $I_i(\vec{x}_{-i})$  nearest to the declared location of agent  $i$ . Formally, for any agent  $i$ , all instances  $\vec{x}$ , and all locations  $y$ ,  $d(y, M(\vec{x}_{-i}, y)) = \inf_{a \in I_i(\vec{x}_{-i})} \{d(y, a)\}$ .

Some care is due, because we consider mechanisms that need to be strategyproof only for  $\gamma$ -stable instances  $(\vec{x}_{-i}, y)$ . The image set of such a mechanism  $M$  is well defined (possibly by assuming that all facilities are placed to essentially  $+\infty$ ), whenever  $(\vec{x}_{-i}, y)$  is not  $\gamma$ -stable. Moreover, the requirement that  $M$  places a facility at the location in  $I_i(\vec{x}_{-i})$  nearest to the declared location  $y$  of agent  $i$  holds only if the resulting instance  $(\vec{x}_{-i}, y)$  is stable. We should underline that all instances considered in the proof of Theorem 3 are stable (and the same holds for the proofs of the propositions adapted from [27, Section 2.2]).

Any (open) interval in the complement of an image set  $I \equiv I_i(\vec{x}_{-i})$  is called a *hole* of  $I$ . Given a location  $y \notin I$ , we let  $l_y = \sup_{a \in I} \{a < y\}$  and  $r_y = \inf_{a \in I} \{a > y\}$  be the locations in  $I$  nearest to  $y$  on the left and on the right, respectively. Since  $I$  is a collection of closed intervals,  $l_y$  and  $r_y$  are well-defined and satisfy  $l_y < y < r_y$ . For convenience, given a  $y \notin I$ , we refer to the interval  $(l_y, r_y)$  as a  $y$ -hole in  $I$ .

**Well-Separated Instances.** Given a deterministic strategyproof mechanism  $M$  with a bounded approximation  $\rho \geq 1$  for  $k$ -Facility Location, an instance  $\vec{x}$  is  $(x_1 | \dots | x_{k-1} | x_k, x_{k+1})$ -*well-separated* if  $x_1 < \dots < x_k < x_{k+1}$  and  $\rho d(x_{k+1}, x_k) < \min_{i \in \{2, \dots, k\}} \{d(x_{i-1}, x_i)\}$ . We call  $x_k$  and  $x_{k+1}$  the *isolated pair* of the well-separated instance  $\vec{x}$ .

Hence, given a  $\rho$ -approximate mechanism  $M$  for  $k$ -Facility Location, a well-separated instance includes a pair of nearby agents at distance to each other less than  $1/\rho$  times the distance between any other pair of consecutive agents. Therefore, any  $\rho$ -approximate mechanism serves the two

nearby agents by the same facility and serve each of the remaining “isolated” agents by a different facility. We remark that well-separated instances are also  $\rho$ -stable.

We are now ready to adapt some useful properties of well-separated instances from [27, Section 2.2]. It is not hard to verify that the proofs of the auxiliary lemmas below apply to  $\sqrt{2}$ -stable instances, either without any change or with some minor modifications (see also [27, Appendix A]). For completeness, we give the proofs of the lemmas below in Appendix D.

**Lemma 4 (Proposition 2.2, [27]).** *Let  $M$  be any deterministic strategyproof mechanism with a bounded approximation ratio  $\rho \geq 1$ . For any  $(x_1 | \dots | x_{k-1} | x_k, x_{k+1})$ -well-separated instance  $\vec{x}$ ,  $M_k(\vec{x}) \in [x_k, x_{k+1}]$ .*

**Lemma 5 (Proposition 2.3, [27]).** *Let  $M$  be any deterministic strategyproof mechanism with a bounded approximation ratio  $\rho \geq 1$ , and let  $\vec{x}$  be a  $(x_1 | \dots | x_{k-1} | x_k, x_{k+1})$ -well-separated instance with  $M_k(\vec{x}) = x_k$ . Then, for every  $(x_1 | \dots | x_{k-1} | x'_k, x'_{k+1})$ -well-separated instance  $\vec{x}'$  with  $x'_k \geq x_k$ ,  $M_k(\vec{x}') = x'_k$ .*

**Lemma 6 (Proposition 2.4, [27]).** *Let  $M$  be any deterministic strategyproof mechanism with a bounded approximation ratio  $\rho \geq 1$ , and let  $\vec{x}$  be a  $(x_1 | \dots | x_{k-1} | x_k, x_{k+1})$ -well-separated instance with  $M_k(\vec{x}) = x_{k+1}$ . Then, for every  $(x_1 | \dots | x_{k-1} | x'_k, x'_{k+1})$ -well-separated instance  $\vec{x}'$  with  $x'_{k+1} \leq x_{k+1}$ ,  $M_k(\vec{x}') = x'_{k+1}$ .*

## 6.2 The Proof of the Impossibility Result

We are now ready to establish the main result of this section. The proof of the following builds on the proof of [27, Theorem 3.7]. However, we need some additional ideas and to be way more careful with the agent deviations used in the proof, since our proof can only rely on  $\sqrt{2}$ -stable instances.

**Theorem 3.** *For every  $k \geq 3$  and any  $\delta > 0$ , any deterministic anonymous strategyproof mechanism for  $(\sqrt{2} - \delta)$ -stable instances of  $k$ -Facility Location on the real line with  $n \geq k + 1$  agents has an unbounded approximation ratio.*

*Proof.* We only consider the case where  $k = 3$  and  $n = 4$ . It is not hard to verify that the proof applies to any  $k \geq 3$  and  $n \geq k + 1$ . To reach a contradiction, let  $M$  be any deterministic anonymous strategyproof mechanism for  $(\sqrt{2} - \delta)$ -stable instances of 3-Facility Location with  $n = 4$  agents and with an approximation ratio of  $\rho \geq 1$ .

We consider a  $(x_1 | x_2 | x_3, x_4)$ -well-separated instance  $\vec{x}$ . For a large enough  $\lambda \gg \rho$  and a very large (practically infinite)  $B \gg 6\rho\lambda$ , we let  $\vec{x} = (0, \lambda, 6B + \lambda, 6B + \lambda + \varepsilon)$ , for some small enough  $\varepsilon > 0$  ( $\varepsilon \ll \lambda/\rho$ ). By choosing  $\lambda$  and  $\varepsilon$  appropriately, becomes the instance  $\vec{x}$   $\gamma$ -stable, for  $\gamma \gg \sqrt{2}$ .

By Lemma 4,  $M_3(\vec{x}) \in [x_3, x_4]$ . Wlog, we assume that  $M_3(\vec{x}) \neq x_3$  (the case where  $M_3(\vec{x}) = x_3$  is fully symmetric and requires Lemma 5). Then, by moving agent 4 to  $M_3(\vec{x})$ , which results in a well-separated instance and, by strategyproofness, requires that  $M$  keeps a facility there, we can assume wlog. that  $M_3(\vec{x}) = x_4$ .

Since  $\vec{x}$  is well-separated and  $M$  is  $\rho$ -approximate, both  $x_3$  and  $x_4$  are served by the facility at  $x_4$ . Hence, there is a  $x_3$ -hole  $h = (l, r)$  in the image set  $I_3(\vec{x}_{-3})$ . Since  $M(\vec{x})$  places a facility at  $x_4$  and not in  $x_3$ , the right endpoint  $r$  of  $h$  lies between  $x_3$  and  $x_4$ , i.e.  $r \in (x_3, x_4]$ . Moreover, since  $M$  is  $\rho$ -approximate and strategyproof for  $(\sqrt{2} - \delta)$ -stable instances, agent 3 should be served by a facility at distance at most  $\rho\lambda$  to her, if she is located at  $4B$ . Hence, the left endpoint of the hole  $h$  is  $l > 3B$ . We distinguish two cases based on the distance of the left endpoint  $l$  of  $h$  to  $x_4$ .

**Case 1:**  $x_4 - l > \sqrt{2}\lambda$ . We consider the instance  $\vec{y} = (\vec{x}_{-3}, a)$ , where  $a > l$  is arbitrarily close to  $l$  (i.e.,  $a \gtrsim l$ ) so that  $d(a, x_4) = \sqrt{2}\lambda$ . Since  $d(x_1, x_2) = \lambda$ ,  $d(x_2, a)$  is quite large, and  $d(a, x_4) = \sqrt{2}\lambda$ ,

---

**Mechanism 3: RANDOM**

---

**Result:** An allocation of  $k$ -facilities

**Input:** A  $k$ -Facility Location instance  $\vec{x}$ .

- 1 Find the optimal clustering  $\vec{C} = (C_1, \dots, C_k)$  of  $\vec{x}$ .
  - 2 **if** there are two consecutive clusters  $C_i$  and  $C_{i+1}$  with  $1.6 \cdot \max\{D(C_i), D(C_{i+1})\} \geq d(C_i, C_{i+1})$  **then**  
    | **Output:** "FACILITIES ARE NOT ALLOCATED".
  - 3 **for**  $i \in \{1, \dots, k\}$  **do**
  - 4 |   Allocate the facility to an agent  $c_i$  selected uniformly at random from the agents of cluster  $C_i$
  - 5 **end**
  - Output:** The  $k$ -facility allocation  $\vec{c} = (c_1, \dots, c_k)$ .
- 

the instance  $\vec{y}$  is  $(\sqrt{2} - \delta)$ -stable, for any  $\delta > 0$ . By strategyproofness,  $M(\vec{y})$  must place a facility at  $l$ , since  $l \in I_3(\vec{x}_{-3})$ .

Now, we consider the instance  $\vec{y}' = (\vec{y}_{-4}, l)$ . Since we can choose  $a > l$  so that  $d(l, a) \ll \lambda$ , the instance  $\vec{y}'$  is  $(x_1|x_2|l, a)$ -well-separated and  $(\sqrt{2} - \delta)$ -stable. Hence, by strategyproofness,  $M(\vec{y}')$  must keep a facility at  $l$ , because  $l \in I_4(\vec{y}_{-4})$ .

Then, by Lemma 6,  $y'_4 = a \in M(\vec{y}')$ , because for the  $(x_1|x_2|x_3, x_4)$ -well-separated instance  $\vec{x}$ ,  $M_3(\vec{x}) = x_4$ , and  $\vec{y}'$  is a  $(x_1|x_2|l, a)$ -well-separated instance with  $y'_4 \leq x_4$ . Since both  $l, a \in M(\vec{y}')$ , either agents 1 and 2 are served by the same facility of  $M(\vec{y}')$  or agent 2 is served by the facility at  $l$ . In both cases, the social cost of  $M(\vec{y}')$  becomes arbitrarily larger than  $a - l$ , which is the optimal social cost of the 3-Facility Location instance  $\vec{y}'$ .

**Case 2:**  $x_4 - l \leq \sqrt{2}\lambda$ . This case is similar to Case 1, but it requires a bit more careful further case analysis. The details can be found in Appendix E.  $\square$

## 7 A Randomized Mechanism with Constant Approximation

In this section, we show that for an appropriate stability, a simple randomized mechanism is strategyproof, can deal with singleton clusters and achieves an approximation ratio of 2.

The intuition is that the ALMOSTRIGHTMOST mechanism can be easily transformed to a randomized mechanism, using the same key properties to guarantee strategyproofness, but achieving an  $O(1)$ -approximation, as opposed to  $O(n)$ -approximation of ALMOSTRIGHTMOST. Specifically, RANDOM (see also Mechanism 3) again finds the optimal clusters, but then places a facility at the location of an agent selected uniformly at random from each optimal cluster. We again use cluster-separation property, as a necessary condition for stability of the optimal clustering. The stability properties required to guarantee strategyproofness are very similar to those required by ALMOSTRIGHTMOST, because the set of possible profitable deviations is very similar for ALMOSTRIGHTMOST and RANDOM. Finally, notice that the cluster-separation property step of RANDOM (step 2) now makes use that due to Lemma 1, it must be  $1.6 \cdot \max\{D(C_i), D(C_{i+1})\} < d(C_i, C_{i+1})$  for 5-stable instances.

**Theorem 4.** RANDOM (Mechanism 3) is strategyproof and achieves an approximation ratio of 2 for 5-stable instances of  $k$ -Facility Location on the line.

*Proof (Sketch).* We present here the outline of the proof. The full proof can be found in Appendix F. The approximation guarantee is straightforward to verify. As mentioned, the proof of strategyproofness is similar to the proof of Theorem 2. In general, we need to cover the key deviation cases, which include the following:

**Case 1:** why agent deviating agent  $x \in C_i$  cannot gain by becoming a member of another cluster,



**Case 2:** or by becoming a self serving center,

**Case 3:** or by merging or splitting  $C_i$ .

Cases 2 and 3 can be immediately derived from the proof of Theorem 2.

The most interesting case is Case 1:  $x_i$  deviates to  $x'$  to be clustered together with agents from a different cluster of  $\vec{C}$ , in order to gain, without splitting  $C_i$  (again consider  $\vec{C} = (C_1, \dots, C_k)$  the optimal clustering of original instance  $\vec{x}$  and  $\vec{C}'$  the optimal clustering of instance  $\vec{x}' = (\vec{x}_{-i}, x')$ ).

By analyzing the expected value of agent  $x_i$  in both clusterings  $\vec{C}$  and  $\vec{C}'$  we show that in order for her to be able to gain from such a deviation, it must be  $d(x', x_i) < D(C_i)$  and  $x'$  is clustered together with agents in  $C_{i-1}$  or  $C_{i+1}$ , suppose  $C_{i-1}$  w.l.o.g. Since agents in  $C_i \setminus x_i$  are not split in clustering  $\vec{C}'$ , we know they form cluster  $C'_{i'} \in \vec{C}'$ . Hence, in this case  $x \in C'_{i'-1}$ . The key to the proof is to show that since  $d(x', x_i) < D(C_i)$  then clustering  $\vec{C}'$  on instance  $\vec{x}'$  violates the cluster separation property verification step, either between clusters  $C'_{i'}$  and  $C'_{i'-1}$  or between clusters  $C'_{i'-1}$  and  $C'_{i'-2}$ . This is also the reason why in this case the cluster separation property verification step needs to be more precise, for 5-stable instances, as mentioned in the description of the algorithm.  $\square$

## References

1. Agarwal, P., Chang, H., Munagala, K., Taylor, E., Welzl, E.: Clustering under perturbation stability in near-linear time. In: Proc. of the 40th IARCS Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2020). LIPIcs, vol. 182, pp. 8:1–8:16 (2020)
2. Alon, N., Feldman, M., Procaccia, A., Tennenholtz, M.: Strategyproof approximation of the minimax on networks. *Mathematics of Operations Research* **35**(3), 513–526 (2010)
3. Angelidakis, H., Makarychev, K., Makarychev, Y.: Algorithms for stable and perturbation-resilient problems. In: Proc. of the 49th ACM Symposium on Theory of Computing (STOC 2017). pp. 438–451 (2017)
4. Archer, A., Kleinberg, R.: Truthful germs are contagious: A local-to-global characterization of truthfulness. In: Proc. of the 9th ACM Conference on Electronic Commerce (EC '08). pp. 21–30 (2008)
5. Arthur, D., Vassilvitskii, S.: k-means++: the advantages of careful seeding. In: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2007). pp. 1027–1035. SIAM (2007)
6. Auletta, V., Prisco, R.D., Penna, P., Persiano, G.: The power of verification for one-parameter agents. *Journal of Computer and System Sciences* **75**, 190–211 (2009)
7. Awasthi, P., Blum, A., Sheffet, O.: Center-based clustering under perturbation stability. *Inf. Process. Lett.* **112**(1-2), 49–54 (2012)
8. Babaioff, M., Dobzinski, S., Oren, S.: Combinatorial auctions with endowment effect. In: Proc. of the 2018 ACM Conference on Economics and Computation (EC 2018). pp. 73–90 (2018)
9. Balcan, M., Haghtalab, N., White, C.:  $k$ -Center Clustering Under Perturbation Resilience. In: Proc. of the 43rd International Colloquium on Automata, Languages and Programming (ICALP 2016). LIPIcs, vol. 55, pp. 68:1–68:14 (2016)
10. Balcan, M.F., Blum, A., Gupta, A.: Clustering under approximation stability. *Journal of the ACM* **60**(2) (2013)
11. Balcan, M., Liang, Y.: Clustering under perturbation resilience. *SIAM Journal on Computing* **45**(1), 102–155 (2016)
12. Bilu, Y., Linial, N.: Are Stable Instances Easy? In: Proc. of the 1st Symposium on Innovations in Computer Science (ICS 2010). pp. 332–341. Tsinghua University Press (2010)
13. Bilu, Y., Daniely, A., Linial, N., Saks, M.E.: On the practically interesting instances of MAXCUT. In: Portier, N., Wilke, T. (eds.) Proceedings of the 30th International Symposium on Theoretical Aspects of Computer Science (STACS 2013). LIPIcs, vol. 20, pp. 526–537. Schloss Dagstuhl - Leibniz-Zentrum für Informatik (2013)
14. Caragiannis, I., Elkind, E., Szegedy, M., Yu, L.: Mechanism design: from partial to probabilistic verification. In: Proc. of the 13th ACM Conference on Electronic Commerce (EC '12). pp. 266–283 (2012)
15. Carroll, G.: When are local incentive constraints sufficient? *Econometrica* **80**(2), 661–686 (2012)
16. Chen, Z., Fong, K.C., Li, M., Wang, K., Yuan, H., Zhang, Y.: Facility location games with optional preference. *Theoretical Computer Science* **847**, 185–197 (2020)
17. Dokow, E., Feldman, M., Meir, R., Nehama, I.: Mechanism design on discrete lines and cycles. In: Proc. of the 13th ACM Conference on Electronic Commerce (EC '12). pp. 423–440 (2012)

18. Düetting, P., Feldman, M., Kesselheim, T., Lucier, B.: Prophet Inequalities Made Easy: Stochastic Optimization by Pricing Non-Stochastic Inputs. In: Proc. of the 58th Symposium on Foundations of Computer Science (FOCS 2017). pp. 540–551 (2017)
19. Escoffier, B., Gourvès, L., Thang, N., Pascual, F., Spanjaard, O.: Strategy-proof mechanisms for Facility Location games with many facilities. In: Proc. of the 2nd International Conference on Algorithmic Decision Theory (ADT '11). LNAI, vol. 6992, pp. 67–81 (2011)
20. Ezra, T., Feldman, M., Friedler, O.: A general framework for endowment effects in combinatorial markets. In: Proc. of the 2020 ACM Conference on Economics and Computation (EC 2020) (2020)
21. Feigenbaum, I., Li, M., Sethuraman, J., Wang, F., Zou, S.: Strategic facility location problems with linear single-dipped and single-peaked preferences. *Autonomous Agents and Multi-Agent Systems* **34**(2), 49 (2020)
22. Feldman, M., Gravin, N., Lucier, B.: Combinatorial Auctions via Posted Prices. In: Proc. of the 26th ACM-SIAM Symposium on Discrete Algorithms. pp. 123–135 (2014)
23. Feldman, M., Wilf, Y.: Randomized strategyproof mechanisms for Facility Location and the mini-sum-of-squares objective. *CoRR abs* **1108.1762** (2011)
24. Fikioris, G., Fotakis, D.: Mechanism design for perturbation stable combinatorial auctions. In: Proceedings of the 13th International Symposium on Algorithmic Game Theory (SAGT 2020). Lecture Notes in Computer Science, vol. 12283, pp. 47–63. Springer (2020)
25. Filimonov, A., Meir, R.: Strategyproof facility location mechanisms on discrete trees. *CoRR abs/2102.02610* (2021), <https://arxiv.org/abs/2102.02610>
26. Fotakis, D., Tzamos, C.: Winner-imposing strategyproof mechanisms for multiple facility location games. *Theoretical Computer Science* **472**, 90–103 (2013)
27. Fotakis, D., Tzamos, C.: On the power of deterministic mechanisms for facility location games. *ACM Transactions on Economics and Computation* **2**(4), 15:1–15:37 (2014)
28. Fotakis, D., Tzamos, C.: Strategyproof facility location for concave cost functions. *Algorithmica* **76**(1), 143–167 (2016)
29. Fotakis, D., Tzamos, C., Zampetakis, M.: Mechanism design with selective verification. In: Proceedings of the 2016 ACM Conference on Economics and Computation (EC 2016). pp. 771–788. ACM (2016)
30. Fotakis, D., Zampetakis, E.: Truthfulness flooded domains and the power of verification for mechanism design. *ACM Transactions on Economics and Computation* **3**(4), 20:1–20:29 (2015)
31. Goel, S., Hann-Caruthers, W.: Coordinate-wise median: Not bad, not bad, pretty good. *CoRR abs/2007.00903* (2020), <https://arxiv.org/abs/2007.00903>
32. Golomb, I., Tzamos, C.: Truthful facility location with additive errors. *CoRR abs/1701.00529* (2017), <http://arxiv.org/abs/1701.00529>
33. Green, J., Laffont, J.: Partially verifiable information and mechanism design. *Review of Economic Studies* **53**(3), 447–456 (1986)
34. Kyropoulou, M., Ventre, C., Zhang, X.: Mechanism design for constrained heterogeneous facility location. In: Proceedings of the 12th International Symposium on Algorithmic Game Theory (SAGT 2019). Lecture Notes in Computer Science, vol. 11801, pp. 63–76. Springer (2019)
35. Li, M., Lu, P., Yao, Y., Zhang, J.: Strategyproof mechanism for two heterogeneous facilities with constant approximation ratio. In: Proceedings of the 29th International Joint Conference on Artificial Intelligence (IJCAI 2020). pp. 238–245 (2020)
36. Lu, P., Sun, X., Wang, Y., Zhu, Z.: Asymptotically Optimal Strategy-Proof Mechanisms for Two-Facility Games. In: Proc. of the 11th ACM Conference on Electronic Commerce (EC '10). pp. 315–324 (2010)
37. Lu, P., Wang, Y., Zhou, Y.: Tighter bounds for facility games. In: Proc. of the 5th Workshop on Internet and Network Economics (WINE '09). LNCS, vol. 5929, pp. 137–148 (2009)
38. Mei, L., Li, M., Ye, D., Zhang, G.: Facility location games with distinct desires. *Discrete Applied Mathematics* **264**, 148–160 (2019)
39. Meir, R.: Strategyproof facility location for three agents on a circle. In: Proceedings of the 12th International Symposium on Algorithmic Game Theory (SAGT 2019). Lecture Notes in Computer Science, vol. 11801, pp. 18–33. Springer (2019)
40. Miyagawa, E.: Locating libraries on a street. *Social Choice and Welfare* **18**, 527–541 (2001)
41. Nissim, K., Smorodinsky, R., Tennenholtz, M.: Approximately optimal mechanism design via differential privacy. In: Innovations in Theoretical Computer Science (ITCS 2012). pp. 203–213. ACM (2012)
42. Procaccia, A., Tennenholtz, M.: Approximate mechanism design without money. In: Proc. of the 10th ACM Conference on Electronic Commerce (EC '09). pp. 177–186 (2009)
43. Roughgarden, T.: Beyond Worst-Case Analysis. *Communications of the ACM* **62**(3), 88–96 (2019)
44. Roughgarden, T.: Beyond the Worst-Case Analysis of Algorithms. Cambridge University Press (2020)
45. Roughgarden, T., Talgam-Cohen, I.: Approximately Optimal Mechanism Design. *CoRR* (2018), <http://arxiv.org/abs/1812.11896>

- 46. Roughgarden, T.: Lecture 6: Perturbation-stable clustering. CS264: Beyond Worst-Case Analysis (2017), <http://timroughgarden.org/w17/l/16.pdf>
- 47. Schummer, J., Vohra, R.: Mechanism design without money. *Algorithmic Game Theory* **10**, 243–299 (2007)
- 48. Serafino, P., Ventre, C.: Heterogeneous facility location without money. *Theoretical Computer Science* **636**, 27–46 (2016)
- 49. Sui, X., Boutilier, C., Sandholm, T.: Analysis and optimization of multi-dimensional percentile mechanisms. In: *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI 2013)*. pp. 367–374. IJCAI/AAAI (2013)

## A The Proof of Lemma 1

*Proof.* It suffices to establish the lemma for two consecutive clusters  $C_i$  and  $C_{i+1}$ . We recall that  $d(C_i, C_{i+1}) = d(x_{i,r}, x_{i+1,l})$ . Moreover, by symmetry, we can assume wlog. that  $D(C_i) \geq D(C_{i+1})$ .

If  $C_i$  is a singleton,  $D(C_i) = 0$  and the lemma holds trivially. If  $|C_i| = 2$ , wlog. we can only consider the case where  $x_{i,l}$  is  $C_i$ 's center. Otherwise, i.e., if  $x_{i,r}$  is  $C_i$ 's center in optimal clustering  $(C_1, \dots, C_i, \dots, C_k)$  with centers  $(c_1, \dots, x_{i,r}, \dots, c_j)$ , the same clustering  $(C_1, \dots, C_i, \dots, C_k)$  with centers  $(c_1, \dots, x_{i,l}, \dots, c_j)$  is also optimal for the  $\gamma$ -stable instance  $\vec{x}$  (and should still be optimal after a  $\gamma$  perturbation of  $\vec{x}$ , due to the stability of the instance). We then have:

$$D(C_i) = d(x_{i,l}, x_{i,r}) = d(c_i, x_{i,r}) < \frac{1}{(\gamma - 1)} d(x_{i,r}, x_{i+1,r}) = \frac{1}{(\gamma - 1)} d(C_i, C_{i+1}) \Rightarrow \\ d(C_i, C_{i+1}) > (\gamma - 1)D(C_i)$$

where the first inequality follows from Proposition 1. The lemma then follows by noticing that for any  $\gamma \geq 1$ :

$$\gamma - 1 \geq \frac{\gamma^2 + 1}{2\gamma} - 1$$

The most interesting case is where  $|C_i| \geq 3$  and  $x_{i,l} < c_i \leq x_{i,r}$ . Suppose  $d(x_{i,l}, c_i) = \beta D(C_i)$ , for some  $\beta \in (0, 1]$  and hence  $d(c_i, x_{i,r}) = (1 - \beta)D(C_i)$  (i.e.,  $\beta$  quantifies how close  $c_i$  is to  $C_i$ 's extreme points and to the closest point of  $C_{i+1}$ .) We recall that  $d(C_i, C_{i+1}) = d(x_{i,r}, x_{i+1,l})$ .

We start with a tighter analysis of the equivalent of Proposition 1 for  $x_{i,l}$  and  $x_{i+1,l}$ , taking into account their specific ordering on the line:

$$\begin{aligned} d(x_{i,l}, x_{i+1,l}) &\geq d(x_{i,l}, c_{i+1}) - d(x_{i+1,l}, c_{i+1}) \\ &> \gamma d(x_{i,l}, c_i) - \frac{d(x_{i+1,l}, c_i)}{\gamma} \\ &= \gamma d(x_{i,l}, c_i) - \frac{d(x_{i+1,l}, x_{i,l}) - d(x_{i,l}, c_i)}{\gamma} \Rightarrow \\ d(x_{i,l}, x_{i+1,l}) &> \frac{\gamma^2 + 1}{\gamma + 1} d(x_{i,l}, c_i) \end{aligned}$$

Where the second inequality stands due to the  $\gamma$ -center proximity property of  $\gamma$  stable instances and the equality stands because  $x_{i,l} < c_i < x_{i+1,l}$ . Since  $d(C_i, C_{i+1}) = d(x_{i,r}, x_{i+1,l}) = d(x_{i,l}, x_{i+1,l}) - D(C_i)$ , and by  $d(x_{i,l}, c_i) = \beta D(C_i)$ , we get that:

$$d(C_i, C_{i+1}) > \left( \frac{\beta(\gamma^2 + 1)}{\gamma + 1} - 1 \right) D(C_i) \quad (7)$$

Furthermore, by Proposition 1, we have that  $d(x_{i,r}, x_{i+1,l}) > (\gamma - 1)d(x_{i,r}, c_i)$ . Hence, by  $d(c_i, x_{i,r}) = (1 - \beta)D(C_i)$ , we get that:

$$d(C_i, C_{i+1}) > (1 - \beta)(\gamma - 1)D(C_i) \quad (8)$$

So, by (7) and (8) we have that it must be:

$$d(C_i, C_{i+1}) > \max \left\{ \frac{\beta(\gamma^2 + 1)}{\gamma + 1} - 1, (1 - \beta)(\gamma - 1) \right\} D(C_i) \quad (9)$$

We now observe that for any fixed  $\gamma > 1$ , the first term of the max in (9),  $\frac{\beta(\gamma^2+1)}{\gamma+1} - 1$ , is increasing for all  $\beta > 0$ , while the second term,  $(1-\beta)(\gamma-1)$ , is decreasing for all  $\beta \in (0, 1]$ . Hence, for any fixed  $\gamma > 1$ , the minimum value of the max in (9) is achieved when  $\beta$  satisfies:

$$\frac{\beta(\gamma^2+1)}{\gamma+1} - 1 = (1-\beta)(\gamma-1)$$

Solving for  $\beta$ , we get that:

$$\beta = \frac{1}{2} + \frac{1}{2\gamma}, \quad (10)$$

with  $\beta \in (1/2, 1]$ , when  $\gamma \geq 1$ .

We conclude the proof by substituting the value of  $\beta$  in (10) to (9).  $\square$

## B The Proof of Lemma 2

*Proof.* We establish the lemma for the leftmost agent  $x_{i,l}$  as the deviating agent. Specifically, we show that  $x_{i,l}$  needs to move by at least  $d(x_{i,l}, c_i)$  to the left in order to become a singleton cluster. The property then follows for the rest of the agents.

Suppose  $x_{i,l}$  can create a singleton cluster by deviating less than  $d(x_{i,l}, c_i)$  to the left. I.e., for some  $x'$  such that  $d(x', x_{i,l}) < d(x_{i,l}, c_i)$  the optimal clustering of  $\vec{x}' = (\vec{x}_{-x_{i,l}}, x')$  is such that the agent location at  $x'$  becomes a singleton cluster. We call this clustering (that is optimal for  $\vec{x}'$ )  $\vec{C}'$ . Notice that since  $d(x', x_{i,l}) < d(x_{i,l}, c_i)$ ,  $x'$  is in the gap between clusters  $C_{i-1}$  and  $C_i$  as by 3-perturbation stability we have  $d(x_{i-1,r}, x_{i,l}) > 2d(x_{i,l}, c_i)$ . This means that in order for this case to be feasible, no agents from  $C_{i-1}$  can be clustered together with agents in  $C_i$  in  $(\vec{x}', \vec{C}')$ , because  $x'$  lies between them and is a singleton cluster.

Consider now the instance  $\vec{x}_{-x_{i,l}}$ . We know that  $\text{cost}(\vec{x}_{-x_{i,l}}, \vec{C}') \geq \text{cost}(\vec{x}_{-x_{i,l}}, \vec{C})$ . That is, since otherwise the optimal clustering for  $\vec{x}$  would make  $x_{i,l}$  a singleton cluster and serve the rest of the agents as in  $\vec{C}'$ . Let  $\text{diff}$  be the difference in the total cost agents in  $\vec{x}_{-x_{i,l}}$  experience between clusterings  $\vec{C}$  and  $\vec{C}'$ . I.e.  $\text{diff} = \text{cost}(\vec{x}_{-x_{i,l}}, \vec{C}') - \text{cost}(\vec{x}_{-x_{i,l}}, \vec{C})$ . As before, since  $x_{i,l}$  is not a singleton cluster in  $(\vec{x}, \vec{C})$  we know that  $d(x_{i,l}, c_i) < \text{diff}$  (or else setting  $x_{i,l}$  as a singleton would have a lower cost in  $\vec{x}$  than  $\vec{C}$ ).

But we can perform a 3-perturbation in  $\vec{x}$  in the following way: Scale down all distances between agents from  $x_1$  up to  $x_{i-1,r}$  and all distances between agents from  $x_{i,l+1}$  to  $x_n$  ( $x_n$  being the rightmost agent of the instance) by 3. Call this instance  $\vec{x}_{\text{per}}$ . Since agents of clusters  $C_{i-1}$  and  $C_i$  are not clustered together neither in  $\vec{C}$  nor in  $\vec{C}'$  we have that

$$\text{diff}_{\text{per}} \leq \frac{\text{cost}(\vec{x}_{-x_{i,l}}, \vec{C}') - \text{cost}(\vec{x}_{-x_{i,l}}, \vec{C})}{3}.$$

So  $\text{diff}_{\text{per}} \leq \text{diff}/3$ . Since  $d(x_{i,l}, c_i)$  is unaffected in the perturbation and by stability the optimal clustering of  $\vec{x}_{\text{per}}$  must remain the same (as  $\vec{x}$ ) we have that it must be  $d(x_{i,l}, c_i) < \text{diff}/3$  (1).

Finally, the least amount of extra social cost suffered between  $\text{cost}(\vec{x}, \vec{C})$  and the case of setting  $x_{i,l}$  as a center that serves only itself and serve the remaining agents of the instance as on  $\vec{C}'$  (i.e. as they would be served should  $x'$  gets a facility that served only herself), will be  $\text{diff} - d(x_{i,l}, c_i)$ . This means that the optimal clustering algorithm would only choose this solution when  $d(x', c_i) > \text{diff} - d(x, c_i)$ . So the agent must deviate by at least  $\text{diff} - 2d(x, c_i)$ . But from (1) we have

$$\text{diff} - 2d(x, c_i) > 3d(x, c_i) - 2d(x_{i,l}, c_i) = d(x_{i,l}, c_i),$$

which concludes the proof of the lemma.  $\square$

## C The Proof of Lemma 3

We first present the outline of the proof and then the proof follows. We do this because despite the mostly relatively straight forward arguments used in the proof, due to the delicate formalization required in order to formally describe all the mentioned conditions, the proof gains a good amount of descriptive length. We consider random agent  $x_i \in C_i$  of instance  $\vec{x}$  with optimal clustering  $\vec{C} = (C_1, \dots, C_k)$ , deviating to location  $x'$  creating instance  $\vec{x}' = (x_{-i}, x')$ .

Initially we show that due to the large distance between clusters  $C_i$  and  $C_j$  with  $i \neq j$ , guaranteed by Lemma 1 for 5-stable instances, we need only study the cases where  $x' \in (x_{i-1,r}, x_{i,l})$  and  $x' \in (x_{i,r}, x_{i+1,l})$  and in the optimal clustering  $\vec{C}'$  of instance  $\vec{x}'$  no agent in  $\vec{x}' \setminus C_i$  is served together with any agent in  $C_i$ <sup>7</sup>, as in all other cases either  $x'$  is not a singleton in  $\vec{C}'$  or  $d(x', x_i) > D(C_i)$ .

The rest of the proof follows the logic of the proof of Theorem 1 (which follows), tailored to this specific case. More specifically, given the observation above, we notice the following: In alternative clustering  $\vec{C}''$  in which we forcefully place two facilities serving only agents in  $C_i$  (optimally with regards to serving agents in  $C_i$ ), and serve the remaining agents,  $\vec{x} \setminus C_i$ , optimally with  $k - 2$  facilities, the cost agents in  $\vec{x} \setminus C_i$  experience in clustering  $\vec{C}''$  is the same cost agents in  $\vec{x}' \setminus C_i \cup x'$  experience in clustering  $\vec{C}$  (notice that the sets  $\vec{x} \setminus C_i$  and  $\vec{x}' \setminus C_i \cup x'$ ). Now, the cost of agents in  $C_i$  in clustering  $\vec{C}''$  is at least  $D(C_i)/2$  smaller than it is in  $\vec{C}'$  (that is since we can always place the facility to the edge agent further from  $c_i$  - see proof of Theorem 1). But since  $\vec{C}''$  is not optimal for  $\vec{x}$  this means that agents in  $\vec{x} \setminus C_i$  experience an increase in cost larger than  $D(C_i)/2$  in clustering  $\vec{C}''$  when compared to clustering  $\vec{C}$ . For brevity we symbolize this cost increase as  $cst$ , so we say  $cst > D(C_i)/2$ .

We now we consider the 4-perturbation of instance  $\vec{x}$  in which all distances among agents to the left and to the right of  $C_i$  are shrunk by a factor of 4. By stability we know that the optimal clustering of the perturbed instance should be the same as the optimal clustering of the original! But in the perturbed instance all costs of agents in  $\vec{x} \setminus C_i$  are divided by 4 in both clusterings  $\vec{C}$  and  $\vec{C}''$  while the costs of agents in  $C_i$  remain the same. So, in order for  $\vec{C}''$  to be sub-optimal in the perturbed instance it must be  $cst/4 > D(C_i)/2$  which means  $cst > 2D(C_i)$ . But serving agent  $x'$  of  $\vec{x}'$  by  $c_i$  has cost at most  $2D(C_i)$  if  $d(x', x_i) < D(C_i)$  since  $d(x_i, c_i) < D(C_i)$ . This means that clustering  $\vec{C}'$  cannot be optimal for  $\vec{x}$ .

*Proof.* We want to show the lemma for any  $\gamma$ -stable instance for  $\gamma \geq 5$ .

We prove the lemma for random agent  $x_j \in C_i$  for some cluster  $C_i$  in optimal clustering  $\vec{C} = (C_1, \dots, C_k)$  of the  $\gamma$ -stable instance  $\vec{x}$ . Consider that the agent declares false location  $x'$  providing input profile  $\vec{x}' = (\vec{x}_{-j}, x')$  to the mechanism in order to become a singleton cluster. That is, if the optimal clustering of instance  $\vec{x}'$  is  $\vec{Y}$  then  $x'$  is a single agent cluster in  $\vec{Y}$ .

We first study the case where  $|C_i| = 2$ . But then, from Lemma 2 we know that for any  $\gamma$ -stable instance for  $\gamma \geq 3$  agent  $x_j \in C_i$  of optimal clustering  $\vec{C}$  must deviate by at least his distance to  $C_i$ 's center in order to become a singleton cluster in  $\vec{Y}$ . I.e. it must be  $d(x', x_j) > d(x_j, c_i) = D(C_i)$ , so the lemma stands for this case.

For the most general case,  $|C_i| \geq 3$  we start with some observations. By Lemma 1 we know that for any two clusters  $C_i$  and  $C_j$  of optimal clustering  $\vec{C} = (C_1, \dots, C_k)$  of  $\vec{x}$  we have  $d(C_i, C_j) > \left(\frac{\gamma-1}{2\gamma}\right) \max\{D(C_i), D(C_j)\}$ . For  $\gamma \geq 5$  that is:

$$d(C_i, C_j) > 1.6 \max\{D(C_i), D(C_j)\}. \quad (11)$$

<sup>7</sup> Note here that we refer to the group of agents that belong in cluster  $C_i$  of the optimal clustering of instance  $\vec{x}$ . This group is well defined for instance  $\vec{x}'$  as well.

Too begin, we notice the following claim:

**Claim 1** *Agent  $x_i$  cannot declare a false location  $x'$  with  $x_{i,l} \leq x' \leq x_{i,r}$  in such a way that  $x'$  is a singleton cluster in  $\vec{Y}$ .*

We can easily see the validity of the claim, since by optimality (also see proof of Theorem 1)  $x_j \in C_i$  cannot change the optimal clustering by deviating within the bounds of cluster  $C_i$ , i.e. if  $x_{i,l} \leq x' \leq x_{i,r}$ . Hence it must be  $x' \neq [x_{i,l}, x_{i,r}]$ . Even so, for completeness, we provide a proof of the claim, tailored to the case of 5-stable instances, after the proof of the lemma.

In addition, we notice that if  $x' \leq x_{i-1,r}$  or  $x' \geq x_{i+1,l}$  then the lemma trivially stands, again by Equation 11 (I.e. in this case it would be  $d(x', x_j) > 1.6D(C_i)$ ).

This means that we need only study the cases where  $x' \in (x_{i-1,r}, x_{i,l})$  or  $x' \in (x_{i,r}, x_{i+1,l})$  and  $d(x', x_j) < D(C_i)$  (and show that  $x'$  cannot become a singleton cluster in  $\vec{Y}$  in these cases).

Suppose, contrary to the lemma's claim, that agent declares location  $x' \in (x_{i-1,r}, x_{i,l})$  with  $d(x', x_j) < D(C_i)$  such that  $x'$  is a singleton in  $\vec{Y}$  (the other case,  $x' \in (x_{i,r}, x_{i+1,l})$ , is symmetrical). Then we notice the following three properties for optimal clustering  $\vec{Y}$  of instance  $\vec{x}$ :

**Property 1:** In  $\vec{Y}$  there is a facility among agents in  $C_i \setminus x_j$ .

**Property 2:** In  $\vec{Y}$  no agent “to the left” of cluster  $C_i$  (i.e. by an agent in some cluster  $C_l$  for  $l < i$ , of  $\vec{C}$ ) is served by an agent in  $C_i \setminus x_j$ .

**Property 3:** In  $\vec{Y}$  no agent “to the right” of cluster  $C_i$  (i.e. by an agent in some cluster  $C_l$  for  $l > i$ , of  $\vec{C}$ ) is served by an agent in  $C_i \setminus x_j$ .

The imminent conclusion from Properties 1, 2 and 3 is the following: Consider instance  $\vec{x} \setminus C_i$  and it's optimal  $k - 2$ -clustering  $\vec{C}_{-2}$ . Then  $\text{cost}(\vec{x} \setminus C_i, \vec{Y}) = \text{cost}(\vec{x} \setminus C_i, \vec{C}_{-2})$ <sup>8</sup>. We provide short proofs for each one of these three properties right after the proof of the lemma.

We are now ready to complete the proof. In order to do so we bound the extra cost experienced by agents in  $\vec{x} \setminus C_i$  in the possible re-clustering after  $x_i$ 's deviation, i.e.  $\text{cost}(\vec{x} \setminus C_i, \vec{Y}) - \text{cost}(\vec{x} \setminus C_i, \vec{C})$ . We do this by considering the following alternative clustering  $\vec{C}'$  of instance  $\vec{x}$ : serve agents in  $C_i$  using two facilities, optimally and agents in  $\vec{x} \setminus C_i$  using the remaining  $k - 2$  facilities optimally. So in  $\vec{C}'$  we have:

$$\text{cost}(C_i, \vec{C}') \leq \text{cost}(C_i, \vec{C}) - \frac{D(C_i)}{2}, \quad (12)$$

since placing the second facility placed among agents in  $C_i$  to the edge-agent further away from  $c_i$  reduces the cost by at least  $\frac{D(C_i)}{2}$ .

But since  $\vec{C}$  is optimal in  $\vec{x}$  and hence  $\vec{C}'$  is not, it must be:

$$\text{cost}(\vec{x} \setminus C_i, \vec{C}') - \text{cost}(\vec{x} \setminus C_i, \vec{C}) > \frac{D(C_i)}{2} \quad (13)$$

Otherwise it would be  $\text{cost}(\vec{x}, \vec{C}') < \text{cost}(\vec{x}, \vec{C})$ . Now notice that properties 1, 2 and 3 mean that agents in  $\vec{x} \setminus C_i$  are clustered in exactly the same way in  $\vec{C}'$  as in  $\vec{Y}$ . That means that:

$$\text{cost}(\vec{x}' \setminus \{C_i \cup x'\}, \vec{Y}) = \text{cost}(\vec{x} \setminus C_i, \vec{C}') \quad (14)$$

and that no agent to the left of  $C_i$  is clustered together with any agent to the right of  $C_i$  in  $\vec{C}'$ .

<sup>8</sup> For a description of this notation, of the form  $\text{cost}(\vec{x}, \vec{C})$ , see proof of Theorem 1

The last observation means if we consider a 4-perturbation of instance  $\vec{x}$ , instance  $\vec{x}_p$ , in which we divide all distances among agents between  $[x_l, x_{i-1,r}]$  and agents between  $[x_{i+1,l}, x_r]$ , where  $x_l$  and  $x_r$  the leftmost and rightmost agents of the instance equivalently we have that:

$$\text{cost}(\vec{x}_p \setminus C_i, \vec{C}') - \text{cost}(\vec{x}_p \setminus C_i, \vec{C}) = \frac{\text{cost}(\vec{x} \setminus C_i, \vec{C}') - \text{cost}(\vec{x} \setminus C_i, \vec{C})}{4}$$

But in  $x_p$  the distances among agents in  $C_i$  remain unaffected which means that in  $x_p$ , Equation 12 still stands. This means, that since the instance is 5-stable, clustering  $\vec{C}'$  must still be sub-optimal in  $\vec{x}_p$  and hence it must be

$$\begin{aligned} \text{cost}(\vec{x}_p \setminus C_i, \vec{C}') - \text{cost}(\vec{x}_p \setminus C_i, \vec{C}) &> \frac{D(C_i)}{2} \Rightarrow \\ \frac{\text{cost}(\vec{x} \setminus C_i, \vec{C}') - \text{cost}(\vec{x} \setminus C_i, \vec{C})}{4} &> \frac{D(C_i)}{2} \Rightarrow \\ \text{cost}(\vec{x} \setminus C_i, \vec{C}') - \text{cost}(\vec{x} \setminus C_i, \vec{C}) &> 2D(C_i). \end{aligned} \tag{15}$$

Noticing again that by Equation (14),  $\text{cost}(\vec{x}' \setminus \{C_i \cup x'\}, \vec{Y}) = \text{cost}(\vec{x} \setminus C_i, \vec{C}')$  and by Equation (15) and  $\text{cost}(\vec{x} \setminus C_i, \vec{C}) = \text{cost}(\vec{x}' \setminus \{C_i \cup x'\}, \vec{C})$  we have

$$\text{cost}(\vec{x}' \setminus \{C_i \cup x'\}, \vec{Y}) - \text{cost}(\vec{x}' \setminus \{C_i \cup x'\}, \vec{C}) > 2D(C_i).$$

Finally, since  $d(x', x_i) < D(C_i)$ ,

$$\text{cost}(\{C_i \cup x' \setminus x_i\}, \vec{C}) - \text{cost}(\{C_i \cup x' \setminus x_i\}, \vec{Y}) < 2D(C_i),$$

since  $d(x_i, c_i) \leq D(C_i)$ . By adding the last two equations we get that  $\text{cost}(\vec{x}', \vec{Y}) > \text{cost}(\vec{x}', \vec{C})$  which means that  $\vec{Y}$  is not optimal.  $\square$

We now present the proofs of **Claim 1** and **Properties 1, 2 and 3**, used in the main proof of Lemma 3.

*Proof (Of Claim 1).* Consider  $x'_{i,l}$  and  $x'_{i,r}$  to be the leftmost and rightmost agents of  $C_i \setminus x_j$  (i.e. if  $x_j \neq x_{i,r}, x_{i,l}$  then  $x_{i,r} = x'_{i,r}$  and  $x_{i,l} = x'_{i,l}$ ).

Contrary to the claim, suppose  $x'_{i,l} \leq x' \leq x'_{i,r}$  and  $x'$  is a singleton cluster  $\vec{Y}$ . Since  $x'$  is a singleton and  $x'_{i,l}$  and  $x'_{i,r}$  are to her left and right side equivalently,  $x'_{i,l}$  and  $x'_{i,r}$  cannot be served by the same facility in  $\vec{Y}$  (since clustering  $\vec{Y}$  is optimal for  $\vec{x}'$ ). This means that either  $x'_{i,l}$  or  $x'_{i,r}$  is served by an agent in  $\vec{x} \setminus C_i$  or there are two facilities among agents in  $C_i \setminus x_j$  in  $\vec{Y}$ . Both of these cases are infeasible though. For the first one, suppose that  $x_{i,r}$  is not served by an agent in  $C_i \setminus x_j$ . By Equation 11 that means that the cost of serving  $x_{i,r}$  is at-least  $1.6D(C_i)$ . But since  $x'_{i,l} \leq x' \leq x'_{i,r}$ ,  $d(x', x_{i,r}) < D(C_i)$  so  $\vec{Y}$  could not be optimal in  $\vec{x}'$ . For the latter case ( $\vec{Y}$  places two facilities among agents in  $C_i \setminus x_j$ ) we see that if  $\vec{Y}$  is optimal for  $\vec{x}'$  then the optimal  $(k-1)$ -clustering of instance  $(\vec{x}_{-j})$  would place two facilities among agents in  $C_i \setminus x_j$  (since  $x'$  is a singleton removing her and one facility from the instance should yield the exact same clustering for the rest of the agents). But then, since in  $\vec{C}$  there is only one facility among agents in  $C_i$ ,  $\vec{C}$  could not be optimal for instance  $\vec{x}$  (because if the optimal  $(k-1)$ -clustering of instance  $(\vec{x}_{-j})$  places



two facilities among agents in set  $C_i \setminus x_j$  then the optimal  $k$ -clustering of instance  $\vec{x}$  should place at least as many agents in  $C_i$ ), which is a contradiction.

Finally we notice that if  $x_j = x_{i,l}$ ,  $x'$  cannot become a singleton in  $\vec{Y}$  if  $x' \in [x_{i,l}, x'_{i,l}]$  since the cost serving agent  $x_j$  by  $c_i$  in that interval is only decreased (in relation to the cost of serving her by  $c_i$  in  $\vec{x}$  - she's getting closer to her serving facility). Similarly for the case of  $x_j = x_{i,r}$  moving in interval  $[x'_{i,r}, x_{i,r}]$ . The above mean that agent  $x_j$  cannot become a singleton cluster by moving within the bounds of  $C_i$  (i.e. if  $x'$  is a singleton in  $\vec{Y}$  it must be  $x' \notin [x_{i,l}, x_{i,r}]$ ), which is the claim.  $\square$

*Proof (Of Property 1).* We know that  $|C_i \setminus x_j| \geq 2$ . Furthermore, since  $d(x', x_j) < D(C_i)$  we have that  $d(x', c_i) < 2D(C_i)$ . But if there is no facility among agents in  $C_i \setminus x_j$  that means that these agents are all served by a facility placed in a location  $x_l$  with  $x_l \in C_l$  with  $l \neq i$ . But, again by Equation 11 that would mean that

$$\text{cost}(C_i \setminus x_j, \vec{Y}) > 2 * 1.6D(C_i) + \text{cost}(C_i \setminus x_j, \vec{C}) \quad (16)$$

(since  $|C_i \setminus x_j| \geq 2$ ,  $d(C_i, C_l) \geq 1.6D(C_i)$ ). Furthermore, since agents in  $\vec{x} \setminus C_i$  are served by the same number of facilities in  $\vec{Y}$  as in  $\vec{C}$ , but also have to serve agents in  $C_i \setminus x_j$  in  $\vec{Y}$  (i.e. the placement of the  $(k-1)$  facilities among agents in  $\vec{x} \setminus C_i$  is not optimal in  $\vec{Y}$  as it is in  $\vec{C}$ , for these agents) we have

$$\text{cost}(\vec{x} \setminus C_i, \vec{Y}) \geq \text{cost}(\vec{x} \setminus C_i, \vec{C}). \quad (17)$$

Hence, by adding (16) and (17) we have that :

$$\text{cost}(\vec{x} \setminus x_j, \vec{Y}) > 2 * 1.6D(C_i) + \text{cost}(\vec{x} \setminus x_j, \vec{C})$$

By remembering that  $\vec{x}' = (\vec{x}_{-j}, x')$  and in  $\vec{Y}$   $x'$  is a singleton cluster (i.e. has cost 0) the above becomes:

$$\text{cost}(\vec{x}', \vec{Y}) > 2 * 1.6D(C_i) + \text{cost}(\vec{x} \setminus x_j, \vec{C}) \quad (18)$$

But, alternative clustering  $\vec{C}'$  for  $\vec{x}'$  in which we serve all agents as we do in  $\vec{C}$  and also serve location  $x'$  by  $c_i$  has cost

$$\text{cost}(\vec{x}', \vec{C}') \leq \text{cost}(\vec{x} \setminus x_j, \vec{C}) + 2D(C_i), \quad (19)$$

since  $d(x', c_i) < 2D(C_i)$ .

This means that, by (18) and (19)  $\text{cost}(\vec{x}', \vec{Y}) > \text{cost}(\vec{x}', \vec{C}')$  which means that clustering  $\vec{Y}$  would be sup-optimal for instance  $\vec{x}'$ , which is a contradiction.

Notice that by Observation 2, property 1 means that no agents in  $C_i \setminus x_j$  are served by an agent not in  $C_i$  in  $\vec{Y}$ .  $\square$

*Proof (Of Property 2).* Property 2 is trivial: since  $x' \in (x_{i-1,r}, x_{i,l})$  and  $x'$  forms a singleton cluster in  $\vec{Y}$ , by optimality no agent to the left of  $x'$  is clustered together with agents to the right of  $x'$ .  $\square$

*Proof (Of Property 3).* Initially, for property 3 we notice the following: At most 1 agent in  $C_{i+1}$  can be clustered together with agents in  $C_i$  in  $\vec{Y}$ . Otherwise, due to the distance between  $C_i$  and  $C_{i+1}$ , clustering  $\vec{Y}$  would be sub-optimal (using the same reasoning as for property 1). Obviously, due to optimality, this agent can only be  $x_{i+1,l}$ .

We now consider the structure of cluster  $C_i$  in relation to agent  $x_{i+1,l}$ . Specifically, by Equation (11) it must be

$$d(x_{i,r}, x_{i+1,l}) > 1.6 \cdot D(C_i), \quad (20)$$

since  $d(x_{i,r}, x_{i+1,l}) = d(C_i, C_{i+1})$ . By looking at the proof of Lemma 1 we see that the smallest possible distance between  $C_i$  and  $C_{i+1}$  is achieved when  $d(x_{i,l}, c_i) = \frac{D(C_i)}{c}$  for  $c = \frac{2\gamma^2}{\gamma^2 + \gamma} \Rightarrow \frac{1}{c} = 0.6$  for  $\gamma = 5$ . This means that since agent  $x_j$  deviates to the left in this case, by at most  $D(C_i)$ , it must be

$$d(x', c_i) \leq 1.6D(C_i), \quad (21)$$

in the edge case. Furthermore, by Observation 2, since  $x_{i+1,l}$  is not served by an agent in  $C_{i+1}$  there is no facility among agents in  $C_{i+1}$  in  $\vec{Y}$ . I.e. all agents in  $C_{i+1} \setminus x_{i+1,l}$  are served by a facility placed on  $[x_{i+2,l}, x_n]$  where  $x_n$  the rightmost agent location in the instance. But, by Lemma 1, if  $x_{i+1,l}$  is served by  $c_{i+1} \in C_{i+1}$  in  $\vec{C}$ ,  $d(C_{i+1}, C_{i+2}) > 1.6D(C_{i+1}) \geq 1.6d(x_{i+1,l}, c_{i+1})$  and so, it is

$$\text{cost}(x_{i+1,o}, \vec{Y}) \geq d(x_{i+1,o}, x_{i+2,l}) \geq 1.6d(x_{i+1,l}, c_{i+1}), \quad (22)$$

for every  $x_{i+1,o} \in C_{i+1} \setminus x_{i+1,l}$ .

Now we are able to show that clustering  $\vec{Y}$  cannot be optimal for instance  $\vec{x}'$  in the edge case. We will compare it with clustering  $\vec{C}$  (where every agent is served by the same facility as in clustering  $\vec{C}$  and  $x'$  is served by  $c_i$ ). We have the following:

$$\text{cost}(\vec{x}' \setminus \{x' \bigcup C_i \bigcup C_{i+1}\}, \vec{Y}) \geq \text{cost}(\vec{x}' \setminus \{x' \bigcup C_i \bigcup C_{i+1}\}, \vec{C}),$$

by optimality. Furthermore,

$$\text{cost}(C_{i+1}, \vec{Y}) \geq \text{cost}(C_{i+1}, \vec{C}) - d(x_{i+1,l}, c_i) + 1.6d(x_{i+1,l}, c_{i+1}) + 1.6D(C_i),$$

by optimality and equations (20) and (22). Also,

$$\text{cost}(C_i \setminus x, \vec{Y}) \geq \text{cost}(C_i \setminus x, \vec{C}),$$

by optimality. Finally,

$$\text{cost}(x', \vec{Y}) + 1.6D(C_i) > \text{cost}(x', \vec{C}),$$

by equation (21).

By adding we get  $\text{cost}(\vec{x}', \vec{Y}) > \text{cost}(\vec{x}', \vec{C})$  which means that  $\vec{Y}$  is sub-optimal for instance  $\vec{x}'$ . All we need to finalize this observation is realize that as we move away from the edge case, the above inequalities become easier to satisfy. Specifically if  $C_i$  had center  $c'_i < c_i$  we see that factor 1.6 of inequality (21) decreases while  $d(C_i, C_{i+1})$  increases. If  $c'_i > c_i$  the same factor of inequality (21) may increase by  $|c'_i - c_i|$ , but then  $d(C_i, C_{i+1})$  increases by at least  $\frac{\gamma^2+1}{\gamma+1} \cdot |c'_i - c_i| > 4.3|c'_i - c_i|$  (since  $d(x_{i,l}, x_{i+1,l}) > \frac{\gamma^2+1}{\gamma+1}d(x_{i,l}, c_i)$  - see proof of Lemma 1), hence maintaining  $\text{cost}(\vec{x}', \vec{Y}) > \text{cost}(\vec{x}', \vec{C})$ .  $\square$

## D Proofs of Auxiliary Lemmas Used in the Proof of Theorem 3

For completeness, we restate the proofs of the auxiliary lemmas with the properties of well-separated instances adapted from [27] and used in the proof of Theorem 3.

Before we proceed with the proofs of the auxiliary lemmas, we need the following basic fact about the facility allocation of any deterministic strategyproof mechanism.

**Lemma 7 (Proposition 2.1, [27]).** *Let  $M$  be a deterministic strategyproof with a bounded approximation ratio of  $\rho \geq 1$  for  $\sqrt{2}$ -stable instances of  $k$ -Facility location on the line. For any  $(k+1)$ -location instance  $\vec{x}$  with  $x_1 \leq x_2 \leq \dots \leq x_{k+1}$ ,  $M_1(\vec{x}) \leq x_2$  and  $M_k(\vec{x}) \geq x_k$ .*

*Proof.* We show it for  $M_1(\vec{x}) \leq x_2$ , the other case is symmetric. Suppose  $x_2 < M_1(\vec{x})$ . Then the agent in  $x_1$  has the incentive to deviate to location  $x_2$ , since  $M_1(\vec{x}_{-1}, x_2) = x_2$  due to the bounded approximation of  $M$  (i.e., in  $(\vec{x}_{-1}, x_2)$ ,  $M$  allocates  $k$  facilities to  $k$  different locations). Notice that  $(\vec{x}_{-1}, x_2)$  is  $\gamma$ -stable for any  $\gamma \geq 1$ .

## D.1 The Proof of Lemma 4

*Proof.* Since  $M$  has a bounded approximation, the isolated pair  $x_k$  and  $x_{k+1}$  must be served by the same facility  $M_k(\vec{x})$ . By Lemma 7, we know that  $M_k(\vec{x}) \geq x_k$ . Then, it must also be  $M_k(\vec{x}) \leq x_{k+1}$ . Otherwise, like in Lemma 7, agent  $x_k$  could declare location  $x_{k+1}$  and decrease her cost, since  $M_k(\vec{x}_{-k}, x_{k+1}) = x_{k+1}$  by the bounded approximation of  $M$ . Again, the instance  $(\vec{x}_{-k}, x_{k+1})$  is arbitrarily stable.

## D.2 The Proof of Lemma 5

We can now proceed to the proofs of the auxiliary lemmas, Lemma 6 and Lemma 5, which refer to the movement of isolated pairs. We only present the proof of Lemma 5 here. The proof of Lemma 6 is fully symmetric.

The proof shown here, refers to 2-Facility Location on well separated instances with 3 agents. All arguments as well as the stability factor of the instance only depend on the well separated property of the rightmost pairs of agents as well as their distance from the third agent from the right. That is, that since in all instances studied in the proof we only change distance between the agents of the isolated, rightmost pair, in the range  $(0, d(x_1, x_2)/r)$  and only increase the distance between the isolated pair and the leftmost agent  $x_1$ , any instance with a large enough distance between  $x_1$  and  $x_2$ , i.e. for which  $d(x_1, x_2) > \gamma \cdot \rho d(x_2, x_3)$  will be  $\gamma$ -stable in all parts of the proof. In that way it is easy to verify that the arguments presented here extend to  $(x_1 | \dots | x_{k-1} | x_k, x_{k+1})$ -well separated and stable instances of at least a specific minimum distance  $d(x_{k-1}, x_k)$ .

Consider  $M$  to be a deterministic, strategyproof, anonymous and bounded approximation mechanism, with approximation ratio of at most  $\rho$ , for 2-facility location. We will work on instance  $\vec{x}$  with three agents  $x_1 < x_2 < x_3$  which is  $(x_1 | x_2, x_3)$ -well separated.

The proof of Lemma 5 directly follows from the following propositions, originally established in [27, Appendix A].

**Proposition 2.** *Consider  $(x_1 | x_2, x_3)$ -well separated, stable instance  $\vec{x}$  for which  $M_2(\vec{x}) = x_2$ . Then for instance  $\vec{x}' = (\vec{x}_{-2}, x'_2)$  where  $x_2 \leq x'_2 \leq x_3$  it will be  $M_2(\vec{x}') = x'_2$*

*Proof.* Notice that since  $d(x'_2, x_3) < d(x_2, x_3)$  instance  $\vec{x}'$  is still  $(x_1 | x_2, x_3)$ -well separated. Furthermore, since  $x_1$  is allocated a facility (by the  $\rho$ -approximation property of the instance),  $\vec{x}'$  is at least as stable as  $\vec{x}$  since the distance between the isolated pair is shortened and their distance from  $x_1$  has grown. All that needs to be shown is that image set  $I_2(\vec{x}_{-2})$  includes the interval  $[x_2, x_3]$ . Since  $x_2$  is allocated a facility, we know  $x_2 \in I_2(\vec{x}_{-2})$ . Furthermore, by the bounded approximation property of  $M$   $x_3 \in I_2(\vec{x}_{-2})$ . Assume there is a hole  $(l, r) \in I_2(\vec{x}_{-2})$  with  $x_2 \leq l < r \leq x_k$ . Consider location  $y \in (l, r)$  such that  $d(y, l) < d(y, r)$ . By strategyproofness  $l \in M(\vec{x}_{-2}, y)$ . But then, by Lemma 7 we have that  $F_2(\vec{x}_{-j}, y) > y$  which contradicts  $M$ 's bounded approximation ratio, since the two agents of the isolated pair of  $(\vec{x}_{-j}, y)$  are served by different facilities.

**Proposition 3.** Consider  $(x_1|x_2, x_3)$ -well separated stable instance  $\vec{x}$  for which  $M_2(\vec{x}) = x_2$ . Then for every  $(x_1|x_2, x'_3)$ -well separated instance  $\vec{x}' = (\vec{x}_{-3}, x'_3)$ , if  $\vec{x}'$  is also well separated,  $M_2(\vec{x}') = x_2$ .

We notice that in that case, the distance between the agents of the isolated pair might grow a from  $\vec{x}$  to  $\vec{x}'$ . Since the proof of this proposition uses instances where the distance of the isolated pair varies from  $\epsilon$  to  $d(x_1, x_2)/\rho$  the proposition stands for stable instances only if all possible  $(x_1|x_2, x'_3)$ -well separated instances  $\vec{x}' = (\vec{x}_{-3}, x'_3)$  are well separated. It is easy to see, that since in all these instances it must be  $d(x_2, x_3) < d(x_1, x_2)/\rho$  then for a large enough distance  $d(x_1, x_2)$  (i.e.  $d(x_1, x_2) > \gamma \cdot \rho d(x_2, x_3)$ )  $\vec{x}'$  is always stable. We show the following proof considering that we have made this assumption.

*Proof.* Since  $M_2(\vec{x}) < x_3$ , we know that  $x_3 \notin I_3(x_{-3})$ . So, there is a  $x_3$ -hole  $(l, r) \in I_3(x_{-3})$ . Since  $M_2(\vec{x}) = x_2$ ,  $l = x_2$  and  $r > 2x_3 - x_2$  (by strategyproofness). By strategyproofness, if  $x'_3 < (r + l)/2$  (for  $x_2 < x'_3$  for well separated instance  $\vec{x}'$ ),  $M_2(\vec{x}') = x_2$ .

To finish, we show that there are no  $(x_1|x_2, x'_3)$ -well separated instances  $\vec{x}' = (\vec{x}_{-3}, x'_3)$  with  $x'_3 \geq (r + l)/2$  and  $M_2(\vec{x}') \neq x_2$ . Again, we reach a contradiction by assuming that there is a point  $y \geq (r + l)/2$  for which  $(x_{-3}, y)$  is a  $(x_1|x_2, y)$ -well separated instance with  $M_2((x_{-3}, y)) \neq x_2$ . If such a  $y$  exists, then there exists  $x'_k \in [(r + l)/2, r)$  for which  $\vec{x}' = (\vec{x}_{-3}, x'_3)$  is a  $(x_1|x_2, x'_3)$ -well separated. But then,  $M_2(\vec{x}') = r > x'_3$  (by strategyproofness, because  $x'_3$  is closer to  $r$  than to  $l$ ). Since  $\vec{x}'$  is  $(x_1|x_2, x'_3)$ -well separated this contradicts lemma 4 which dictates that it must be  $M_2(\vec{x}') = x'_3$ .

**Proposition 4.** Consider  $(x_1|x_2, x_3)$ -well separated stable instance  $\vec{x}$  for which  $M_2(\vec{x}) = x_2$ . Then for every  $(x_1|x'_2, x'_3)$ -well separated instance  $\vec{x}' = (\vec{x}_{-\{2,3\}}, x'_2, x'_3)$ , with  $x_2 < x'_2 < (x_2 + x_3)/2$ , if  $\vec{x}'$  is also well separated,  $M_2(\vec{x}') = x_2$ .

Note that, as for proposition 3 the restriction that  $\vec{x}'$  is also  $\gamma$ -stable is equivalent to  $d(x_1, x_2) > \gamma \cdot \rho d(x_2, x_3)$ .

*Proof.* Since  $x'_2 \in [x_2, x_3]$  we have that  $M_2(x_{-2}, x'_2) = x'_2$ , by proposition 2. But since  $d(x'_2, x_3) < d(x_2, x_3)$ ,  $(x_{-2}, x'_2)$  is  $(x_1|x'_2, x_3)$ -well separated. Hence, by proposition 3, for  $(x_1|x'_2, x'_3)$ -well separated instance  $\vec{x}' = (\vec{x}_{-\{2,3\}}, x'_2, x'_3)$ ,  $M_2(\vec{x}') = x'_2$ .

**Proposition 5.** Consider  $(x_1|x_2, x_3)$ -well separated stable instance  $\vec{x}$  for which  $M_2(\vec{x}) = x_2$ . Then for every  $(x_1|x'_2, x'_3)$ -well separated instance  $\vec{x}' = (\vec{x}_{-\{2,3\}}, x'_2, x'_3)$ , with  $x_2 \leq x'_2$ , if  $\vec{x}'$  is also well separated,  $M_2(\vec{x}') = x_2$ .

*Proof.* We will inductively use proposition 4 to create instance  $\vec{x}'$ . Consider  $d = d(x'_2, x_2)$ ,  $\delta = d(x_3, x_2)/2$  and  $\kappa = \lceil d/\delta \rceil$ . Then for every  $\lambda = 1, 2, 3, \dots, \kappa$  consider instance  $\vec{x}_\lambda = (\vec{x}_{-\{2,3\}}, x_2 + (\lambda - 1)\delta, x_3 + (\lambda - 1)\delta)$ . Now observe that  $\vec{x}_\lambda$  is well separated since for it's rightmost pair,  $x'_2 = x_2 + (\lambda - 1)\delta$  and  $x'_3 = x_3 + (\lambda - 1)\delta$  it is  $d(x'_2, x'_3) > 2\delta$  while  $d(x_1, x'_2) > d(x_1, x_2)$ . By iteratively applying proposition 4 to  $\vec{x}_\lambda$ , we have that for every  $(\vec{x}_{-\{2,3\}}, y_2, y_3)$  well separated instance with  $x_2 + (\lambda - 1)\delta \leq y_2 \leq x_2 + \lambda\delta$ ,  $M_2(\vec{x}_{-\{2,3\}}, y_2, y_3) = y_2$ . For  $\lambda = \kappa$  we get  $M_2(\vec{x}_{-\{2,3\}}, x'_2, x'_3) = x'_2$ .

## E Missing Details from the Proof of Theorem 3: Case 2

Next, we present a detailed proof of Case 2 in the proof of Theorem 3.

**Case 2:**  $x_4 - l \leq \sqrt{2}\lambda$ . Let  $m = (r + l)/2$  be the midpoint of the  $x_3$ -hole  $(l, r)$  in  $I_3(\vec{x}_{-3})$ . We consider the instance  $\vec{y} = (\vec{x}_{-3}, a)$ , where  $a < m$  is arbitrarily close to  $m$  (i.e.,  $a \lesssim m$ ) so that

$a - l < r - a$  and  $d(a, x_4) \lesssim \sqrt{2}\lambda/2$ . The latter is possible since  $x_3$  is already arbitrarily close to  $x_4$  and the right endpoint  $r$  of the hole  $h = (l, r)$  lies in  $(x_3, x_4]$ . Since  $d(x_1, x_2) = \lambda$ ,  $d(x_2, a)$  is quite large, and  $d(a, x_4) \lesssim \sqrt{2}\lambda/2$ , the instance  $\vec{y}$  is  $(\sqrt{2} - \delta)$ -stable, for any  $\delta > 0$ . By strategyproofness,  $M(\vec{y})$  must place a facility at  $l$ , since  $l \in I_3(\vec{x}_{-3})$  and  $l$  is the nearest endpoint of the hole  $h = (l, r)$  to  $a$ .

As before, we now consider the instance  $\vec{y}' = (\vec{y}_{-4}, l)$ . Since  $d(x_1, x_2) = \lambda$ ,  $d(x_2, a)$  is quite large, and  $d(a, l) < d(a, r) \leq \sqrt{2}\lambda/2$ , the instance  $\vec{y}'$  is  $(\sqrt{2} - \delta)$ -stable, for any  $\delta > 0$ . Hence, by strategyproofness,  $M(\vec{y}')$  must keep a facility at  $l$ , because  $l \in I_4(\vec{y}'_{-4})$ .

To conclude the proof, we need to construct a  $(x_1|x_2|l', l' + \varepsilon)$ -well-separated instance  $\vec{z}$  with  $l' \in M(\vec{z})$ . Then, we can reach a contradiction to the hypothesis that  $M$  has a bounded approximation ratio, by applying Lemma 6, similarly to Case 1.

To this end, we consider the image set  $I_4(\vec{y}'_{-4})$  of agent 4 in  $\vec{y}'_{-4} = (x_1, x_2, a)$ . Since  $l \in M(\vec{y}')$ ,  $l \in I_4(\vec{y}'_{-4})$ . If  $a - \varepsilon \in I_4(\vec{y}'_{-4})$ , the instance  $\vec{z} = (\vec{y}'_{-4}, a - \varepsilon)$  is  $(x_1|x_2|a - \varepsilon, a)$ -well-separated (and thus,  $(\sqrt{2} - \delta)$ -stable, for any  $\delta > 0$ ). Moreover, by strategyproofness,  $M(\vec{z})$  must place a facility at  $a - \varepsilon$ , because  $a - \varepsilon \in I_4(\vec{y}'_{-4})$ . Otherwise, there must be a hole  $h' = (l', r')$  in the image set  $I_4(\vec{y}'_{-4})$ , with  $l' > l$  (because  $l \in I_4(\vec{y}'_{-4})$ ) and  $r' < a - \varepsilon$  (because of the hypothesis that  $a - \varepsilon \notin I_4(\vec{y}'_{-4})$ ). We consider the instance  $\vec{z}' = (\vec{y}'_{-4}, l' + \varepsilon) = (x_1, x_2, l' + \varepsilon, a)$ . Since  $l' + \varepsilon \in (l, a)$ ,  $d(a, l' + \varepsilon) < d(a, l) < \sqrt{2}\lambda/2$  and the instance  $\vec{z}'$  is  $(\sqrt{2} - \delta)$ -stable, for any  $\delta > 0$ . Therefore, by strategyproofness and since  $l' \in I_4(\vec{y}'_{-4})$ ,  $M(\vec{z}')$  must place a facility at  $l'$ . We now consider the instance  $\vec{z} = (\vec{z}'_{-3}, l') = (x_1, x_2, l', l' + \varepsilon)$ , which is  $(x_1|x_2|l', l' + \varepsilon)$ -well-separated (and thus,  $(\sqrt{2} - \delta)$ -stable, for any  $\delta > 0$ ). Moreover, by strategyproofness and since  $l' \in M(\vec{z}')$ , and thus,  $l' \in I_3(\vec{z}'_{-3})$ ,  $M(\vec{z})$  must place a facility at  $l'$ .

Therefore, starting from the  $(\sqrt{2} - \delta)$ -stable instance  $\vec{y}'$ , with  $l \in M(\vec{y}')$ , we can construct a  $(x_1|x_2|l', l' + \varepsilon)$ -well-separated instance  $\vec{z}$  with  $l' \in M(\vec{z})$ . Then, by Lemma 6,  $z_4 = l' + \varepsilon \in M(\vec{z})$ , because for the  $(x_1|x_2|x_3, x_4)$ -well-separated instance  $\vec{x}$ ,  $M_3(\vec{x}) = x_4$ , and  $\vec{z}$  is a  $(x_1|x_2|l', l' + \varepsilon)$ -well-separated instance with  $z_4 \leq x_4$ . Since both  $l', l' + \varepsilon \in M(\vec{z})$ , the social cost of  $M(\vec{z})$  is arbitrarily larger than  $\varepsilon$ , which is the optimal social cost of the 3-Facility Location instance  $\vec{z}$ .  $\square$

## F The Proof of Theorem 4

*Proof.* The approximation guarantee easily follows from the fact that since a facility is uniformly at random placed over each optimal cluster, the expected cost of the sum of the cost of the agents in each cluster is 2 times their cost in the optimal clustering.

As is it always with our mechanisms, agent  $x_i \in C_i$  cannot gain by moving within the range of  $C_i$  (this would only increase her utility).

Since the analysis of RANDOM is so similar to the analysis of the mechanism in Section 5, we skip the detailed case analysis and mention only the key deviation cases that need be covered. Specifically these include:

**Case 1:** why agent  $x_i \in C_i$  cannot gain by becoming a member of another cluster,

**Case 2:** or by becoming a self serving center

**Case 3:** or by merging or splitting  $C_i$ .

Without loss of generality, consider the deviating agent to be the edge agent  $x_{i,l} \in C_i$ , declaring location  $x'$  creating instance  $\vec{x}' = (\vec{x}_{-x_{i,l}}, x')$  with optimal clustering  $\vec{C}'$ . If our results stand for her, they easily transfer to all agents in  $C_i$ .  $C_i$  contains  $n$  agents, including  $x_{i,l}$ . For simplicity, without loss of generality we index these agents from left to right, excluding  $x_{i,l}$ , such as  $x_{i,l} \leq x_{i,1} \leq \dots \leq x_{i,n-1}$ , where  $x_{i,1} = x_{i,l+1}$  and  $x_{i,n-1} = x_{i,r}$ . Now for simplicity, we represent  $d(x_{i,l}, x_{i,j})$

by  $d_{i,j}$ . Of course  $d_{i,l} = 0$ . We define as  $X_i$  the discrete random variable that takes values from sample space  $\{d_{i,1}, d_{i,1}, d_{i,2}, \dots, d_{i,n-1}\}$  uniformly at random. That is,  $X_i$  represents the cost agent  $x_{i,l}$  experiences if she is served by the facility placed in  $C_i$  by the mechanism. Then, the expected cost of  $x_{i,l}$  should she not deviate is:

$$\mathbb{E}(X_i) = \frac{0 + d_{i,1} + \dots + d_{i,n-1}}{n}$$

That is, since for any agent  $x_j \notin C_i$ ,  $d(x_j, x_{i,l}) > D(C_i) = d_{i,n-1}$  by Lemma 1.

Now, for Case 1, “why agent  $x \in C_i$  cannot gain by becoming a member of another cluster”. Notice that this is the case where agents in  $C_i$  are not merged or splitted in  $\bar{C}'$ . With some abuse of notation, this allows us to refer to the cluster containing agents in  $C_i \setminus x_{i,l}$  in  $\bar{C}'$  of  $\vec{x}$  as  $C'_i$ .  $C'_{i-1}$  then is the set of agents belonging to the cluster immediately to the left of  $C'_i$  (i.e. the rightmost agent of  $C'_{i-1}$ , excluding  $x'$ , is  $x_{i-1,r}$ ). Consider a deviation  $x'$  that places the deviating agent in cluster  $C'_{i-1}$  after step 1 of the mechanism. Again for simplicity consider  $d(x', x_{i,l}) = c$  and we index agents in  $C_{i-1}$  inversely, such that  $x_{i-1,\hat{1}} \geq x_{i-1,\hat{2}} \geq \dots \geq x_{i-1,\hat{n}'}$  (meaning that now  $x_{i-1,r} = x_{i-1,\hat{1}}$ ,  $x_{i-1,r-1} = x_{i-1,\hat{2}}$  etc.) where  $|C_{i-1}| = n'$ . Equivalently we set  $d(x_{i,l}, x_{i-1,j}) = d_{i-1,j}$ . By Corollary 1, we have  $d_{i,1} \leq d_{i,2} \leq \dots \leq d_{i,n-1} \leq d_{i-1,1} \leq \dots \leq d_{i-1,n'}$ . Now we define uniform random variable  $X'_i$  with sample space  $\{d_{i,1}, \dots, d_{i,n-1}\}$  (see that  $d_{i,l}$  is now absent) and random variable  $X'_{i-1}$  with sample space  $\{c, d_{i-1,1}, \dots, d_{i-1,n'}\}$ . Now  $X'_i$  represents the cost of  $x_{i,l}$  should she be served by the facility placed in  $C'_i$  of the changed instance (which now doesn't include her) and  $X'_{i-1}$  her cost should she be served by the facility placed at  $C'_{i-1}$  (which now includes her false declared location). The expected cost of  $x_{i,l}$  now becomes  $\mathbb{E}(\min\{X'_i, X'_{i-1}\})$ .

But, since  $d_{i,1} \leq d_{i,2} \leq \dots \leq d_{i,n-1} \leq d_{i-1,1} \leq \dots \leq d_{i-1,n'}$ , unless  $d(x', x_{i,l}) < d_{i,n-1} = D(C_i)$ , we have that:

$$\mathbb{E}(\min\{X'_i, X'_{i-1}\}) = \mathbb{E}(X'_i) = \frac{d_{i,1} + \dots + d_{i,n-1}}{n-1} > \mathbb{E}(X_i)$$

That means that  $x_{i,l}$  cannot gain by this deviation unless  $x'$  both belongs in  $C'_{i-1}$  and  $d(x', x_{i,l}) < D(C_i)$ . All we need to show now is that any such situation would result in a violation of the inter-cluster distance between  $C'_{i-1}$  and  $C'_i$  or between  $C'_{i-1}$  and  $C'_{i-2}$ , guaranteed by the cluster-separation property and hence it would be caught by the mechanism's cluster-separation property verification step.

Specifically consider the distance of  $x_{i,l}$  to her center  $c_i$  of  $C_i$  in the optimal clustering. We know that it must be  $d(C_{i-1}, C_i) \geq D(C_i) \cdot 1.6$ , by Lemma 1, for the given stability factor of 5. But in order for this distance to be tight, it must be that  $d(x_{i,l}, c_i) = 0.4 \cdot D(C_i)$  (see factor  $c$  of proof of Lemma 1 -due to stability properties, if  $d(x_{i,l}, c_i) < 0.4 \cdot D(C_i)$  or  $> 0.4 \cdot D(C_i)$ ,  $d(C_{i-1}, C_i)$  grows larger than  $D(C_i) \cdot 1.6$ ). Furthermore, in order for this distance to be tight, it must also be  $d(c_{i-1}, x_{i-1,r}) < 0.4 \cdot D(C_i)$  (since by stability it must be  $d(C_{i-1}, C_i) = d(x_{i-1,r}, x_{i,l}) > (\gamma - 1)d(x_{i-1,r}, c_{i-1})$ ).

Now, since it must be  $d(x', x_{i,l}) < D(C_i)$  it will be  $d(x', c_i) < 1.4D(C_i)$  and  $d(x', x_{i-1,r}) > 0.6D(C_i)$  (since  $d(C_i, C_{i-1}) > 1.6D(C_i)$  by Lemma 1). Finally we distinguish between two cases:

**Case 1:**  $c_{i-1} \in C'_{i-1}$ . Now notice that  $d(x_{i,l}, c_{i-1}) > 5d(x_{i,l}, c_i)$  so  $d(x_{i,l}, c_{i-1}) > 2D(C_i)$ . Then  $D(C'_{i-1}) \geq d(c_{i-1}, x') > D(C_i)$  (since  $d(x', x_{i,l}) < D(C_i)$ ). But then  $d(C'_{i-1}, C'_i) \leq d(x', c_i) \leq 1.4 \cdot D(C_i)$  which means that the cluster separation verification property of step 2 would be violated.

**Case 2:**  $c_{i-1} \notin C'_{i-1}$ . Then, in this edge case we notice it would be  $d(C'_{i-1}, C'_{i-2}) \leq d(c_{i-1}, x_{i-1,r}) \leq 0.4D(C_i)$ . But  $D(C'_{i-1}) \geq d(x_{i-1,r}, x') \geq 0.6D(C_i)$ . Hence the verification property of step 2 is again violated between  $C'_{i-1}$  and  $C'_{i-2}$ .

All we have to do to finish, is note that as  $c_i$  moves to the right or to the left,  $d(C_{i-1}, C_i)$  grows by a multiplicative factor  $\gamma - 1$  ( $=4$ ) of  $d(x_{i,l}, c_i)$  (see proof of Lemma 1) and  $d(x_{i,l}, c_{i-1})$  by a multiplicative factor of 5 (remember, it must be both  $d(x_{i,l}, c_{i-1}) > 5d(x_{i,l}, c_i)$  and  $d(x_{i,r}, c_{i-1}) > 5d(x_{i,r}, c_i)$ ). Which means that the above inequalities will still hold.<sup>9</sup>

For Case 2, why agent  $x_{i,l}$  cannot gain by becoming a self serving cluster, we simply notice the following: her cost, should she not deviate, is at most  $D(C_i)$  (see expected value from previous case). But, from Lemma 3 we know that  $x_{i,l}$  must deviate by at-least  $\geq D(C_i)$ , for a stability factor of 5. So she cannot gain from this deviation<sup>10</sup>.

For Case 3, it is not hard to see that by merging all the agents in  $C_i$  with agents  $\notin C_i$ , her expected cost can only increase. Furthermore, splitting the agents in  $C_i$  would cause the cluster-separation property verification step to identify the split (see the proof of the strategyproofness of the ALMOSTRIGHTMOST mechanism, in Section 5) and remove all agents of  $C_i$  from the game.

---

<sup>9</sup> Notice here that while this property was a must-have for ALMOSTRIGHTMOST to work i.e. the mechanism wouldn't work if  $x'$  both belongs in  $C'_{i-1}$  and  $d(x', x_{i,l}) < D(C_i)$ , here this might not be the case. We can easily see this guarantees strategyproofness, but it might not be necessary which means the mechanism may work for smaller stability factors.

<sup>10</sup> Again, while this property guarantees strategyproofness, it might not be necessary for example, we see that in one of the bad edge cases, where all agents of  $C_i$  are gathered on  $x_{i,r}$ , with  $c_i = x_{i,r}$  a stability of 3 would suffice to guarantee that  $x_{i,l}$  needs to deviate by at-least  $D(C_i)$  to become a self-serving cluster.