MAC Advice for Facility Location Mechanism Design

Zohar Barak
School of Computer Science
Tel Aviv University
zoharbarak@mail.tau.ac.il

Anupam Gupta
New York University and Google Research
anupam.g@nyu.edu

Inbal Talgam-Cohen
School of Computer Science
Tel Aviv University
inbaltalgam@gmail.com

Abstract

Algorithms with predictions are gaining traction across various domains, as a way to surpass traditional worst-case bounds through (machine-learned) advice. We study the canonical problem of k-facility location mechanism design, where the n agents are strategic and might misreport their locations. We receive a prediction for each agent's location, and these predictions are crucially allowed to be only "mostly" and "approximately" correct (MAC for short): a δ-fraction of the predicted locations are allowed to be arbitrarily incorrect, and the remainder of the predictions are required to be correct up to an ε-error. Moreover, we make no assumption on the independence of the errors. Can such "flawed" predictions allow us to beat the current best bounds for strategyproof facility location?

We show how natural robustness of the 1-median (also known as the geometric median) of a set of points leads to an algorithm for single-facility location with MAC predictions. We extend our results to a natural "balanced" variant of the k-facility case, and show that without balancedness, robustness completely breaks down even for $k = 2$ facilities on a line. As our main result, for this "unbalanced" setting we devise a truthful random mechanism, which outperforms the best known mechanism (with no predictions) by Lu et al. [2010]. En route, we introduce the problem of "second" facility location, in which the first facility location is already fixed. Our robustness findings may be of independent interest, as quantitative versions of classic breakdown-point results in robust statistics.

1 Introduction

Algorithms with predictions is a popular field of study in recent years within the paradigm of beyond the worst case analysis of algorithms — see Mitzenmacher and Vassilvitskii (2022) for a comprehensive survey. Motivated by developments in machine learning, this area assumes that algorithms can access predictions regarding the input or solution, thus leveraging predictability in typical computational problems. Recently, Agrawal et al. (2022) and Xu and Lu (2022) proposed to study predictions in the context of mechanism design, where they have the potential to transform existing designs by adding information about agents' private knowledge (see Balkanski et al. (2023a)).

Our focus in this work is on the canonical problem of facility location. The (utilitarian) k-facility location problem is as follows: Consider a multi-set of n points $X \subset \mathbb{R}^d$ and k facility locations; the *cost* of point $x \in X$ is the minimum distance between x and any of the k locations. The goal is to compute k locations that minimize the utilitarian cost, which is the total cost of points in X. In the context of mechanism design, the points X are the privately-known true locations of strategic agents,

38th Conference on Neural Information Processing Systems (NeurIPS 2024).

and the mechanism receives location reports (see Chan et al. (2021) for a recent survey). Facility location mechanism design with predictions has been studied by Agrawal et al. (2022); Xu and Lu (2022).[1] We diverge from previous work by assuming a prediction for each point — that is, the advice X' consists of a prediction x_i' for every $x_i \in X$. Importantly, our measure of prediction error allows for very large errors (outliers) for a δ-fraction of points (hence we call it the "*mostly* approximately correct" model). Our work shows how to get improved bounds despite these large errors, and can thus be of interest beyond facility location.

Standard worst-case prediction error. In the majority of the literature, the goal of (algorithm or mechanism) design with predictions centers around two seemingly-conflicting properties: *robustness* to erroneous predictions, and *consistency* with predictions that are non-erroneous. A fairly common way of achieving both consistency and robustness is interpolating between the approaches of completely following the predictions, and applying the best worst-case algorithm while disregarding the predictions. An ideal design is one that gracefully transitions — as the prediction error increases — from optimal performance when the prediction is correct, to the best-known worst-case performance (Lykouris and Vassilvitskii, 2021; Purohit et al., 2018; Agrawal et al., 2022). By definition, achieving such graceful degradation hinges on how the prediction error is measured.

The most common measure of prediction error is arguably the distance between the predicted and actual values. If $X = \{x_1, \ldots, x_n\}$ contains the actual values and $X' = \{x_1', \ldots, x_n'\}$ is the prediction, the error is defined as $\eta = \ell_p(X - X')$, where ℓ_p represents either the ℓ_1 norm ($\ell_1(t) = \sum_i |t_i|$) or the ℓ_∞ norm ($\ell_\infty(t) = \max_i |t_i|$). This is a "worst case" measure of prediction error, since a single deviation in one of the n entries can significantly inflate the error. Prediction errors measured in this way appear in the context of ski rental Purohit et al. (2018), online caching Lykouris and Vassilvitskii (2021); Rohatgi (2020), the secretary problem Antoniadis et al. (2020), single-item auctions Xu and Lu (2022), and many other problems. In what follows, we address this type of prediction error as the "worst case" error model.

Motivation for an alternative measure. We demonstrate the need for an alternative measure of prediction error combined with new algorithms through the following k-facility location instance, in which a single error causes the optimal solution for prediction X' to perform badly for the actual dataset X.[2]

Example 1 (Sensitivity of 2-facility location on a line)**.** *Let $k = 2$, and assume the n agents' actual locations X are divided evenly between 0 and 1 on the real line. Let X' be the prediction, with all coordinates predicted accurately except for a single coordinate i, which is predicted erroneously to be located at $x_i' = M \gg n$. While the optimal solution for X is to place the $k = 2$ facilities at 0 and 1 for a total cost of zero, in the optimal solution for X' one of the facilities moves to M. This solution performs badly for X — the approximation ratio compared to the optimal solution is unbounded.*

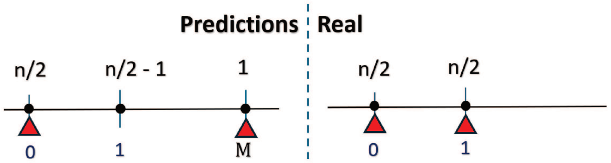

Figure 1: Illustration of Example 1; the predictions X' are on the left side and the real points X are on the right side. The triangles are the optimal facility locations for the points. A single bad prediction translates into an unbounded "worst case" error and approximation ratio.

In many ML contexts, it is very possible that a small fraction of points have a large "worst case" prediction error — even excellent predictors may err on a certain fraction of the dataset. This motivates alternative measures of prediction accuracy. Concretely, we ask:

Question 1. *For 2-facility location on a line (as in Example 1), can we design a mechanism that uses agent reports to get a good solution despite a large "worst case" prediction error?*

We answer this question positively, proving that, using agent reports, it is possible to design resilient mechanisms against adversarial "worst case" errors. The main take-away is that predictions can still

[1] Predictions have also been studied for the online non-strategic version of facility location — see Appendix A.

[2] Similarly, bad performance persists for any algorithm that does not get location reports from the agents.

be valuable despite high "worst case" prediction error, if the fraction of such grossly-erroneous points is small. In other words, a high "worst case" prediction error may not always justify "throwing out" the predictions and settling for the best worst-case guarantee (although it is currently standard in design with predictions). Our new error model and results thus suggest a new pathway for effectively utilizing predictive advice in strategic settings, despite significant uncertainties in the predictions.

2 Overview of Contributions

2.1 Our Model

We introduce MAC predictions where MAC stands for *Mostly Approximately Correct*. This formulation captures the essential features of some closely related previous models, e.g., the model (Azar et al., 2022) formulated for online metric problems, and the PMAC model of (Balcan and Harvey, 2018) formulated in the context of learning submodular functions. A vector of predictions X' is $\text{MAC}(\varepsilon, \delta)$ with respect to dataset X if at least a $(1-\delta)$-fraction (i.e., most) of the predictions are approximately correct up to an additive ε-error. Intuitively, if we do not think it is likely that most of our predictions will be close to being correct —we would typically refrain from relying on them in decision-making processes. Throughout, $\delta \in [0, 0.5)$ is assumed to be small (but can still be a constant). Crucially, when a prediction is not approximately correct (i.e., it belongs to the δ-fraction of incorrect predictions), the error is allowed to be arbitrarily large. Also, the δ-fraction of wrong predictions is allowed to be arbitrary among all possible subsets of that size (there is no assumption of independence). We show that despite their adversarial nature, and the fact that the prediction error is unbounded, such predictions can be used, in combination with strategic reports, to produce a solution with better cost compared to the best no-prediction solution.

The MAC prediction model is suitable for capturing errors from a variety of sources including ML-generated predictions, data that accommodates changes over time, data corrupted by a malicious user, and expert advice (see Appendix B for more details). The accuracy parameters ε, δ of the MAC model can also be viewed as *confidence parameters*. The use of confidence parameters is prevalent in the algorithms with predictions literature (see, e.g., Agrawal et al. (2022)). Arguably, ε, δ are highly *explainable* compared to most confidence measures in the literature, and it is also quite natural to estimate them from data.

2.2 Our Results

We design both deterministic and randomized strategyproof anonymous mechanisms for facility location problems, with MAC predictions of the n agent locations. A preliminary simplifying observation is that in the context of facility location, it suffices to consider mostly-correct (i.e., $\text{MAC}(0, \delta)$) predictions, since handling a nonzero approximation parameter ε follows directly (see Appendix C).

Randomized mechanism design. As our main technical result, for 2*-facility location on a line* we design a randomized mechanism that guarantees an expected approximation ratio of $3.6 + O(\delta)$ (see Algorithm 4 and Theorem 9). For sufficiently small (but still constant) δ, this improves upon the best-known no-prediction expected guarantee of 4 by Lu et al. (2010). (The mechanism of Lu et al. (2010) works for any metric space.) This result provides a positive answer to Question 1, justifying the take-away that predictions with high "worst case" error are useful. We give a brief description of the mechanism in Section 2.3. Note that 2-facility location on a line was also studied by Procaccia and Tennenholtz (2013), and they provided a *deterministic* mechanism with an $(n-2)$ approximation ratio, which is tight (Lu et al., 2010; Fotakis and Tzamos, 2014).

Deterministic mechanism design. For *single-facility location in* $\mathbb{R}^d$, we design a mechanism that guarantees an approximation ratio of $\min\{1 + \frac{4\delta}{1-2\delta}, \sqrt{d}\}$ (see Algorithm 1 and Theorem 5). For sufficiently small (but still constant) δ, this improves upon the no-prediction guarantee of $\sqrt{d}$ by Meir (2019), which is tight for $d = 2$. For β*-balanced* k*-facility location* with constant k (in any metric space), which is a natural extension of facility location where at least a β-fraction of the n points must be assigned to each of the k facilities (for cost-sharing or load-balancing reasons), we design a simple strategyproof mechanism which guarantees an approximation ratio of $c + O(\frac{\delta}{\beta})$, where c is a constant (see Algorithm 7 and Theorem 13). For a sufficiently small (but still constant) δ, this greatly improves upon the best-known no-prediction guarantee of $n/2 - 1$ by Aziz et al. (2020) (we remark that Aziz et al. (2020) study a variant called *capacitated*, to which our result applies).[3]

While we have not optimized the constants, our conceptual contribution is in showing that even if the worst case error of the predictions is unbounded, it can still provide useful information for improving the performance of facility location mechanisms when coupled with strategic reports. In other words, the MAC model allows achieving robustness to outliers while still beating the no-prediction performance. Our single-facility result demonstrates how the accuracy parameter of the model, δ, can serve as a trust parameter which has natural explainability. Another interesting feature of our results is that utilizing MAC predictions seems to suggest a richer family of mechanisms than standard interpolations of "no-predictions" and "complete trust" in the predictions.

Below we give a summary of the results in tables.

Deterministic Mechanism Design		
Problem	**Best known "no predictions" approximation ratio**	**Approximation ratio obtained using MAC predictions**
Single facility location in $\mathbb{R}^d$	$\sqrt{d}$	$1 + O(\delta)$
β-balanced k facilities in $\mathbb{R}^d$	Linear (O(n))*	A constant depending on δ, β

Table 1: Deterministic Mechanism Design.
*The linear appoximation ratio is of Aziz et al. (2020) for the capacitated facility location variant where there is an additional constraint of a maximum cluster size. The β-balanced variant of the problem was not studied without predictions, but is comparable to the capacitated variant, since a minimum cluster size (as in the balanced variant) implies a maximum cluster size (and for $k = 2$, a maximum cluster size implies a minimum cluster size as well).

Random Mechanism Design		
Problem	**Best known "no predictions" approximation ratio**	**Approximation ratio obtained using MAC predictions**
2 facility location on a line	4	$3.6 + O(\delta)$

Table 2: Random Mechanism Design

2.3 Our Techniques and Roadmap

We break the problem into two natural conceptual steps. First, we establish what can be done by just using the predictions – this requires us to show some stability properties of algorithms for k-medians. For some problems, such as the single facility location, we show that using only the predictions is enough to get a good approximation ratio. For other problems, such as 2-facility location, it is not enough, and thus we need to include the input in the computation. For this purpose, we show how to choose the first facility by just using the predictions, and a way to choose the second facility by taking the first facility and the input into account. Then, we show how the interplay between these parts allows us to get a better approximation. Let us now give a few more details for these steps.

Robust statistics and the robustness of the geometric median and k-medians. We first develop quantitative versions of classic results in robust statistics on *breakdown points*, which may be of possible independent interest. Consider a *location estimator*, a function that gets n points in a metric space and returns k locations. Its breakdown point is the proportion δ of adversarially-modified points it can handle before returning an arbitrarily far result from the original estimator. It is well-known that the mean has a breakdown point of $\delta = 1/n$ (consider changing one point to an arbitrarily large value) and that the median is robust with a breakdown point of $\delta = 1/2$. However, the notion of breakdown point is qualitative, and does not reveal the relation between the fraction δ of modified points, and how much the estimator changes. In Section 5 we fill this gap by defining measures of robustness (*distance robustness* (Definition 9) and *approximation robustness* (Definition 10)). Next, we show robustness properties of the median and its generalizations, showing that the cost of the geometric median

[3] β-balanced means every cluster is of size at least βn, and hence at most $n - (k-1)\beta n$.

behaves smoothly as the fraction of modified points increases. These smooth robustness results prove crucial to obtaining improved mechanisms for facility locations using predictions. Moreover, the robustness guarantees have interesting interpretations in the context of robust statistics and robust clustering in machine learning, and might also be of independent interest as explained in Appendix D.

Second facility location. In Section 6.2.1 we introduce the second facility location problem. It is often the case that facilities are created incrementally and not all at once. Hence, it is natural to ask how can we design a mechanism for choosing the location of a second facility (say a second hospital) given an existing first one. We show that the random proportional choice (second step) used in the *proportional mechanism* by Lu et al. (2010) achieves a (tight) approximation ratio of 3 for the line metric, improving over the factor of 4 in the general case for the second facility location problem.

The Robust Half Technique: Beyond interpolation. We generate half of the solution (the first facility location) from the predictions, and use the result together with the input (agent reports) to generate the other half of the solution (the second facility location). To generate the first half of the solution, we define and analyze a robust location estimator we name BIG-CLUSTER-CENTER (which turns out to be a non-trivial task); Hence the name: "Robust Half". To generate the second half of the solution from the input (given the first half) we define and analyze the second facility location on the line problem. Finally, to combine the results we consider two cases: the input either induces a balanced clustering or an unbalanced clustering. We utilize the robustness results (both for balanced k-medians and for Big-Cluster-Center) to show that in both cases we get a good approximation ratio. Our mechanism differs from many existing designs which interpolate between the two approaches of trusting the predictions and using the no-predictions algorithm.

3 Related Work

Facility location mechanism design with predictions. In the “worst case” error model, Agrawal et al. (2022) show a mechanism for *single-facility location* on the real plane $\mathbb{R}^2$. Their prediction consists of a single value representing advice for where to locate the facility. Given a trust parameter $c \in [0, 1]$, their mechanism guarantees an approximation ratio of at most $\frac{\sqrt{2c^2+2}}{1+c}$ if the prediction is perfect, and $\frac{\sqrt{2c^2+2}}{1-c}$ in the worst case. In theory, we could reduce the MAC model to the one studied by Agrawal et al. (2022), by setting their single prediction to be the geometric median of the MAC predictions. Then c can be determined according to δ. However, knowing δ and the δ-robustness of 1-Median, we can directly choose either the no-predictions mechanism or the predictions’ 1-Median, rather than interpolating between these two solutions.

Mechanism design for the 2-*facility location* on the line with predictions is studied by Xu and Lu (2022), where the predictions are for the locations of the agents. While they do not define an error parameter, their deterministic mechanism is in line with the “worst case” prediction error model: it achieves an approximation ratio of $\frac{n}{2}$ if the predictions are perfect, and $2n$ if they are arbitrarily bad.

Two-facility location on a line mechanism design without predictions. In the setting of no-predictions, the best strategyproof deterministic anonymous mechanism achieves a tight approximation ratio of $n - 2$ Procaccia and Tennenholtz (2013); Fotakis and Tzamos (2014). In comparison, in our MAC model under the additional assumption that the number of agents assigned to each facility in the optimal solution is of minimum size $\Omega(\delta n)$, the deterministic mechanism in Algorithm 7 achieves a constant approximation ratio. Under no further assumption, our randomized mechanism in Algorithm 4 achieves a constant expected approximation ratio for small δ.

Additional related work is deferred to Appendix A.

4 Formal Definitions

Let (V, d) be a metric space and $k, n \in \mathbb{N}$. The input to the facility location problem is a multi-set $X = \{x_1, \ldots, x_n\} \in V^n$. We are also given a *prediction*, which is another multi-set $X' = \{x'_1, \ldots, x'_n\} \in V^n$. The distance between a point u and a multi-set W is $d(u, W) = \min_{w \in W} d(u, w)$. The *Hausdorff distance* between multi-sets $U, W \in V^n$ is

$$d_H(U, W) := \max\{\max_{u \in U} d(u, W), \max_{w \in W} d(w, U)\},$$

that is, the maximum distance between a point in one multi-set and the other multi-set.

Definition 1. *Fix $\varepsilon \geq 0$. For $x, x' \in V$, we say x' is an ε-*correct *prediction of x if $d(x, x') \leq \varepsilon$, otherwise we say it is ε-*incorrect.

Definition 2. *The* (ε, r)-neighborhood *of multi-set $X \in V^n$ is defined to be all predictions $X' \in V^n$ such that $|\{i \in [n] \mid x'_i$ is an ε-incorrect prediction of $x_i\}| \leq r$.*

The $(0, r)$-neighborhood of X is often just referred to as the *r-neighborhood* of X.

Definition 3 (Mostly Approximately Correct (MAC) Predictions)**.** *Fix $\varepsilon \geq 0$, $\delta \in [0, 0.5)$. The point set X' is an (ε, δ)-*MAC prediction *for X if X' belongs to the $(\varepsilon, \delta|X|)$-neighborhood of X.*

Definition 4 (Location Estimator)**.** *For $n, k \in \mathbb{N}$, a function $f(X) : V^n \to V^k$ is called a location estimator.*

Some common examples of location estimators with $k = 1$ for points on the real line are the minimum, maximum, mean, and median; for general metric spaces, k-means and k-medians are well-studied examples (k-means is similar to k-medians with squared distances). For $k = 1$ and $V = \mathbb{R}^d$, the k-medians solution is also called the GEOMETRIC-MEDIAN which is a generalization of the median to higher dimensions. A different such median generalization is the COORDINATE-WISE-MEDIAN:

Definition 5. *For a multi-set of n points $X \subseteq \mathbb{R}^d$:* COORDINATE-WISE-MEDIAN(X) $:=$ $(l_1, \ldots, l_d)$ *where where l_j is the median of the multi-set of the j'th coordinates of X for all $j \in [d]$.*

Definition 6 (k-medians cost function)**.** *Given multi-sets $X \in V^n, F \in V^k$, the k-*medians cost function *is*

$$\text{med}_k(X, F) := \sum_{x \in X} d(x, F). \tag{1}$$

In the context of the facility location problem, this function is also known as the social cost function or the utilitarian goal function.

The *k-medians problem* takes as input a multi-set $X \in V^n$ and outputs a multi-set F that minimizes this cost function. The minimizer $F^* \in V^k$ the problem is called the k-Medians *solution.* k-Medians can be viewed as a *location estimator* of X that optimizes the k-medians objective function:

Definition 7 (Center-Induced Partitions)**.** *A collection of centers $F = \{f_1, \ldots, f_k\} \subseteq V^k$* induces a partition $\mathscr{C} = \{C_1, \ldots, C_k\}$ *of the dataset X if $C_i \subseteq \{x \in X \mid d(x, f_i) \leq d(x, f_j) \, \forall j\}$. This partition is called the clustering of X induced by F. This partition (or clustering) is β-*balanced *if $|C_i| \geq \beta|X|$ for all $i \in [k]$.*

Note that a collection $\mathscr{C}$ need not be unique, since points in X can "choose" between any of their closest centers in F.

Definition 8 (β-balanced k-median)**.** *The* β-balanced k-medians problem *takes a dataset $X \in V^n$ as input, and outputs centers $F \in V^k$ along with an induced β-balanced partition $\mathscr{C}$ of X; the goal is to minimize the cost function* $\text{med}_k(X, F) = \sum_{x \in X} d(x, F)$.

This problem is similar to the *lower-bounded k-median* problem where we want to output $F = \{f_1, \ldots, f_k\} \in V^k$ as well as a partition $\{C_1, \ldots, C_k\}$ of X such that $|C_i| \geq \beta|X|$, in order to minimize $\sum_{x \in C_i} d(x, f_i)$ (see Han et al. (2020), Wu et al. (2022)). A significant difference between that definition and ours is that we require the partition to be induced by F (in the sense of Definition 7), which is not required by lower-bounded k-median.

5 δ-Robustness of Location Estimators

In this section we define notions of robustness with respect to changing a δ-fraction of the points (hence referred to as δ-robustness) and analyze the δ-robustness of 1-Median and k-Medians. These results will be very useful for the mechanism design for facility location with predictions. We consider two robustness notions for location estimators: The first notion measures the movement in the solution (as measured by the Hausdorff distance), and the second measures the change in an objective function applied to the solution.

Definition 9 (Distance Robustness)**.** *Let $\rho \geq 0$, $\delta \in [0, 0.5)$. For location estimators $f, \widehat{f} : V^n \to V^k$, we say that $\widehat{f}$ is* (ρ, δ)-distance-robust *with respect to f if for any $X \in V^n$ and any X' in the $\delta|X|$-neighborhood of X,*

$$d_H\Big(f(X), \widehat{f}(X')\Big) \leq \rho.$$

If $\widehat{f} = f$, we say that f is (ρ, δ)-distance-robust.

In the next definition, cost function F is evaluated for the same dataset $X \in V^n$ with respect to two different solutions $f(X)$ and $\widehat{f}(X')$.

Definition 10 (Approximation Robustness)**.** *Let $\gamma \geq 1$, $\delta \in [0, 0.5)$ and let $F : V^n \times V^k \to \mathbb{R}_{\geq 0}$ be a cost measure. For location estimators $f, \widehat{f} : V^n \to V^k$, we say that $\widehat{f}$ is a (γ, δ)-*approximation-robust solution *for cost function F with respect to f if for any $X \in V^n$ and any X' in the $\delta|X|$-neighborhood of X,*

$$F(X, \widehat{f}(X')) \leq \gamma \cdot F(X, f(X)). \tag{2}$$

If $\widehat{f} = f$, we say that f is a (γ, δ)-approximation-robust solution for F. If Eq. (2) *only holds for datasets $X \in \mathscr{Y}$ for some $\mathscr{Y} \subseteq V^n$, we say that f is a (γ, δ)-approximation-robust solution for F restricted to instances $\mathscr{Y}$.*

We give the the following δ-robustness results. For the 1-Median:

Theorem 2. *For $\delta < 1/2$, the* 1-Median *location estimator is $\left(\frac{2}{(1-2\delta)} \cdot \frac{\mathrm{med}_1(X)}{|X|}, \delta\right)$-distance-robust, where $\mathrm{med}_1(X)$ is the optimal cost of the geometric median, that is: $\mathrm{med}_1(X) = \sum_{x_i \in X} d(x_i, \textsc{Geometric-Median}(X))$.*

Corollary 3. *For $\delta < 1/2$, the* 1-Median *estimator is $(1 + \frac{4\delta}{1-2\delta}, \delta)$-approximation-robust for* med_1.

Recall Example 1 from Section 1. In this example, no choice of $k = 2$ centers simultaneously achieves a good (bounded) approximation for the 2-medians problem on both X and X'. Given this negative example, we turn to a *balanced* version of k-medians for which we are able to give a robustness guarantee. The following theorem shows that computing the best "slightly less balanced" solution on the predictions has a cost (on the original dataset X) that is close to the optimal one.

Theorem 4. *Let $b > 2k + 2$. Consider the algorithm $\mathcal{B}$ that computes the optimal $(b-1)\delta$-balanced k-medians solution on its input. For any instance $X \in V^n$ of the $b\delta$-balanced k-medians problem with optimal solution G, let $X' \in V^n$ belong to the δ-neighborhood of X. The algorithm $\mathcal{B}$, when given X', returns a solution H such that*

$$d_H(G, H) \leq \frac{2k}{\delta|X| \cdot (b - 2 - 2k)} \cdot \sum_{x \in X} d(x, G).$$

Moreover, the k-medians cost $\mathrm{med}_k(X, H) \leq (1 + \frac{4k}{b-2-2k}) \mathrm{med}_k(X, G)$, and H induces a $(b-2)\delta$-balanced partition $\mathscr{C}_H$ of X.

For a fixed value of k, this theorem gives a solution H that is $(b-2)\delta$-balanced and also $(1+O(1/b), \delta)$-approximation-robust with respect to the best $b\delta$-balanced solution.

Missing proofs are provided in Appendix E. The intuition behind the proofs in this section is that since most points $(1 - \delta)$ are shared between the actual instances and the predictions, any cost-minimizing solution for the predictions will be close (in terms of Hausdorff distance) to the cost-minimizing solution for the real points, thus resulting in a similar cost as implied by Lemma 10 in Appendix E.

6 Mechanism Design for Facility Location with Mac Predictions

6.1 Warmup: Deterministic Mechanism Design for Facility Location

In this section we show how to design deterministic mechanisms for facility location problems that utilize the Mac predictions to get better approximation guarantees. We do so by utilizing the δ-robustness results from Section 5. In the mechanism design setting, there are n agents with (true) locations $X = \{x_i \mid i \in [n]\} \in \mathbb{R}^d$, where d is the dimension of the facility location problem. The mechanism has access to location reports by the strategic agents, and also to $X' = \{x'_1, \ldots, x'_n\}$, which are the $\textsc{Mac}(\varepsilon, \delta)$ predictions of the true locations X. The mechanism only has access to the reported locations $(\tilde{X})$ and the predictions (X'). For any strategyproof mechanism, the reported locations will be the same as real ones $(X = \tilde{X})$ as illustrated in Fig. 2.

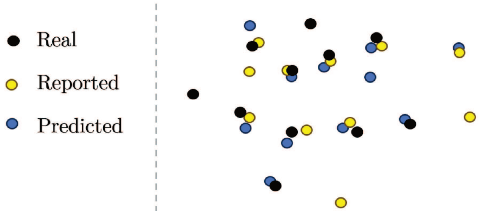

Figure 2: There are three sets of points: The real locations X (black), the reported locations $\tilde{X}$ (yellow), and the predicted locations X'(blue). The algorithm can only access $\tilde{X}, X'$. For any strategyproof mechanism, the yellow and black points overlap.

As explained in Appendix C, for the sake of the analysis we assume that $\varepsilon = 0$. All proofs can be found at Appendix F.

For the single facility location problem $(k = 1)$**:** In this setting we assume δ (or an upper bound value) is known and thus we can simply check in advance if $1 + \frac{4\delta}{1-2\delta} \leq \sqrt{d}$ or not, and use the algorithm that guarantees us the best approximation ratio.

Algorithm 1: Best-Choice-Single-Facility-Loc(X, X', δ)

```
if ($1 + \frac{4\delta}{1-2\delta} \leq \sqrt{d}$)
 then return GEOMETRIC-MEDIAN($X'$)
else return COORDINATE-WISE-MEDIAN($X$);
```

See Appendix F.2 for a clarification on the computation of the geometric median.

Theorem 5. *Algorithm 1 is strategyproof and gets an approximation ratio of min*$\left(1 + \frac{4\delta}{1-2\delta}, \sqrt{d}\right)$.

Supporting the weaker setting of high probability (only w.p. 1 - $o(1/n)$**) MAC predictions**: In Appendix F.3 we show how to achieve the same approximation ratio as in Theorem 5.

For the k **facility location problem:** In a similar way to the single facility, we can also get a deterministic mechanism for the k facility location problem with predictions, under some balancedness condition (see Appendix G). The balancedness condition roughly translates to a minimum cluster size in the optimal solution. The resulting approximation ratio is bounded by $1 + \frac{4k}{b-2-2k}$ where b is a constant that depends on the minimum cluster size and δ.

6.2 Randomized Mechanism Design for 2-Facility Location on a Line

We now consider the 2-facility-location problem, where the metric is given by the real line $\mathbb{R}$. Formally, our metric is over the point set $V = \mathbb{R}$, with $d(x, y) = |x - y|$. There are n agents with locations on the real line $X = \{x_1, \ldots, x_n\}$, and the cost of a solution $G \in V^2$ is $\text{med}_k(X, G) = \sum_{x_i \in X} d(x_i, G)$. In the mechanism-design setting, the agent locations X are unknown, but we are given (a) MAC predictions X' for these locations, as well as (b) the agents' reports for their locations. Unlike in the β-balanced problem variation, we do not assume the balancedness of the optimal solution, but handle the unconstrained case where the optimal solution might induce either a balanced or an unbalanced clustering of X. In this section, we present a strategy-proof radnom mechanism with a small (expected) approximation ratio.

Before doing so, we solve two other problems. After we solve those in Sections 6.2.1 and 6.2.2, we use the solutions in our mechanism for the 2-facility-location on a line problem in Section 6.2.3.

6.2.1 The Second Facility Location Problem

First, we define another independent mechanism design problem (the *second facility location*) for which we show a randomized mechanism. We later use to solve the 2-facility-location problem.

Definition 11 (Second Facility Location)**.** *Given a metric space* (V, d)*, a dataset* $X \in V^n$*, and a single facility* $h_S \in V$*, the goal is to find another facility* h_T *to minimize the* 2*-medians cost* $\text{med}_2(X, \{h_S, h_T\})$.

We consider the following Second-Proportional-Mechanism, which is the second step of the random mechanism for the 2 facility location on the line problem proposed by Lu et al. (2010).

Algorithm 2: Second-Proportional-Mechanism (X, h_S)

```
$a_i \leftarrow d(x_i, h_S)$;
Pick $h \in X$ such that $Pr[h = x_i] = a_i / \sum_{j \in [n]} a_j$;
```

Lemma 6. Second-Proportional-Mechanism *is strategyproof for any metric space.*

Theorem 7 (Second Facility Results)**.** *Let $V = \mathbb{R}$. Fix any dataset X and first facility $\{g_S\}$. For any second facility $g_T^* \in V$; let (S, T) be a partition of the dataset induced by $F^* = \{g_S, g_T^*\}$. The expected cost of* Second-Proportional-Mechanism *given X and g_S is:*

$$\mathbb{E}[\mathrm{med}_2(X, \{g_S, g_T\})] \leq 2\,\mathrm{med}_1(S, \{g_S\}) + 3\,\mathrm{med}_1(T, \{g_T^*\}), \tag{3}$$

where g_T is the second facility location chosen by Second-Proportional-Mechanism.

We defer the proofs of Lemma 6 and Theorem 7 to Appendix I. See Remark 29 for a tight example.

The algorithm prioritizes farther locations (belonging to agents who pay more if their location is not chosen) over agents that are near the first facility. The intuition for the strategyproofness is that if the agent deviates to a different location, it might be chosen with a higher probability, but the agent will also pay more. The largest technical difference in the analysis from Lu et al. (2010) is our introducing of a more delicate analysis for the line metric. We stress, however, that the better approximation ratio of 3 (rather than the ratio of 4 of Lu et al. (2010)) is due to the different nature of the problem (rather then the more delicate analysis) as both algorithms are tight for the line metric. One intuitive explanation for why the similar approach works better for the second facility problem is that the first facility is fixed and so we do not have to pay the cost of "guessing it" as done in Lu et al. (2010).

6.2.2 The δ-Robustness of the Big Cluster Center

Our randomized mechanism uses the predictions to estimate one facility, and randomly chooses the second facility. For this purpose, we come up with an algorithm to "estimate" the first facility location. We formally define what it means to "estimate" one of the two facilities, and then we quantify the distance and approximation robustness of the estimator. This is the last piece of the puzzle that we need for our randomized mechanism. The idea here is that even though Example 1 shows we cannot simply compute the optimal solution, most of the predictions are correct and contain useful information: We can still get a good estimation for one of the two centers of the optimal solution. In this section, we show an algorithm (location estimator), Big-Cluster-Center, that "estimates" of the center of the bigger[4] cluster center out of the two clusters induced by 2-$\mathrm{Medians}$.

Definition 12 (Big-Cluster-Cost)**.** *For any $X = \{x_1, \ldots, x_n\} \in V^n$, let F be its optimal 2-median, and let $\mathscr{C} = \mathscr{C}(X, F)$ be the induced clustering. Let* Big-Cluster(X) *be the cluster in clustering $\mathscr{C}$ with at least $|X|/2$ points (if both clusters have the same size* Big-Cluster(X) *can be the one containing x_1). For $t \in V$, define*

$$\text{Big-Cluster-Cost}(X, t) := \mathrm{med}_1(\text{Big-Cluster}(X), t).$$

Definition 13 (Big Cluster Location Problem)**.** *Given n points $X \in V^n$, the objective is to find a point $t \in V$ to minimize the cost function* Big-Cluster-Cost(X, t).

We specifically focus on the line metric, and only consider $X \in \mathbb{R}^n$ instances where the clustering induced by 2-$\mathrm{Medians}(X)$ are $b\delta$-unbalanced for some (some constants) $b > 1$, $\delta \in [0, 1/2)$.

Algorithm 3: Big-Cluster-Center(X)

```
$(h_L, h_R) = (b-1)\delta$-Balanced 2-Median$(X)$ ;
$L' = \{x_i \in X \mid d(x_i, h_L) \leq d(x_i, h_R)\}$;
$R' = X \setminus L'$;
Return $h_L$ if $|L'| \geq |R'|$ and $h_R$ otherwise;
```

[4] For any clustering induced by any 2 points in space, there is always one bigger cluster (that is, the cluster that contains at least half the points).

Theorem 8. *Let $\mathscr{Y}$ be the collection of all datasets in $\mathbb{R}$ for which the optimal 2-medians induces no $b\delta$-balanced clusterings. Then for a small enough (constant) δ, Algorithm* BIG-CLUSTER-CENTER *is $(1.8 + O(\delta), \delta)$-approximation-robust for the cost function* BIG-CLUSTER-COST *restricted to instances in $\mathscr{Y}$.*

The intuition for the proof of Theorem 8 is that one cluster has a big portion of the points, and most of them also appear in the predictions. Thus, in the cost minimizing solution computed on the predictions, there must be a big center "near" these points (since any solution far away from these points would result in a higher cost). The analysis is close to tight (see Remark 25 in Appendix H).

6.2.3 The 2-Facility Location Algorithm

Our randomized algorithm below first uses the BIG-CLUSTER-CENTER procedure on the predictions to approximate the "big" cluster center, and then uses the SECOND-PROPORTIONAL-MECHANISM to approximate the second cluster center using only the agents' reports. (We set $b \in \mathbb{R}_{>1}$ such that the approximation robustness in Theorem 4 equals 1.2.) The proof is at Appendix J.

```
Algorithm 4: Robust-Half-Mechanism($X, X'$)
$h_1$ = BIG-CLUSTER-CENTER($X', b, \delta$);
$h_2$ = SECOND-PROPORTIONAL-MECHANISM($X, h_1$);
Return $H = (h_1, h_2)$;
```

Theorem 9. *For a small enough (but still constant) δ: Algorithm 4 is strategyproof and has an expected approximation ratio of $3.6 + O(\delta)$.*

The proof uses most of the tools we came up with so far, and it considers two main scenarios: when clustering is $b\delta$-balanced and when it is unbalanced. In both cases, by leveraging the δ-robustness of 1-median, 2-medians, big-cluster-center and the (randomized) strategic placement of the second facility, the mechanism ensures that the solution approaches the optimal cost.

7 Conclusion and Future Directions

We explore strategyproof facility location within our introduced MAC predictions error model. Our model integrates the trust level into the error model, handles outliers, and leads to new mechanism designs (that do not seem to be an interpolation between completely trusting the predictions and resorting to a "no-predictions" method). To analyze our designs, we utilize distance and approximation robustness notions established in Section 5.

An immediate question arising from our results is whether there is a *deterministic* mechanism with MAC predictions for 2-facility location on a line that yields a constant approximation. Another avenue for future research is improving the approximation ratio by randomized mechanisms for this problem. One promising strategy entails refining the selection process for the first facility in Algorithm 4, which can lead to a superior approximation ratio. Our current mechanisms treat the multi-sets of predictions and reported values as independent entities; an intriguing direction for further investigation is to devise mechanisms that capitalize on the matching between predictions and agent-reported values, potentially yielding a more accurate approximation ratio. Lastly, exploring the application of the MAC model to problems beyond mechanism design for facility location may yield insights into the efficacy of such predictions in diverse contexts.

Acknowledgements

We are grateful to Batya Berzack for comments on an earlier draft of the paper, and to Haim Kaplan, Michal Feldman, Amos Fiat and Yossi Azar for discussions about a previous version of the model. We would like to thank the NeurIPS 2024 anonymous reviewers for their comments on this paper. This work is funded by the European Union (ERC grant No. 101077862, project: ALGOCONTRACT, PI: Inbal Talgam-Cohen), a Google Research Scholar award, the Israel Science Foundation (ISF grant No. 3331/24), the Binational Science Foundation (BSF grant No. 2021680), and the National Science Foundation (NSF grant Nos. CCF-1955785 and CCF-2006953).

References

Gagan Aggarwal, Rina Panigrahy, Tomás Feder, Dilys Thomas, Krishnaram Kenthapadi, Samir Khuller, and An Zhu. 2010. Achieving anonymity via clustering. *ACM Transactions on Algorithms (TALG)* 6, 3 (2010), 1–19.

Akanksha Agrawal, Tanmay Inamdar, Saket Saurabh, and Jie Xue. 2023. Clustering what matters: Optimal approximation for clustering with outliers. *Journal of Artificial Intelligence Research* 78 (2023), 143–166.

Priyank Agrawal, Eric Balkanski, Vasilis Gkatzelis, Tingting Ou, and Xizhi Tan. 2022. Learning-augmented mechanism design: Leveraging predictions for facility location. In *Proceedings of the 23rd ACM Conference on Economics and Computation*. 497–528.

Matteo Almanza, Flavio Chierichetti, Silvio Lattanzi, Alessandro Panconesi, and Giuseppe Re. 2021. Online facility location with multiple advice. *Advances in Neural Information Processing Systems* 34 (2021), 4661–4673.

Idan Amir, Idan Attias, Tomer Koren, Yishay Mansour, and Roi Livni. 2020. Prediction with corrupted expert advice. *Advances in Neural Information Processing Systems* 33 (2020), 14315–14325.

Keerti Anand, Rong Ge, Amit Kumar, and Debmalya Panigrahi. 2021. A regression approach to learning-augmented online algorithms. *Advances in Neural Information Processing Systems* 34 (2021), 30504–30517.

Keerti Anand, Rong Ge, Amit Kumar, and Debmalya Panigrahi. 2022. Online algorithms with multiple predictions. In *International Conference on Machine Learning*. PMLR, 582–598.

Keerti Anand, Rong Ge, and Debmalya Panigrahi. 2020. Customizing ML predictions for online algorithms. In *International Conference on Machine Learning*. PMLR, 303–313.

Antonios Antoniadis, Themis Gouleakis, Pieter Kleer, and Pavel Kolev. 2020. Secretary and online matching problems with machine learned advice. *Advances in Neural Information Processing Systems* 33 (2020), 7933–7944.

Yossi Azar, Debmalya Panigrahi, and Noam Touitou. 2022. Online graph algorithms with predictions. In *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. SIAM, 35–66.

Yossi Azar, Debmalya Panigrahi, and Noam Touitou. 2023. Discrete-Smoothness in Online Algorithms with Predictions. In *Thirty-seventh Conference on Neural Information Processing Systems*. https://openreview.net/forum?id=DDmH3H78iJ

Haris Aziz, Hau Chan, Barton Lee, Bo Li, and Toby Walsh. 2020. Facility location problem with capacity constraints: Algorithmic and mechanism design perspectives. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34. 1806–1813.

Maria-Florina Balcan, Siddharth Prasad, and Tuomas Sandholm. 2023. Bicriteria Multidimensional Mechanism Design with Side Information. *CoRR* abs/2302.14234 (2023). https://doi.org/10.48550/ARXIV.2302.14234 arXiv:2302.14234

Maria-Florina Balcan and Nicholas JA Harvey. 2018. Submodular functions: Learnability, structure, and optimization. *SIAM J. Comput.* 47, 3 (2018), 703–754.

Eric Balkanski, Vasilis Gkatzelis, and Xizhi Tan. 2023a. Mechanism Design with Predictions: An Annotated Reading List. *SIGecom Exchanges* 21, 1 (2023), 54–57.

Eric Balkanski, Vasilis Gkatzelis, and Xizhi Tan. 2023b. Strategyproof Scheduling with Predictions. In *14th Innovations in Theoretical Computer Science Conference, ITCS 2023, January 10-13, 2023, MIT, Cambridge, Massachusetts, USA (LIPIcs, Vol. 251)*, Yael Tauman Kalai (Ed.). Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 11:1–11:22. https://doi.org/10.4230/LIPICS.ITCS.2023.11

Eric Balkanski, Vasilis Gkatzelis, Xizhi Tan, and Cherlin Zhu. 2023c. Online Mechanism Design with Predictions. *CoRR* abs/2310.02879 (2023). `https://doi.org/10.48550/ARXIV.2310.02879` arXiv:2310.02879

Amir Beck and Shoham Sabach. 2015. Weiszfeld's method: Old and new results. *Journal of Optimization Theory and Applications* 164 (2015), 1–40.

Ben Berger, Michal Feldman, Vasilis Gkatzelis, and Xizhi Tan. 2023. Optimal Metric Distortion with Predictions. *CoRR* abs/2307.07495 (2023). `https://doi.org/10.48550/ARXIV.2307.07495` arXiv:2307.07495

Giulia Bernardini, Alexander Lindermayr, Alberto Marchetti-Spaccamela, Nicole Megow, Leen Stougie, and Michelle Sweering. 2022. A universal error measure for input predictions applied to online graph problems. *Advances in Neural Information Processing Systems* 35 (2022), 3178–3190.

Anup Bhattacharya, Dishant Goyal, and Ragesh Jaiswal. 2020. Hardness of Approximation of Euclidean k-Median. *arXiv preprint arXiv:2011.04221* (2020).

Hau Chan, Aris Filos-Ratsikas, Bo Li, Minming Li, and Chenhao Wang. 2021. Mechanism Design for Facility Location Problems: A Survey. In *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21*, Zhi-Hua Zhou (Ed.). International Joint Conferences on Artificial Intelligence Organization, 4356–4365. `https://doi.org/10.24963/ijcai.2021/596` Survey Track.

Moses Charikar, Samir Khuller, David M Mount, and Giri Narasimhan. 2001. Algorithms for facility location problems with outliers. In *SODA*, Vol. 1. Citeseer, 642–651.

Michael B Cohen, Yin Tat Lee, Gary Miller, Jakub Pachocki, and Aaron Sidford. 2016. Geometric median in nearly linear time. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*. 9–21.

Ilias Diakonikolas, Vasilis Kontonis, Christos Tzamos, Ali Vakilian, and Nikos Zarifis. 2021. Learning online algorithms with distributional advice. In *International Conference on Machine Learning*. PMLR, 2687–2696.

Paul Dütting, Silvio Lattanzi, Renato Paes Leme, and Sergei Vassilvitskii. 2021. Secretaries with advice. In *Proceedings of the 22nd ACM Conference on Economics and Computation*. 409–429.

Ulrich Eckhardt. 1980. Weber's problem and Weiszfeld's algorithm in general spaces. *Mathematical Programming* 18 (1980), 186–196.

Yuval Emek, Yuval Gil, Maciej Pacut, and Stefan Schmid. 2023. Online Algorithms with Randomly Infused Advice. *arXiv preprint arXiv:2302.05366* (2023).

Dimitris Fotakis, Evangelia Gergatsouli, Themis Gouleakis, and Nikolas Patris. 2021. Learning augmented online facility location. *arXiv preprint arXiv:2107.08277* (2021).

Dimitris Fotakis and Christos Tzamos. 2014. On the power of deterministic mechanisms for facility location games. *ACM Transactions on Economics and Computation (TEAC)* 2, 4 (2014), 1–37.

Vasilis Gkatzelis, Kostas Kollias, Alkmini Sgouritsa, and Xizhi Tan. 2022. Improved Price of Anarchy via Predictions. In *EC '22: The 23rd ACM Conference on Economics and Computation*. 529–557.

Sumit Goel and Wade Hann-Caruthers. 2023. Optimality of the coordinate-wise median mechanism for strategyproof facility location in two dimensions. *Social Choice and Welfare* 61, 1 (2023), 11–34.

Anupam Gupta, Debmalya Panigrahi, Bernardo Subercaseaux, and Kevin Sun. 2022. Augmenting Online Algorithms with epsilon Accurate Predictions. *Advances in neural information processing systems* (2022).

Lu Han, Chunlin Hao, Chenchen Wu, and Zhenning Zhang. 2020. Approximation algorithms for the lower-bounded k-median and its generalizations. In *International Computing and Combinatorics Conference*. Springer, 627–639.

Gabriel Istrate and Cosmin Bonchis. 2022. Mechanism Design With Predictions for Obnoxious Facility Location. *CoRR* abs/2212.09521 (2022). `https://doi.org/10.48550/ARXIV.2212.09521` arXiv:2212.09521

Shaofeng H-C Jiang, Erzhi Liu, You Lyu, Zhihao Gavin Tang, and Yubo Zhang. 2021a. Online facility location with predictions. *arXiv preprint arXiv:2110.08840* (2021).

Zhihao Jiang, Pinyan Lu, Zhihao Gavin Tang, and Yuhao Zhang. 2021b. Online selection problems against constrained adversary. In *International Conference on Machine Learning*. PMLR, 5002–5012.

Misha Khodak, Maria-Florina F Balcan, Ameet Talwalkar, and Sergei Vassilvitskii. 2022. Learning predictions for algorithms with predictions. *Advances in Neural Information Processing Systems* 35 (2022), 3542–3555.

Akshay Krishnamurthy, Thodoris Lykouris, and Chara Podimata. 2020. Corrupted Multidimensional Binary Search: Learning in the Presence of Irrational Agents. *CoRR* abs/2002.11650 (2020). arXiv:2002.11650 `https://arxiv.org/abs/2002.11650`

Ravishankar Krishnaswamy, Shi Li, and Sai Sandeep. 2018. Constant approximation for k-median and k-means with outliers via iterative rounding. In *Proceedings of the 50th annual ACM SIGACT symposium on theory of computing*. 646–659.

Kevin A. Lai, Anup B. Rao, and Santosh Vempala. 2016. Agnostic Estimation of Mean and Covariance. arXiv:1604.06968 [cs.DS]

Thomas Lavastida, Benjamin Moseley, R Ravi, and Chenyang Xu. 2020. Learnable and instance-robust predictions for online matching, flows and load balancing. *arXiv preprint arXiv:2011.11743* (2020).

Hendrik P Lopuhaa and Peter J Rousseeuw. 1991. Breakdown points of affine equivariant estimators of multivariate location and covariance matrices. *The Annals of Statistics* (1991), 229–248.

Pinyan Lu, Xiaorui Sun, Yajun Wang, and Zeyuan Allen Zhu. 2010. Asymptotically optimal strategy-proof mechanisms for two-facility games. In *Proceedings of the 11th ACM conference on Electronic commerce*. 315–324.

Pinyan Lu, Yajun Wang, and Yuan Zhou. 2009. Tighter bounds for facility games. In *Internet and Network Economics: 5th International Workshop, WINE 2009, Rome, Italy, December 14-18, 2009. Proceedings 5*. Springer, 137–148.

Thodoris Lykouris, Vahab S. Mirrokni, and Renato Paes Leme. 2018. Stochastic bandits robust to adversarial corruptions. *CoRR* abs/1803.09353 (2018). arXiv:1803.09353 `http://arxiv.org/abs/1803.09353`

Thodoris Lykouris and Sergei Vassilvitskii. 2021. Competitive caching with machine learned advice. *Journal of the ACM (JACM)* 68, 4 (2021), 1–25.

Nimrod Megiddo and Kenneth J Supowit. 1984. On the complexity of some common geometric location problems. *SIAM journal on computing* 13, 1 (1984), 182–196.

Reshef Meir. 2019. Strategyproof facility location for three agents on a circle. In *International symposium on algorithmic game theory*. Springer, 18–33.

Michael Mitzenmacher and Sergei Vassilvitskii. 2022. Algorithms with predictions. *Commun. ACM* 65, 7 (2022), 33–35. `https://doi.org/10.1145/3528087`

Ariel D Procaccia and Moshe Tennenholtz. 2013. Approximate mechanism design without money. *ACM Transactions on Economics and Computation (TEAC)* 1, 4 (2013), 1–26.

Manish Purohit, Zoya Svitkina, and Ravi Kumar. 2018. Improving online algorithms via ML predictions. *Advances in Neural Information Processing Systems* 31 (2018).

Dhruv Rohatgi. 2020. Near-optimal bounds for online caching with machine learned advice. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*. SIAM, 1834–1845.

Bo Sun, Jerry Huang, Nicolas Christianson, Mohammad Hajiesmaili, and Adam Wierman. 2023. Online Algorithms with Uncertainty-Quantified Predictions. *arXiv preprint arXiv:2310.11558* (2023).

Xiaoliang Wu, Feng Shi, Yutian Guo, Zhen Zhang, Junyu Huang, and Jianxin Wang. 2022. An approximation algorithm for lower-bounded k-median with constant factor. *Science China Information Sciences* 65, 4 (2022), 140601.

Chenyang Xu and Pinyan Lu. 2022. Mechanism Design with Predictions. In *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI-22*, Lud De Raedt (Ed.). International Joint Conferences on Artificial Intelligence Organization, 571–577. `https://doi.org/10.24963/ijcai.2022/81` Main Track.

Chenyang Xu and Benjamin Moseley. 2022. Learning-augmented algorithms for online steiner tree. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 36. 8744–8752.

Emmanouil Zampetakis and Fred Zhang. 2023. Bayesian Strategy-Proof Facility Location via Robust Estimation. In *International Conference on Artificial Intelligence and Statistics*. PMLR, 4196–4208.

Organization of the Appendices

Appendix A contains more related work.

Appendix B contains more examples for sources of MAC predictions.

Appendix C contains the explanation for the effect of $\varepsilon > 0$.

Appendix D contains an interpretation of the established robust statistics δ-robustness results.

Appendix E contains the proofs for Section 5.

Appendix F and Appendix G contain proofs and supplementary material for Section 6.1.

Appendix H, Appendix I and Appendix J contain the missing proofs for section Section 6.2 (for the analysis of Algorithm 4 - "Robust Half" mechanism). Namely:

- Appendix I contains the proof for Algorithm 2 properties: Lemma 6, Theorem 7.
- Appendix H contains the proof of Theorem 8.
- Appendix J contains the proof of Theorem 9.

A Additional Related Work

Models of prediction accuracy. Azar et al. (2022) study online graph algorithms with a notion of error they call *metric error with outliers*, applied to problems where the predictions are locations in a metric space. The error parameters of this model are D and Δ—roughly, D is the minimum cost matching between the predictions and the actual locations, where Δ outliers may be excluded from the matching. These parameters are similar to the (ε, δ) parameters of the MAC model. They then give a general framework for solving online graph problems in their model. Interestingly, their matching cost D captures a notion of average error, whereas we focus on individual error; while our results can be extended to fit within their model, it would be interesting to investigate a version of MAC where ε captures some notion of average error.

Gkatzelis et al. (2022) generalize the outlier model of Azar et al. (2022) and apply it to improve the price of anarchy of cost-sharing mechanisms. Xu and Moseley (2022) study online Steiner tree problems with predictions where the error parameter is the number of erroneous predictions. Bernardini et al. (2022) introduce a model named the "cover error model" that generalizes the model of Azar et al. (2022) to achieve more precision for online problems.

Gupta et al. (2022) study predictions that are incorrect independently with probability δ. Their work focuses on online algorithms rather than mechanism design. Our model is related to their model since independent errors imply MAC predictions w.h.p. by Chernoff. The MAC model is more general in the sense that it does not assume independence of erroneous predictions. Emek et al. (2023) also consider a similar model to the one of Gupta et al. (2022), assuming access to predictions that are correct with some probability or contain random bits otherwise. Dütting et al. (2021) consider a signaling scheme advice model for the secretary, which for binary ML-advice roughly translates into measures of accuracy of the signals (named recall and specificity).

Jiang et al. (2021b) consider a prediction guaranteed to be inside a prediction interval. While their error model is closer to the "worst case" error model, their motivation is confidence intervals from statistics, where it is reasonable to assume that a confidence interval is achieved with probability $1 - \delta$ for some $\delta \in [0, 1]$. Sun et al. (2023) consider an approach for "uncertainty quantified" prediction. One class of predictions they focus on, the "probabilistic interval prediction", can also be seen as a a confidence interval, as in the motivation of Jiang et al. (2021b). Like our MAC model, they "model in" the amount of trust in the advice. Sun et al. (2023) focus on settings where the algorithm receives a single prediction (such as the ski rental problem). They use online learning for tuning the trust parameters, and introduce a regret minimization analysis for learning the trust parameter online (over many instances of the same problem). The MAC model addresses problems that necessitate the quantification of uncertainty for multiple predictions simultaneously within the same input instance, such as in the k-facility location problem.

Mechanism design for facility location problems without predictions.

Mechanism design for the single facility location problem. In the single facility location in $\mathbb{R}^d$ problem, the task is to return a single facility minimizing the social cost function. For the case of no-predictions, Meir (2019) shows a deterministic strategyproof mechanism that gets an approximation ratio of $\sqrt{d}$ which is optimal for $d = 2$ (Goel and Hann-Caruthers (2023)). The optimal deterministic strategyproof mechanism is computing the COORDINATE-WISE-MEDIAN of the points. In short, the $\sqrt{d}$ ratio follows from the fact that COORDINATE-WISE-MEDIAN is the optimal solution is the optimal solution w.r.t the $l1$ norm, which is at most $sqrt(d)$ times $l2$ norm. The goal of introducing predictions is to get something better than $\sqrt{d}$ in this context.

Mechanism design for the two facility location on the line. In this variant the goal is to return two facilities to minimize the social cost, where all points lie on the real line $\mathbb{R}$. We want to find a strategyproof mechanism that given the agent reported locations returns the two locations for facilities in $\mathbb{R}$, s.t. the social cost is minimal. A deterministic $n - 2$ approximation ratio mechanism (called the "Two Extremes" mechanism) was given by Procaccia and Tennenholtz (2013), and a lower bound of $n/2 - 1$ was given by Lu et al. (2010) which was later improved to (the tight) lower bound of $n - 2$ by Fotakis and Tzamos (2014). A randomized strategyproof mechanism with an expected approximation ratio of 4 was given by Lu et al. (2010) (which also works for any metric space), while the currently best known lower bound for random mechanisms (Lu et al. (2009)) is 1.045.

k-facility location. For $k > 2$, Fotakis and Tzamos (2014) show that there is no deterministic anonymous strategyproof mechanism with a bounded approximation ratio for k-facility location on the line for the general case (not just balanced) for any $k \geq 3$, even for simple instances with $k+1$ agents. Moreover, they show that there do not exist any deterministic strategyproof mechanisms with a bounded approximation ratio for 2-Facility Location on more general metric spaces, which is true even for simple instances with 3 agents located in a star.

Mechanism design for the capacitated facility location is a variant of the problem studied by Aziz et al. (2020) where each facility has a maximum capacity, limiting the number of points that can be assigned to it. For the utilitarian cost function, they have shown a $n/2 - 1$ approximation ratio.

Robust k medians and facility location. The robustness of the (offline non-strategic) k medians and k facility location has been studied under different variations. As Example 1 demonstrates, it is not possible to get any bounded approximation ratio for the optimal solution. The approach of Charikar et al. (2001) is to look at different variants of the problems with less restrictive objectives. In one variant they consider, the problem is to place facilities so as to minimize the service cost to any subset of facilities of size at least p for some parameter p. Another variant they consider allows denial of service for some of the clients with the additional cost of some penalty for each such denied client. For work in these kind of models see Krishnaswamy et al. (2018); Agrawal et al. (2023).

Other models of advice/predictions. In the recent literature of algorithms with predictions there are existing other models for algorithms with predictions which are different from the models we already discussed. We now mention some of them.

Some other ML advice is to focus on values that are learned via classical PAC learnability and focus on the learnability and analysis of sample complexity bounds. The assumption is that there is some distribution generating the input for the algorithm. Then the approach is to learn the distribution in terms of classic supervised learning. This kind of modeling is detailed in Anand et al. (2020), Lavastida et al. (2020), Anand et al. (2021).

Diakonikolas et al. (2021) assumes that the advice is not given deterministically, but that they can access the distributions of the inputs and sample from these: they consider access to the instances of interest, and the goal is to learn a competitive algorithm given access to i.i.d. samples. They provide sample complexity bounds for the underlying learning tasks.

Online variants of the facility location problem with different notions of prediction were studied e.g. by Almanza et al. (2021); Fotakis et al. (2021); Jiang et al. (2021a); Azar et al. (2022); Gupta et al. (2022); Azar et al. (2023); Anand et al. (2022). The online setting of the problem differs significantly from our mechanism design setting: the real points arrive in sequence where each time an irrevocable decision must be made by the online algorithm (unlike the mechanism design setting where all the input is given to the mechanism at once, but is reported by strategic agents).

Learning the trust parameter. In the context of online algorithms with prediction, Khodak et al. (2022), show that the confidence parameter can be estimated under certain conditions via online learning. Sun et al. (2023) also have online learning analysis to estimate their different type of confidence parameter.

Other algorithmic game theory with prediction work Istrate and Bonchis (2022) study mechanism design with predictions for obnoxious facility location where agents prefer to be as far as possible from the facilities. Research in algorithmic game theory incorporating predictions, beyond the previously discussed work, has been explored by Balkanski et al. (2023b,c); Berger et al. (2023); Balcan et al. (2023).

Models outside of algorithms with predictions literature. One way to view the MAC model is to view the predictions as data with with corruptions. Designing algorithms to try and handle the corruptions was studied before. Multi-Arm-Bandit with corruptions is such a setting (Lykouris et al. (2018); Krishnamurthy et al. (2020); Amir et al. (2020)).

Zampetakis and Zhang (2023) propose the use of robust statistical estimators for strategyproof mechanism design in a *Bayesian* setting. They show how to use a location estimator which is robust to corrupting δ fraction of the data drawn from some known distribution, to get a strategyproof mechanism for the same location problem. This is conceptually very similar to our result for the single facility case. Since there is no robust estimator for the two-facility problem, their approach cannot apply without making Bayesian assumptions; nonetheless, in our work we show how to use predictions and agent reports to get improvements on the worst-case approximation guarantees.

B MAC Prediction Sources

The MAC prediction model is suitable for capturing errors from a variety of sources, mainly:

- *Machine learning.* In practice, useful ML models produce predictions which are mostly accurate, and the accuracy parameters ϵ, δ can be estimated through testing. For example, we often have a predictor with the guarantee that each prediction (independently) will be ε-correct with probability at least $1 - \delta$. Thus, through concentration bounds such as Chernoff, with high probability the number of predictions which are ε-incorrect is at most $\approx \delta n$. This was the motivation in Balcan and Harvey (2018). Gupta et al. (2022) also consider independent predictions, each correct w.p. δ. Note that unlike this example, the MAC model does not require that the predictor's errors are independent (and in that sense is a generalization, having a weaker assumption).
- *Data with corruptions.* In a setting in which the algorithm's input itself might contain errors, we can treat the corrupted input as a prediction of the true input and apply the MAC model. Corruptions can be a result of adversarial changes to the algorithm's input, e.g., in a malicious attempt to affect the output — in which case the adversary may change only a δ-fraction of the input to avoid getting caught. Corruptions can also capture *outliers* in the data, e.g., as a result of rare but arbitrary measurement errors. This was the motivation in Azar et al. (2022). As another example, the input data may have been collected at a certain point in time, while the ground truth continues to evolve — in the context of facility location, a population census may form the prediction, with errors occurring since a small fraction of the population has relocated since being surveyed.
- *Experts advice.* The data can also originate from expert advice, where experts are usually correct, but can be significantly inaccurate when they are wrong.

C Supporting Small $\varepsilon > 0$ Values

Throughout the analysis we assume that $\varepsilon = 0$, meaning each prediction is either wrong, or exactly correct. As we predict values in space (such as real values) it is hard to expect exact precision in the predictions. Thus, we can weaken this assumption by allowing a prediction to be "correct" if it is ε-accurate for some small $\varepsilon > 0$ value. Consider, for example, the scenario where we want to locate a hospital close to where people live where the locations are the house addresses. An "approximately correct" prediction that equals to the location of the house next door to the actual house of a person would be a reasonable prediction but would result in a small $\varepsilon > 0$. The prediction (and the cost) is almost the same, but it is not exactly right.

Consider, for example, the case of the single facility location problem. The effect of ε on the approximation ratio is bounded by an additive term of at most εn, simply because each of the n points "contributes" another additive ε term. If εn is small in comparison to OPT, or equivalently, if ε is small compared to ${}^{OPT}\!/_{n}$ then the effect on the approximation ratio will be small; also, it simply means that every agent pays at most ε more on average. On the other hand, if ε is big in comparison to ${}^{OPT}\!/_{n}$, then we have no chance to compute a good solution. Indeed, ${}^{OPT}\!/_{n}$ is the *average deviation* of the points from their true geometric median, which can be viewed as a notion of the variance of the input; this is a measure of the "scale" of the cost of the problem, and a noise level above this value swamps the signal in the data.

We drop the ε from the calculations to avoid "dragging" the additive ε-dependent term along each result.

D Robust Statistics and Clustering Results Interpretation

In Section 2.3 we've mentioned how the δ-robustness notion generalizes the breakdown point notion from robust statistics. The existing literature does have a quantitative approach, but mostly for specific distributions or for distributions under some constraints, and mostly for mean and covariance robust estimation. For example, Lai et al. (2016) shows that the distance between the mean of a Gaussian distribution in $\mathbb{R}^d$ from the geometric median is $\Omega(\delta\sqrt{d})$. In section Section 5 we give the definitions of δ-robustness, to consider the robust estimation of general location estimators (such as the mean, median and their generalizations) for data chosen adversarially (thus fitting any distribution, with no further assumption). Then, we use these to study the robustness of the median and its generalizations to adversarial corruptions.

In Appendix E.1, we show that 1-Median (also known as the geometric median) has a distance robustness of $O_\delta(1) \cdot \mathrm{MAD}(X)$ where $O_\delta(1)$ is a small δ-dependent constant, and MAD stands for Mean Average Deviation, a known notion of variance defined as $\mathrm{MAD}(X) := \frac{1}{n}\sum_{x_i \in X} d(x_i, \textsc{Geometric-Median}(X))$. An interpretation of our result is that given a small δ, the robustness of the geometric median depends on the MAD of the points, and the quantification of this relation is given by Theorem 2 and Corollary 3. When we calculate the (geometric) median we often want a single point representing the dataset. The geometric median represents the dataset in the sense that it is the point closest to the dataset points. Intuitively, for data with high MAD (or variance), there is no single point that represents it in a good way. Consider the examples in Fig. 3:

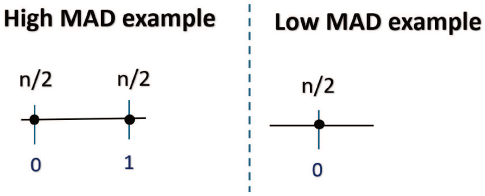

Figure 3: Illustration of instances with different MAD; In the LHS example, MAD is very high unlike RHS where $\mathrm{MAD} = 0$.

On the left example (high MAD example) half the points at 0 and half the points at 1: no single point can represent the data in a good way — any point returned would misrepresent half the data. In the second example (low MAD example) all points are co-located at 0. In this case $\mathrm{MAD} = 0$ and so the geometric median (at 0) represents the dataset in the best possible way — it equals to every point in the dataset. So the interpretation of the results here is that the geometric median is robust in the sense that it still returns a good representation of the dataset: a point that gets a representation of the dataset, which is almost as good as the real "best" representation of the dataset.

In Appendix E.2, we show that for $k > 1$ there is no hope for robustness of k-medians. We thus turn to β-balanced k-medians, which are k medians that induce clusters of size at least βn, where n is the total number of points (see Definition 7). Clustering with the additional constraint of minimum cluster size has applications such as data privacy (Aggarwal et al. (2010)), or clustering with a minimum/maximum cluster size (capacitated variation). The robustness of β-balanced k-medians implies that balanced clustering algorithms can be robust to adversarial corruptions, outliers or missing values.

In our context, we show the usefulness of these δ-robustness results for all of the mechanisms for the $\textsc{Mac}$ predictions settings.

E δ-Robustness of the Median and Its Generalizations: Proofs and More Details

First, we show a lemma connecting the distance robustness to the approximation robustness notions, in the context of k-Medians.

Lemma 10. *Consider location estimators $f, \widehat{f} : V^n \to V^k$ that satisfy the following two properties:*

1. *For any $X \in V^n$ and X' in the $\delta|X|$-neighborhood of X,*
$$\mathrm{med}_k(X', \widehat{f}(X')) \leq \mathrm{med}_k(X', f(X));$$
2. *$\widehat{f}$ is (ρ, δ)-distance-robust with respect to f.*

Then $\widehat{f}$ is a $\left(1 + \frac{2\delta|X|\rho}{\mathrm{med}_k(X, f(X))}, \delta\right)$-approximation-robust solution for $F = \mathrm{med}_k$ with respect to f.

Proof. Let $n = |X|$, let $A := \{i \in [n] \mid x_i = x'_i\}$ be the indices where X, X' are the same, and $B := [n] \setminus A$ be the remaining indices. Define $G := f(X)$ and $H := \widehat{f}(X')$. Using this notation we can write $\mathrm{med}_k(X, \widehat{f}(X'))$ as

$$\begin{aligned}
\sum_i d(x_i, H) &= \sum_i d(x'_i, H) + \sum_{i \in B} (d(x_i, H) - d(x'_i, H)) \\
&\leq \sum_i d(x'_i, G) + \sum_{i \in B} (d(x_i, H) - d(x'_i, H)) \qquad (4) \\
&= \sum_i d(x_i, G) + \sum_{i \in B} (d(x'_i, G) - d(x_i, G)) + \sum_{i \in B} (d(x_i, H) - d(x'_i, H)) \\
&= \mathrm{med}_k(X, f(X)) + \sum_{i \in B} (d(x'_i, G) - d(x'_i, H)) + \sum_{i \in B} (d(x_i, H) - d(x_i, G)), \qquad (5)
\end{aligned}$$

where (4) uses the first assumption of the theorem.

We claim that each of the differences in Eq. (5) is at most $d_H(G, H)$. Indeed, for any $x \in V$, let its closest points in G and H be g and h respectively. Then $d(x, g) - d(x, h) \leq d(g, h)$ by the triangle inequality, and this is at most the Hausdorff distance; the case of $d(x, h) - d(x, g)$ is identical. Therefore,

$$\mathrm{med}_k(X, \widehat{f}(X')) \leq \mathrm{med}_k(X, f(X)) + 2|B| \cdot d_H(G, H).$$

Finally, using the definition of (ρ, δ)-distance-robustness implies that $d_H(G, H)$ is at most ρ. The fact that $|B| \leq \delta|X|$ completes the proof. □

Now, we move onto proving the robustness of 1-Median (which is GEOMETRIC-MEDIAN for a Eucledian space).

E.1 The δ-Robustness of 1-Median

In this section, we quantify the distance robustness (and hence approximation robustness, by Lemma 10) of the 1-median location estimator. We show that changing any δ-fraction of a dataset X where $\delta < 1/2$ can move the 1-Median by only $\frac{2}{1-2\delta}$ times the average cost of a point in the optimal solution of X. Hence, this changes the total 1-Median cost by only a $1 + O(\frac{\delta}{1-2\delta})$ factor.

This is captured by Theorem 2, Corollary 3.

E.1.1 Proof of Theorem 2

Proof. Let $X = \{x_1, \ldots, x_n\}, X' = \{x'_1, \ldots, x'_n\} \in V^n$ where X' is in the $\delta|X|$-neighborhood of X. Let $n = |X|$, let $A := \{i \in [n] \mid x_i = x'_i\}$ be the indices where X, X' are the same, and $B := [n] \setminus A$ be the remaining indices. Let $m := \arg\min_{v \in V} \mathrm{med}_1(X, \{v\})$ and $m' := \arg\min_{v \in V} \mathrm{med}_1(X', \{v\})$. It follows that $\sum_{i \in [n]} d(x'_i, m') \leq \sum_{i \in [n]} d(x'_i, m)$. Using the triangle inequality and the fact that $x'_i = x_i$ for all $i \in A$, we get that

$$\begin{aligned}
\sum_{i \in A} (d(m', m) - d(m, x_i)) + \sum_{i \in B} (d(x'_i, m) - d(m', m)) &\leq \sum_{i \in A} d(x_i, m) + \sum_{i \in B} d(x'_i, m) \\
\implies d(m', m) \cdot (|A| - |B|) \leq 2 \sum_{i \in A} d(x_i, m) &\leq 2 \sum_{i \in [n]} d(x_i, m) = 2\ \mathrm{med}_1(X).
\end{aligned}$$

Since $|B| \leq \delta n$ and $|A| + |B| = n$, we get $d(m', m) \leq \frac{2}{(1-2\delta)n} \mathrm{med}_1(X)$. □

The quantity $\frac{\mathrm{med}_1(X)}{|X|}$ is called the *mean absolute deviation* and can be viewed as a notion of variance. Hence, Theorem 2 essentially says that since $\delta < 1/2$, the 1-Median only changes by $O_\delta(1)$ (a constant depending on δ) times the mean absolute deviation.

To get Corollary 3 we simply substituting the distance-robustness parameter of $\rho = \frac{2}{1-2\delta} \frac{\mathrm{med}_1(X)}{|X|}$ into Lemma 10.

Corollary 3 says that if X' is obtained by perturbing a δ-fraction of the points in X, then the geometric median of X' is a $(1+\frac{4\delta}{1-2\delta})$-approximately optimal solution for the original point set X (with respect to the 1-Median cost objective). This result can be viewed as a quantitative version of a classical result of Lopuhaa and Rousseeuw (1991), which shows that the *breakdown point* for the geometric median is $1/2$. This gives us an understanding not only of the fraction of perturbations required to "break" the estimator (i.e., make it arbitrarily far from the real location), but also of the deterioration in the estimation before breaking it. The quality of estimation as a function of δ is captured by our result of $\rho = \frac{2}{1-2\delta}\frac{\text{med}_1(X)}{|X|}$, showing that a smaller δ corresponds to better estimation.

We conclude the discussion of δ-robustness of the 1-Median with two remarks:

1. Results similar to Theorem 2 hold even for the case where X' is obtained by adding or removing points; in fact, the approximation factors can be improved slightly (the proof is almost the same).
2. The quantitative bounds are tight; consider the example where X comprises of $(0.5-\delta)n+1$ points at 0, and $(0.5+\delta)n-1$ points at 1 and X' is obtained from X by moving δn points from 1 to 0. A calculation shows that in this case the approximation robustness of the geometric median is $(1+\frac{4\delta}{1-2\delta}-\Theta(\frac{1}{n}),\delta)$.

E.2 The δ-Robustness of β-Balanced k-Medians

Proof of Theorem 4:

Proof. The claimed algorithm $\mathcal{B}$ is simple: compute the optimal $(b-1)\delta$-balanced k-medians solution on X'. Since X admits a $b\delta$-balanced solution $G=\{g_1,\ldots,g_k\}$, and X' differs from X in only $\delta|X|$ points, this solution G is $(b-1)\delta$-balanced for X'. Now let $H=\{h_1,\ldots,h_k\}$ be a $(b-1)\delta$-balanced k-medians solution of least cost for X' computed by $\mathcal{B}$: this means

$$\sum_{x'\in X'} d(x',H) \le \sum_{x'\in X'} d(x',G).$$

Let $\mathscr{C}'_H=\{C'_1,\ldots,C'_k\}$ be the balanced partition of X' induced by H. Similarly, since X is in the δ-neighborhood of X', the centers in H induce a $(b-2)\delta$-balanced partition of the original dataset X; call this $\mathscr{C}_H=\{C_1,\ldots,C_k\}$. Finally, let $\mathscr{C}_G=\{C^*_1,\ldots,C^*_k\}$ be the $b\delta$-balanced partition of X induced by G. Recall that $n=|X|$; define

$$OPT := \text{med}_k(X,G) = \sum_{x\in X} d(x,G) = \sum_{i=1}^{k}\sum_{x\in C^*_i} d(x,g_i) \quad \text{and}$$

$$ALG := \text{med}_k(X,H) = \sum_{x\in X} d(x,H) = \sum_{j=1}^{k}\sum_{x\in C_j} d(x,h_j).$$

We now want to prove that $\max_i d(g_i,H)$ and $\max_j d(h_j,G)$ are both small. To show the former, fix any i and let C^*_i be the cluster corresponding to g_i. Since the clustering is $b\delta$-balanced, $|C^*_i|\ge b\delta n$. By averaging, there exists some cluster $C_j\in\mathscr{C}_H$ such that $|C^*_i\cap C_j|\ge b\delta n/k$; choose this j, and consider the corresponding center $h_j\in H$. Now

$$d(g_i,H)\le d(g_i,h_j) \overset{(\star)}{\le} \frac{1}{|C^*_i\cap C_j|}\sum_{x\in C^*_i\cap C_j}\big[d(x,g_i)+d(x,h_j)\big] \le \frac{ALG+OPT}{b\delta n/k},$$

where $(\star)$ uses the triangle inequality. A similar argument shows:

$$d(h_j,G)\le \frac{ALG+OPT}{(b-2)\delta n/k}.$$

Hence the Hausdorff distance between the two solutions is $d_H(G,H)\le\frac{ALG+OPT}{(b-2)\delta n/k}$. Using Lemma 10 with this bound on the Hausdorff distance, we get

$$ALG \le OPT + 2\delta n\cdot d_H(G,H) \le OPT + 2k\cdot\frac{ALG+OPT}{(b-2)}.$$

Simplifying, we get $ALG \leq (1 + \frac{4k}{b-2-2k})OPT$, and that $d_H(G, H) \leq \frac{2k}{b-2-2k} \cdot \frac{OPT}{\delta n}$, as claimed. □

F Proofs and Additional Analysis for Section 6.1

F.1 Proof of Theorem 5

Proof. By Corollary 3, computing the 1-Median on X' results in a $1 + \frac{4\delta}{1-2\delta}$-approximation ratio: this is because the metric space is $\mathbb{R}^d$, and hence the 1-Median is in fact the geometric median (which is also known as the spatial median, L_1 median or Fermat point). Since the geometric median computation depends solely on the predictions it is strategyproof. Thus, by the the strategyproofness of COORDINATE-WISE-MEDIAN, and by the fact that the decision of which mechanism to use does not depend on the reported locations, Algorithm 1 is strategyproof. Due to the $\sqrt{d}$ approximation ratio of COORDINATE-WISE-MEDIAN (Meir (2019)), we get a $\min\Big(1 + \frac{4\delta}{1-2\delta}, \sqrt{d}\Big)$ approximation ratio for Algorithm 1. □

F.2 Remark on Computing the Geometric Median

For general dimensions d and number of points n, there is no known formula or algorithm to find the geometric median exactly. However, since the problem is a convex optimization problem, there are optimization algorithms that find solutions with arbitrary precision efficiently. By using such an optimization algorithm we get an additional factor of $\varepsilon' > 0$ where ε' is arbitrarily small. For more information on efficiently computing the geometric median, see Eckhardt (1980), Beck and Sabach (2015) Cohen et al. (2016).

F.3 Combined Strategy for High Probability MAC(ε, δ)

In this section we show how we can use both the predictions and the reports of the agents to gain good performance even in the case of MAC(ε, δ) only in high probability.

The MAC(ε, δ) model requires that at most a δ fraction of the points have error more than ε; one can extend this to the setting where this requirement holds only with high probability. For instance, consider the setting where the number of prediction errors can be larger than a δ fraction with probability at most $o(1/n)$. In this case the computed geometric median might be arbitrarily far away with this small probability, and hence the expected approximation ratio for our mechanism would be unbounded. To avoid this problem, we can use the agent reports. Intuitively the agents have the incentive to make sure that the returned facility location is not infinitely far away from them.

One way to do this is by using the MIN-BOUNDING-BOX mechanism, introduced by (Agrawal et al., 2022, Mechanism 2) for the egalitarian cost function for $d = 2$; The mechanism naturally extends to any $d \geq 1$ (Algorithm 6). The mechanism calculates the minimum bounding box B containing all the input points; then, given a single prediction o, it returns the point $\hat{o} \in B$ closest to o. We show in Appendix F.4 that the approximation ratio of the generalized MIN-BOUNDING-BOX mechanism for the utilitarian cost function (med_1) is $O(n)$.

Therefore, by composing MIN-BOUNDING-BOX with Algorithm 1 we get a strategyproof mechanism which has an approximation ratio of $min\Big(1 + \frac{4\delta}{1-2\delta}, \sqrt{d}\Big)$ with high probability, and has an $O(n)$ approximation ratio with probability $o(1/n)$; this gives an expected approximation ratio of at most $min\Big(1 + \frac{4\delta}{1-2\delta}, \sqrt{d}\Big) + o(1)$. The form of the resulting mechanism is simple:

```
Algorithm 5: Bounded-Best-Choice-Single-Facility-Loc
Input: $\delta \in [0, \frac{1}{2})$ and $X, X' \subseteq \mathbb{R}^d$ where $X$ is reported by strategic agents and $X'$ contains the
    MAC predictions for $X$
Output: The facility location in $\mathbb{R}^d$
return MIN-BOUNDING-BOX($X$, Best-Choice-Single-Facility-Loc($X, X', \delta$))
```

F.4 Generalized Minimum Bounding Box Mechanism

In this section we properly define the generalized MIN-BOUNDING-BOX mechanism and prove that for points in $\mathbb{R}^d$ (for any $d \geq 1$) it has a tight approximation ratio of $O(n)$ for the med_1 cost function.

For any $y \in \mathbb{R}^d$ let $[y]_j$ denote the j'th coordinate of y.

```
Algorithm 6: MIN-BOUNDING-BOX
Input: $X \subset \mathbb{R}^d$, $o \in R^d$
Output: A facility location in $\mathbb{R}^d$ which is inside the minimum bounding box of $X$
for $j \in [d]$ do
    $o'_j = \text{MINMAXP}(([x_1]_j, \ldots, [x_n]_j), o_j)$
return $o'$
```

Algorithm 6 simply computes, for each coordinate j, the closest point to o_j that is inside the minimum closed interval that contains all of the j'th coordinates of the points of X (which is what MINMAXP does, see Mechanism 1 of Agrawal et al. (2022)). Let us now prove an approximation ratio for the med_1 cost.

Theorem 11. *Algorithm 6 is strategyproof and has a tight $O(n)$ approximation ratio for the* med_1 *cost function.*

Proof. Let us assume, w.l.o.g, that the geometric median g is at $g = 0$. Let B be the minimum bounding box of X. Formally:

$B = \left\{ y \in \mathbb{R}^d \mid \forall j \in [d] : y_j \in \left[\min_{i\in[n]} [x_i]_j, \max_{i\in[n]} [x_i]_j\right] \right\}$. Let h be the point returned by Algorithm 6 for $X \subset \mathbb{R}^d$, $o \in \mathbb{R}^d$ (thus h must be inside of B). Let a_j be the side length of the box B in each coordinate. As usual, we denote OPT to be the cost of the optimal solution and ALG the cost of the algorithm.

$$OPT^2 = \left(\sum_{i\in[n]} \|x_i\|\right)^2 \geq \sum_{i\in[n]} \sum_{j\in d} ([x_i]_j)^2 = \sum_{j\in[d]} \left(\sum_{i\in[n]} ([x_i]_j)^2\right) \overset{(\star)}{\geq} \sum_{j\in[d]} \frac{a_j^2}{4} \implies$$
$$OPT \geq \frac{1}{2}\sqrt{\sum_j a_j^2} \overset{(\star\star)}{\geq} \frac{1}{2}\sqrt{\sum_j h_j^2} = \frac{1}{2}\|h\|. \tag{6}$$

$(\star)$ is due to the fact that for any $j \in [d]$ there is some x_i s.t. $|[x_i]_j| \geq \frac{a_j}{2}$ since otherwise we would get that the j'th side length is smaller than a_j.

$(\star\star)$ explanation: It is a known property of the geometric median that it lies inside the convex hull of the points. Thus, g lies inside B (since the convex hull of the points lies inside the bounding box of the points). So by the fact that $h \in B$ we get that $|h_j| = |h_j - 0| = |h_j - g_j| \leq |a_j|$.

Finally:

$$ALG = \sum_{i\in[n]} \|x_i - h\| \overset{\textit{triangle inequality}}{\leq} \sum_{i\in[n]} \|x_i\| + n\|h\| \overset{Eq.\ (6)}{\leq} OPT + 2n\, OPT = O(n)\, OPT.$$

We deduce that Algorithm 6 has an approximation ratio of at most $O(n)$.

To see that the result is tight, consider the instance where X is the multi-set of n points where $n-1$ points are at $a = (-1, \ldots, -1) \in \mathbb{R}^d$ and one point is at $b = (1, \ldots, 1) \in \mathbb{R}^d$. Let $o = b$. The optimal solution puts the facility at a and has a cost of $OPT = d(a, b) = 2\sqrt{d}$. Algorithm 6 returns b and therefore the cost of the algorithm is $ALG = (n-1) \cdot d(a, b) = (n-1) \cdot 2\sqrt{d}$.

Overall we get: $\frac{ALG}{OPT} = \frac{(n-1)\cdot 2\sqrt{d}}{2\sqrt{d}} = n - 1 = \Omega(n)$.

The strategyproofness of the mechanism is similar to the one proof given by Agrawal et al. (2022) for the egalitarian cost function. □

One could wonder why use the minimum bounding box rather than the convex hull of the points. After all, the convex hull is always contained in the minimum bounding box, and always contains the geometric median. Unfortunately, such a mechanism is not strategyproof:

Remark 12. *The mechanism obtained by replacing the minimum bounding box in Algorithm 6 with the convex hull is not a strategyproof. The example showing this is for $d = 2$: Given $X = \{(-0.5, 0), (0.5, 0), (0, 1)\}$, a computation shows that the agent at $(0, 1)$ can lower its cost by reporting $(1/2, 1)$.*

We can further generalize the mechanism for k facilities by simply using Algorithm 6 for each facility independently.

G A Deterministic Mechanism for Balanced k-Facility Location in General Metric Spaces

For $\beta \in [0, 1]$, the β-balanced k-facility location problem considers n agents with locations $X = \{x_1, \ldots, x_n\} \subseteq V$. Each agent reports a location to the mechanism, and the goal is to return a multi-set $H = (h_1, \ldots, h_k)$ containing k points from V that minimizes $\mathrm{med}_k(X, H)$, such that the clustering induced by H is β-balanced (according to Definition 7).

For a small enough δ, by choosing $b := \beta/\delta$ the balancedness condition translates into a minimum cluster size of $b\delta$. Hence, we can utilize the δ-robustness results of $(b-1)\delta$ -BALANCED k -MEDIAN (Theorem 4) to immediately obtain the deterministic mechanism in Algorithm 7 with the following guarantee:

Theorem 13. *Algorithm 7 is a deterministic strategyproof mechanism with a constant (k-dependent) approximation ratio of at most*

$$1 + \frac{4k}{b - 2 - 2k}.$$

```
Algorithm 7: Balanced-k-Facility-Loc($X, X', b, \delta$)
Input: $X, X' \subseteq V$, and $b > 1$, $\delta \in [0, 0.5)$ where $X'$ are the MAC predictions for $X$
Output: The $k$ facility locations in $V$
return $(b-1)\delta$ -BALANCED $k$ -MEDIAN($X'$)
```

We do not specify how $(b-1)\delta$ -BALANCED k -MEDIAN is implemented. If all points lie on the real line, then there exists an $O(n^k)$-time algorithm for it (computing the minimum cost of all $\binom{n}{k}$ options).

In Euclidean spaces of dimensions greater than one, the k-medians problem (and its balanced or lower bounded variant) is classified as NP-hard Megiddo and Supowit (1984); Bhattacharya et al. (2020) necessitating the use of approximation algorithms. Any approximation algorithm with approximation ratio c can be applied, incurring an additional multiplicative cost factor of c. This approach allows for practical solutions within the constraints of computational complexity, ensuring that we can still achieve near-optimal placements of facilities even in high-dimensional contexts. The effect of ε values and handling the case of high probability MAC predictions is similar to the single facility location results.

H Full Proof of BIG-CLUSTER-CENTER δ-Approximation Robustness

In this section we show that Algorithm 3 indeed has "good" approximation-robustness for unbalanced clusters by providing the full proof for Theorem 8.

Proof. Let $G = (g_L, g_R)$ be the $2 - median$ solution, where $L = \{i \mid d(x_i, g_L) \leq d(x_i, g_R)\}$ and $R = [n] \setminus L$. W.l.o.g $g_L \leq g_R$.

We have a slight abuse of notation throughout for the sake of not introducing more variables, referring to L, R as multisets of points and sometimes as the sets of indices of the points in X.

We assume that the clusters are $b\delta$-unbalanced. That is, at least one of the clusters L, R are of size less than $b\delta n$. W.l.o.g $|R| < b\delta n$ and so $|L| \geq (1 - b\delta)n$.

We also assume that δ is small enough s.t. $b\delta < \frac{1}{4}$ and so $|L| > \frac{3n}{4}$.

Let $H = (h_L, h_R) = (b-1)\delta$-Balanced-2-Medians$(X')$ (as in the first line of Algorithm 3). W.l.o.g $h_L \leq h_R$.

Let L', R' be the multi-set of the elements closest to h_L and h_R the remaining. So $L' = \{i \mid d(x'_i, h_L) \leq d(x'_i, h_R)\}$ and $R' = [n] \setminus L'$.

Let $A := \{i \in [n] \mid x_i = x'_i\}$ the shared points between X, X' (the "correct" predictions), and $B := [n] \setminus A$ the remaining.

Let $m = \frac{g_L + g_R}{2}$ be the middle point between g_L and g_R. Let $m' = \frac{h_L + h_R}{2}$ be the middle point between h_L and h_R. By the definition of m: All points of L lie to the left of m and all points of R lie to the right of m. Similarly, all points of L' lie to the left of m' and all points of R' lie to the right of m'.

We begin by proving the following simple lemma:

Lemma 14. *Let $S \subseteq \mathbb{R}^l$ (for some $l \in \mathbb{N}$) be a multi-set of n points. For any $\beta \in [0,1]$, $S \subseteq V$, let $M = (m_1, \ldots, m_k) := \beta - balanced - k - median(S)$, s.t. $S_1, \ldots, S_k$ are the disjoint clusters induced by $m_1, \ldots, m_k$ of S. Then for all $i \in [k]$: m_i is the $1 - median$ of S_i*

Proof. Consider the implementation of $\beta - balanced - k - median$ which is obtained by taking the cost-minimizing balanced partition into k $\beta - balanced$ clusters. That is, the implementation considers all of the partitions of S into k $\beta - balanced$ clusters and then computes the location that minimizes the center of each cluster S_i. For each such center S_i, the minimizing location is (by definition) $1 - median(S_i)$. Thus, the implementation always returns points that are the $1 - median$ of their clusters. □

We divide the proof into 4 cases as follows (also illustrated in Fig. 4):

Case 1 : $h_L \leq g_L$ and $h_R \geq g_R$.

Case 2 : $h_L \geq g_L$ and $h_R \leq g_R$.

Case 3 : $h_L \geq g_L$ and $h_R \geq g_R$.

Case 4 : $h_L \leq g_L$ and $h_R \leq g_R$.

Cases 1 to 3 are similar and have short proofs. The hard case is Case (4).

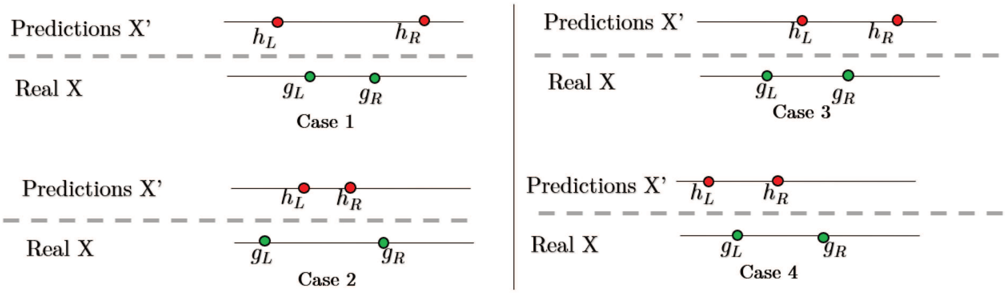

Figure 4: Illustration of the 4 cases. On the top of each case drawing we have the "estimated"/"predicted" locations H, computed on X'. On the bottom we have the "real" locations, G, computed on X.

Case 1: $h_L \leq g_L$ and $h_R \geq g_R$. In this case we can image the two centers moving away from one another. That is, g_L "moves to the left" to h_L, and g_R "moves to the right" to h_R.

If $m' \geq m$ then L' contains all of the points in $L \cap A$. and $|L \cap A| \geq |L| - \delta n \geq n - (b+1)\delta n$. Thus, for a small enough δ: $n - (b+1)\delta n \geq \frac{n}{2}$ and thus the algorithm returns h_L. From Corollary 3 we get an approximation-robustness of $1+O(b\delta) = 1+O(\delta)$ (since L' is obtained from L by change or add of at most $(b+1)\delta$ elements).

Otherwise, $m' \leq m$. The number of elements to the right of g_R in R is $\frac{|R|}{2}$ since it is the median of R (by Lemma 14).

Since $h_R \geq g_R$ then the number of elements to the right of h_R in R' is at most $\frac{|R|}{2} + |B| \leq \frac{|R|}{2} + \delta n$.

h_R is the median of R' (by Lemma 14) and thus $|R'| \leq 2(\frac{|R|}{2} + \delta n) = |R| + 2\delta n \leq (b+2)\delta n$. Thus again by Corollary 3 we get an approximation-robustness of $1 + O(\delta)$.

Case 2: $h_L \geq g_L$ and $h_R \leq g_R$. In this case we can image the two centers moving towards one another. That is, g_L "moves to the right" to h_L, and g_R "moves to the left" to h_R.

If $m' \geq m$ then L' contains all of $L \cap A$ plus at most $(b+1)\delta n$ points and thus just like the previous case we get (from the 1-median approximation robustness) a approximation-robustness of $1 + O(\delta)$.

Otherwise $m' < m$. By Lemma 14 we know that h_L is the median of L' and that g_L is the median of L. Since $h_L \geq g_L$, there are at least $\frac{|L|}{2} - \delta n$ points in L' to the left of h_L. Thus: $|L'| \geq 2(\frac{|L|}{2} - \delta n) \geq |L| - 2\delta n \geq (1-(b+2)\delta)n$, and we can utilize Corollary 3 again to get the desired.

Case 3: $h_L \geq g_L$ and $h_R \geq g_R$. In this case we can image both centers "moving to the right".

In this case, $m' = \frac{h_L + h_R}{2} \geq \frac{g_L + g_R}{2} = m$.

Just like in the previous case: L' contains all of $L \cap A$ plus at most $(b+1)\delta n$ points and thus (from 1-median approximation robustness) a approximation-robustness of $1 + O(\delta)$.

Case 4: $h_L \leq g_L$ and $h_R \leq g_R$. In this case we can image both centers "moving to the left".

By the definition of m': $m' = \frac{h_L + h_R}{2} \leq \frac{g_L + g_R}{2} = m$.

Until now we had an approximation-robustness result of $1 + O(\delta)$. This case is the most difficult case, as for this one we will get a $1.8 + O(\delta)$ approximation result. Since we also have a lower bound of $\approx 1.667 + \Omega(\delta)$ we get that this is indeed "strictly" the hard case.

If $h_R \geq m$ then the number of elements in X to its right is at most $|R|$, and thus the number of elements to its right on X' is at most $|R| + \delta n$. Since h_R median of R' it must be the case that $|R'| \leq 2(|R| + \delta n) \leq 2(b+1)\delta n = O(\delta)\ n$. Thus $|L'| = n - |R'| \geq (1 - O(\delta))n$ and therefore: (a) The algorithm returns h_L (since $|L'| \geq \frac{n}{2}$) and (b) From the $1 - median$ robustness (Corollary 3) we get the required $1 + O(\delta)$ approximation-robustness.

Thus, let us assume that $h_R < m$.

We first handle the case where $m' \geq g_L$, and then we will show a reduction from the case $m' > g_L$ to this one.

H.1 Sub-case: $m' \geq g_L$

Since m' is the middle point between h_L and h_R we get that: $m' - h_L = h_R - m'$ but $m' - h_L = m' - g_L + g_L - h_L \geq g_L - h_L$ and so: $g_L - h_L \leq h_R - m'$. But $h_R - m' \leq h_R - g_L$ and therefore:

$$g_L - h_L \leq h_R - g_L. \tag{7}$$

We introduce the following notations (also see Fig. 5): First, we denote the left and right parts of L', R', L, R:

- Let ${L'}_l = \{i \in L' \mid x'_i \leq h_L\}, {L'}_r = L' \setminus {L'}_l$.
- Let ${R'}_l = \{i \in R' \mid x'_i \leq h_R\}, {R'}_r = R' \setminus {R'}_l$.
- Let $L_l = \{i \in L \mid x_i \leq g_L\}, L_r = L \setminus L_l$.
- Let $R_l = \{i \in R \mid x_i \leq g_R\}, R_r = R \setminus R_l$.

Next, we denote the partition of L into 4 disjoint multi-sets S, T, U, V by the 3 points h_L, m', h_R. So $S = \{i \in L \mid x_i \leq h_L\}$, $T = \{i \in L \mid i \notin S, x_i \leq m'\}$, $U = \{i \in L \mid i \notin S \cup T, x_i \leq h_R\}$, $V = \{i \in L \mid i \notin S \cup T \cup U\}$.

Let $m'' = \frac{g_L + h_R}{2}$ be the middle point between g_L and h_R. W.l.o.g we assume that $m'' > m'$ (otherwise the proof is similar). We define U_l, U_r to be the as follows: $U_l = \{i \in U \mid x_i \leq m''\}$, $U_r = U \setminus U_l$.

Finally, we introduce the notion of α-approximately-equal: $\approx^\alpha$:

Definition 14. *For any $a, b, \alpha \in \mathbb{R}$:*

$$a \approx^\alpha b \iff b - \alpha \leq a \leq b + \alpha$$

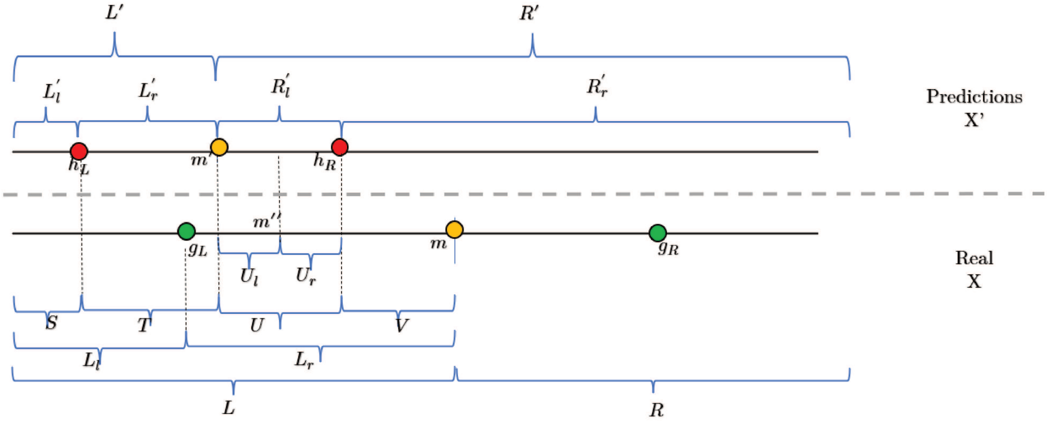

Figure 5: Illustration of case (4) where $m' \geq g_L$. On the top we have h_L, h_R, computed on the "predicted" locations X'. On the bottom we have the g_L, g_R, the 2-Medians of the "real" locations X. L' and R' are the disjoint partitions of X' into two multi-sets of points: those closer to h_L and those closer to h_R (respectively). Similarly L and R are the disjoint multi-sets of g_L, g_R. L'_l, L'_r are the disjoint partitions of L' into two multi-sets: all points to the left of h_L and all points to the right of h_R. In a similar manner R', L, R are partitioned into their left and right parts ($R'_l, R'_r, L_l, L_r, R_l, R_r$). We can also see S, T, U, V which is the disjoint partition of L determined by the points h_L, m', h_R. Finally, we have U_l and U_r which are the left and right parts of U.

For every $M \subseteq X$ we define OPT_M to be the cost that the optimal solution (G) pays for the multi-set of points M. That is: $OPT_M := \text{med}_2(M, G)$ (we sometimes use this notation where $M \subseteq [n]$ in which case OPT_M is a slight abuse of notation for $OPT_{\{x_i | i \in M\}}$).

Let $\beta, \gamma \in [0,1]$ be the ratio between $|S|, |V|$ and $|L|$. That is: $\beta := {}^{|S|}/_{|L|}$ and $\gamma := {}^{|V|}/_{|L|}$.

Claim 15.

$$\text{med}_1(L, h_L) = OPT_S - OPT_{T\cap L_l} + OPT_{L_r} + (g_L - h_L)(|L|(1-2\beta)).$$

Proof.

$$\begin{aligned}
\text{med}_1(L, h_L) &= \sum_{i\in L} |x_i - h_L| = \sum_{i \in S} h_L - x_i + \sum_{i \in T\cup U \cup V} x_i - h_L \\
&\overset{(\star)}{=} \sum_{i\in S} g_L - x_i - |S|(g_L - h_L) - \sum_{i \in T\cap L_l} g_L - x_i + |T \cap L_l|(g_L - h_L) \\
&+ \sum_{i\in (T\cap L_r)\cup U\cup V} x_i - g_L + (|T\cap L_r| + |U| + |V|)(g_L - h_L) \\
&\overset{(\star\star)}{=} OPT_S - OPT_{T\cap L_l} + OPT_{L_r} + (g_L - h_L)(|L| - 2|S|) \\
&= OPT_S - OPT_{T\cap L_l} + OPT_{L_r} + (g_L - h_L)(|L|(1-2\beta)).
\end{aligned}$$

$(\star)$ - triangle inequality

$(\star\star)$ - due to the fact that $L_r = T \cap L_r \cup U \cup V$ and $|L| = |S| + |U| + |V| + |T|$. □

If the algorithm returns h_L then we want to show that $\text{med}_1(L, h_L) \le (1.8 + O(\delta))\,\text{med}_1(L, g_L)$.

From Claim 15 we get the following equivalent condition:

$$\begin{aligned}
&\text{med}_1(L, h_L) \le (1.8 + O(\delta))OPT_L \iff \\
&OPT_S - OPT_{T\cap L_l} + OPT_{L_r} + (g_L - h_L)(|L|(1-2\beta)) \le (1.8 + O(\delta))OPT_L \iff \\
&(g_L - h_L)(|L|(1-2\beta)) \le 0.8OPT_S + 2.8OPT_{T\cap L_l} + 0.8OPT_{L_r} + O(\delta)OPT_L \iff \\
&(g_L - h_L) \le \frac{0.8OPT_S + 2.8OPT_{T\cap L_l} + 0.8OPT_{L_r} + O(\delta)OPT_L}{|L|(1-2\beta)}. \qquad (8)
\end{aligned}$$

Thus, all we have to do is find the above bound for $g_L - h_L$. Indeed we will show that this is the case. We start by proving a few claims that will help us bound $g_L - h_L$.

Claim 16. $(\beta + \gamma)|L| \approx^{O(\delta n)} |L| - \frac{n}{2}$

Proof. Since h_L is the median of L', and the points in $S \cup T$ differ from those in L' by at most $O(\delta n)$, we get that $|T| \approx^{O(\delta n)} \beta|L|$. Similarly, since h_R is the median of R' and since $|R| = n - |L|$: $|U| \approx^{O(\delta n)} (\gamma - 1)|L| + n$. By using the fact that $|L| = |S| + |T| + |U| + |V|$ we get:

$$\begin{aligned}
&|L| \approx^{O(\delta n)} \beta|L| + \beta|L| + (\gamma - 1)|L| + n + \gamma|L| \implies \\
&2(\beta+\gamma)|L| \approx^{O(\delta n)} 2|L| - n \implies \\
&(\beta + \gamma)|L| \approx^{O(\delta n)} |L| - \frac{n}{2}.
\end{aligned}$$

□

Claim 17.

$$(g_L - h_L)|S| \le OPT_S. \qquad (9)$$

Proof.

$$OPT_S = \sum_{i\in S} g_L - x_i \overset{\forall x_i \in S:\, x_i \le h_L}{\ge} \sum_{i \in S} g_L - h_L = |S|(g_L - h_L).$$

□

Claim 18.

$$(h_R - g_L)|U_r| \leq 2OPT_{U_r}. \tag{10}$$

Proof.

$$OPT_{U_r} = \sum_{i \in U_r} x_i - g_L \overset{\forall x_i \in U_r \colon x_i \geq m''}{\geq} \sum_{i \in U_r} m'' - g_L = |U_r|(m'' - g_L) = \frac{|U_r|}{2}(h_R - g_L).$$

where the last equality is due to the fact that m'' is the middle point between g_L and h_R.

Thus we get:

$(h_R - g_L)|U_r| \leq 2OPT_{U_r}$ □

Claim 19.

$$(h_R - g_L)|V| \leq OPT_V. \tag{11}$$

Proof.

$$OPT_V = \sum_{i \in V} x_i - g_L \overset{\forall x_i \in V \colon x_i \geq h_R}{\geq} \sum_{i \in S} h_R - g_L = |V|(h_R - g_L).$$

□

Consider the clustering induced by $H' = (g_L, h_R)$. If this clustering is unbalanced then Algorithm 3 returns h_L since $L_l \cap A$ is contained in the left cluster and since $|L_l \cap A| \geq \frac{L}{2} - \delta n \geq \frac{n}{2} - O(\delta n)$ it must be that the right cluster is the smaller one. And so the left cluster has at least $n-(b-1)\delta n = n - O(\delta n)$ elements and thus it is obtained from L by modifying or dropping at mots $O(\delta n)$ elements and thus the algorithm returns h_L and like before we get a $(1 + O(\delta), \delta)$ approximation-robustness.

Otherwise, the clustering induced by H' is balanced, and we get the following claim:

Claim 20. $(h_R - g_L)(|U_l| - O(\delta n)) \leq 2(OPT_{T \cap L_l} + OPT_{U_l})$.

Proof. For convenience, for any multi-set of the real points $M \subseteq X$ we denote $\hat{M} \subseteq X'$ to be the multi-set of the estimated points X' that correspond to the same partition of the real line as M. So we define the following multi-sets: $\hat{L_l} = \{x'_i \in X' \mid x'_i \leq g_L\}$, $\hat{L_r} = \{x'_i \in X' \setminus \hat{L_l} \mid x'_i \leq m\}$, $\hat{S} = L'_l$, $\hat{T} = L'_r$, $\hat{U} = R'_l$, $\hat{V} = \{i \in R'_r \mid x'_i \in [h_R, m]\}$, $\hat{U_l} = \{x'_i \in R'_l \mid x'_i \leq m''\}$, $\hat{U_r} = \{x'_i \in R'_l \mid x'_i > m''\}$. The reason we use this notation is that we have an inequality in terms of the points in X' and we want to later move on to the inequality in terms of the points in X. This notation will help us see the connection between the points in X' and the points in X. Each such $M, \hat{M}$ contain the same points up to at most δn points.

By definition of H we know that $\mathrm{med}_2(X', H) \leq \mathrm{med}_2(X', T)$ for any $T = (t_1, t_2) \in \mathbb{R}^2$ that induces a $(b-1)\delta$-balanced clustering of X'. Therefore, $\mathrm{med}_2(X', H) \leq \mathrm{med}_2(X', (g_L, h_R))$.

By the definition of med_2:

$$\begin{gathered}\sum_{x'_i \in L'_l} h_L - x'_i + \sum_{x'_i \in L'_r} x'_i - h_L + \sum_{x'_i \in R'_l} h_R - x'_i + \sum_{x'_i \in R'_r} x'_i - h_R \leq \\ \sum_{x'_i \in \hat{S} \cup (\hat{T} \cap \hat{L_l})} g_L - x'_i + \sum_{x'_i \in (\hat{T} \cap \hat{L_r}) \cup \hat{U_l}} x'_i - g_L + \sum_{x'_i \in \hat{U_r}} h_R - x'_i + \sum_{x'_i \in R'_r} x'_i - h_R.\end{gathered}$$

Which implies:

$$\begin{aligned}&\sum_{x'_i \in L_{h_l}} h_L - x'_i + \sum_{x'_i \in L'_r} x'_i - h_L + \sum_{x'_i \in R'_l} h_R - x'_i \\ &\leq \sum_{x'_i \in \hat{S} \cup (\hat{T} \cap \hat{L_l})} g_L - x'_i + \sum_{x'_i \in (\hat{T} \cap \hat{L_r}) \cup \hat{U_l}} x'_i - g_L + \sum_{x'_i \in \hat{U_r}} h_R - x'_i.\end{aligned} \tag{12}$$

Let us observe LHS:

$$
\begin{aligned}
&\sum_{x'_i \in L_{h_l}} h_L - x'_i + \sum_{x'_i \in {L'}_r} x'_i - h_L + \sum_{x'_i \in {R'}_l} h_R - x'_i \\
&= |{L'}_l|(h_L - g_L) + \sum_{x'_i \in L_{h_l}} g_L - x'_i + \sum_{x'_i \in {L'}_r} x'_i - g_L + |{L'}_r|(g_L - h_L) + |{R'}_l|(h_R - g_L) + \sum_{x'_i \in {R'}_l} g_L - x'_i \\
&\overset{|{L'}_l| = |{L'}_r|}{=} \sum_{x'_i \in L_{h_l}} g_L - x'_i + \sum_{x'_i \in {L'}_r} x'_i - g_L + |{R'}_l|(h_R - g_L) + \sum_{x'_i \in {R'}_l} g_L - x'_i \\
&= \sum_{x'_i \in \hat{S}} g_L - x'_i + \sum_{x'_i \in \hat{T}} x'_i - g_L + |\hat{U}|(h_R - g_L) + \sum_{x'_i \in \hat{U}} g_L - x'_i \\
&= \sum_{x'_i \in \hat{S}} g_L - x'_i + \sum_{x'_i \in \hat{T} \cap \hat{L_l}} x'_i - g_L + \sum_{x'_i \in \hat{T} \cap \hat{L_r}} x'_i - g_L + |\hat{U}|(h_R - g_L) + \sum_{x'_i \in \hat{U}} g_L - x'_i.
\end{aligned}
$$

So by plugging this in (12) we get:

$$
\begin{aligned}
&\sum_{x'_i \in \hat{S}} g_L - x'_i + \sum_{x'_i \in \hat{T} \cap \hat{L_l}} x'_i - g_L + \sum_{x'_i \in \hat{T} \cap \hat{L_r}} x'_i - g_L + |\hat{U}|(h_R - g_L) + \sum_{x'_i \in \hat{U}} g_L - x'_i \\
&\leq \sum_{x'_i \in \hat{S} \cup (\hat{T} \cap \hat{L_l})} g_L - x'_i + \sum_{x'_i \in (\hat{T} \cap \hat{L_r}) \cup \hat{U_l}} x'_i - g_L + \sum_{x'_i \in \hat{U_r}} h_R - x'_i. \\
&\implies \\
&|\hat{U}|(h_R - g_L) + \sum_{x'_i \in \hat{U}} g_L - x'_i \leq 2 \sum_{x'_i \in \hat{T} \cap \hat{L_l}} g_L - x'_i + \sum_{x'_i \in \hat{U_l}} x'_i - g_L + \sum_{x'_i \in \hat{U_r}} h_R - x'_i \\
&= 2 \sum_{x'_i \in \hat{T} \cap \hat{L_l}} g_L - x'_i + \sum_{x'_i \in \hat{U_l}} x'_i - g_L + |\hat{U_r}|(h_R - g_L) + \sum_{x'_i \in \hat{U_r}} g_L - x'_i \implies \\
&|\hat{U_l}|(h_R - g_L) \leq 2 \sum_{x'_i \in \hat{T} \cap \hat{L_l}} g_L - x'_i + 2 \sum_{x'_i \in \hat{U_l}} x'_i - g_L. \qquad (13)
\end{aligned}
$$

By the definition of $\hat{T}, \hat{L_l}$: $\hat{T} \cap \hat{L_l}$ and $T \cap L_l$ differ by at most δn elements, and for any $x'_i \in \hat{T} \cap \hat{L_l}$: $x'_i \geq h_L$. Similarly $\hat{U_l}$ and U_l differ by at most δn elements and for any $x'_i \in \hat{U_l}$: $x'_i \leq m''$.
By plugging this in the above (13) we get:

$$
\begin{aligned}
(|U_l| - \delta n)(h_R - g_L) &\leq |\hat{U_l}|(h_R - g_L) \\
&\leq 2 \sum_{x_i \in T \cap L_l} g_L - x_i + 2\delta n(g_L - h_L) + 2 \sum_{x_i \in U_l} x_i - g_L + 2\delta n(m'' - g_L).
\end{aligned}
$$

Which implies:

$$
(h_R - g_L)|U_l| \leq 2(OPT_{T \cap L_l} + OPT_{U_l}) + \delta n(2(g_L - h_L) + 2(m'' - g_L) + (h_R - g_L)).
$$

Since $m'' - g_L = {}^{(h_R - g_L)}/_2$ (by the definition of m'') and from Eq. (7) we get:

$$
(h_R - g_L)|U_l| \leq 2(OPT_{T \cap L_l} + OPT_{U_l}) + O(\delta n)(h_R - g_L).
$$

$$
(h_R - g_L)(|U_l| - O(\delta n)) \leq 2(OPT_{T \cap L_l} + OPT_{U_l}). \qquad (14)
$$

□

Now we are ready to bound $g_L - h_L$.

Lemma 21. *If the algorithm returns h_L then:*

$$(g_L - h_L) \leq \frac{0.8OPT_S + 2.8OPT_{T\cap L_l} + 0.8OPT_{L_r} + O(\delta)OPT_L}{|L|(1-2\beta)}.$$

Proof. By Claim 17, Claim 18, Claim 20, Claim 19 and the fact that $g_L - h_L \leq h_R - g_L$ we get the following four inequalities:

$$\begin{aligned}(g_L - h_L)|U_r| &\leq 2OPT_{U_r},\\ (g_L - h_L)(|U_l| - O(\delta n)) &\leq 2(OPT_{T\cap L_l} + OPT_{U_l}),\\ (g_L - h_L)2|V| &\leq 2OPT_V,\\ (g_L - h_L)2|S| &\leq 2OPT_S,\end{aligned}$$

By summing the above inequalities:

$$\begin{aligned}&(g_L - h_L)(2|S| + |U| + 2|V| - O(\delta n)) \leq 2(OPT_S + OPT_{T\cap L_l} + OPT_U + OPT_V) \implies\\ &g_L - h_L \leq \frac{(OPT_S + OPT_{T\cap L_l} + OPT_U + OPT_V)}{(1-2\beta)|L|} \cdot \frac{2(1-2\beta)|L|}{2|S| + |U| + 2|V| - O(\delta n)}. \qquad (15)\end{aligned}$$

Since $|L'_l| = |L'_r|$ we get that $|T| \approx^{\delta n} |S| = \beta|L|$ and so

$$\begin{aligned}&2|S| + |U| + 2|V| - O(\delta n)\\ &\qquad \geq |S| + |T| + |U| + 2|V| - O(\delta n)\\ &\qquad = (|S| + |T| + |U| + |V|) + (|S| + |V|) - |S| - O(\delta n)\\ &\qquad \overset{(\star)}{=} |L| + (\beta+\gamma)|L| - \beta|L| - O(\delta n)\\ &\qquad \overset{Claim\ 16}{\geq} |L| + |L| - \frac{n}{2} - \beta|L| - O(\delta n)\\ &\qquad = 2|L| - \beta|L| - \frac{n}{2} - O(\delta n).\end{aligned}$$

Where $(\star)$ is due to: $|L| = |S| + |T| + |U| + |V|$, $|V| = \gamma|L|$, $|S| = \beta|L|$.

By plugging this inequality into Eq. (15):

$$g_L - h_L \leq \frac{(OPT_S + OPT_{T\cap L_l} + OPT_U + OPT_V)}{(1-2\beta)|L|} \cdot \left(\frac{2(1-2\beta)|L|}{2|L| - \beta|L| - \frac{n}{2} - O(\delta n)}\right). \qquad (16)$$

We will bound the second term in RHS. First, we know that $|L| \approx^{O(\delta n)} n$ and thus:

$$\frac{2(1-2\beta)|L|}{2|L| - \beta|L| - \frac{n}{2} - O(\delta n)} \leq \frac{2(1-2\beta)(n - O(\delta n))}{\frac{3n}{2} - \beta n - O(\delta n)} \leq \frac{2(1-2\beta)}{\frac{3}{2} - \beta}(1 + O(\delta))$$

We denote: $f(\beta) := \frac{2(1-2\beta)}{\frac{3}{2}-\beta}$.

Since the algorithm returns h_L we deduce that $|L'| \geq \frac{n}{2}$. Thus, $|L'_l|, |L'_r| \geq \frac{n}{4}$. But S and L'_l differ by at most δn elements, and so $|S| \geq \frac{n}{4} - \delta n$. Since $|S| = \beta|L|$ and $|L| \leq n$ we get: $\beta n \geq \frac{n}{4} - \delta n$ which implies $\beta \geq \frac{1}{4} - \delta$.

Also obviously $\beta \leq 0.5$ since $h_L \leq g_L$ by the definition of S.

So $f : [\frac{1}{4} - \delta, \frac{1}{2}] \to \mathbb{R}$ is a well defined function. We find its maximum:

$$f'(\beta) = 2\frac{-2(\frac{3}{2} - \beta) - (1-2\beta)(-1)}{(\frac{3}{2} - \beta)^2} = 2\frac{-3 + 2\beta + 1 - 2\beta}{(\frac{3}{2} - \beta)^2} = \frac{-4}{(\frac{3}{2} - \beta)^2} < 0.$$

So f gets its maximum at the left boundary where $\beta = \frac{1}{4} - \delta$:

$$f(\frac{1}{4}-\delta) = 2\frac{\frac{1}{2}+2\delta}{\frac{5}{4}+\delta} = \frac{4}{5}(1+O(\delta)).$$

By plugging the above into Eq. (16) we get:

$$g_L - h_L \leq \frac{4}{5}\frac{(OPT_S + OPT_{T\cap L_l} + OPT_U + OPT_V + O(\delta)OPT_L)}{(1-2\beta)|L|}.$$

□

What we have shown is that if the algorithm returns h_L we are done. If the algorithm returns h_R we can similarly show that in this case $\mathrm{med}_1(L, h_R) \leq (1.8 + \Theta(\delta))\,\mathrm{med}_1(L, g_L)$.

H.1.1 Handling the Case where the Algorithm Returns h_R

In this case the cost $\mathrm{med}_1(L, h_R)$ will be:

Claim 22.

$\mathrm{med}_1(L, h_R) = OPT_S + OPT_{T\cap L_l} - OPT_{T\cap L_r} - OPT_U + OPT_V + (|L|(1-2\gamma))(h_R - g_L).$

Proof.

$$\begin{aligned}
\mathrm{med}_1(L, h_R) &= \sum_{i\in S\cup T\cup U} h_R - x_i + \sum_{i\in V} x_i - h_R \\
&= \sum_{i\in S\cup T\cup U} (h_R - g_L) + (g_L - x_i) + \sum_{i\in V}(x_i - g_L) - (h_R - g_L) \\
&= OPT_S + OPT_{T\cap L_l} - OPT_{T\cap L_r} - OPT_U + OPT_V + (|S| + |T| + |U| - |V|)(h_R - g_L) \\
&\overset{|S|+|T|+|U|+|V|=|L|}{=} OPT_S + OPT_{T\cap L_l} - OPT_{T\cap L_r} - OPT_U + OPT_V + (|L| - 2|V|)(h_R - g_L) \\
&= OPT_S + OPT_{T\cap L_l} - OPT_{T\cap L_r} - OPT_U + OPT_V + (|L|(1-2\gamma))(h_R - g_L).
\end{aligned}$$

□

To get $\mathrm{med}_1(L, h_R) \leq (1.8+O(\delta))\,\mathrm{med}_1(L, g_L)$ we need to show (by the above claim) the following lemma:

Lemma 23. *If the algorithm returns h_R then:*

$$(h_R - g_L) \leq \frac{0.8(OPT_S + OPT_{T\cap L_l} + OPT_V) + 2.8(OPT_{T\cap L_r} + OPT_U) + O(\delta)OPT_L}{(1-2\gamma)|L|}.$$

Proof. We begin by showing the following claim:

Claim 24.

$$(h_R - g_L)|S| \leq (h_R - g_L)|S| \leq OPT_S + \frac{|S|}{|U|}OPT_U.$$

Proof. $h_R - g_L = m' - h_L = (m' - g_L) + (g_L - h_L)$ and so $g_L - h_L = (h_R - g_L) - (m' - g_L)$. Together with Claim 17 we get:

$$(h_R - g_L)|S| \leq OPT_S + |S|(m' - g_L). \tag{17}$$

Also, $OPT_U = \sum_{i\in U} x_i - g_L \geq \sum_{i\in U} m' - g_L = |U|(m' - g_L)$ which implies

$$m' - g_L \leq \frac{OPT_U}{|U|}. \tag{18}$$

By Eq. (17) and Eq. (18) we get the desired inequality. □

From Claim 24, Claim 18, Claim 20, Claim 19:

we get the following four inequalities:

$$\begin{gathered}(h_R - g_L)2|S| \leq 2OPT_S + 2\frac{|S|}{|U|}OPT_U, \\ (h_R - g_L)|U_r| \leq 2OPT_{U_r}, \\ (h_R - g_L)(|U_l| - O(\delta n)) \leq 2(OPT_{T\cap L_l} + OPT_{U_l}), \\ (h_R - g_L)2|V| \leq 2OPT_V,\end{gathered}$$

By summing these inequalities:

$$(h_R - g_L)(2|S| + |U| + 2|V| - O(\delta n)) \leq \frac{OPT_S + OPT_U + OPT_{T\cap L_l} + OPT_V + \frac{2|S|}{|U|}OPT_U}{(1-2\gamma)|L|} \cdot 2(1-2\gamma)|L|. \tag{19}$$

Since the algorithm returns h_R it must be the case that $|R'| \geq \frac{n}{2}$ and so $|R'{}_l| \geq \frac{n}{4}$ which means that $|U| \geq \frac{n}{4} - \delta n$. This also means that $|L'| \leq \frac{n}{2}$ and thus similarly $|S| \leq \frac{n}{4} + \delta n$. So together:

$$2\frac{|S|}{|U|} \leq \frac{\frac{n}{2} + 2\delta n}{\frac{n}{4} - \delta n} = 2 + O(\delta).$$

By plugging this back into Eq. (19) we get:

$$(h_R - g_L)(2|S| + |U| + 2|V| - O(\delta n)) \leq \frac{OPT_S + (3 + O(\delta))OPT_U + OPT_{T\cap L_l} + OPT_V}{(1-2\gamma)|L|} \cdot 2(1-2\gamma)|L|. \tag{20}$$

Since $|L'{}_l| = |L'{}_r|$ we get that $|T| \approx^{\delta n} |S|$ and so

$$\begin{aligned} & 2|S| + |U| + 2|V| - O(\delta n) \\ & \qquad \geq |S| + |T| + |U| + 2|V| - O(\delta n) \\ & \qquad = (|S| + |T| + |U| + |V|) + |V| - O(\delta n) \\ & \qquad \overset{(\star)}{=} (1+\gamma)|L| - O(\delta n)\end{aligned}$$

Where $(\star)$ is due to: $|L| = |S| + |T| + |U| + |V|$, $|V| = \gamma|L|$.

Let us use this fact in Eq. (20) and get:

$$h_R - g_L \leq \frac{OPT_S + (3 + O(\delta))OPT_U + OPT_{T\cap L_l} + OPT_V}{(1-2\gamma)|L|} \cdot \left(2\frac{(1-2\gamma)|L|}{(1+\gamma)|L| - O(\delta n)}\right). \tag{21}$$

The second term on RHS is (up to $1 + O(\delta)$ factor): $h(\gamma) := 2\frac{1-2\gamma}{1+\gamma}$. Let us find a bound f or h.

Since $|R'| \geq \frac{n}{2}$:

$$|V| + |R| \approx =^{O(\delta n)} |R'{}_r| \geq \frac{n}{4}.$$

and since $|R| \leq b\delta n$: $\gamma|L| = |V| \geq \frac{n}{4} - b\delta n$. We deduce:

$\gamma \geq \frac{1}{4} - O(\delta)$. The nominator is decreasing in γ and the denominator is increasing in γ and thus the maximum is received on the left end point: $\gamma = \frac{1}{4} - O(\delta)$. The maximum value of h is thus: $2\frac{\frac{1}{2} - O(\delta)}{\frac{5}{4} - O(\delta)} = \frac{4}{5} + O(\delta)$.

So we found a bound for the RHS of Eq. (21) and therefore:

$$h_R - g_L \leq \frac{\frac{4}{5}OPT_S + \frac{12}{5}OPT_U + \frac{4}{5}OPT_{T\cap L_l} + \frac{4}{5}OPT_V + O(\delta)OPT_L}{(1-2\gamma)|L|}.$$

□

H.2 Sub-case: $m' < g_L$.

To handle this sub-case we use a reduction to the previous sub-case and show that we pay another multiplicative factor of at most $1 + O(\delta)$. We show a bound of $\text{med}_1(h_L, L)$ in case the algorithm returns h_L. In the case the algorithm returns h_R a bound for $\text{med}_1(h_R, L)$ is achieved similarly.

Let $\tilde{g_L} = m'$. For any multi-set of points $M \subseteq L$ let $\widetilde{OPT_M} = \sum_{i \in M} |x_i - \tilde{g_L}|$.

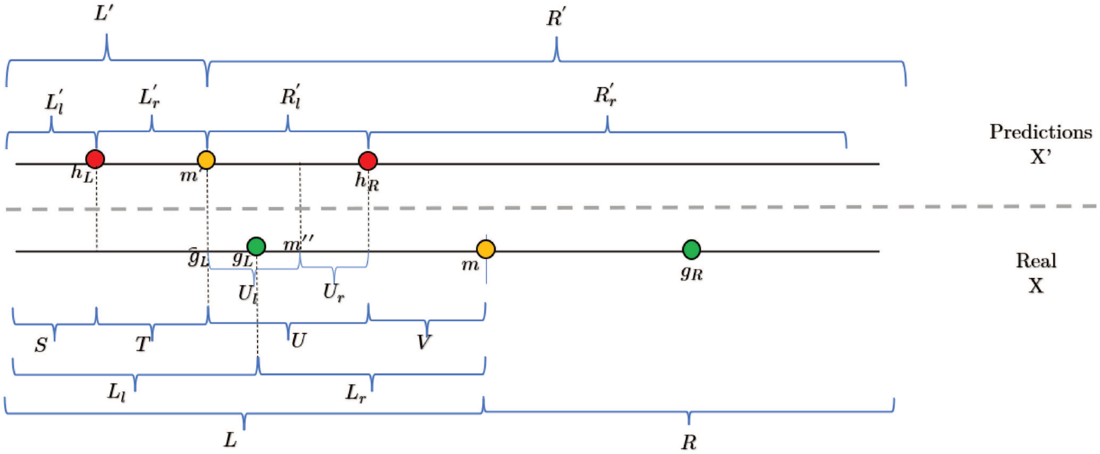

Figure 6: Illustration of case (4) where $m' < g_L$. On the top we have the "estimated" locations h_L, h_R, computed on X'. On the bottom we have the "real" locations, g_L, g_R, computed on X. $L', R', L'_l, L'_r, R'_l, R'_r, S, T, U, V, U_l, U_r$ are all the same as before. The difference is that m' is now to the left of g_L.

By the exact same analysis of the case where $m' \geq g_L$, now we have that $m' \geq \tilde{g_L}$ and thus, just like before, $\tilde{g_L} - h_L \leq \frac{4}{5}\frac{\widetilde{OPT_S}+\widetilde{OPT_T}+\widetilde{OPT_{L_r}}}{(1-2\beta)|L|}$.

But $\widetilde{OPT_S} + \widetilde{OPT_T} = OPT_S + OPT_T - (g_L - m')(|S| + |T|)$ and $\widetilde{OPT_{L_r}} = OPT_{L_r} + (g_L - m')|L_r|$.

So together we have:

$$\begin{aligned} \tilde{g_L} - h_L &\leq \frac{4}{5}\frac{OPT_S + OPT_T - (g_L - m')(|S| + |T|) + OPT_{L_r} + (g_L - m')(|U| + |V|)}{(1-2\beta)|L|} && (22) \\ &= \frac{4}{5}\frac{(OPT_S + OPT_T + OPT_{L_r} + (g_L - m')(-|S| - |T| + |U| + |V|))}{(1-2\beta)|L|}. && (23) \end{aligned}$$

Since $|L'| \geq \frac{n}{2}$ it must be the case that $|S| + |T| \geq \frac{n}{2} - \delta n$ since otherwise there would be strictly less than $\frac{n}{2}$ elements in $|L'|$. For this reason also $|U| + |V| \leq \frac{n}{2}$

So $-|S| - |T| + |U| + |V| \leq \delta n$. By plugging this in Eq. (22) we get:

$$\tilde{g_L} - h_L \leq \frac{4}{5}\frac{(OPT_S + OPT_T + OPT_{L_r} + (g_L - m')(\delta n)}{(1-2\beta)|L|}.$$

Note that $\tilde{g_L} - h_L = \tilde{g_L} - g_L + g_L - h_L = g_L - h_L - (g_L - \tilde{g_L}) = (g_L - h_L) - (g_L - m')$.

By plugging this in the above we get:

$$g_L - h_L \leq \frac{4}{5}\frac{\left(OPT_S + OPT_T + OPT_{L_r} + \frac{10}{4}(g_L - m')\delta n\right)}{(1-2\beta)|L|}.$$

Note that $OPT_T = \sum_{i \in T} g_L - x_i \geq \sum_{i \in T} g_L - m'$ so $g_L - m' \leq \frac{OPT_T}{|T|}$.

$$\begin{aligned} \tilde{g_L} - h_L &\leq \frac{4}{5}\frac{\left(OPT_S + OPT_T + OPT_{L_r} + 2.5 OPT_T \frac{\delta n}{|T|}\right)}{(1-2\beta)|L|} \\ &\leq \frac{4}{5}\frac{\left(OPT_S + OPT_T + OPT_{L_r} + 2 OPT_T \frac{\delta n}{n/2 - \delta n}\right)}{(1-2\beta)|L|} \\ &= \frac{4}{5}\frac{(OPT_S + (1 + \Theta(\delta))OPT_T + OPT_{L_r})}{(1-2\beta)|L|}. && (24) \end{aligned}$$

and so we get the desired bound (due to the equivalent condition stated in Eq. (8).

□

We conclude with a final remark:

Remark 25. *Algorithm 3 does not have (c, δ) approximation robustness for any $c < \frac{5}{3} \approx 1.667$. This lower bound is obtained by observing the following instance: Let $\varepsilon > 0$ be a small enough value, $X \in \mathbb{R}^n$ where $n/2$ points are at $x = 0$, $n/4$ points are at $x = -0.5 - \varepsilon$, $n/4$ points at $x = -1$, and one point at $M \gg n$. Now let X' be obtained from X by moving the point at $x = M$ to $x = -1$. Calculation shows that the above instance leads to a $\frac{5}{3}$ approximation ratio.*

I Second Facility Location: SECOND-PROPORTIONAL-MECHANISM Proofs

In this section we show the strategyproofness and the expected approximation ratio of Algorithm 2.

Proof of Lemma 6. We follow the proof of (Lu et al., 2010, Thm 4.1). In the proof they show that given that the first facility is at x_k for any $k \in [n]$, then for any $X' =< x'_i, X_{-i} >$ that is gained from X by agent i deviating from x_i to x'_i, the expected cost of agent i only increases by deviating to x'_i. The proof does not depend at all on the location of the first facility. Since the first facility is entirely independent of the agent reported values, then the above implies that the choice of the second facility is also strategy proof. □

Proof of Theorem 7. The proof is similar in structure to the one given by Lu et al. (2010) for the Proportional-Mechanism; the main difference is a careful (and tight) analysis of Lemma 27 for the real line which leads to an approximation ratio of at most 3 rather than 4.

Let $G := \{g_S, g_T\}$, $S = \{i \mid x_i \in X, d(x_i, g_S) \leq d(x_i, g_T)\}$ and $T = [n] \setminus S$.

For any $M \subseteq X$ let $OPT_M = \text{med}_2(M, G)$. So $OPT_S = \text{med}_1(S, g_S), OPT_T = \text{med}_1(T, g_T)$ and thus $OPT = \text{med}_2(X, G) = OPT_S + OPT_T$.

Let $H = (g_S, h)$ be the algorithms solution for the problem, and $ALG = \text{med}_2(X, H)$. We denote ALG_S, ALG_T to be $\text{med}_2(S, H), \text{med}_2(T, H)$ respectively, so that $ALG = ALG_S + ALG_T$ and therefore $\mathbb{E}[ALG] = \mathbb{E}[ALG_S] + \mathbb{E}[ALG_T]$.

Since $ALG_S \leq \text{med}_1(S, g_S) = OPT_S$, we get that $\mathbb{E}[ALG_S] \leq OPT_S$, and we only need to find a bound for $\mathbb{E}[ALG_T]$.

$$\begin{aligned}
\mathbb{E}[ALG_T] &= \sum_{i \in [n]} \mathbb{E}[ALG_T \mid h = x_i] Pr(h = x_i) && (25)\\
&= \sum_{i \in T} \mathbb{E}[ALG_T \mid h = x_i] Pr(h = x_i) + \sum_{i \in T} \mathbb{E}[ALG_T \mid h = x_i] Pr(h = x_i) && (26)
\end{aligned}$$

In the summation above the first term is the expected cost due to choosing the second facility to be a location in S and the second term is the expected cost due to choosing the second facility to be a location in T. We will bound each one individually.

For all $i \in [n]$ let $a_i = d(x_i, g_S)$, $p_i = \frac{a_i}{\sum_{j \in [n] a_j}}$ and $b_i = d(x_i, g_T)$. So $\sum_{i \in S} a_i = OPT_S$, $\sum_{i \in T} b_i = OPT_T$ and $\sum_{i \in [n]} p_i = 1$. For each $i, j \in [n]$ let $d_{i,j} = d(x_i, x_j)$.

Lemma 26. $\sum_{i \in S} \mathbb{E}[ALG_T \mid h = x_i] Pr(h = x_i) \leq OPT_S$

Proof.

$$\begin{aligned}
\sum_{i \in S} \mathbb{E}[ALG_T \mid h = x_i] Pr(h = x_i) &= \sum_{i \in S} \left(\sum_{t \in T} \min\{a_t, d_{i,t}\} \right) \cdot p_i \\
\overset{(\star)}{\leq} \sum_{i \in S} \left(\sum_{t \in T} a_t \right) \cdot p_i &= \sum_{i \in S} \left(\frac{\sum_{t \in T} a_t}{\sum_{t \in T} a_t + \sum_{s \in S} a_s} \right) \cdot a_i \leq \sum_{i \in S} a_i = OPT_S
\end{aligned}$$

Where inequality $(\star)$ is gained by ignoring the second facility. □

Lemma 27. $\sum_{i \in T} \mathbb{E}[ALG_T \mid h = x_i] Pr(h = x_i) \leq 3OPT_T$

Proof. Let $D = d(g_A, g_B)$.

$$\begin{aligned}
&\sum_{i \in T} \mathbb{E}[ALG_T \mid h = x_i] Pr(h = x_i) = \sum_{i \in T} \left[\left(\sum_{t \in T} \min\{a_t, d(x_t, x_i)\} \right) \cdot p_i \right] \\
&\overset{\textit{triangle inequality}}{\leq} \sum_{i \in T} \left[\left(\sum_{t \in T} \min\{a_t, b_t + b_i\} \right) \cdot p_i \right] = \sum_{i \in T} \left[\left(\sum_{t \in T} b_t + \sum_{t \in T} \min\{a_t - b_t, b_i\} \right) \cdot p_i \right] \\
&= \sum_{i \in T} \left[\left(\sum_{t \in T} b_t \right) \cdot p_i \right] + \sum_{i \in T} \left[\left(\sum_{t \in T} \min\{a_t - b_t, b_i\} \right) \cdot p_i \right]
\end{aligned}$$

$$= \sum_{i\in T}\left[\left(\sum_{t\in T} b_t\right)\cdot p_i\right] + \sum_{i\in T}\left[\left(\sum_{t\in T}\min\{a_t-b_t,b_i\}\right)\cdot\frac{b_i}{\sum_{j\in[n]}a_j}\right]$$
$$+\sum_{i\in T}\left[\left(\sum_{t\in T}\min\{a_t-b_t,b_i\}\right)\cdot\frac{a_i-b_i}{\sum_{j\in[n]}a_j}\right]$$

We bound each of these 3 terms individually by OPT_T, and thus get the desired bound.

The first term bound:

$$\sum_{i\in T}\left[\left(\sum_{t\in T} b_t\right)\cdot p_i\right] = \left(\sum_{t\in T} b_t\right)\sum_{i\in T} p_i \le OPT_T \tag{27}$$

The second term bound:

$$\sum_{i\in T}\left[\left(\sum_{t\in T}\min\{a_t-b_t,b_i\}\right)\cdot\frac{b_i}{\sum_{j\in[n]}a_j}\right] \le \sum_{i\in T}\left[\left(\sum_{t\in T}a_t-b_t\right)\cdot\frac{b_i}{\sum_{j\in[n]}a_j}\right]$$
$$\le \sum_{i\in T}\left[\left(\sum_{t\in T}a_t\right)\cdot\frac{b_i}{\sum_{j\in[n]}a_j}\right] = \sum_{i\in T} b_i = OPT_T$$

The third term bound:

$$\sum_{i\in T}\left[\left(\sum_{t\in T}\min\{a_t-b_t,b_i\}\right)\cdot\frac{a_i-b_i}{\sum_{j\in[n]}a_j}\right] \le \sum_{i\in T}\left[\left(\sum_{t\in T}b_i\right)\cdot\frac{a_i-b_i}{\sum_{j\in[n]}a_j}\right] = \sum_{i\in T}\left[|T|b_i\cdot\frac{a_i-b_i}{\sum_{j\in[n]}a_j}\right] \tag{28}$$

We finish the third term bound by showing the following claim

Claim 28.

$$\sum_{i\in T}\left[|T|b_i\cdot\frac{a_i-b_i}{\sum_{j\in[n]}a_j}\right] \le OPT_T$$

Proof. Let T_a be the part of T closer to g_S and T_b be the other part of T.

So $T_a = \{i \mid i\in T, d(x_i,g_S)\le d(x_i,g_T)\}$, and $T_b = T\setminus T_a$. Since the points are all on the line metric:

For all $i\in T_a$: $a_i = d(g_S,x_i) = d(g_S,g_T) - b_i = D - b_i$.

For all $i\in T_b$: $a_i = d(g_S,x_i) = d(g_S,g_T) + b_i = D + b_i$.

Let us use these facts in Eq. (28) to get:

$$\sum_{i\in T}\left[|T|b_i\cdot\frac{a_i-b_i}{\sum_{j\in[n]}a_j}\right] = \sum_{i\in T_a}\left[|T|b_i\cdot\frac{D-b_i-b_i}{\sum_{j\in[n]}a_j}\right] + \sum_{i\in T_b}\left[|T|b_i\cdot\frac{D+b_i-b_i}{\sum_{j\in[n]}a_j}\right]$$
$$= \frac{|T|D}{\sum_{j\in[n]}a_j}\left(\sum_{i\in T}b_i\right) - \frac{2|T|\sum_{i\in T_a}{b_i}^2}{\sum_{j\in[n]}a_j} = \frac{|T|D\cdot OPT_T - 2|T|\sum_{i\in T_a}{b_i}^2}{\sum_{j\in[n]}a_j}$$
$$\overset{(\star)}{\le} \frac{|T|D\cdot OPT_T - 2(OPT_{T_a})^2}{\sum_{j\in[n]}a_j} \overset{(\star\star)}{\le} \frac{|T|D - 2\frac{(OPT_{T_a})^2}{OPT_T}}{|T|D + OPT_{T_b} - OPT_{T_a}}\cdot OPT_T \tag{29}$$

$(\star)$ is due to $QM - AM$ inequality since

$$\sum_{i\in T_a}{b_i}^2 \ge \frac{\left(\sum_{i\in T_a}b_i\right)^2}{|T_a|} \overset{|T_a|\le|T|}{\ge} (OPT_{T_a})^2\frac{1}{|T|}.$$

($\star\star$) explanation:

$$\sum_{j\in[n]} a_j = \sum_{j\in S} a_j + \sum_{j\in T_a} a_j + \sum_{j\in T_b} a_j = OPT_S + \sum_{j\in T_a} D - b_j + \sum_{j\in T_b} D + b_j$$
$$= OPT_S + D|T| + OPT_{T_b} - OPT_{T_a} \geq |T|D + OPT_{T_b} - OPT_{T_a}$$

We now show that

$$\frac{|T|D - 2\frac{(OPT_{T_a})^2}{OPT_T}}{|T|D + OPT_{T_b} - OPT_{T_a}} \leq 1.$$

Indeed, if $OPT_{T_b} \geq OPT_{T_a}$ then $|T|D + OPT_{T_b} - OPT_{T_a} \geq |T|D$ and thus:

$$\frac{|T|D - 2\frac{(OPT_{T_a})^2}{OPT_T}}{|T|D + OPT_{T_b} - OPT_{T_a}} \leq \frac{|T|D - 2\frac{(OPT_{T_a})^2}{OPT_T}}{|T|D} \leq \frac{|T|D}{|T|D} = 1.$$

Otherwise $OPT_{T_a} - OPT_{T_b} > 0$ and thus:

$OPT_{T_a} > OPT_{T_b} \implies$
$OPT_T = OPT_{T_a} + OPT_{T_b} < 2OPT_{T_a} \implies$
$(OPT_{T_a} - OPT_{T_b})OPT_T \leq 2(OPT_{T_a} - OPT_{T_b})OPT_{T_a} = 2(OPT_{T_a})^2 - 2OPT_{T_b}OPT_{T_a} \leq 2(OPT_{T_a})^2 \implies$

$$OPT_{T_a} - OPT_{T_b} \leq \frac{2(OPT_{T_a})^2}{OPT_T} \implies \frac{|T|D - 2\frac{(OPT_{T_a})^2}{OPT_T}}{|T|D + OPT_{T_b} - OPT_{T_a}} \leq \frac{|T|D + OPT_{T_b} - OPT_{T_a}}{|T|D + OPT_{T_b} - OPT_{T_a}} = 1$$

So from Eq. (29): $\sum_{i\in T}\left[|T|b_i \cdot \frac{a_i - b_i}{\sum_{j\in[n]} a_j}\right] \leq OPT_T$.

□

To summarize: $\sum_{i\in T} \mathbb{E}[ALG_T \mid h = x_i] Pr(h = x_i) \leq OPT_T + OPT_T + OPT_T = 3OPT_T$. □

From plugging Lemma 26, Lemma 27 into Eq. (26) we get:

$$\mathbb{E}[ALG] \leq 2OPT_S + 3OPT_T \leq 3OPT.$$

□

Remark 29. *The upper bound of* 3 *in the analysis of Theorem 7 is tight.*

To show tightness, consider an instance with n points on the real line: $n/2$ points at $x = 0$, $n/2 - 1$ points at $x = 1$, and a single point at $x = 2$. Assume the given first facility is at 0. The optimal solution puts the second facility at 1 and pays the cost of 1. On the other hand $\mathbb{E}[ALG] = \frac{1}{\frac{n}{2}+1}\left(\frac{n-2}{2} + 2\frac{n-2}{2}\right) = 3\frac{n/2-1}{n/2+1} \approx 3$.

J Proof of Theorem 9

Proof. The strategyproofness is due to the fact that the first facility is chosen based only using the predictions, and the second facility choice is strategyproof due to Lemma 6.

Let $G = (g_L, g_R)$ be the optimal cluster centers for X, and let L, R be the respective corresponding clusters. We consider two cases: when this optimal clustering is $b\delta$-balanced, and when it is not; in both cases we show that the algorithm achieves low expected cost. Let $H = (h_1, h_2)$ be the solution returned by Algorithm 4.

The first case is when (L, R) is a $b\delta$-balanced clustering of X. In this case we claim that

$$\mathbb{E}[\mathrm{med}_2(X, H)] \leq (3.6 + O(\delta))\,\mathrm{med}_2(X, G).$$

Indeed, let $T = (t_L, t_R)$ be the two centers of the $(b-1)\delta$-balanced 2-medians algorithm on X'. Since G induces a $b\delta$-balanced clustering, Theorem 4 and our choice of b ensures that

$$\operatorname{med}_2(X, T) \leq 1.2 \operatorname{med}_2(X, G).$$

Since Algorithm 3 returns one of the centers of the $(b-1)\delta$-balanced 2-medians algorithm on X' as h_1, let us assume w.l.o.g. that BIG-CLUSTER-CENTER returns $h_1 = t_L$. Let (X_L, X_R) and (X'_L, X'_R) be the clusterings of X and X' induced by $T = \{t_L, t_R\}$. From Theorem 7:

$$\begin{aligned} E[\operatorname{med}_2(X, H)] &\underset{(\star)}{=} \mathbb{E}[\operatorname{med}_2(X, \{t_L, h_2\})] \leq 2 \operatorname{med}_1(X_L, t_L) + 3 \operatorname{med}_1(X_R, median(X_R)) \\ &\leq 2 \operatorname{med}_1(X_L, t_L) + 3(1 + O(\delta)) \operatorname{med}_1(X_R, t_R) \\ &\leq (3 + O(\delta)) \operatorname{med}_2(X, T) \leq (3.6 + O(\delta)) \operatorname{med}_2(X, G), \end{aligned}$$

where the inequality $(\star)$ uses the fact that X_R and X'_R differ on at most $\delta|X|$ points, and hence we can sue $(1 + O(\delta), \delta)$ approximation-robustness of 1-Median from Corollary 3.

Now for the other case: suppose (L, R) is a $b\delta$-unbalanced clustering. Theorem 8 implies that either $\operatorname{med}_1(L, h_1) \leq (1.8 + O(\delta)) \operatorname{med}_1(L, g_L)$ or $\operatorname{med}_1(R, h_1) \leq (1.8 + O(\delta)) \operatorname{med}_1(R, g_R)$. W.l.o.g., consider the first option. Then from Theorem 7:

$$\begin{aligned} E[\operatorname{med}_2(X, H)] &\leq 2 \operatorname{med}_1(L, h_1) + 3 \operatorname{med}_1(R, g_R) \\ &\leq 2((1.8 + O(\delta)) \operatorname{med}_1(L, g_L)) + 3 \operatorname{med}_1(R, g_R) \leq (3.6 + O(\delta)) OPT. \end{aligned}$$

Combining the two cases completes the proof. □

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: See Section 7, where we list possible directions to improve what we have done in this work, and directions for future work.

3. **Theory Assumptions and Proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

4. **Experimental Result Reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [NA]

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [NA]

6. **Experimental Setting/Details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

   Answer: [NA]

7. **Experiment Statistical Significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [NA]

8. **Experiments Compute Resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [NA]

9. **Code Of Ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

10. **Broader Impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [NA]

    Justification: It is a theoretic paper discussing beyond worst case analysis of algorithms with predictions.

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

    Answer: [NA]

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [NA]

13. **New Assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [NA]

14. **Crowdsourcing and Research with Human Subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [NA]

15. **Institutional Review Board (IRB) Approvals or Equivalent for Research with Human Subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [NA]