On Voting and Facility Location

Michal Feldman, Amos Fiat and Iddan Golomb

Abstract

"We would all like to vote for the best man, but he is never a candidate"

— Kin Hubbard

We study mechanisms for candidate selection that seek to minimize the social cost, where voters and candidates are associated with points in some underlying metric space. The social cost of a candidate is the sum of its distances to each voter. Some of our work assumes that these points can be modeled on the real line, but other results of ours are more general.

A question closely related to candidate selection is that of minimizing the sum of distances for facility location. The difference is that in our setting there is a fixed set of candidates, whereas the large body of work on facility location seems to consider every point in the metric space to be a possible candidate. This setting gives rise to three types of candidate selection mechanisms which differ in the granularity of their input space (single candidate, ranking and location mechanisms). We study the relationships between these three classes of mechanisms.

While it may seem that Black's 1948 median algorithm is optimal for candidate selection on the line, this is not the case. We give matching upper and lower bounds for a variety of settings. In particular, when candidates and voters are on the line, our universally truthful *spike* mechanism gives a [tight] approximation of two. When assessing candidate selection mechanisms, we seek several desirable properties: (a) efficiency (minimizing the social cost) (b) truthfulness (dominant strategy incentive compatibility) and (c) simplicity (a smaller input space). We quantify the effect that truthfulness and simplicity impose on the efficiency.

1 Introduction

The Hotelling-Downs model ([13], [20]) used to study political strategies, assumes that individual voters occupy some point along the real line. Non-principled political parties (or ice cream vendors) strategically position themselves at a point along the left-right axis (or along a beach) so as to garner the greatest number of supporters (clients). Implicitly, voters will vote for the closest candidate.

We consider an analogous problem to the Hotelling-Downs model, where candidates are principled (*i.e.*, non-strategic) whereas the voters have preferences but may misrepresent them in order to achieve what is a better outcome from their perspective. In this model, in which both voters and candidates are represented by points in the metric space, a closer candidate is preferable to one further away.

Examples for candidate selection:

- A municipality plans to erect a public library on a street, and every resident seeks to be as close as possible to the proposed library. However, the new library can only be built on suitable locations (the candidates).

- Social choice issues in which the distance is not physical: there is a set of policies ranging from left to right, and several political candidates stand for election, each one advocating a different policy. Every voter is associated with a point along the real line. An example of a collective decision problem which does not revolve around the political sphere yet may also fit this setting is the task of determining the temperature of an air

conditioner in a room, where each individual has a different ideal point along the scale of temperatures (a line). There are many additional settings of relevant candidate selection problems, e.g., in the realms of recommendation systems and computational economics. While our results do not necessarily apply to all social choice settings, there are many such problems for which they do apply (whether in entirety or partially).

Assuming quasi-linear utilities, and allowing payments — the well known Vickrey-Clarke-Groves (VCG) mechanism is truthful and can achieve the optimal social cost (see, e.g., [23]). However, in many real-life situations we restrict the use of money due to ethical, legal or other considerations, e.g, in democratic elections and in examples previously mentioned.

We study deterministic truthful mechanisms with no payments for the candidate selection problem. In such mechanisms, no agent can benefit from misreporting her location, regardless of the reported locations of the other agents. Such mechanisms are also known as dominant strategy incentive compatible mechanisms. We also consider randomized truthful mechanisms, both universally truthful (ex-post Nash) and truthful in expectation.

Given a set of candidate and voter locations, it is polytime to find the candidate that minimizes the social cost.

When restricted to deterministic truthful mechanisms, we show that the optimal candidate cannot be selected in the general case. Moreover, we show that the cost may be as bad as three times the optimal cost (matching lower and upper bounds). When considering randomized mechanisms on the line, the approximation factor drops to two (matching lower and upper bounds).

There are other reasons that an optimal candidate may not be chosen. In particular, this depends on the amount of information the agents supply to the mechanism. We formulate three different types of mechanisms, based on the information each agent submits to the mechanism. We note that all three mechanism types are *candidate selection mechanisms*, that is – their output is a single candidate.

- Single Candidate [vote] mechanisms, in which every agent votes for one of the candidates.
- Ranking [vote] mechanisms, in which every agent states ordinal preferences over the candidates (a permutation).
- Location [vote] mechanisms, in which every agent sends a position.

Clearly, knowing the true location of an agent allows one to infer the ranking preferred by that agent, which in turn unravels the favorite candidate of the agent (up to tie-breaking).

In the vast majority of previous work done on the facility location problem every point in the metric space was considered to be a candidate. Therefore there was no difference between these three mechanism types.

The social choice literature mostly considers social choice functions (which are ranking mechanisms that are not necessarily truthful). Note that Arrow's impossibility theorem does not hold when assuming the preferences are single-peaked.

The more information an agent transmits, the more tools the mechanism has to devise an accurate solution. Albeit, this information comes at a cost, since it might disclose more private information which the agents wish to keep confidential. Furthermore, behavioral economists have long argued that the agents cannot fully acquire their utility function, or that obtaining this information requires a high cognitive cost. Additionally, sending more information also casts a higher burden on the mechanism itself. For all of these reasons deploying a simple mechanism[1] which requires less information from agents is beneficial.

[1] We use the term "simplicity" from the perspective of the voters, who have a smaller action space, i.e, less options to choose from. The mechanism itself can act in an arbitrarily complex fashion.

Indeed, in many practical scenarios, single candidate mechanisms are used rather than ranking mechanisms. Generally there is a trade-off between the accuracy of a mechanism and its simplicity.

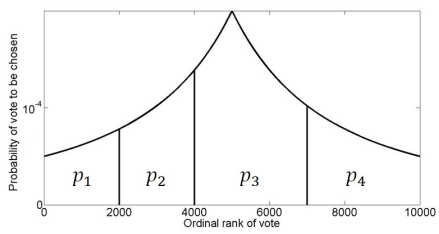
1.1 Our Contributions

In the paper, we show the following:

- In Section 3 we formulate a framework of reductions that compare the various mechanism types. We utilize this framework to show the relations (equivalence or strict containment) between the three classes of *truthful* mechanisms – single candidate, ranking and location (see Figure 2 in the appendix). Furthermore, we show that for the case of two candidates, the set of truthful in expectation location mechanisms is equivalent to the set of truthful in expectation single candidate mechanisms. These results provide a significant step towards a full characterization of truthful mechanisms.

- In Section 4 we define a family of universally truthful single candidate mechanisms on the line called weighted percentile single candidate (WPSC) mechanisms, which choose the ith vote with some predetermined probability p_i. We introduce the *spike* mechanism, which is a WPSC mechanism that carefully crafts the probability distribution to account for misreports by any agent - whether they are near the center or close to the extremes (see Figure 1). We then use backwards induction to show that spike achieves an approximation ratio of two (Theorem 7).

Figure 1: The density function of the spike mechanism, which gives rise to the mechanism's name (the cumulative distribution function is given explicitly in Definition 3). In this example there are 10000 agents and 4 candidates. The candidates, when ordered from left to right, receive 2000, 2000, 3000 and 3000 votes respectively. The graph depicts probability of choosing each vote – votes are chosen with higher probability when they are closer to the 50th percentile. The area beneath the graph represents the probability that each candidate will be elected, e.g., the probability of choosing the second candidate (p_2) is the integral of the function between 2000 and 4000.



- In Section 5 we show additional bounds for randomized mechanisms – On the line there is a lower bound of two, even for location mechanisms, which shows that the result for spike is tight. Furthermore, when combining this understanding with the results of Section 3, it can be concluded that two is also the tight approximation ratio for truthful in expectation mechanisms (single candidate, ranking or location) and for universally truthful single candidate mechanisms.

  We move on to show bounds for randomized mechanisms for more general metric spaces[2] (see Figure 5 in the appendix). An easy observation is that the random

dictator mechanism achieves an upper bound of three for any metric space. Theorem 11 shows a lower bound of $3 - \frac{2}{d}$ for any single candidate mechanism in $\Re^d$ (by using a counterexample based on a regular simplex). This is enough to conclude that on an arbitrary metric space, the bound of three is tight for single candidate mechanisms. Theorem 14 displays a lower bound of 7=3 for any ranking mechanism in $\Re^2$ (which also holds in any higher dimension Euclidean space $\Re^d$).

- In Section 6 we present deterministic bounds on the line – a lower bound of three is met by a matching upper bound due to the median mechanism. All the results on the line, deterministic or randomized, are displayed in the table in Figure 6 in the appendix.

  We highlight the following surprising phenomenon apparent in Figure 6 in the appendix. In both deterministic and randomized cases, imposing any constraint in either information or truthfulness, yields the same ratio as taking the both of these constraints simultaneously — when insisting on truthful mechanisms (in the strategic case), there is no trade-off between high and low information settings, and one can enjoy the benefits of minimal information mechanisms (single candidate mechanisms) without incurring any additional cost to the approximation ratio. Similarly, when deciding to reduce the information requirements to anything less than location mechanisms, it is possible to devise a truthful [single candidate] mechanism, without increasing the approximation ratio.

Due to space constraints, almost all of the proofs and figures are deferred to the appendix.

1.2 Related Work

Procaccia and Tennenholtz introduced game theoretical aspects to the facility location problem [25]. As mentioned before, their setting is similar to the one in this paper, except that the location of the facility is not restricted to a set of candidates. This model was extended in many different ways. The metric space was extended from the line to cycles ([1], [2]), trees ([1], [14]) and general graphs ([1]). Many papers consider building several facilities (or electing a committee of candidates), e.g. [16], [17], [21]. The goal of the majority of these papers is to optimize over some global target function. The most popular target functions are the social cost and the maximal cost of an agent, but additional target functions like the L_2 norm were also considered [14]. Some papers deal with "obnoxious facility location" — a setting in which agents want to be as far away as possible from the facility, e.g., when selecting a location for a central garbage dump (e.g., [11]).

[28] proposed deterministic percentile [location] mechanisms for locating multiple facilities in $\Re^d$. [27] showed that even when the location of the facilities are constrained to a set of candidates (as in our paper), these percentile mechanisms are group-strategyproof on $\Re$.

[12] characterize deterministic truthful mechanisms for locating a facility on the discrete line and the discrete cycle. In their model agents must be located precisely on candidates, and the distance between neighbor candidates is constant.

When constraining the outcome to a set of candidates, the facility location setting resembles social choice problems. The seminal Gibbard-Satterthwaite theorem (e.g, [18]) shows that in a general setting the only onto truthful deterministic mechanisms are dictatorships.

[2] We do not present results for deterministic mechanisms in general metric spaces, since in these cases the incentive compatibility constraints take a significant toll on the approximation ratio – according to [3] in the non-strategic setting it is possible to reach a constant ratio in any metric space, while due to the characterization of [26] there exist metric spaces (e.g., cycles) in which the approximation ratio is $\Omega(n)$ even in the continuous model.

However, if there are limitations on the rankings, then the impossibility theorem of Gibbard-Satterthwaite does not hold. In many cases the rankings can be limited to single-peaked preferences, a notion used as early as 1948 by Black [6], and was later fully characterized ([22] and [26]).

There has been extensive work describing numerous candidate selection schemes (e.g, [9]). These schemes typically have no assumptions on the preferences of the agents, and according to Gibbard-Sattethwaite they are not truthful. [18] further characterizes truthful randomized mechanisms under arbitrary preferences. Some work on social choice makes use of randomized schemes in order to elicit truthfulness (e.g., [7] and [5]).

The advantages of simple mechanisms (in which each voter has less options to choose from) have been widely acknowledged. For instance, aiming for simplicity is a major reason for which the vast majority of the candidate selection schemes above accept ordinal rankings rather than cardinal rankings as input. Truthfulness is also a very common and desired trait of mechanisms at large. Nonetheless, we do not know of any work formally describing the three types of mechanisms, or any similar framework for uncovering relationships between these mechanism types, as in Theorem 1.

Since in the lack of cardinal costs no global objective functions can be measured (e.g, the social cost), the focus of many of the aforementioned schemes is on achieving some desirable axiomatic properties. However, the use of utilitarianism in the realm of social choice has firm and ancient roots (e.g, [15] and [19]).

Moreover, in recent years a line of work in computational social choice regarded *distortion*, a measure for assessing social choice functions (i.e, ranking mechanisms) which also refers to the utilitarian goal of minimizing social cost. The term distortion was coined by Procaccia and Rosenschein in [24], and was followed by several other papers (e.g., [10], [8], [3] and [4]). Roughly speaking, the distortion is the worst case ratio between the social cost of the candidate elected and the social cost of the optimal candidate. Note that while the distortion stems from an information deficiency (access only to ordinal rankings of the agents), the approximation ratio in this paper is greater than one *both* because of this information deficiency (for ranking and single candidate mechanisms), *and* because of incentive compatibility constraints. Computing the approximation ratio and the distortion can quantify the affect of these two deficiencies in various settings.

Caragiannis and Procaccia deal with a setting in which the utility functions of agents are more general than ours, which leads to a higher deterministic lower bound on the distortion [10].

Anshelevich et. al. provide a deterministic lower bound of 3 on the distortion in a general metric space, and show that both Copeland and uncovered set reach a distortion of 5 [3].

Spike is a truthful mechanism which achieves an approximation ratio of 2 on the line. Independently from us, Anshelevich et. al. recently showed a mechanism which also achieves a distortion of 2 on $\Re$, albeit it is not truthful [4].

2 Model

Let $\mathsf{N} = \{1; : : : \mathsf{n}\}$ be a set of agents, where each agent $\mathsf{i} \in \mathsf{N}$ is located at some point x_i. We refer to the location of agent i as agent i's *type*. Let $\mathbf{x} = (\mathsf{x}_1; : : : ; \mathsf{x}_n)$ be the *location profile* of the agents. There exist m candidates located at publicly known points $\mathbf{y} = (\mathsf{y}_1; : : : ; \mathsf{y}_m)$ (we refer to y_i as the ith candidate and as the location of the ith candidate interchangeably). The agents and candidates are located on some metric space. Significant parts of the paper deal with specific metric spaces, and these parts will be noted. When the metric space is $\Re$, we assume that the agents and the candidates are both numbered in ascending order based

on their locations (otherwise they could be renamed).

A *deterministic mechanism* M, is a function which maps an *action profile* $\mathbf{a} = (\mathsf{a}_1; \dots; \mathsf{a}_n) \in \mathcal{A}^n$ to a candidate, that is: $\mathsf{M} : \mathcal{A}^n \to \mathbf{y}$. We consider three classes of mechanisms that differ in the input they accept, *i.e.*, in the action space $\mathcal{A}$ of the agents:

- *Single candidate mechanisms*, in which each agent casts a vote for a candidate, that is: $\mathsf{a}_i \in \mathbf{y}$.
- *Ranking mechanisms*, in which every agent reports ordinal preferences over all the m candidates. The notation $\mathsf{y}_j \succeq \mathsf{y}_k$ indicates a preference of candidate y_j over candidate y_k (or indifference between the two). In ranking mechanisms $\mathsf{a}_i \in \Pi$, where Π is the set of all permutations of the candidates $\mathbf{y}$. These mechanisms are sometimes referred to in the literature as social choice functions.
- *Location mechanisms*, in which every agent reports a location, that is a_i is some point in the metric space.

Given a joint action profile $\mathbf{a}$, the *cost* of point x is its distance to the facility, that is: $\text{cost}_x(\mathsf{M}; \mathbf{a}) = |\mathsf{x} - \mathsf{M}(\mathbf{a})|$. For agent $\mathsf{i} \in \mathsf{N}$ located at point x_i, we refer to $\text{cost}_{x_\mathsf{i}}(\mathsf{M}; \mathbf{a})$ as the *cost of agent* i. The goal of each agent is to minimize their cost.

Truthful mechanisms are usually defined in the context of direct revelation mechanisms. Since in ranking and single candidate mechanisms the action space does not coincide with the type space, we extend this notion in the following trivial manner. For an agent in location x_i and for any mechanism (location, ranking or single candidate), let $\mathcal{A}(\mathsf{x}_i)$ be the set of *true actions* of this agent — the actions which convey the real preferences of this agent. For instance, in single candidate mechanisms $\mathcal{A}(\mathsf{x}_i)$ is the set of candidates closest to x_i, which we refer to as the *favorite candidates* of x_i (this might be a set since there may be ties). An agent reporting $\mathsf{a}_i \in \mathcal{A}(\mathsf{x}_i)$ is said to be reporting *truthfully*, and an action profile $\mathbf{a}$ in which all agents report truthfully is called a *truthful profile*. The set of truthful profiles is denoted $\mathcal{A}(\mathbf{x})$. A *truthful mechanism* M is one in which no agent can suffer from reporting truthfully, regardless of the actions of the other agents:

$$\forall \mathsf{i} \in \mathsf{N}; \forall \mathsf{x}_i; \forall \mathsf{a}_i \in \mathcal{A}(\mathsf{x}_i); \forall \mathbf{a}_{-i} \in \mathcal{A}^{n-1}; \forall \mathsf{a}'_i \in \mathcal{A} : \text{cost}_{x_\mathsf{i}}(\mathsf{M}; (\mathsf{a}_i; \mathbf{a}_{-i})) \leq \text{cost}_{x_\mathsf{i}}(\mathsf{M}; (\mathsf{a}'_i; \mathbf{a}_{-i}))$$

A *randomized mechanism* is a mapping from an action profile to a distribution over the candidates, that is: $\mathsf{M} : \mathcal{A}^n \to \Delta(\mathbf{y})$. The cost of agent i is the expected cost of this agent according to the probability distribution returned by the mechanism, that is: $\text{cost}_{x_\mathsf{i}}(\mathsf{M}; \mathbf{a}) = \mathbb{E}_{y_\mathsf{j} \sim M(\mathbf{a})} |\mathsf{x}_i - \mathsf{y}_j|$.

Two different notions of randomized truthful mechanisms have been studied in the literature, and we extend them naturally based on our definitions of truthful reports:

- *Truthful in expectation (TIE)* mechanisms — where the expected cost of an agent reporting truthfully is never higher than any other action. That is: $\forall \mathsf{i} \in \mathsf{N}$, $\forall \mathsf{a}_i \in \mathcal{A}(\mathsf{x}_i)$, $\forall \mathbf{a}_{-i} \in \mathcal{A}^{n-1}$, $\forall \mathsf{a}'_i \in \mathcal{A}$: $\text{cost}_{x_\mathsf{i}}(\mathsf{M}; (\mathsf{a}_i; \mathbf{a}_{-i})) \leq \text{cost}_{x_\mathsf{i}}(\mathsf{M}; (\mathsf{a}'_i; \mathbf{a}_{-i}))$. In these mechanisms the agent may regret her action ex-post for some of the instances.
- *Universally truthful* mechanisms are mechanisms which can be expressed as a probability distribution over deterministic truthful mechanisms. In these mechanisms an agent never regrets reporting truthfully, even after the random outcome is unraveled.

Clearly, every universally truthful mechanism is truthful in expectation, but not necessarily vice versa. Throughout the paper, in the randomized setting we use the term "truthful" to refer to truthful in expectation mechanisms, unless otherwise stated.

The *social cost* of a mechanism is the sum of the agents' costs. For a location profile $\mathbf{x}$ and an action profile $\mathbf{a}$ the social cost is: $\text{SC}(\mathsf{M}; \mathbf{x}; \mathbf{a}) = \sum_i \text{cost}_{x_\mathsf{i}}(\mathsf{M}; \mathbf{a})$. The cost of a

candidate is the cost of the mechanism which locates the facility on that candidate, that is: $\mathrm{SC}(y_j; \mathbf{x}) = \sum_{i \in N} |y_j - x_i|$. Given a location profile $\mathbf{x}$, the *optimal mechanism*, denoted $\mathrm{OPT}(\mathbf{x})$, is one which chooses a candidate that minimizes the social cost (y_{opt}). When there are when there are several optimal candidates, we break ties consistently (e.g., when the metric space is $\Re$, we refer to the leftmost among them as y_{opt}). For any truthful in expectation mechanism M (including universally truthful mechanisms), the *social cost of* M *given a location profile* $\mathbf{x}$ is the maximal social cost it yields by any truthful action profile $\mathbf{a}$, that is: $\mathrm{SC}(M; \mathbf{x}) = \max_{\mathbf{a} \in \mathcal{A}(\mathbf{x})} \mathrm{SC}(M; \mathbf{x}; \mathbf{a})$. The *approximation ratio* of a truthful in expectation mechanism M is the maximal ratio for any location profile $\mathbf{x}$, between social cost of M given $\mathbf{x}$ and the optimal social cost given $\mathbf{x}$: $\max_{\mathbf{x}} \frac{\mathrm{SC}(M;\mathbf{x})}{\mathrm{SC}(\mathrm{OPT};\mathbf{x})}$.

For single candidate mechanisms when the metric space is a line:

- For the line, let be a permutation on indices $1; \ldots; n$ such that

  $$a_{(1)} \leq a_{(2)} \leq \cdots \leq a_{(n)}$$

  Note that there are many permutations satisfying the above, each of which represents a different version of breaking ties amongst votes for the same candidate. is an arbitrary such permutation. Let $z_j = a_{(j)}$ for $j = 1; \ldots; n$. *I.e.*, z_j is the [location of the] reported ideal candidate for voter (j).

- A *percentile mechanism* is a mechanism specified by an index $1 \leq i \leq n$, which chooses candidate z_i.

- A *weighted percentile single candidate (WPSC)* mechanism is specified by a vector of probabilities $p_1; \ldots; p_n$, such that $\sum_j p_j = 1$, and chooses y_i with probability $\sum_{j: z_j = y_i} p_j$.

  This can be interpreted as follows: a mechanism is WPSC if and only if there exists some as described above, such that for every profile $\mathbf{a}$ voter (j) determines the winning candidate with probability p_j.

In single candidate mechanisms, the set of candidates y induces a partition of the metric space in the following manner — the *candidate zone* of candidate y_i, denoted $\mathcal{Z}_i$, is the set of points whose favorite candidate is y_i: $\mathcal{Z}_i = \{x : \forall y_j : |x - y_i| \leq |x - y_j|\}$. The candidate zones are bounded by *candidate borders*. For example, when the metric space is $\Re$, there are $n - 1$ borders, which are the midpoints between two consecutive candidates: $b_i = \frac{y_i + y_{i+1}}{2}$ (see Figure 7). When the metric space is $\Re^d$, the candidate zones form a Voronoi diagram. A candidate which receives at least one vote is called *active*.

In ranking mechanisms, y induces a partition which divides the metric space into *ranking zones*. All points in some ranking zone $\mathcal{R}_i$ share some ranking $_i$. In this case, we say that the ranking $_i$ is consistent with ranking zone $\mathcal{R}_i$. The ranking zones bounded by *ranking borders*. For example, when the metric space is $\Re$ the ranking borders are the midpoints between any two candidates: $b_{i;j} = \frac{y_i + y_j}{2}$.

3 Classes of Mechanisms

In this section we go over the containment hierarchy of various classes of truthful mechanisms (e.g., Figure 2). We start with some intuition, then define some necessary terms, and finally present the main theorem of this section.

Intuitively, for any mechanism M, there exists a mechanism M' which receives a "richer" input than M, and acts identically to M. For instance, for some arbitrary single candidate

mechanism M, there obviously exists a ranking mechanism M' which disregards all of the preferences except the top choice of each agent, and behaves essentially just like M does.

We generalize this notion in the following informal definition — a mechanism M (whether location/ranking/or single candidate) is said to be *reducible* to a mechanism M' (location/ranking/or single candidate) if for every location profile $\mathbf{x}$ and true reports, the output of M is identical to the output of M' (a formal definition, which is based on M simulating M', is deferred to appendix A.2).

As pointed out, it is clear that every single candidate mechanism M is reducible to some ranking mechanism M' (or some location mechanism M'). In these cases, if M is truthful then so is M', since M' only uses the information which is inputted to M, so any misreports to M' which would not change the input of M do not affect the outcome at all. Note that the same reasoning also shows that every ranking mechanism is reducible to some location mechanism, and that any single candidate mechanism is reducible to some location mechanism.

On the other hand, it is not true that every location mechanism is reducible to some ranking mechanism. Somewhat surprisingly, we will soon show that when we restrict ourselves to deterministic truthful mechanisms this does hold, that is — every *deterministic truthful* location mechanism is reducible to some deterministic truthful ranking mechanism.

Two sets of mechanisms, A and B, are said to be *equivalent* if every $\mathsf{a} \in \mathsf{A}$ is reducible to some $\mathsf{b} \in \mathsf{B}$, and every $\mathsf{b} \in \mathsf{B}$ is reducible to some $\mathsf{a} \in \mathsf{A}$.

A set of mechanisms A is said to be *strictly contained* in a set of mechanisms B if every mechanism $\mathsf{a} \in \mathsf{A}$ is reducible to some mechanism $\mathsf{b} \in \mathsf{B}$, yet not every mechanism $\mathsf{b} \in \mathsf{B}$ is reducible to some mechanism $\mathsf{a} \in \mathsf{A}$. This is a slight abuse of terminology since the sets A and B may be disjoint, as their input space may be different.

The following theorem shows several claims regarding relations (equivalence or strict containment) between sets of truthful mechanisms. Notice that not only does this theorem show the hierarchy of the different classes, but it also provides notions relevant to a full characterization of truthful mechanisms. For instance, the second claim proves that no mechanism can use any information regarding the location of the agents beyond their ranking, while maintaining truthfulness. In addition, in the claims showing strict containment, the examples in the proofs portray the expressiveness that the additional information gives the mechanism.

Theorem 1. *The following claims hold in the Euclidean metric space $\Re^d$ (for any $\mathsf{d} \in \mathbb{N}$):*

1. *The set of truthful deterministic ranking mechanisms strictly contains the set of truthful deterministic single candidate mechanisms.*

2. *The set of truthful deterministic location mechanisms is equivalent to the set of truthful deterministic ranking mechanisms.*

3. *The set of truthful in expectation randomized ranking mechanisms strictly contains the set of truthful in expectation randomized single candidate mechanisms.*

4. *The set of truthful in expectation randomized location mechanisms strictly contains the set of truthful in expectation randomized ranking mechanisms.*

5. *The set of truthful in expectation randomized single candidate mechanisms strictly contains the set of universally truthful randomized single candidate mechanisms.*

6. *When there are two candidates, the set of truthful in expectation randomized location mechanisms is equivalent to the set of truthful in expectation randomized single candidate mechanisms.*

4 The Spike Mechanism

In the upcoming sections we will prove that both the median mechanism and the random dictator mechanism achieve an approximation ratio of three on $\Re$. However, the source for this ratio in these two cases is different - for median it is due to an instance which is costly for the median agent, while for random dictator it is due to a bad instance for an agent in one of the extremes. The spike mechanism was devised with the objective of being resistant to costly instances of *any* agent.

This section contains foundations needed for the introduction of the spike mechanism, the definition of spike, and the theorem showing that spike achieves an approximation ratio of 2. The reductions in Section 3 show that this positive result extends to ranking and location mechanisms as well. In the entirety of this section, the metric space is $\Re$ and the mechanisms are single candidate mechanisms.

Lemma 2. *Any weighted percentile single candidate (WPSC) mechanism* M *on* $\Re$ *is universally truthful.*

Definition 3 (Spike Mechanism)**.** *Let* $\mathsf{P}(\mathsf{j})$ *be the following function for any* $0 \leq \mathsf{j} \leq \mathsf{n}$*:*

$$\mathsf{P}(\mathsf{j}) = \begin{cases} 0 & \text{if } \mathsf{j} = 0 \\ \frac{j}{2(n-j)} & \text{if } 0 < \mathsf{j} \leq \mathsf{n}/2 \\ 1 - \frac{n}{2j} & \text{if } \mathsf{j} > \mathsf{n}/2 \end{cases}$$

Let $\mathsf{p}_j = \mathsf{P}(\mathsf{j}) - \mathsf{P}(\mathsf{j}-1)$.
The spike mechanism chooses candidate y_i *with probability* $\sum_{j:z_j=y_i} \mathsf{p}_j$.

Equivalently, the spike mechanism chooses voter (j) with probability p_j, and then locates the facility on $\mathsf{a}_{(j)}$.

The mechanism is named after the shape of the function that p_j creates (see Figure 1). We note that the result of the mechanism depends on the number of votes that each candidate received and on the order of the candidate along the line, but not on the distances between the candidates,

Observation 4. *Spike induces a symmetric distribution:* $\forall \mathsf{i} : \mathsf{F}(\mathsf{i}) = 1 - \mathsf{F}(\mathsf{n}-\mathsf{i})$.

Proof. Without loss of generality let $1 \leq \mathsf{i} \leq \mathsf{n}/2$, then:

$$\mathsf{F}(\mathsf{i}) = \frac{\mathsf{i}}{2(\mathsf{n}-\mathsf{i})}$$

$$1 - \mathsf{F}(\mathsf{n}-\mathsf{i}) = 1 - 1 - \frac{\mathsf{n}}{2(\mathsf{n}-\mathsf{i})} = \frac{\mathsf{n}}{2(\mathsf{n}-\mathsf{i})} - \frac{1}{2} = \frac{\mathsf{n}-(\mathsf{n}-\mathsf{i})}{2(\mathsf{n}-\mathsf{i})} = \frac{\mathsf{i}}{2(\mathsf{n}-\mathsf{i})}$$

□

We now define a few terms needed for the proof of the approximation ratio. Recall that y_{opt} is uniquely defined for a location profile $\mathbf{x}$, since ties are broken in favor of the leftmost candidate. Denote the set of borders $\{\mathsf{b}_i\}_{i=1}^{m-1}$ by B.

Definition 5 (Tight profile of $\mathbf{x}$, see Figure 12)**.** *Given a location profile* $\mathbf{x}$*, the profile* $\mathbf{x}'$ *is said to be the tight profile of* $\mathbf{x}$ *if it moves all agents who are not on a border as close as possible to* y_{opt} *within their zones, that is:*

$$\forall \mathsf{i} : \mathsf{x}'_i = \begin{cases} \mathsf{x}_i & \text{if } \mathsf{x}_i \in \mathsf{B} \\ \mathsf{y}_{\text{opt}} & \text{if } \mathsf{x}_i \in \mathcal{Z}_{\text{opt}} \setminus \mathsf{B} \\ \mathsf{b}_j & \text{if } \mathsf{x}_i \in \mathcal{Z}_j \setminus \mathsf{B} \text{ and } \mathsf{j} < \mathsf{y}_{\text{opt}} \\ \mathsf{b}_{j-1} & \text{if } \mathsf{x}_i \in \mathcal{Z}_j \setminus \mathsf{B} \text{ and } \mathsf{j} > \mathsf{y}_{\text{opt}} \end{cases}$$

Definition 6 (Left-compressed profile of $\mathbf{x}$)**.** *Given a tight location profile* $\mathbf{x}$*, a left-compressed profile of* $\mathbf{x}$ *moves all the agents on the leftmost border to their neighboring border on the right, if this border is left of* y_{opt}*. Formally: let the location of the leftmost agent be* $x_1 = b_j$*, then the left-compressed profile of* $\mathbf{x}$ *is:*

$$\forall i : x'_i = \begin{cases} b_{j+1} & \text{if } (x_i = b_j) \wedge (b_{j+1} < y_{\text{opt}}) \\ x_i & \text{otherwise.} \end{cases}$$

Left-compressed profiles can be seen in the transition between Figures 15 and 16, and in Figure 13. Note that the left-compressed profile of a tight profile is also a tight profile. The right-compressed profile of $\mathbf{x}$ is defined in a completely symmetrical fashion.

After compressing location profiles, there are likely to be locations in which there are many agents. We therefore use the following notation: the location profile is written as $\mathbf{x} = \{(\hat{x}_1; n_1); :::; (\hat{x}_k; n_k)\}$, which means that for each $j : 1 \leq j \leq k$, there are n_j agents located at $\hat{x}_j$ (see, e.g, Figure 16).

We now use these definitions to prove the main result of this section:

Theorem 7. *The spike mechanism is universally truthful, and it achieves an approximation ratio of 2 on* $\Re$*.*

Proof. Spike is a WPSC mechanism, so from Lemma 2 it is universally truthful.

The analysis of the approximation ratio is more involved, and is based on backwards induction which follows these steps (see Figures 14 through 17):

1. Figure 14: Start with an arbitrary location profile $\mathbf{x}$, and compute its optimal candidate, y_{opt}.

2. Figure 15: Let $\mathbf{x}^{(1)}$ be the tight profile of $\mathbf{x}$. We show that the transition from $\mathbf{x}$ to $\mathbf{x}^{(1)}$ cannot reduce the approximation ratio (Lemma 8).

3. Figure 16: Let $\mathbf{x}^{(2)}$ be the left-compression of $\mathbf{x}^{(1)}$. We show that if the ratio of $\mathbf{x}^{(2)}$ is not higher than 2, then so is the ratio of $\mathbf{x}^{(1)}$ (Lemma 9).

4. Figure 17: Repeat left and right compressions until we can no longer compress. At this stage, the profile is tight with at most 3 active candidates, and we note this profile $\mathbf{x}^{(3)}$. We show that the approximation ratio of $\mathbf{x}^{(3)}$ is not above 2 (Lemma 10).

Proving these steps is sufficient to complete the proof of the theorem, since in Lemma 10 we show that the ratio of $\mathbf{x}^{(3)}$ is not higher than 2 (the base case). According to Lemma 9, this implies that the ratio of $\mathbf{x}^{(1)}$ (prior to all of the compressions) is also not higher than 2 (the induction steps). Since the ratio of the $\mathbf{x}$ is not higher than that of $\mathbf{x}^{(1)}$, this means that the approximation ratio of $\mathbf{x}$ is not above 2, as needed.

Notice that throughout this process y_{opt} remains the optimal candidate, since it was optimal in the original profile $\mathbf{x}$, and in each step all agents move towards it, so the cost of any other candidate can decrease by no more than what the cost of y_{opt} decreases.

Since spike is a single candidate mechanism and agents may be on borders, truthful reports are not necessarily unique. In cases of ties, we show that the worst-case ratio always occurs when the agents vote for the candidate located farther away from y_{opt} (Lemma 27, whose statement and proof appear in Appendix A.3).

We now present the aforementioned lemmas formally. Their proofs, which are given in the appendix, prove the backwards induction and conclude the proof of the theorem.

Lemma 8. *Let* $\mathbf{x}$ *be an arbitrary location profile on* $\Re$*, let* $\mathbf{x}'$ *be the tight profile of* $\mathbf{x}$ *and let* M *be an arbitrary WPSC mechanism. Then the approximation ratio of* M *given* $\mathbf{x}'$ *is not lower than that of* M *given* $\mathbf{x}$*:*

$$\frac{\mathrm{SC}(\mathsf{M};\mathbf{x})}{\mathrm{SC}(\mathrm{OPT};\mathbf{x})} \leq \frac{\mathrm{SC}(\mathsf{M};\mathbf{x}')}{\mathrm{SC}(\mathrm{OPT};\mathbf{x}')}.$$

The previous lemma holds for any WPSC mechanism, and in particular for spike.

Lemma 9. *Let* $\mathbf{x}$ *be a tight location profile on* $\Re$*, let* $\mathbf{x}'$ *be the left-compressed profile of* $\mathbf{x}$ *and let* S *be the spike mechanism. Then if the approximation ratio of* S *given* $\mathbf{x}'$ *is not higher than 2, then so is that of* S *given* $\mathbf{x}$*:*

$$\frac{\mathrm{SC}(\mathsf{S};\mathbf{x}')}{\mathrm{SC}(\mathrm{OPT};\mathbf{x}')} \leq 2 \Rightarrow \frac{\mathrm{SC}(\mathsf{S};\mathbf{x})}{\mathrm{SC}(\mathrm{OPT};\mathbf{x})} \leq 2$$

Observation 4 shows that the cumulative function defining the spike mechanism is symmetrical, so the lemma can be trivially extended to right-compressions as well.

After reapplying compressions on both sides, the resulting profile has agents in three locations at most (see Figure 17). The last lemma in the proof states that in this final stage, the ratio is not higher than 2:

Lemma 10. *Let* $\mathbf{x}$ *be a tight location profile on* $\Re$ *in which there are at most 3 active candidates:* $\mathsf{y}_{\mathrm{opt}-1} < \mathsf{y}_{\mathrm{opt}} < \mathsf{y}_{\mathrm{opt}+1}$*. The ratio of the spike mechanism* S *given* $\mathbf{x}$ *is not higher than 2:* $\frac{\mathrm{SC}(\mathsf{S};\mathbf{x})}{\mathrm{SC}(\mathrm{OPT};\mathbf{x})} \leq 2$*.*

□

5 Additional Results for Randomized Mechanisms

5.1 Lower Bounds

This section shows lower bounds of randomized mechanisms in different settings. When the network is $\Re$, we present a lower bound of 2 for any truthful in expectation mechanism, even if it is a location mechanism. By the hierarchy from Theorem 1, this lower bound holds for truthful in expectation ranking and single candidate mechanisms as well. Additionally, we show a lower bound of 2 for any randomized ranking mechanism, even when the mechanism need not be truthful (in the non-strategic setting). These results prove that the ratio achieved by spike is tight (see Figure 6).

For more general metric spaces the lower bound changes — In $\Re^{\mathsf{d}}$ we show a lower bound of $3 - \frac{2}{\mathsf{d}+1}$ for truthful single candidate mechanisms. We also present a lower bound of 7=3 for any truthful ranking mechanism in $\Re^2$ (this bound also holds for $\Re^{\mathsf{d}}$, for any $\mathsf{d} > 2$).

Theorem 11. *In the* d *dimensional real space* $\Re^{\mathsf{d}}$*, any truthful in expectation single candidate mechanism has an approximation ratio of at least* $3 - \frac{2}{\mathsf{d}+1}$*.*

Observation 12. *Any truthful in expectation location mechanism has an approximation ratio of at least 2, even on the line.*

Observation 13. *The bound of 3 by random dictator is tight for a general metric space.*

Theorem 14. *In* $\Re^2$*, any truthful in expectation ranking mechanism has an approximation ratio of at least* 7=3*.*

Lemma 15. *No randomized ranking mechanism can achieve an approximation ratio strictly below 2, even if the metric is* $\Re$ *and even if there are no truthfulness requirements from the mechanism (non-strategic case).*

5.2 Upper Bound

We previously showed that spike achieves an approximation ratio of 2 on the line. We now show that for a general metric space, random dictator has a ratio of 3 (ergo, the upper bound is 3). Recall that random dictator locates the facility on vote a_i with probability $1=n$ for all $i \in N$, and achieves an approximation ratio of 2 in the continuous model.

Lemma 16. *On any metric space, random dictator yields a* 3 *approximation of the optimal social cost.*

We note that in contradiction to the continuous model, in our candidate model random dictator is not group-strategyproof (see Appendix A.4.2).

Corollary 17. *The combination of the lower bound in Theorem 11 with the upper bound of random dictator show that for a general metric space the bound of 3 is tight for single candidate mechanisms.*

6 Deterministic Mechanisms on the Line

In the continuous model, choosing the location of the median agent is both truthful and optimal [25]. The following theorem shows that in the candidate model "median" results in a ratio of 3, and that this is the best one can hope for with any deterministic mechanism (even location mechanisms). The proof, as well as the formal definition of the median mechanism, are deferred to Appendix 6.

Theorem 18. *The following claims hold:*

1. *No deterministic truthful mechanism (location, ranking or single candidate) has an approximation ratio strictly below* 3 *for the social cost, even on* $\Re$.
2. *Median is truthful on* $\Re$ *and results in a 3 approximation of the social cost.*

7 Discussion and Open Problems

We defined three types of truthful mechanisms, and showed the relations between these sets of truthful mechanisms. Then, we gave bounds on the approximation ratio of these mechanism types in various settings. In particular, we introduced the spike mechanism, a truthful single candidate mechanism which achieves a [tight] bound of 2 on $\Re$.

We believe that there are many possible manifestations of this setting, though it has barely been investigated, so there is plenty of room for future work, for instance:

- Electing a committee of multiple candidates, i.e., locating multiple facilities.
- Closing the gap in the bounds for ranking and location mechanisms in $\Re^d$.
- More generally, we studied the affects of information deficiency and truthfulness in the context of voting. These affects can be addressed for many additional problems.

Acknowledgements. The work of M. Feldman and I. Golomb was partially supported by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement number 337122.

References

[1] Noga Alon, Michal Feldman, Ariel D Procaccia, and Moshe Tennenholtz. Strategyproof approximation mechanisms for location on networks. *arXiv preprint arXiv:0907.2049*, 2009.

[2] Noga Alon, Michal Feldman, Ariel D Procaccia, and Moshe Tennenholtz. Walking in circles. *Discrete Mathematics*, 310(23):3432–3435, 2010.

[3] Elliot Anshelevich, Onkar Bhardwaj, and John Postl. Approximating optimal social choice under metric preferences. 2014.

[4] Elliot Anshelevich and John Postl. Randomized social choice functions under metric preferences. *arXiv preprint arXiv:1512.07590*, 2015.

[5] Haris Aziz, Florian Brandl, and Felix Brandt. On the incompatibility of efficiency and strategyproofness in randomized social choice. In *AAAI*, pages 545–551. Citeseer, 2014.

[6] Duncan Black. On the rationale of group decision-making. *The Journal of Political Economy*, pages 23–34, 1948.

[7] Anna Bogomolnaia and Hervé Moulin. A new solution to the random assignment problem. *Journal of Economic theory*, 100(2):295–328, 2001.

[8] Craig Boutilier, Ioannis Caragiannis, Simi Haber, Tyler Lu, Ariel D Procaccia, and Or Sheffet. Optimal social choice functions: A utilitarian view. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, pages 197–214. ACM, 2012.

[9] Felix Brandt, Vincent Conitzer, and Ulle Endriss. Computational social choice. *Multi-agent systems*, pages 213–283, 2012.

[10] Ioannis Caragiannis and Ariel D Procaccia. Voting almost maximizes social welfare despite limited communication. *Artificial Intelligence*, 175(9):1655–1671, 2011.

[11] Yukun Cheng, Wei Yu, and Guochuan Zhang. Strategy-proof approximation mechanisms for an obnoxious facility game on networks. *Theoretical Computer Science*, 497:154–163, 2013.

[12] Elad Dokow, Michal Feldman, Reshef Meir, and Ilan Nehama. Mechanism design on discrete lines and cycles. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, pages 423–440. ACM, 2012.

[13] Anthony Downs. An economic theory of political action in a democracy. *The journal of political economy*, pages 135–150, 1957.

[14] Michal Feldman and Yoav Wilf. Strategyproof facility location and the least squares objective. In *Proceedings of the fourteenth ACM conference on Electronic commerce*, pages 873–890. ACM, 2013.

[15] Marcus Fleming. A cardinal concept of welfare. *The Quarterly Journal of Economics*, pages 366–384, 1952.

[16] Dimitris Fotakis and Christos Tzamos. Strategyproof facility location for concave cost functions. In *Proceedings of the fourteenth ACM conference on Electronic commerce*, pages 435–452. ACM, 2013.

[17] Dimitris Fotakis and Christos Tzamos. On the power of deterministic mechanisms for facility location games. *ACM Transactions on Economics and Computation*, 2(4):15, 2014.

[18] Allan Gibbard. Manipulation of voting schemes: a general result. *Econometrica: journal of the Econometric Society*, pages 587–601, 1973.

[19] J Harsanyi. Icardinal welfare, individualistic ethics, and interpersonal comparisons of welfare. *J. Polit. Econ*, 63, 1955.

[20] Harold Hotelling. *Stability in competition*. Springer, 1990.

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

[22] Hervé Moulin. On strategy-proofness and single peakedness. *Public Choice*, 35(4):437–455, 1980.

[23] Noam Nisan. Introduction to mechanism design (for computer scientists). *Algorithmic game theory*, 209:242, 2007.

[24] Ariel D Procaccia and Jeffrey S Rosenschein. The distortion of cardinal preferences in voting. In *Cooperative Information Agents X*, pages 317–331. Springer, 2006.

[25] Ariel D Procaccia and Moshe Tennenholtz. Approximate mechanism design without money. In *Proceedings of the 10th ACM conference on Electronic commerce*, pages 177–186. ACM, 2009.

[26] James Schummer and Rakesh V Vohra. Strategy-proof location on a network. *Journal of Economic Theory*, 104(2):405–428, 2002.

[27] Xin Sui and Craig Boutilier. Approximately strategy-proof mechanisms for (constrained) facility location. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, pages 605–613. International Foundation for Autonomous Agents and Multiagent Systems, 2015.

[28] Xin Sui, Craig Boutilier, and Tuomas W Sandholm. Analysis and optimization of multi-dimensional percentile mechanisms. AAAI, 2013.

A Appendix

A.1 Figures from the Introduction

Figure 2: Relationships between classes of mechanisms(Theorem 1): For deterministic truthful mechanisms, the set of ranking mechanisms strictly contains the set of single candidate (SC) mechanisms, yet the set of location mechanisms is equivalent to the set of ranking mechanisms.
In the randomized case, there is a hierarchy of strict containment in the following order - truthful in expectation (TIE) location mechanisms, TIE ranking mechanisms, TIE single candidate (SC) mechanisms and universally truthful (UT) single candidate (SC) mechanisms. Refer to Section 3 for formal definitions of equivalence and strict containment in our setting.

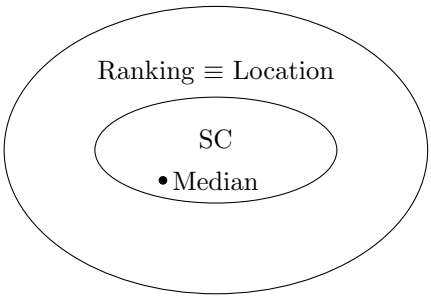


Figure 3: **Deterministic truthful** mechanisms

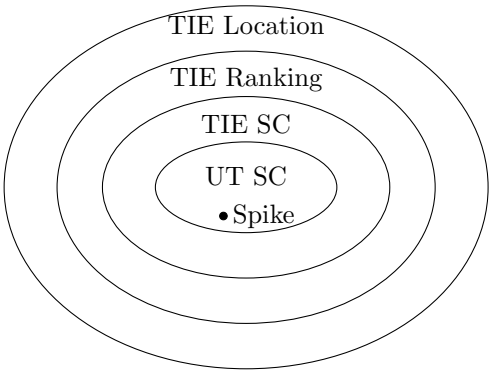


Figure 4: **Randomized** truthful mechanisms

Figure 5: Summary of our results for randomized mechanisms in $\Re^d$. The columns correspond to the truthfulness constraints, while the rows show the information constraints and are further divided into lower and upper bounds. Note that for non-strategic location mechanisms the result is always optimal by definition, since there are neither information nor strategic constraints.
Most of the results here are rather straightforward, except for the upper bounds of $3 - \frac{2}{d+1}$ and of 7=3, which are more involved.

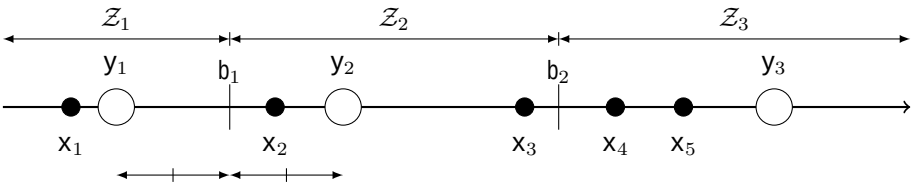
		Strategic		Non-Strategic	
Single Candidate (low information)	LB	$3 - \frac{2}{d+1}$	(Thm. 11)	2	(Lemma 15)
	UB	3	(Lemma 16)	3	(Lemma 16)
Ranking	LB	7=3	(Thm. 14)	2	(Lemma 15)
	UB	3	(Lemma 16)	3	(Lemma 16)
Location (high information)	LB	2	(Obs. 12)	1	
	UB	3	(Lemma 16)	1	

Figure 6: The approximation ratios of mechanisms on the line ($\Re$) in various settings. All the results in the table are tight. In particular, the randomized upper bound of two in the strategic case holds due to the spike mechanism.

	Deterministic		Randomized	
	Strategic	Non-Strategic	Strategic	Non-Strategic
Single Candidate	3*	3†	2‡	2§
Ranking	3*	3†	2‡	2**
Location	3*	1	2‡	1

*LB and UB: Thm. 18
†LB: Thm. 3 from [3], UB: Thm. 18
‡LB: Obs. 12, UB: Thm.7
§LB: Lm. 15, UB: Thm.7
**Lm. 15, UB: Thm. 18

Figure 7: Illustration of candidates (white circles), agents (black circles), candidate borders and candidate zones when the metric space is $\Re$.
For example, in this case the favorite candidate of x_3 is y_2: $\mathsf{y}(\mathsf{x}_3) = \{\mathsf{y}_2\}$. The candidate borders divide the distance between two consecutive candidates exactly in half, for example: $|\mathsf{b}_1 - \mathsf{y}_1| = |\mathsf{y}_2 - \mathsf{b}_1|$.



A.2 Missing Proofs from Section 3

In this part we aim to define reducibility of some mechanism type (single candidate, ranking or location) to some other mechanism type. The definition of the reduction requires a couple of additional definitions, which we will express formally as well as explain intuitively.

Intuitively, we say that a mechanism type $\hat{A}$ is of finer granularity than mechanism type $\hat{B}$, if the information of a true action in $\hat{A}$ can determine a true action in $\hat{B}$. For instance, a location determines a ranking (or several rankings, if a point is on a border), therefore location mechanisms are of finer granularity than ranking mechanisms. Similarly, ranking mechanisms are of finer granularity than single candidate mechanisms, and location mechanisms are also of finer granularity than single candidate mechanisms. However, ranking mechanisms are not of finer granularity than location mechanisms, since a ranking does not determine a location. That is, for a ranking $_1$ there exist different locations $x_1; x_2$ whose true ranking is $_1$. Formally, a mechanism type $\hat{A} : A \to \mathbf{y}$ is of *finer granularity* than mechanism type $\hat{B} : B \to \mathbf{y}$ if for any point x and any $a \in A$ there exists some $b \in B$ such that: if a is a true action of an agent at point x under $\hat{A}$, then b is a true action of point x under $\hat{B}$. In this case we denote $\hat{A} \succ \hat{B}$.

We now utilize this notion of granularity to define consistent functions. Intuitively, we would like to define functions which map between inputs of different mechanism types in a "consistent" manner. For instance, when mapping from rankings (the input to ranking mechanisms) to votes (candidates, that is - the input to single candidate mechanisms), we search for functions which map each ranking to the top candidate in that ranking. When mapping from votes to rankings, we seek functions which map a vote for a candidate to some ranking in which this candidate is first. Formally, given mechanism types $\hat{A} : A \to \mathbf{y}$ and $\hat{B} : B \to \mathbf{y}$ such that $\hat{A} \succ \hat{B}$, a function $f : A \to B$ is called *consistent* if for any point x and for any $a \in A$, then if a is a true action under $\hat{A}$, then $f(a)$ is a true action under $\hat{B}$. Notice that in these cases a function f is unique (except for the cases in which $\hat{A}$ is a location mechanism and x is a point on a border). If $\hat{B} \succ \hat{A}$ then we define $f : A \to B$ to be consistent if for any point x, if $f(a)$ is a true action for x under $\hat{B}$ then a is a true action for x under $\hat{A}$. In these cases the function f is not unique (for example, there are several rankings in which a specific candidate is first).

The function f may be randomized, as long as it is a randomization over deterministic consistent functions. For example, a consistent function f mapping locations (the input of location mechanisms) to candidates (the inputs of single candidate mechanisms) must map every point which is not on a border to their favorite candidate. On the other hand, for a point x on some border, f may randomize the output of x arbitrarily over the set of favorite candidates of x.

A candidate selection mechanism M (whether location, ranking or single candidate) is said to be *reducible* to a candidate selection mechanism M' (location, ranking, or single candidate) if there exists a consistent function f mapping every action profile $\mathbf{a}$, which is an input of M, to some action profile $f(\mathbf{a}) = \mathbf{a}'$ which is the input of M', such that the distribution over the candidates, $M(\mathbf{a})$, is identical to the distribution over candidates $M'(\mathbf{a}')$ (see Figure 8).

For example, every single candidate mechanism, M, is reducible to some ranking mechanism M'. The reduction is as follows: Let f be the consistent function describes previously – it receives a vector of n candidates, and outputs a vector of n rankings (permutations), where for each i, the ith candidate is ranked first in the ith ranking. We choose a ranking mechanism M' which ignores all entries in the rankings but the first, and simulates the single candidate mechanism M on the top entries of the rankings. By definition, M is reducible to M'. Moreover, note that M was an arbitrary single candidate mechanism, so we conclude that indeed every single candidate mechanism is reducible to some ranking mechanism. This

logic further shows that for mechanism types $\hat{\mathsf{A}}; \hat{\mathsf{B}}$ such that $\hat{\mathsf{B}} \succ \hat{\mathsf{A}}$, any mechanism M of type $\hat{\mathsf{A}}$ is reducible to some mechanism M' of type $\hat{\mathsf{B}}$.

As written in Section 3, two sets of mechanisms, S_1 and S_2, are said to be *equivalent* if every $\mathsf{M}_1 \in \mathsf{S}_1$ is reducible to some $\mathsf{M}_2 \in \mathsf{S}_2$, and every $\mathsf{M}_2 \in \mathsf{S}_2$ is reducible to some $\mathsf{M}_1 \in \mathsf{S}_1$. A set of mechanisms S_1 is said to be *strictly contained* in a set of mechanisms S_2 if every mechanism $\mathsf{M}_1 \in \mathsf{S}_1$ is reducible to some mechanism $\mathsf{M}_2 \in \mathsf{S}_2$, yet not every mechanism $\mathsf{M}_2 \in \mathsf{S}_2$ is reducible to some mechanism $\mathsf{M}_1 \in \mathsf{S}_1$. This is a slight abuse of terminology since the sets S_1 and S_2 may be disjoint, as their input space may be different.

The following lemma will be of use in the main theorem of this section.

Lemma 19. *In any metric space, let $\hat{\mathsf{A}}; \hat{\mathsf{B}}$ be mechanism types such that $\hat{\mathsf{B}} \succ \hat{\mathsf{A}}$. Let $\mathsf{S}_1; \mathsf{S}_2$ be the sets of truthful mechanisms of type $\hat{\mathsf{A}}; \hat{\mathsf{B}}$ respectively. Then for any $\mathsf{M}_1 \in \mathsf{S}_1$ there exists some $\mathsf{M}_2 \in \mathsf{S}_2$ such that M_1 is reducible to M_2.*

Proof. By the fact that $\hat{\mathsf{B}}$ is of finer granularity than $\hat{\mathsf{A}}$ and by the definition of the consistent function f, it is eminent that M_1 is reducible to some M_2 of type $\hat{\mathsf{B}}$ since M_2 can completely disregard any input beyond any information in $\hat{\mathsf{A}}$ and simulate M_1 (as explained previously for the example in which a ranking mechanism disregards any candidate except for the top candidate in each ranking).

In order to complete the proof, it is left to show that there exists such a mechanism M_2 which is truthful. Since the only reports which change the outcome of M_2 are consistent with reports which would change the outcome of M_1, then if M_2 weren't truthful this would contradict the truthfulness of M_1. □

Observation 20. *Notice that this reasoning also holds for truthful in expectation mechanisms (that is if $\mathsf{S}_1; \mathsf{S}_2$ are defined as the sets of truthful in expectation mechanisms of types $\hat{\mathsf{A}}; \hat{\mathsf{B}}$).*

We move on to proving the main theorem of this section:

Proof. of Theorem 1: The proof of each claim is given separately:

Claim 21. *The class of truthful deterministic ranking mechanisms strictly contains the class of truthful deterministic single candidate mechanisms.*

Proof. According to Lemma 19 any truthful single candidate mechanism is reducible to some truthful ranking mechanism. We exhibit a deterministic truthful ranking mechanism M which is not reducible to any single candidate mechanism (even on $\Re$). We show this by exhibiting an example in which M acts differently under two ranking profiles which are mapped by any consistent function to the same candidate zone.

Figure 8: A graphic sketch of a mechanism M which is reducible to a mechanism M'.

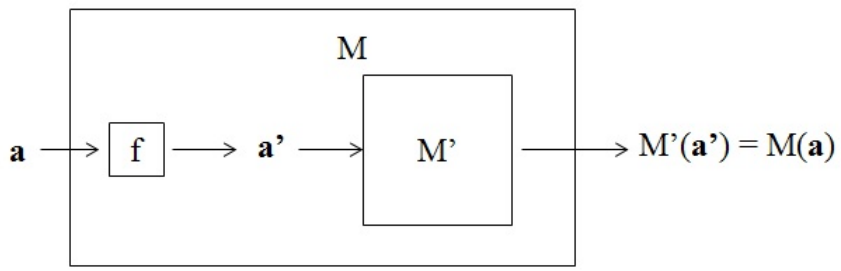
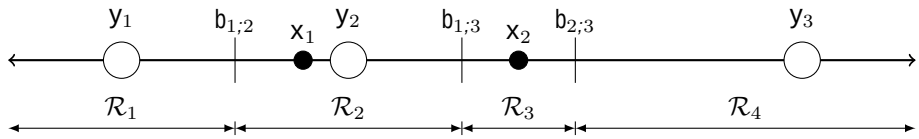


Figure 9: The 4 ranking zones. Every border $\mathsf{b}_{i;j}$ is the midpoint between candidates $\mathsf{y}_i;\mathsf{y}_j$. The points $\mathsf{x}_1;\mathsf{x}_2$ have different rankings – the ranking of x_1 is $_2 = \mathsf{y}_2 \succeq \mathsf{y}_1 \succeq \mathsf{y}_3$ whereas the ranking of x_2 is $_3 = \mathsf{y}_2 \succeq \mathsf{y}_3 \succeq \mathsf{y}_1$. However, both strictly prefer y_2 over any other candidate, therefore for any consistent function f: $\mathsf{f}(\ _2) = \mathsf{f}(\ _3) = \mathsf{y}_2$.



Let there be 3 candidates, and denote the ranking zones $\mathcal{R}_1;\mathcal{R}_2;\mathcal{R}_3;\mathcal{R}_4$, which match the permutations over candidates $_1;\ _2;\ _3;\ _4$ respectively (see Figure 9). Let there be 2 agents. M acts as follows: If $\mathsf{a}_1 \in\ _1 \cup\ _2$ and $\mathsf{a}_2 \in\ _1 \cup\ _2$, then choose y_1. Otherwise, choose y_3.

M is truthful — if y_1 is chosen, then both agents prefer it over y_3 and have no incentive to misreport. If y_3 is chosen, then at least one of the agents is in zones $\mathcal{R}_3$ or $\mathcal{R}_4$ - this agent has no incentive to misreport since she prefers y_3 over y_1. The other agent has no influence over the outcome, therefore also has no incentive to misreport.

M is not reducible to any single candidate mechanism — any consistent function f must map both $_2$ and $_3$ to y_2. However, whilst M acts differently under inputs $(\ _1;\ _2)$ and $(\ _1;\ _3)$, then if there were a reduction, then both of these would have been mapped to the same location $\mathsf{M}'(\mathsf{y}_1;\mathsf{y}_2)$ in contradiction. □

Claim 22. *The set of truthful deterministic location mechanisms is equivalent to the set of truthful deterministic ranking mechanisms.*

Proof. According to Lemma 19, every truthful deterministic ranking mechanism is reducible to some truthful deterministic location mechanism. It is left to show that every truthful deterministic location mechanism M in $\Re^d$ is reducible to a truthful deterministic ranking mechanism M'.

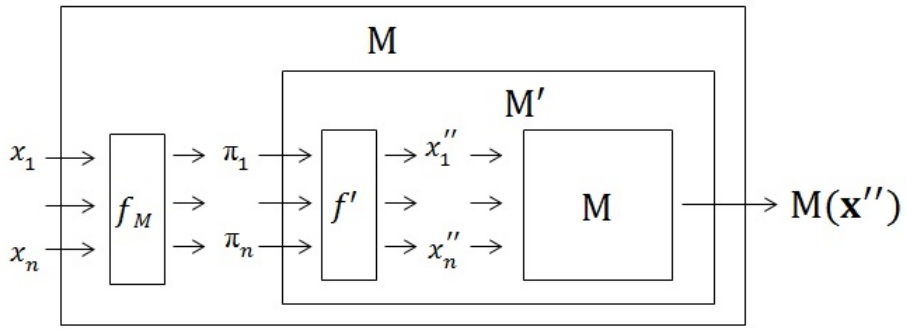
The proof consists of several parts. For an arbitrary location profile $\mathbf{x}$ and an arbitrary truthful deterministic mechanism M, we define a location profile $\mathbf{x}'$ which has no locations on borders, and show that it necessarily holds that $\mathsf{M}(\mathbf{x}) = \mathsf{M}(\mathbf{x}')$. We then define a different location profile $\mathbf{x}''$ and show the same: $\mathsf{M}(\mathbf{x}) = \mathsf{M}(\mathbf{x}'')$. The profile $\mathbf{x}''$ is special in the sense that it is uniquely defined by some ranking profile . Finally, we show that given some input $= \mathsf{f}_M(\mathbf{x})$ a consistent function f_M (a function which depends on M, and will be defined later), there exists a ranking mechanism M' which simulates M on $\mathbf{x}''$ (see Figure 10). The result is a constructive reduction which maintains the same output as the original location mechanism $\mathsf{M}(\mathbf{x})$, as needed.

Let M be an arbitrary truthful deterministic location mechanism, and let $\mathbf{x}$ be an arbitrary location profile. Let $\mathsf{M}(\mathbf{x}) = \mathsf{y}_j$ for some candidate y_j. Denote the ranking borders as B.

Informally, we define $\mathbf{x}'$ as a location profile which moves all agents in $\mathbf{x}$ which are on a border, an infinitesimal distance towards the chosen candidate y_j. The resulting profile $\mathbf{x}'$ has no agents on borders. We now define this formally: let be some small positive number. For all i such that $\mathsf{x}_i \in \mathsf{B}$, let $\sim_i$ be a vector of size in direction $\mathsf{y}_j - \mathsf{x}_i$ [5]. For any i such that $\mathsf{x}_i \notin \mathsf{B}$, let $|\sim_i| = 0$. Let $\mathbf{x}' = (\mathsf{x}_1 +\ _1;\dots\mathsf{x}_n +\ _n)$. We choose an sufficiently small such that for each i, x'_i remains in the ranking zone of x_i.

[5] If $x_\mathrm{i} = y_\mathrm{j}$ then the direction of $\vec{\epsilon}_\mathrm{i}$ can be set arbitrarily.

Figure 10: A graphic sketch of the reductions in the proof of Claim 22.



We now show that $\mathsf{M}(\mathbf{x}) = \mathsf{M}(\mathbf{x}')$ by moving agents from $\mathbf{x}$ to $\mathbf{x}'$ one by one, and showing that if any of these transitions were to change the chosen candidate, this would lead to a violation of truthfulness of M. Let:

$$\begin{aligned}
&\mathsf{w}_0 = (\mathsf{x}_1; \dots; \mathsf{x}_n) = \mathbf{x} \\
&\mathsf{w}_1 = (\mathsf{x}_1'; \mathsf{x}_2; \dots; \mathsf{x}_n) \\
&\dots \\
&\mathsf{w}_i = (\mathsf{x}_1'; \dots; \mathsf{x}_i'; \mathsf{x}_{i+1}; \dots; \mathsf{x}_n) \\
&\dots \\
&\mathsf{w}_n = (\mathsf{x}_1'; \dots; \mathsf{x}_n') = \mathbf{x}'
\end{aligned}$$

Assume towards a contradiction that $\mathsf{M}(\mathsf{w}_0) \neq \mathsf{M}(\mathsf{w}_n)$. Let i be the minimal index such that $\mathsf{M}(\mathsf{w}_{i-1}) \neq \mathsf{M}(\mathsf{w}_i)$. It is known that $\mathsf{M}(\mathsf{w}_{i-1}) = \mathsf{M}(\mathbf{x}) = \mathsf{y}_j$. Let $\mathsf{M}(\mathsf{w}_i) = \mathsf{y}_l$. There are two options:

- If $|\mathsf{y}_l - \mathsf{x}_i| < |\mathsf{y}_j - \mathsf{x}_i|$, then in profile w_{i-1}, x_i has an incentive to misreport to x_i'.
- If $|\mathsf{y}_l - \mathsf{x}_i| \geq |\mathsf{y}_j - \mathsf{x}_i|$, then $|\mathsf{y}_j - \mathsf{x}_i'| = |\mathsf{y}_j - \mathsf{x}_i| - \quad < |\mathsf{x}_i' - \mathsf{y}_l|$. Therefore in profile w_i, x_i' has an incentive to misreport to x_i.

Hence, $\mathsf{M}(\mathbf{x}) = \mathsf{M}(\mathbf{x}')$ as needed.

Intuitively, we create the location profile $\mathbf{x}''$ by moving all agents in $\mathbf{x}'$ to some specific point within their ranking zone. Since $\mathbf{x}'$ contained no agents on borders, each agent in $\mathbf{x}'$ is located in exactly one ranking zone, hence $\mathbf{x}''$ is well defined. We now define this formally: For any ranking zone $\mathcal{R}_i$ such that $\mathcal{R}_i \setminus \mathsf{B} \neq \emptyset$, let $\hat{\mathsf{x}}_i$ be some point in $\mathcal{R}_i \setminus \mathsf{B}$ (for instance, the centroid of the ranking zone $\mathcal{R}_i$)[6]. Denote the ranking zone which contains x_i' as $\mathcal{R}_j$. For all $\mathsf{i} \in \mathsf{N}$, let $\mathsf{x}_i'' = \hat{\mathsf{x}}_j$. Let $\mathbf{x}'' = (\mathsf{x}_1''; \dots; \mathsf{x}_n'')$.

We now show that $\mathsf{M}(\mathbf{x}') = \mathsf{M}(\mathbf{x}'')$ in a similar fashion as we showed that $\mathsf{M}(\mathbf{x}) = \mathsf{M}(\mathbf{x}')$

[6]We can safely disregard ranking zones which do not have any points which are not on a border, as no point in $\mathbf{x}'$ will be in such a zone, since $\mathbf{x}'$ does not contain any points on borders.

previously. Let:

$$\begin{aligned}
&\mathsf{h}_0 = (\mathsf{x}'_1; \dots; \mathsf{x}'_n) = \mathbf{x}' \\
&\mathsf{h}_1 = (\mathsf{x}''_1; \mathsf{x}'_2; \dots; \mathsf{x}'_n) \\
&\dots \\
&\mathsf{h}_i = (\mathsf{x}''_1; \dots; \mathsf{x}''_i; \mathsf{x}'_{i+1}; \dots; \mathsf{x}'_n) \\
&\dots \\
&\mathsf{h}_n = (\mathsf{x}''_1; \dots; \mathsf{x}''_n) = \mathbf{x}''
\end{aligned}$$

Assume towards a contradiction that $\mathsf{M}(\mathsf{h}_0) \neq \mathsf{M}(\mathsf{h}_n)$. Let i be the minimal index such that $\mathsf{M}(\mathsf{h}_{i-1}) \neq \mathsf{M}(\mathsf{h}_i)$. Let $\mathsf{M}(\mathsf{h}_i) = \mathsf{y}_m$. Since $\mathsf{x}'_i; \mathsf{x}''_i$ are in the same ranking zone and not on a border, there are two options:

- If $|\mathsf{y}_m - \mathsf{x}'_i| < |\mathsf{y}_j - \mathsf{x}'_i|$, then in profile h_{i-1}, x'_i has an incentive to misreport to x''_i.
- If $|\mathsf{y}_m - \mathsf{x}'_i| > |\mathsf{y}_j - \mathsf{x}'_i|$, then it also holds that $|\mathsf{y}_m - \mathsf{x}''_i| > |\mathsf{y}_j - \mathsf{x}''_i|$, and in profile h_i, x''_i has an incentive to misreport to x'_i.

Therefore, $\mathsf{M}(\mathbf{x}') = \mathsf{M}(\mathbf{x}'') \Rightarrow \mathsf{M}(\mathbf{x}) = \mathsf{M}(\mathbf{x}'')$.

It is left to show that it is possible to perform the "nested reductions" as shown in Figure 10. Let f_M be the consistent function which breaks ties just like M does. That is, f_M simulates M on input $\mathbf{x}$, finds the candidate y_j and breaks ties in favor of rankings closer to y_j. The output of f_M is a ranking profile denoted by . Given , there exists some M' that simulates M on $\mathbf{x}''$[7] — let f' be a consistent function which maps every ranking $_i$ (consistent with ranking zone $\mathcal{R}_i$) to the point $\hat{\mathsf{x}}_i$. Therefore, such a reduction exists, and the output is $\mathsf{M}(\mathbf{x}'') = \mathsf{M}(\mathbf{x})$.

□

Claim 23. *The set of truthful in expectation ranking mechanisms strictly contains the set of truthful in expectation single candidate mechanisms.*

Proof. As shown in Observation 20, any truthful in expectation single candidate mechanism is reducible to some truthful in expectation ranking mechanism. The proof of Claim 21 exhibits a truthful in expectation ranking mechanism which is not reducible to any single candidate mechanism. □

Claim 24. *The set of truthful in expectation randomized location mechanisms strictly contains the set of truthful in expectation randomized ranking mechanisms.*

Proof. As shown in Observation 20, any truthful in expectation ranking mechanism is reducible to some truthful in expectation location mechanism.

We will show a truthful in expectation location mechanism M which is not reducible to any ranking mechanism: Let there be 3 candidates at points $\mathsf{y}_1 = 0; \mathsf{y}_2 = 3; \mathsf{y}_3 = 4$. M acts as follows:

Choose an agent i uniformly at random. Choose the candidates with the following probabilities:

$$\mathsf{M}(\mathbf{a}) = \begin{cases} \mathsf{y}_1 = 1/3; \mathsf{y}_2 = 1/3; \mathsf{y}_3 = 1/3 \text{ if } \mathsf{a}_i \leq 1 \\ \mathsf{y}_1 = 1/4; \mathsf{y}_2 = 1/2; \mathsf{y}_3 = 1/4 \text{ otherwise.} \end{cases}$$

[7] To avoid a circular definition, one can think of M' simulating a location mechanism which acts precisely like M does.

M is not reducible to any ranking mechanism — any consistent function f must map both points $\mathsf{x}_1 = 0{:}75$ and $\mathsf{x}_2 = 1{:}25$ to $_1 = \mathsf{y}_1 \succ \mathsf{y}_2 \succ \mathsf{y}_3$. However, mechanism M treats these two inputs differently.

It is left to show that M is truthful in expectation. We do so by assessing all possibilities of misreports. Obviously, the mechanism is not affected by any agents except the one who was chosen. Since there are only two possible outcomes, it is sufficient to compare truthful reports a_i with misreports a'_i such that a'_i changes the outcome. Let $\mathbf{a} = (\mathsf{a}_i; \mathbf{a}_{-i})$ and $\mathbf{a}' = (\mathsf{a}'_i; \mathbf{a}_{-i})$.

- If $\mathsf{x}_i \leq 0$ it holds that

$$\begin{aligned} \text{cost}_{\mathsf{x}_i}(\mathsf{M}; \mathbf{a}) &= \frac{1}{3}[-\mathsf{x}_i + (3 - \mathsf{x}_i) + (4 - \mathsf{x}_i)] = -\mathsf{x}_i + 7{=}3 \\ \text{cost}_{\mathsf{x}_i}(\mathsf{M}; \mathbf{a}') &= \frac{1}{4}[-\mathsf{x}_i + (4 - \mathsf{x}_i)] + \frac{1}{2}(3 - \mathsf{x}_i) = -\mathsf{x}_i + 5{=}2: \end{aligned}$$

  Therefore: $\text{cost}_{\mathsf{x}_i}(\mathsf{M}; \mathbf{a}) \leq \text{cost}_{\mathsf{x}_i}(\mathsf{M}; \mathbf{a}')$.

- If $0 < \mathsf{x}_i < 3$ it holds that

  - If the outcome is $\mathsf{y}_1 = 1{=}3; \mathsf{y}_2 = 1{=}3; \mathsf{y}_3 = 1{=}3$, the cost of agent i is:

$$\frac{1}{3}[\mathsf{x}_i + (3 - \mathsf{x}_i) + (4 - \mathsf{x}_i)] = -\mathsf{x}_i{=}3 + 7{=}3:$$

  - If the result is $\mathsf{y}_1 = 1{=}4; \mathsf{y}_2 = 1{=}2; \mathsf{y}_3 = 1{=}4$ the cost of agent i is:

$$\frac{1}{4}(\mathsf{x}_i + 4 - \mathsf{x}_i) + \frac{1}{2}(3 - \mathsf{x}_i) = -\mathsf{x}_i{=}2 + 5{=}2:$$

    It holds that $-\mathsf{x}_i{=}3 + 7{=}3 \geq -\mathsf{x}_i{=}2 + 5{=}2 \Leftrightarrow \mathsf{x}_i \geq 1$. Therefore the first outcome is preferable to agents for which $0 < \mathsf{x}_i \leq 1$ and the second is better for agents for which $1 < \mathsf{x}_i < 3$, and the mechanism is truthful in expectation in this interval.

- If $3 \leq \mathsf{x}_i < 4$ it holds that

$$\begin{aligned} \text{cost}_{\mathsf{x}_i}(\mathsf{M}; \mathbf{a}) &= \frac{1}{4}(\mathsf{x}_i + 4 - \mathsf{x}_i) + \frac{1}{2}(\mathsf{x}_i - 3) = \mathsf{x}_i{=}2 - 1{=}2 \\ \text{cost}_{\mathsf{x}_i}(\mathsf{M}; \mathbf{a}') &= \frac{1}{3}[\mathsf{x}_i + (\mathsf{x}_i - 3) + (4 - \mathsf{x}_i)] = \mathsf{x}_i{=}3 + 1{=}3: \end{aligned}$$

  Therefore: $\text{cost}_{\mathsf{x}_i}(\mathsf{M}; \mathbf{a}) \leq \text{cost}_{\mathsf{x}_i}(\mathsf{M}; \mathbf{a}') \Leftrightarrow \mathsf{x}_i \leq 5$, therefore agent i cannot benefit from misreporting.

- If $\mathsf{x}_i \geq 4$ it holds that

$$\begin{aligned} \text{cost}_{\mathsf{x}_i}(\mathsf{M}; \mathbf{a}) &= \frac{1}{4}(\mathsf{x}_i + \mathsf{x}_i - 4) + \frac{1}{2}(\mathsf{x}_i - 3) = \mathsf{x}_i - 5{=}2 \\ \text{cost}_{\mathsf{x}_i}(\mathsf{M}; \mathbf{a}') &= \frac{1}{3}[\mathsf{x}_i + (\mathsf{x}_i - 3) + (\mathsf{x}_i - 4)] = \mathsf{x}_i - 7{=}3: \end{aligned}$$

  Therefore: $\text{cost}_{\mathsf{x}_i}(\mathsf{M}; \mathbf{a}) \leq \text{cost}_{\mathsf{x}_i}(\mathsf{M}; \mathbf{a}')$.

□

Claim 25. *The set of truthful in expectation randomized single candidate mechanisms strictly contains the set of universally truthful randomized single candidate mechanisms.*

Proof. Any universally truthful single candidate mechanism is reducible to a truthful in expectation single candidate mechanism using the identity function f (which is consistent).

We exhibit a truthful in expectation (TIE) single candidate mechanism M which is not reducible to any universally truthful mechanism. Let there be 2 candidates. M chooses an agent i uniformly at random, and chooses a_i with probability 0.9 and the other candidate $\mathbf{y} \setminus \mathsf{a}_i$ with probability 0.1.

M is truthful in expectation since for any agent j, if they are chosen, they are better off receiving their favorite candidate with probability 0.9 than with probability 0.1. M is not universally truthful, since for each agent i there exist cases in which reporting truthfully would lead to choosing their less favorite candidate, while there exist cases in which reporting non-truthfully would lead to choosing the favorite candidate. Clearly, no composition with a consistent function f can transform M to a universally truthful mechanism (for instance, let $\mathsf{a}_i = \mathsf{y}_j$ for some j. From consistency, $\mathsf{f}(\mathsf{a}_i) = \mathsf{y}_j$, so f does not change the outcome at all). □

Claim 26. *When there are two candidates, the set of truthful in expectation randomized location mechanisms is equivalent to the set of truthful in expectation randomized single candidate mechanisms.*

Proof. As shown in Observation 20, any truthful in expectation single candidate mechanism is reducible to some truthful in expectation location mechanism. This also holds for single candidate mechanisms for two candidates. We now show that any truthful in expectation location mechanism with two candidates is reducible to some truthful in expectation single candidate mechanism. The proof follows similar lines as the proof of Claim 22.

Let $\mathbf{x}$ be an arbitrary location profile, let M be an arbitrary truthful in expectation location mechanism, and let B be the border between y_1 and y_2. Define $\mathbf{x}'$ as the location profile which moves all agents which are not on borders to their favorite candidate, that is:

$$\mathsf{x}'_i = \begin{cases} \mathsf{y}_1 & \text{if } \mathsf{x}_i \in \mathcal{Z}_1 \setminus \mathsf{B} \\ \mathsf{y}_2 & \text{if } \mathsf{x}_i \in \mathcal{Z}_2 \setminus \mathsf{B} \\ \mathsf{x}_i & \text{if } \mathsf{x}_i \in \mathsf{B} \end{cases}$$

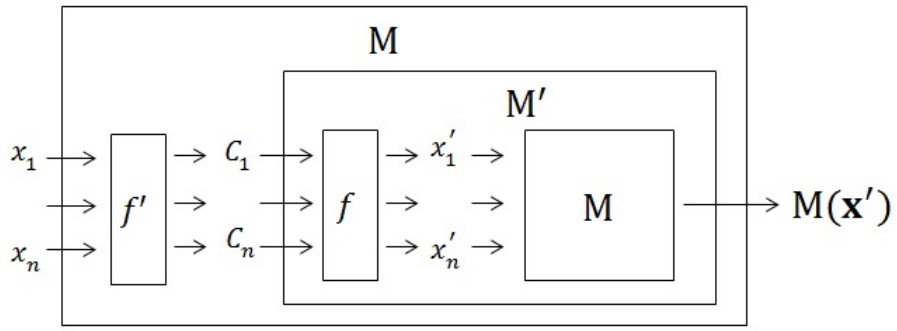
We now show that $\mathsf{M}(\mathbf{x}) = \mathsf{M}(\mathbf{x}')$ by using a hybrid argument. Define:

$$\begin{aligned} \mathsf{w}_0 &= (\mathsf{x}_1; \dots; \mathsf{x}_n) = \mathbf{x} \\ \mathsf{w}_1 &= (\mathsf{x}'_1; \mathsf{x}_2; \dots; \mathsf{x}_n) \\ &\dots \\ \mathsf{w}_i &= (\mathsf{x}'_1; \dots; \mathsf{x}'_i; \mathsf{x}_{i+1}; \dots; \mathsf{x}_n) \\ &\dots \\ \mathsf{w}_n &= (\mathsf{x}'_1; \dots; \mathsf{x}'_n) = \mathbf{x}' \end{aligned}$$

Assume towards a contradiction that $\mathsf{M}(\mathsf{w}_0) \neq \mathsf{M}(\mathsf{w}_n)$. Then there exists some index j such that $\Pr[\mathsf{M}(\mathsf{w}_j) = \mathsf{y}_1] \neq \Pr[\mathsf{M}(\mathsf{w}_{j-1}) = \mathsf{y}_1]$. If this is the case then necessarily $\mathsf{x}_j \in \mathsf{B}$ since that would imply that w_{j-1} and w_j are precisely the same profile. Assume without loss of generality that $\Pr[\mathsf{M}(\mathsf{w}_j) = \mathsf{y}_1] > \Pr[\mathsf{M}(\mathsf{w}_{j-1}) = \mathsf{y}_1]$. There are 2 options:

- If $\mathsf{x}_j; \mathsf{x}'_j \in \mathcal{Z}_1$, then under location profile w_j, agent j has an incentive to misreport to x'_j.
- If $\mathsf{x}_j; \mathsf{x}'_j \in \mathcal{Z}_2$, then under location profile w_{j-1}, agent j has an incentive to misreport to x_j.

Figure 11: A graphic sketch of the reductions in the proof of Claim 26.



Therefore is necessarily holds that $\mathsf{M}(\mathbf{x}) = \mathsf{M}(\mathbf{x}')$.

We now use $\mathbf{x}'$ to show the reduction: Let f be a function which maps single candidate profiles to location profiles, by mapping each vote to candidate y_i to location y_i. This function is clearly consistent. Let M' be a single candidate mechanism which receives a single candidate profile, translates it to a location mechanism using the consistent function f, and then simulates M on the output of f (see Figure 11).

For cases in which no agent is on the border, then the function f' mapping location profiles to single candidate profiles is uniquely defined, and it holds that $\mathsf{M}(\mathbf{x}) = \mathsf{M}'(\mathsf{f}'(\mathbf{x})) = \mathsf{M}(\mathsf{f}(\mathsf{f}'(\mathbf{x})))$. It is left to show that in cases of agents on borders, there exists some consistent function f' which breaks ties in the same manner that M does.

Let $\mathbf{x}'^{(1)}$ be a location profile with n_1 agents at y_1, n_2 agents somewhere on the border B and n_3 agents at y_2. In short, we note $\mathbf{x}'^{(1)} = (\mathsf{n}_1; \mathsf{n}_2; \mathsf{n}_3)$. We remark that $\mathbf{x}'^{(1)}$ is a general location profile, after moving agents to their favorite candidates. Using these amounts, define the following two location profiles:

- Let $\mathbf{x}'^{(2)}$ be the profile in which there are $\mathsf{n}_1 + \mathsf{n}_2$ agents at y_1 and n_3 agents at y_2 (that is $\mathbf{x}'^{(2)} = (\mathsf{n}_1 + \mathsf{n}_2; 0; \mathsf{n}_3)$).
- Let $\mathbf{x}'^{(3)}$ be the profile in which there are n_1 agents at y_1 and $\mathsf{n}_2 + \mathsf{n}_3$ agents at y_2 (that is $\mathbf{x}'^{(3)} = (\mathsf{n}_1; 0; \mathsf{n}_2 + \mathsf{n}_3)$).

Let $\mathsf{p}_1 = \Pr[\mathsf{M}(\mathbf{x}'^{(1)}) = \mathsf{y}_1]$, and similarly: $\mathsf{p}_2 = \Pr[\mathsf{M}(\mathbf{x}'^{(2)}) = \mathsf{y}_1]$ and $\mathsf{p}_3 = \Pr[\mathsf{M}(\mathbf{x}'^{(3)}) = \mathsf{y}_1]$. Under these definitions, we show that: $\mathsf{p}_3 \leq \mathsf{p}_1 \leq \mathsf{p}_2$:

- $\mathsf{p}_1 \leq \mathsf{p}_2$: Start with the profile $\mathbf{x}'^{(1)}$, and move agents on the border one by one to y_1. If in each step the probability of choosing y_1 does not decrease then $\mathsf{p}_1 \leq \mathsf{p}_2$ as needed. Otherwise, there exists a profile $\hat{\mathbf{x}} = (\mathsf{n}_i; \mathsf{n}_j; \mathsf{n}_3)$ for which the probability of choosing y_1 is smaller than in the profile with $\hat{\mathbf{x}}' = (\mathsf{n}_i - 1; \mathsf{n}_j + 1; \mathsf{n}_3)$. If this were the case, then the agents on y_1 in profile $\hat{\mathbf{x}}$ would benefit from misreporting to the point on a border, in contradiction to truthfulness.
- $\mathsf{p}_3 \leq \mathsf{p}_1$ is proved in the exact symmetrical manner, by moving agents from B to y_2 one by one.

By definition, a consistent function can map agents on borders to either of the two candidates, and can also choose any probabilities over the two agents. Therefore, for any $0 \leq \mathsf{q} \leq 1$, there exists a consistent function which takes n agents on the border and maps all of them to y_1 with probability q and maps all of them to y_2 with probability $1 - \mathsf{q}$. For

Figure 12: A tight profile – The first figure shows the original profile $\mathbf{x}$. For instance, b_1 is the inner border of x_1 since it is it's neighbor border closer to y_{opt}. The second figure shows the tight profile of $\mathbf{x}$ (where the arrows display the movements).

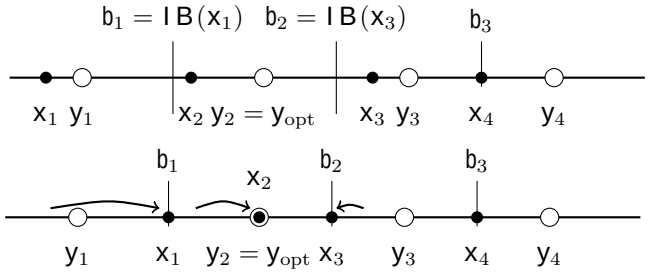
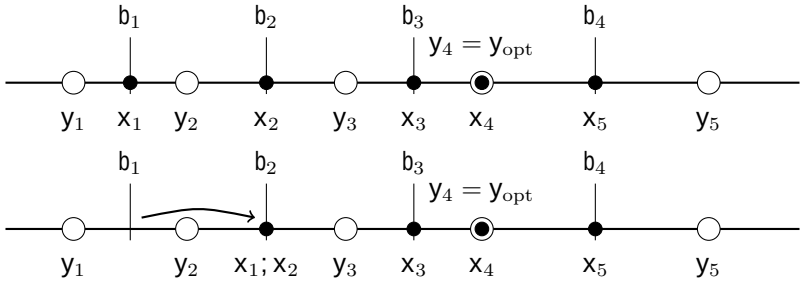


Figure 13: A left-compressed profile – The first figure shows the original tight profile $\mathbf{x}$, and the second shows the left-compressed profile of $\mathbf{x}$ (after x_1 moves to b_2).



any p_1, we choose the consistent function f' which uses a q such that $p_2 \cdot q + p_3 \cdot (1-q) = p_1$. Under this function f', the reduction does not change the outcome of the mechanism, as needed.

□

□

A.3 Missing Proofs from Section 4

We start by showing that any weighted percentile single candidate mechanism M is universally truthful:

Proof. of Lemma 2: [27] show that percentile mechanisms are truthful on $\Re$. WPSC mechanisms take a given distribution over truthful mechanisms, and are therefore universally truthful.

□

We now move on to the lemmas regarding the spike mechanism, from Theorem 7.

Figure 14: Initial state: A general location profile $\mathbf{x}$

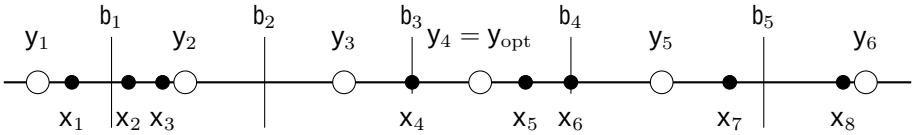


Figure 15: $\mathbf{x}^{(1)}$: The tight profile of $\mathbf{x}$, using the notation in which there are n_i agents at point $\hat{\mathsf{x}}_i$.

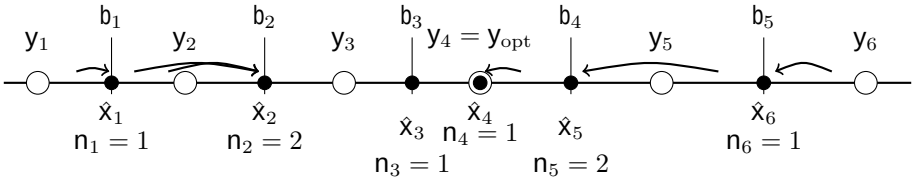


Figure 16: $\mathbf{x}^{(2)}$: The left-compressed profile of $\mathbf{x}^{(1)}$ (moves n_1 agents from b_1 to b_2).

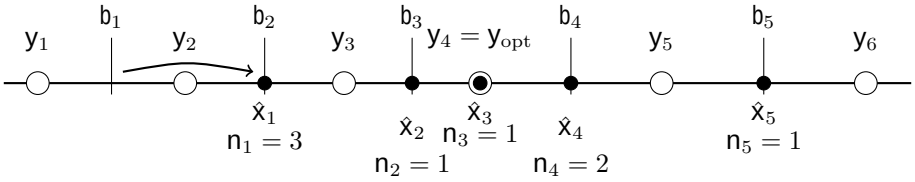
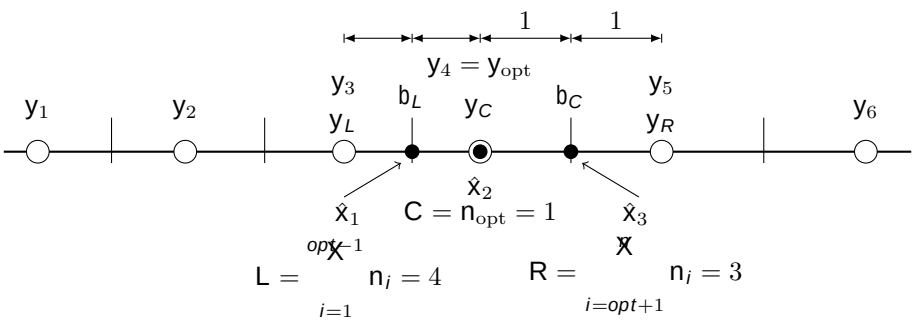


Figure 17: $\mathbf{x}^{(3)}$: The final profile after reapplying left and right compressions on $\mathbf{x}^{(2)}$ repeatedly. The active candidates are denoted y_L, y_C and y_R, the number of agents as L, C and R respectively, and we scale the distances by $\frac{b_C - y_C}{2} = 1$ and $\frac{y_C - b_L}{2} =$.



Proof. of Lemma 8: . For any candidate j define $p_j = \Pr[M(\mathbf{a}) = y_j]$ and $p'_j = \Pr[M(\mathbf{a}') = y_j]$. Define Δ_j as the difference in the cost of candidate j under profile $\mathbf{x}$ and their cost under $\mathbf{x}'$, that is: $\Delta_j = \sum_{i=1}^{n} |y_j - x_i| - \sum_{i=1}^{n} |y_j - x'_i|$. Since $\mathbf{x}'$ was defined by moving all agents towards y_{opt}, then: $\forall j : \Delta_{\text{opt}} \geq \Delta_j$. As noted previously, this means that y_{opt} remains the optimal candidate under $\mathbf{a}'$. According to Lemma 27, the worst-case ratio occurs when all agents on borders vote outwards (farther from y_{opt}), so the votes (and therefore the probabilities) given $\mathbf{a}'$ remain the same as under $\mathbf{a}$: $\forall j : p_j = p'_j$.

We now assess the approximation ratio given profile $\mathbf{x}'$:

$$\text{SC}(\text{OPT}; \mathbf{x}') = \sum_i |y_{\text{opt}} - x'_i| = \text{SC}(\text{OPT}; \mathbf{x}) - \Delta_{\text{opt}}$$

The cost of the spike mechanism given $\mathbf{x}'$ is:

$$\begin{aligned}
\text{SC}(M; \mathbf{x}') &= \sum_j p'_j \cdot \left(\sum_i |y_j - x'_i| \right) \\
&= \sum_j p_j \cdot \left(\sum_i |y_j - x_i| - \Delta_j \right) \\
&\geq \sum_j p_j \cdot \left(\sum_i |y_j - x_i| - \Delta_{\text{opt}} \right) \\
&= \text{SC}(M; \mathbf{x}) - \Delta_{\text{opt}}
\end{aligned}$$

Therefore, the approximation ratio is:

$$\frac{\text{SC}(M; \mathbf{x}')}{\text{SC}(\text{OPT}; \mathbf{x}')} = \frac{\text{SC}(M; \mathbf{x}')}{\text{SC}(\text{OPT}; \mathbf{x}) - \Delta_{\text{opt}}} \geq \frac{\text{SC}(M; \mathbf{x}) - \Delta_{\text{opt}}}{\text{SC}(\text{OPT}; \mathbf{x}) - \Delta_{\text{opt}}} \geq \frac{\text{SC}(M; \mathbf{x})}{\text{SC}(\text{OPT}; \mathbf{x})}$$

The last inequality holds since $\frac{\text{SC}(M; \mathbf{x}) - \Delta_{\text{opt}}}{\text{SC}(\text{OPT}; \mathbf{x}) - \Delta_{\text{opt}}} \geq 1$ and since $\Delta_{\text{opt}} \geq 0$. □

Proof. of Lemma 9. We use the same notation as in Figures 15 and 16 where there are n_i agents at point $\hat{x}_i$.

Let $\Delta = \text{SC}(\text{OPT}; \mathbf{x}) - \text{SC}(\text{OPT}; \mathbf{x}') = n_1 \cdot (\hat{x}_2 - \hat{x}_1) > 0$.

It is sufficient to show that in the worst case scenario (according to Lemma 27, when all agents on borders vote outwards) it holds that $\text{SC}(S; \mathbf{x}) - \text{SC}(S; \mathbf{x}') \leq 2\Delta$, since that would imply:

$$\frac{\text{SC}(S; \mathbf{x})}{\text{SC}(\text{OPT}; \mathbf{x})} = \frac{\text{SC}(S; \mathbf{x})}{\text{SC}(\text{OPT}; \mathbf{x}') + \Delta} \leq \frac{\text{SC}(S; \mathbf{x}') + 2\Delta}{\text{SC}(\text{OPT}; \mathbf{x}') + \Delta} \leq \frac{2 \cdot \text{SC}(\text{OPT}; \mathbf{x}')) + 2\Delta}{\text{SC}(\text{OPT}; \mathbf{x}') + \Delta} = 2$$

The last inequality holds since the approximation ratio of spike given $\mathbf{x}'$ is not greater than 2.

Denote the probabilities as follows: $p_i = \Pr[S(\mathbf{a}) = y_i]$, $p'_i = \Pr[S(\mathbf{a}') = y_i]$. Similarly, the costs of the candidates are denoted by: $c_i = \text{SC}(y_i; \mathbf{x}) = \sum_{j=1}^{n} |x_j - y_i|$ and $c'_i = \text{SC}(y_i; \mathbf{x}') = \sum_{j=1}^{n} |x'_j - y_i|$. The worst-case probabilities given profile $\mathbf{x}'$ are:

$$p'_i = \begin{cases} 0 & \text{if } i = 1 \\ p_1 + p_2 & \text{if } i = 2 \\ p_i & \text{if } i \geq 3 \end{cases}$$

Define $\quad = n_1\left(|\hat{x}_2 - y_2| - |y_2 - \hat{x}_1|\right)$, so the costs given profile $\mathbf{x}'$ are:

$$c'_i = \begin{cases} c_2 + n_1\left(|x_2 - y_2| - |y_2 - x_1|\right) = c_2 + & \text{if } i = 2 \\ c_i - \Delta & \text{if } i \geq 3 \end{cases}$$

Therefore, the difference in the cost is:

$$\begin{aligned} SC(S;\mathbf{x}) - SC(S;\mathbf{x}') &= \sum_i \left(p_i c_i - p'_i c'_i\right) \\ &= p_1 c_1 + p_2 c_2 - (p_1 + p_2)\left(c_2 + \quad\right) + \sum_{i\geq 3} p_i\left[c_i - (c_i - \Delta)\right] \\ &= p_1\left(c_1 - c_2 - \quad\right) - p_2 \cdot \quad + \sum_{i\geq 3} p_i \Delta \\ &= p_1\left(c_1 - c_2 - \quad\right) - p_2 \cdot \quad + (1 - p_1 - p_2)\,\Delta \end{aligned}$$

Due to the triangle inequality:

$$\begin{aligned} & |\hat{x}_2 - y_2| \leq |\hat{x}_2 - \hat{x}_1| + |\hat{x}_1 - y_2| \\ \Leftrightarrow \quad & = n_1\left(|\hat{x}_2 - y_2| - |\hat{x}_1 - y_2|\right) \leq n_1|\hat{x}_2 - \hat{x}_1| = |\Delta| \end{aligned}$$

That is: $\quad \leq |\Delta|$.

Therefore:

$$\begin{aligned} & p_1\left(c_1 - c_2 - \quad\right) - p_2 \cdot \quad + (1 - p_1 - p_2)\,\Delta \\ \leq\ & p_1\left(c_1 - c_2 + \Delta\right) + p_2 \cdot \Delta + (1 - p_1 - p_2)\,\Delta \\ =\ & p_1\left(c_1 - c_2\right) + \Delta \end{aligned}$$

Also $c_1 - c_2 = (n - n_1)|y_2 - y_1|$, so together:

$$SC(S;\mathbf{x}) - SC(S;\mathbf{x}') \quad \leq \quad p_1\left[(n - n_1)|y_2 - y_1|\right] + \Delta$$

To conclude the proof it is left to show that $p_1\left[(n - n_1)|y_2 - y_1|\right] \leq \Delta = |\hat{x}_2 - \hat{x}_1|n_1$. Since $\frac{|y_2 - y_1|}{2} = |y_2 - \hat{x}_1| < |\hat{x}_2 - \hat{x}_1|$, it is sufficient to show that: $p_1(n - n_1) \leq \frac{n_1}{2}$. S is a spike mechanism, so we can compute the value of p_1 (and as a result, of $p_1(n - n_1)$):

- If $n_1 \leq n{=}2$ then: $p_1 = \frac{n_1}{2(n - n_1)} \Rightarrow p_1(n - n_1) = \frac{n_1}{2(n - n_1)}(n - n_1) = n_1{=}2$.
- If $n_1 > n{=}2$ then: $p_1 = 1{:}5 - \frac{n}{2n_1} = \frac{3n_1 - n}{2n_1}$, so:

$$\begin{aligned} p_1(n - n_1) &= \frac{3n_1 - n}{2n_1}(n - n_1) = \frac{-n^2 - 3n_1^2 + 4nn_1}{2n_1} = \frac{-n^2 - 4n_1^2 + 4nn_1}{2n_1} + \frac{n_1^2}{2n_1} \\ &= \frac{-(n - 2n_1)^2}{2n_1} + \frac{n_1}{2} \leq \frac{n_1}{2}: \end{aligned}$$

This concludes the proof of the lemma.

□

Proof. of Lemma 10. We use the notations of the location of the left, center and right candidates in the following manner $y_L = y_{\text{opt}-1}$, $y_C = y_{\text{opt}}$, $y_R = y_{\text{opt}+1}$, and the number of agents in $b_L; y_C; b_C$ as $L; C; R$ respectively (see Figure 17). Denote the probabilities of choosing the candidates as $p_L = \Pr(S(\mathbf{x}) = y_L)$, $p_C = \Pr(S(\mathbf{x}) = y_C)$, $p_R = \Pr(S(\mathbf{x}) = y_R)$.

Also, without loss of generality, the distances can be scaled such that $\mathsf{b}_C - \mathsf{y}_C = 1$. Define: $= \mathsf{y}_C - \mathsf{b}_L$.

According to Lemma 27, the worst-case ratio occurs when the agents at $\mathsf{y}_L; \mathsf{y}_R$ choose $\mathsf{y}_L; \mathsf{y}_R$ respectively.

The costs of the different candidates are:

$$\begin{aligned}
\mathrm{SC}(\mathsf{y}_L; \mathbf{x}) &= \quad \cdot (\mathsf{L} + 2\mathsf{C} + 2\mathsf{R}) + \mathsf{R} \\
\mathrm{SC}(\mathsf{y}_C; \mathbf{x}) &= \mathsf{L} \cdot \quad + \mathsf{R} = \mathrm{SC}(\mathrm{OPT}; \mathbf{x}) \\
\mathrm{SC}(\mathsf{y}_R; \mathbf{x}) &= \mathsf{L} \quad + (2\mathsf{L} + 2\mathsf{C} + \mathsf{R})
\end{aligned}$$

Due to the definition of the spike mechanism, the proof is broken into two parts:

1. The median agent is on y_C
2. The median agent is on b_L

Note that the last option (in which the median agent is on b_C) is identical to the second case due to symmetry, therefore proving for these two cases is sufficient.

In the first case, the median agent is at the center, therefore $\mathsf{L} < \mathsf{C} + \mathsf{R}$ and $\mathsf{R} < \mathsf{L} + \mathsf{C}$, and from the definition of the spike mechanism:

$$\begin{aligned}
\mathsf{p}_L &= \frac{\mathsf{L}}{2(\mathsf{C} + \mathsf{R})} \\
\mathsf{p}_R &= \frac{\mathsf{R}}{2(\mathsf{C} + \mathsf{L})} \\
\mathsf{p}_C &= 1 - \mathsf{p}_L - \mathsf{p}_R = 1 - \frac{\mathsf{L}}{2(\mathsf{C} + \mathsf{R})} - \frac{\mathsf{R}}{2(\mathsf{C} + \mathsf{L})}
\end{aligned}$$

Therefore the ratio is:

$$\begin{aligned}
\frac{\mathrm{SC}(\mathsf{M}; \mathbf{x})}{\mathrm{SC}(\mathrm{OPT}; \mathbf{x})} &= \frac{\mathsf{p}_L \mathrm{SC}(\mathsf{y}_L; \mathbf{x}) + \mathsf{p}_C \mathrm{SC}(\mathsf{y}_C; \mathbf{x}) + \mathsf{p}_R \mathrm{SC}(\mathsf{y}_R; \mathbf{x})}{\mathrm{SC}(\mathsf{y}_C; \mathbf{x})} \\
&= \frac{\mathsf{p}_L \mathrm{SC}(\mathsf{y}_L; \mathbf{x}) + \mathsf{p}_R \mathrm{SC}(\mathsf{y}_R; \mathbf{x})}{\mathrm{SC}(\mathsf{y}_C; \mathbf{x})} + \mathsf{p}_C \\
&= \frac{1}{\mathsf{L} \quad + \mathsf{R}} \quad \frac{\mathsf{L}(\quad (\mathsf{L} + 2\mathsf{C} + 2\mathsf{R}) + \mathsf{R})}{2(\mathsf{C} + \mathsf{R})} + \frac{\mathsf{R}(\mathsf{L} \quad + (2\mathsf{L} + 2\mathsf{C} + \mathsf{R}))}{2(\mathsf{L} + \mathsf{C})} \\
&+ \quad 1 - \frac{\mathsf{L}}{2(\mathsf{C} + \mathsf{R})} - \frac{\mathsf{R}}{2(\mathsf{L} + \mathsf{C})} \\
&= \frac{\mathsf{L}(\mathsf{L} \quad + 2\mathsf{C} \quad + 2\mathsf{R} \quad + \mathsf{R})}{2(\mathsf{C} + \mathsf{R})(\mathsf{L} \quad + \mathsf{R})} + \frac{\mathsf{R}(\mathsf{L} \quad + 2\mathsf{L} + 2\mathsf{C} + \mathsf{R})}{2(\mathsf{L} + \mathsf{C})(\mathsf{L} \quad + \mathsf{R})} \\
&+ \quad 1 - \frac{\mathsf{L}(\mathsf{L} \quad + \mathsf{R})}{2(\mathsf{C} + \mathsf{R})(\mathsf{L} \quad + \mathsf{R})} - \frac{\mathsf{R}(\mathsf{L} \quad + \mathsf{R})}{2(\mathsf{L} + \mathsf{C})(\mathsf{L} \quad + \mathsf{R})} \\
&= \quad 1 + \frac{\mathsf{L}(\mathsf{L} \quad + 2\mathsf{C} \quad + 2\mathsf{R} \quad + \mathsf{R})}{2(\mathsf{C} + \mathsf{R})(\mathsf{L} \quad + \mathsf{R})} - \frac{\mathsf{L}(\mathsf{L} \quad + \mathsf{R})}{2(\mathsf{C} + \mathsf{R})(\mathsf{L} \quad + \mathsf{R})} \\
&+ \quad \frac{\mathsf{R}(\mathsf{L} \quad + 2\mathsf{L} + 2\mathsf{C} + \mathsf{R})}{2(\mathsf{L} + \mathsf{C})(\mathsf{L} \quad + \mathsf{R})} - \frac{\mathsf{R}(\mathsf{L} \quad + \mathsf{R})}{2(\mathsf{L} + \mathsf{C})(\mathsf{L} \quad + \mathsf{R})} \\
&= \quad 1 + \frac{\mathsf{L}(2\mathsf{C} \quad + 2\mathsf{R} \quad)}{2(\mathsf{C} + \mathsf{R})(\mathsf{L} \quad + \mathsf{R})} + \frac{\mathsf{R}(2\mathsf{L} + 2\mathsf{C})}{2(\mathsf{L} + \mathsf{C})(\mathsf{L} \quad + \mathsf{R})} \\
&= \quad 1 + \frac{\mathsf{L}}{\mathsf{L} \quad + \mathsf{R}} + \frac{\mathsf{R}}{\mathsf{L} \quad + \mathsf{R}} = 2
\end{aligned}$$

Therefore the ratio cannot exceed 2 in the first case.

In the second case the median agent is at b_L and the probabilities are:

$$\begin{aligned}
\mathsf{p}_L &= 1{:}5 - \frac{\mathsf{L}+\mathsf{C}+\mathsf{R}}{2\mathsf{L}} = 1 - \frac{\mathsf{C}+\mathsf{R}}{2\mathsf{L}} \\
\mathsf{p}_R &= \frac{\mathsf{R}}{2(\mathsf{C}+\mathsf{L})} \\
\mathsf{p}_C &= 1 - \mathsf{p}_L - \mathsf{p}_R = 1 - \quad 1 - \frac{\mathsf{C}+\mathsf{R}}{2\mathsf{L}} \quad - \frac{\mathsf{R}}{2(\mathsf{C}+\mathsf{L})} \\
&= \frac{\mathsf{C}+\mathsf{R}}{2\mathsf{L}} - \frac{\mathsf{R}}{2(\mathsf{C}+\mathsf{L})}
\end{aligned}$$

Therefore the approximation ratio is:

$$\begin{aligned}
& \frac{\mathrm{SC}(\mathsf{M};\mathbf{x})}{\mathrm{SC}(\mathrm{OPT};\mathbf{x})} = \frac{\mathsf{p}_L \mathrm{SC}(\mathsf{y}_L;\mathbf{x}) + \mathsf{p}_R \mathrm{SC}(\mathsf{y}_R;\mathbf{x})}{\mathrm{SC}(\mathsf{y}_C;\mathbf{x})} + \mathsf{p}_C \\
= \; & \frac{1}{\mathsf{L} \quad + \mathsf{R}} \quad 1 - \frac{\mathsf{C}+\mathsf{R}}{2\mathsf{L}} \quad (\mathsf{L} \quad + 2\mathsf{C} \quad + 2\mathsf{R} \quad + \mathsf{R}) + \frac{\mathsf{R}}{2(\mathsf{L}+\mathsf{C})}(\mathsf{L} \quad + 2\mathsf{L} + 2\mathsf{C} + \mathsf{R}) \\
+ \; & \frac{\mathsf{C}+\mathsf{R}}{2\mathsf{L}} - \frac{\mathsf{R}}{2(\mathsf{C}+\mathsf{L})} \\
= \; & \frac{1}{\mathsf{L} \quad + \mathsf{R}} \quad \mathsf{L} \quad + \mathsf{R} + 2\mathsf{C} \quad + 2\mathsf{R} \quad - \frac{(\mathsf{C}+\mathsf{R})(\mathsf{L} \quad + 2\mathsf{C} \quad + 2\mathsf{R} \quad + \mathsf{R})}{2\mathsf{L}} + \frac{\mathsf{R}(2\mathsf{L} + 2\mathsf{C} + \mathsf{L} \quad + \mathsf{R})}{2(\mathsf{L}+\mathsf{C})} \\
+ \; & \frac{\mathsf{C}+\mathsf{R}}{2\mathsf{L}} - \frac{\mathsf{R}}{2(\mathsf{C}+\mathsf{L})} \\
= \; & 1 + \frac{1}{\mathsf{L} \quad + \mathsf{R}} \quad 2 \quad (\mathsf{C}+\mathsf{R}) - \frac{(\mathsf{C}+\mathsf{R})\mathsf{L}}{2\mathsf{L}} - \frac{(\mathsf{C}+\mathsf{R})(2\mathsf{C} \quad + 2\mathsf{R} \quad + \mathsf{R})}{2\mathsf{L}} + \mathsf{R} + \frac{\mathsf{R}(\mathsf{L} \quad + \mathsf{R})}{2(\mathsf{L}+\mathsf{C})} \\
+ \; & \frac{\mathsf{C}+\mathsf{R}}{2\mathsf{L}} - \frac{\mathsf{R}}{2(\mathsf{L}+\mathsf{C})} \\
= \; & 1 + \frac{1}{\mathsf{L} \quad + \mathsf{R}} \quad \frac{3 \quad (\mathsf{C}+\mathsf{R})}{2} - \frac{(\mathsf{C}+\mathsf{R})(2\mathsf{C} \quad + 2\mathsf{R} \quad + \mathsf{R})}{2\mathsf{L}} + \frac{2\mathsf{R}}{2} \quad + \frac{\mathsf{R}}{2(\mathsf{L}+\mathsf{C})} \\
+ \; & \frac{\mathsf{C}+\mathsf{R}}{2\mathsf{L}} - \frac{\mathsf{R}}{2(\mathsf{L}+\mathsf{C})} \\
= \; & 1 + \frac{3 \quad (\mathsf{C}+\mathsf{R}) + 2\mathsf{R}}{2(\mathsf{L} \quad + \mathsf{R})} - \frac{(\mathsf{C}+\mathsf{R})(2\mathsf{C} \quad + 2\mathsf{R} \quad + \mathsf{R})}{2\mathsf{L}(\mathsf{L} \quad + \mathsf{R})} + \frac{(\mathsf{C}+\mathsf{R})}{2\mathsf{L}}
\end{aligned}$$

Now, in order to show this is a 2 approximation:

$$\begin{aligned}
& 1 + \frac{3 \quad (\mathsf{C}+\mathsf{R}) + 2\mathsf{R}}{2(\mathsf{L} \quad + \mathsf{R})} - \frac{(\mathsf{C}+\mathsf{R})(2\mathsf{C} \quad + 2\mathsf{R} \quad + \mathsf{R})}{2\mathsf{L}(\mathsf{L} \quad + \mathsf{R})} + \frac{(\mathsf{C}+\mathsf{R})}{2\mathsf{L}} \leq 2 \\
\Leftrightarrow \quad & \frac{3 \quad (\mathsf{C}+\mathsf{R}) + 2\mathsf{R}}{2(\mathsf{L} \quad + \mathsf{R})} - \frac{(\mathsf{C}+\mathsf{R})(2\mathsf{C} \quad + 2\mathsf{R} \quad + \mathsf{R})}{2\mathsf{L}(\mathsf{L} \quad + \mathsf{R})} + \frac{(\mathsf{C}+\mathsf{R})(\mathsf{L} \quad + \mathsf{R})}{2\mathsf{L}(\mathsf{L} \quad + \mathsf{R})} \leq 1
\end{aligned}$$

And by multiplying both sides by the common denominator $2\mathsf{L}(\mathsf{L} \quad + \mathsf{R})$:

$$\begin{aligned}
& \mathsf{L}[3 \quad (\mathsf{C}+\mathsf{R}) + 2\mathsf{R}] + (\mathsf{C}+\mathsf{R})(-2\mathsf{C} \quad - 2\mathsf{R} \quad - \mathsf{R} + \mathsf{L} \quad + \mathsf{R}) \leq 2\mathsf{L}(\mathsf{L} \quad + \mathsf{R}) \\
\Leftrightarrow \quad & \mathsf{L}[3 \quad (\mathsf{C}+\mathsf{R}) + 2\mathsf{R}] + (\mathsf{C}+\mathsf{R})(\mathsf{L} \quad - 2\mathsf{C} \quad - 2\mathsf{R} \quad) \leq 2\mathsf{L}(\mathsf{L} \quad + \mathsf{R}) \\
\Leftrightarrow \quad & \mathsf{L}[3 \quad (\mathsf{C}+\mathsf{R})] + \quad (\mathsf{C}+\mathsf{R})(\mathsf{L} - 2\mathsf{C} - 2\mathsf{R}) \leq 2\mathsf{L} \cdot \mathsf{L}
\end{aligned}$$

Since is always positive, it is possible to divide both sides by :

$$
\begin{aligned}
& \mathsf{L}[3(\mathsf{C}+\mathsf{R})] + (\mathsf{C}+\mathsf{R})(\mathsf{L}-2\mathsf{C}-2\mathsf{R}) \leq 2\mathsf{L}^2 \\
\Leftrightarrow\ & 3\mathsf{LC} + 3\mathsf{LR} + \mathsf{LC} - 2\mathsf{C}^2 - 2\mathsf{CR} + \mathsf{LR} - 2\mathsf{CR} - 2\mathsf{R}^2 \leq 2\mathsf{L}^2 \\
\Leftrightarrow\ & 0 \leq 2\mathsf{L}^2 + 2\mathsf{C}^2 + 2\mathsf{R}^2 - 4\mathsf{LC} - 4\mathsf{LR} + 4\mathsf{CR} \\
\Leftrightarrow\ & 0 \leq \mathsf{L}^2 + \mathsf{C}^2 + \mathsf{R}^2 - 2\mathsf{LC} - 2\mathsf{LR} + 2\mathsf{CR} \\
\Leftrightarrow\ & 0 \leq (\mathsf{L}-\mathsf{C}-\mathsf{R})^2
\end{aligned}
$$

This term is indeed always non-negative, so this concludes the proof. □

The following lemma shows that for any WPSC mechanism, when agents on borders vote outwards (farther from the optimal candidate) the approximation ratio cannot decrease.

Lemma 27. *Let* $\mathbf{x}$ *be an arbitrary location profile, and let agent* i *be on a border* b_j *such that* $\mathsf{y}_j < \mathsf{x}_i = \mathsf{b}_j < \mathsf{y}_{j+1} \leq \mathsf{y}_{\text{opt}}$.
Let $\mathbf{a}_1 = (\mathsf{a}_i = \mathsf{y}_j; \mathbf{a}_{-i})$*, let* $\mathbf{a}_2 = (\mathsf{a}_i = \mathsf{y}_{j+1}; \mathbf{a}_{-i})$*, and let* M *be some WPSC mechanism. Then* $\mathrm{SC}(\mathsf{M}; \mathbf{x}; \mathbf{a}_1) \geq \mathrm{SC}(\mathsf{M}; \mathbf{x}; \mathbf{a}_2)$.

Proof. Let $\mathsf{p}_i = \Pr[\mathsf{M}(\mathbf{a}_1) = \mathsf{y}_i]$ and $\mathsf{q}_i = \Pr[\mathsf{M}(\mathbf{a}_2) = \mathsf{y}_i]$. According to the definition of WPSC mechanisms, the change of vote only affects the probabilities of candidates y_j and y_{j+1}, that is: $\forall \mathsf{k} \neq \mathsf{j}; \mathsf{j}+1$: $\mathsf{p}_k = \mathsf{q}_k$. Denote: $\mathsf{p}_j = \mathsf{q}_j +$ and $\mathsf{p}_{j+1} +$ $= \mathsf{q}_{j+1}$ for some > 0. Therefore:

$$
\begin{aligned}
\mathrm{SC}(\mathsf{M}; \mathbf{x}; \mathbf{a}_1) - \mathrm{SC}(\mathsf{M}; \mathbf{x}; \mathbf{a}_2) &= \sum_k \mathsf{p}_k \cdot \mathrm{SC}(\mathsf{y}_k; \mathbf{x}) - \sum_k \mathsf{q}_k \cdot \mathrm{SC}(\mathsf{y}_k; \mathbf{x}) \\
&= \quad [\mathrm{SC}(\mathsf{y}_j; \mathbf{x}) - \mathrm{SC}(\mathsf{y}_{j+1}; \mathbf{x})]
\end{aligned}
$$

Therefore it is sufficient to show that $\mathrm{SC}(\mathsf{y}_j; \mathbf{x}) \geq \mathrm{SC}(\mathsf{y}_{j+1}; \mathbf{x})$. We define the cost function for any point on the line: $\mathsf{f}(\mathsf{x}) = \sum_k |\mathsf{x} - \mathsf{x}_k|$. By definition, for any candidate y_l: $\mathrm{SC}(\mathsf{y}_l; \mathbf{x}) = \mathsf{f}(\mathsf{y}_l)$, therefore we need to show that $\mathsf{f}(\mathsf{y}_j) \geq \mathsf{f}(\mathsf{y}_{j+1})$.

Clearly, $\mathsf{f}(\mathsf{x})$ is single-peaked, with a peak at the median $\mathsf{x}_{\lceil n=2 \rceil}$, since moving in any direction away from the median only increases the distance to at least half of the agents.

We check the different cases:

- If $\mathsf{y}_{\text{opt}} \leq \mathsf{x}_{\lceil n=2 \rceil}$: Then $\mathsf{y}_j < \mathsf{y}_{j+1} \leq \mathsf{y}_{\text{opt}} \leq \mathsf{x}_{\lceil n=2 \rceil}$. From the fact that f's peak is at $\mathsf{x}_{\lceil n=2 \rceil}$, it holds that $\mathsf{f}(\mathsf{y}_j) \geq \mathsf{f}(\mathsf{y}_{j+1})$.
- If $\mathsf{y}_{\text{opt}} > \mathsf{x}_{\lceil n=2 \rceil}$:
  - If $\mathsf{y}_{j+1} = \mathsf{y}_{\text{opt}}$: Then the proof is concluded by definition of optimality.
  - If $\mathsf{y}_{j+1} < \mathsf{y}_{\text{opt}}$, then by definition of optimality and from the fact that f's peak is at $\mathsf{x}_{\lceil n=2 \rceil}$, it holds that $\mathsf{y}_j < \mathsf{y}_{j+1} < \mathsf{x}_{\lceil n=2 \rceil} < \mathsf{y}_{\text{opt}}$. Therefore: $\mathsf{f}(\mathsf{y}_j) \geq \mathsf{f}(\mathsf{y}_{j+1})$.

□

The proof also holds for the symmetrical case in which $\mathsf{y}_{\text{opt}} \leq \mathsf{y}_j < \mathsf{b}_j = \mathsf{x}_i < \mathsf{y}_{j+1}$.

A.4 Missing Proofs from Section 5

A.4.1 Lower Bounds

We begin with a lemma which is used in the proof of Theorem 11

Lemma 28. *For any truthful in expectation ranking mechanism* M *in any metric space, let* $\mathsf{b}_{i;j}$ *be the border between ranking zones* $\mathcal{R}_i;\mathcal{R}_j$*. Let* $_i;\ _j$ *be the rankings consistent with* $\mathcal{R}_i;\mathcal{R}_j$ *respectively. Let agent* l *be located on this border, that is:* $\mathsf{x}_l \in \mathsf{b}_{i;j}$*.*
Then the cost at point x_l *remains the same whether the agent reports* $_i$ *or* $_j$*, that is:*

$$\text{cost}_{\mathsf{x}_l}(\mathsf{M};(\mathsf{a}_l = \ _i;\mathbf{a}_{-l})) = \text{cost}_{\mathsf{x}_l}(\mathsf{M};(\mathsf{a}_l = \ _j;\mathbf{a}_{-l}))$$

Proof. of Lemma 28. Proof via contradiction. Assume $\text{cost}_{\mathsf{x}_l}(\mathsf{M};(\mathsf{a}_l = \ _j;\mathbf{a}_{-l})) = \text{cost}_{\mathsf{x}_l}(\mathsf{M};(\mathsf{a}_l = \ _i;\mathbf{a}_{-l})) + \ $ for some $\ > 0$. Let there be an agent k located in ranking zone $\mathcal{R}_j$ such that $|\mathsf{x}_k - \mathsf{x}_l| = \ < \frac{}{2}$.

Then agent k has an incentive to misreport:

$$\begin{aligned}
\text{cost}_{\mathsf{x}_k}(\mathsf{M};(\mathsf{a}_k = \ _j;\mathbf{a}_{-k})) &\geq \text{cost}_{\mathsf{x}_l}(\mathsf{M};(\mathsf{a}_k = \ _j;\mathbf{a}_{-k})) - \\
&= \text{cost}_{\mathsf{x}_l}(\mathsf{M};(\mathsf{a}_k = \ _i;\mathbf{a}_{-k})) + \ - \\
&> \text{cost}_{\mathsf{x}_l}(\mathsf{M};(\mathsf{a}_k = \ _i;\mathbf{a}_{-k})) + \\
&\geq \text{cost}_{\mathsf{x}_k}(\mathsf{M};(\mathsf{a}_k = \ _i;\mathbf{a}_{-k}))
\end{aligned}$$

The transitions in the first and last rows are due to the triangle inequality (for any location the mechanism may choose), the second row holds by the assumption, and the third row holds since $\ < \frac{}{2}$.

Agent k has an incentive to misreport, contradicting the assumption and completing the proof. □

Observation 29. *Lemma 28 also holds for single candidate mechanisms.*

Proof. of Theorem 11. Let there be $\mathsf{d}+1$ candidates, located on the vertices of a regular simplex H (all $\mathsf{d}+1$ vertices are equally distanced from one another). Let there be $\mathsf{d}+1$ agents, and let M be an arbitrary truthful in expectation single candidate mechanism.

Let $\mathbf{x}$ be the profile in which each agent i is located precisely on candidate y_i. Therefore $\mathbf{a} = (\mathsf{y}_1;\mathsf{y}_2 :: \mathsf{y}_{d+1})$ is the only truthful single candidate profile for $\mathbf{x}$. Denote the probability of choosing candidate i as $\mathsf{p}_i(\mathsf{a})$, that is: $\mathsf{p}_i(\mathsf{a}) = \Pr(\mathsf{M}(\mathbf{a}) = \mathsf{y}_i)$. Clearly there exists some candidate which is chosen by M with probability at least $\frac{1}{d+1}$. Assume without loss of generality that this candidate is y_{d+1}, that is: $\mathsf{p}_{d+1}(\mathsf{a}) \geq \frac{1}{d+1}$.

We move on to define another location profile, $\mathbf{x}'$, which is also consistent with the single candidate profile $\mathbf{a}$. Let H' be the regular simplex in which candidates $\mathsf{y} \setminus \mathsf{y}_{d+1}$ are on the vertices. Let P be the point with equal distance to all d vertices in H' (H' is a regular simplex, so such a point necessarily exists). Denote this distance as t. However, this distance is different from the distance from P to y_{d+1}: $|\mathsf{P} - \mathsf{y}_{d+1}| = \mathsf{u} \neq \mathsf{t}$. Let $\mathbf{x}'$ be the profile in which there are k agents at P and one agent at y_{d+1} (see Figure 18).

According to Lemma 28 (which also holds for single candidate mechanisms, as explained), the cost of an agent at point P should not change under any truthful vote, that is for any vote $\mathsf{y}_j : 1 \leq \mathsf{j} \leq \mathsf{d}$. In particular, this holds when any agent on point P votes for candidate y_1. We make use of this observation several times by changing the votes for each of the points at P to y_1, one at a time, such that the final single candidate profile is $\mathbf{a}' = (\mathsf{y}_1;\mathsf{y}_1 :: \mathsf{y}_1;\mathsf{y}_{d+1})$ (d agents vote for y_1, one agent votes for y_{d+1}).. Due to Observation 29, the cost of point P must remain the same throughout these transitions, that is: $\text{cost}_P(\mathsf{M};\mathsf{a}') = \text{cost}_P(\mathsf{M};\mathsf{a})$.

Therefore:

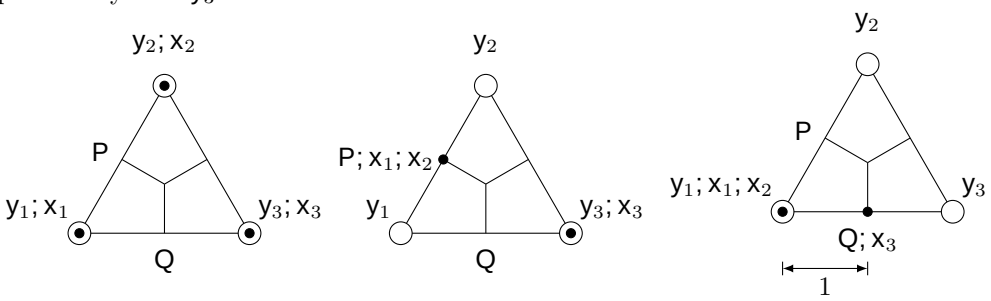
$$\begin{aligned}
& \text{cost}_P(\mathsf{M};\mathsf{a}) = \text{cost}_P(\mathsf{M};\mathsf{a}') \\
\Rightarrow\ & \mathsf{u}\cdot \mathsf{p}_{d+1}(\mathsf{a}) + \mathsf{t}\cdot(1-\mathsf{p}_{d+1}(\mathsf{a})) = \mathsf{u}\cdot \mathsf{p}_{d+1}(\mathsf{a}') + \mathsf{t}\cdot(1-\mathsf{p}_{d+1}(\mathsf{a}')) \\
\Rightarrow\ & \mathsf{t} + (\mathsf{u}-\mathsf{t})\cdot \mathsf{p}_{d+1}(\mathsf{a}) = \mathsf{t} + (\mathsf{u}-\mathsf{t})\cdot \mathsf{p}_{d+1}(\mathsf{a}') \\
\Rightarrow\ & \mathsf{p}_{d+1}(\mathsf{a}') = \mathsf{p}_{d+1}(\mathsf{a}) \\
\Rightarrow\ & \mathsf{p}_{d+1}(\mathsf{a}') \geq \frac{1}{\mathsf{d}+1}
\end{aligned}$$

Denote the midpoint between y_1 and y_{d+1} as Q. Without loss of generality, scale the distances such that $|\mathsf{y}_1 - \mathsf{Q}| = |\mathsf{Q} - \mathsf{y}_{d+1}| = 1$. Examine the following location profile $\mathbf{x}'' = (\mathsf{y}_1;\mathsf{y}_1 :: \mathsf{y}_1;\mathsf{Q})$, which is also consistent with the single candidate profile $\mathbf{a}'$. In this case the cost of y_1, which is the optimal candidate, is: $\text{SC}(\mathsf{y}_1;\mathbf{x}'') = 1$. The cost of y_{d+1} is $\text{SC}(\mathsf{y}_{d+1};\mathbf{x}'') = \mathsf{d}\cdot 2 + 1 = 2\mathsf{d}+1$. Therefore the approximation ratio of M is at least:

$$\begin{aligned}
\frac{\text{SC}(\mathsf{M};\mathbf{x}'')}{\text{SC}(\text{OPT};\mathbf{x}'')} &= \mathsf{p}_{d+1}(\mathsf{a}')(2\mathsf{d}+1) + (1-\mathsf{p}_{d+1}(\mathsf{a}'))(1) \\
&= 2\mathsf{d}\cdot \mathsf{p}_{d+1}(\mathsf{a}') + 1 \\
&\geq (2\mathsf{d})\frac{1}{\mathsf{d}+1} + 1 \\
&= \frac{2\mathsf{d}+2-2}{\mathsf{d}+1} + 1 = 3 - \frac{2}{\mathsf{d}+1}
\end{aligned}$$

□

Figure 18: An illustration of the proof of Theorem 11 for the case of $\mathsf{d} = 2$. The three candidates are on the vertices of an equilateral triangle (a regular simplex with 3 vertices). The lines within the triangle denote the borders between a pair of candidates. The simplex H' is the line between $\mathsf{y}_1;\mathsf{y}_2$, and its midpoint is P.
These three figures, from left to right, show the dynamics of the proof (location profiles $\mathbf{x}$,$\mathbf{x}'$ and $\mathbf{x}''$ respectively). The key observation is that the agents at point P are at equal distance to all candidates except y_3, therefore the transitions do not change the probability that y_3 be chosen.



Proof. of Observation 12. Let there be two candidates on the line. According to Theorem 1 (6), any truthful in expectation location mechanism is equivalent to a single candidate mechanism. One can apply the same proof as in Theorem 11 for the case of $\mathsf{d} = 1$ (in this case, both point P and point Q are the midpoint between the two candidates), to achieve a lower bound of $3 - \frac{2}{d+1} = 2$. □

Proof. Proof of Theorem 14. Let there be 3 candidates located such that they form an equilateral triangle, and let M be a truthful in expectation ranking mechanism. Let $\mathbf{a} = (\mathsf{a}_1; \mathsf{a}_2; \mathsf{a}_3)$ be the following ranking profile:

$$\begin{aligned}
\mathsf{a}_1 &= \mathsf{y}_1 \succeq \mathsf{y}_2 \succeq \mathsf{y}_3 \\
\mathsf{a}_2 &= \mathsf{y}_2 \succeq \mathsf{y}_3 \succeq \mathsf{y}_1 \\
\mathsf{a}_3 &= \mathsf{y}_3 \succeq \mathsf{y}_1 \succeq \mathsf{y}_2
\end{aligned}$$

Let $\mathbf{x}$ be some location profile consistent with $\mathbf{a}$ (Figure 19). Denote $\mathsf{p}_i(\mathbf{a}) = \Pr[\mathsf{M}(\mathbf{a}) = \mathsf{y}_i]$. From symmetry, there exists some candidate chosen with probability at least 1=3. Assume without loss of generality that this is candidate y_3, that is: $\mathsf{p}_3(\mathbf{a}) \geq 1{=}3$.

Let P_1 be a point such that $|\mathsf{P}_1 - \mathsf{y}_1| = |\mathsf{P}_1 - \mathsf{y}_2| = \mathsf{t}_1$, and $|\mathsf{P}_1 - \mathsf{y}_3| = \mathsf{u}_1$, where $\mathsf{t}_1 \neq \mathsf{u}_1$. Let $\mathbf{x}' = (\mathsf{P}_1; \mathsf{x}_2; \mathsf{x}_3)$. Let $\mathsf{a}'_1 = \mathsf{y}_2 \succeq \mathsf{y}_1 \succeq \mathsf{y}_3$ and let $\mathbf{a}' = (\mathsf{a}'_1; \mathsf{a}_2; \mathsf{a}_3)$. Notice that $\mathbf{x}'$ is consistent with both $\mathbf{a}$ and $\mathbf{a}'$, therefore according to Lemma 28 the cost at P_1 should remain the same for $\mathbf{a}; \mathbf{a}'$:

$$\begin{aligned}
& \text{cost}_{P_1}(\mathsf{M}; \mathbf{a}) = \text{cost}_{P_1}(\mathsf{M}; \mathbf{a}') \\
\Rightarrow\ & \mathsf{u}_1 \cdot \mathsf{p}_3(\mathbf{a}) + \mathsf{t}_1 \cdot (1 - \mathsf{p}_3(\mathbf{a})) = \mathsf{u}_1 \cdot \mathsf{p}_3(\mathbf{a}') + \mathsf{t}_1 \cdot (1 - \mathsf{p}_3(\mathbf{a}')) \\
\Rightarrow\ & \mathsf{t}_1 + (\mathsf{u}_1 - \mathsf{t}_1) \cdot \mathsf{p}_3(\mathbf{a}) = \mathsf{t}_1 + (\mathsf{u}_1 - \mathsf{t}_1) \cdot \mathsf{p}_3(\mathbf{a}') \\
\Rightarrow\ & \mathsf{p}_3(\mathbf{a}') = \mathsf{p}_3(\mathbf{a}) \\
\Rightarrow\ & \mathsf{p}_3(\mathbf{a}') \geq \frac{1}{3}
\end{aligned}$$

Let P_2 be a point such that $|\mathsf{P}_2 - \mathsf{y}_2| = |\mathsf{P}_2 - \mathsf{y}_1| = \mathsf{t}_2$, and $|\mathsf{P}_2 - \mathsf{y}_3| = \mathsf{u}_2$, where $\mathsf{t}_2 \neq \mathsf{u}_2$. Let $\mathbf{x}'' = (\mathsf{x}'_1; \mathsf{x}_2; \mathsf{P}_2)$. Let $\mathsf{a}''_3 = \mathsf{y}_3 \succeq \mathsf{y}_2 \succeq \mathsf{y}_1$ and let $\mathbf{a}'' = (\mathsf{a}'_1; \mathsf{a}_2; \mathsf{a}''_3)$. According to Lemma 28 the cost at P_2 should remain the same for $\mathbf{a}'; \mathbf{a}''$:

$$\begin{aligned}
& \text{cost}_{P_2}(\mathsf{M}; \mathbf{a}') = \text{cost}_{P_2}(\mathsf{M}; \mathbf{a}'') \\
\Rightarrow\ & \mathsf{u}_2 \cdot \mathsf{p}_3(\mathbf{a}) + \mathsf{t}_2 \cdot (1 - \mathsf{p}_3(\mathbf{a}')) = \mathsf{u}_2 \cdot \mathsf{p}_3(\mathbf{a}'') + \mathsf{t}_2 \cdot (1 - \mathsf{p}_3(\mathbf{a}'')) \\
\Rightarrow\ & \mathsf{p}_3(\mathbf{a}'') = \mathsf{p}_3(\mathbf{a}') \\
\Rightarrow\ & \mathsf{p}_3(\mathbf{a}'') \geq \frac{1}{3}
\end{aligned}$$

Let Q be the midpoint between $\mathsf{y}_2; \mathsf{y}_3$, and let $\mathbf{x}''' = (\mathsf{y}_2; \mathsf{y}_2; \mathsf{Q})$. Without loss of generality, scale the distances such that $|\mathsf{y}_3 - \mathsf{Q}| = |\mathsf{Q} - \mathsf{y}_2| = 1$. Therefore the cost of y_2, the optimal candidate, is: $\text{SC}(\mathsf{y}_2; \mathbf{x}'''; \mathbf{a}'') = \text{SC}(\text{OPT}; \mathbf{x}'''; \mathbf{a}'') = 1$. The cost of y_3 is: $\text{SC}(\mathsf{y}_3; \mathbf{x}'''; \mathbf{a}'') = 2 \cdot 2 + 1 = 5$. Therefore the approximation ratio of M is at least:

$$\frac{\text{SC}(\mathsf{M}; \mathbf{x}'''; \mathbf{a}'')}{\text{SC}(\text{OPT}; \mathbf{x}'''; \mathbf{a}'')} = \mathsf{p}_3(\mathbf{a}'') \cdot 5 + (1 - \mathsf{p}_3(\mathbf{a}'')) \cdot 1 \quad = \quad 1 + 4 \cdot \mathsf{p}_3(\mathbf{a}'') \geq 1 + \frac{4}{3} = \frac{7}{3}$$

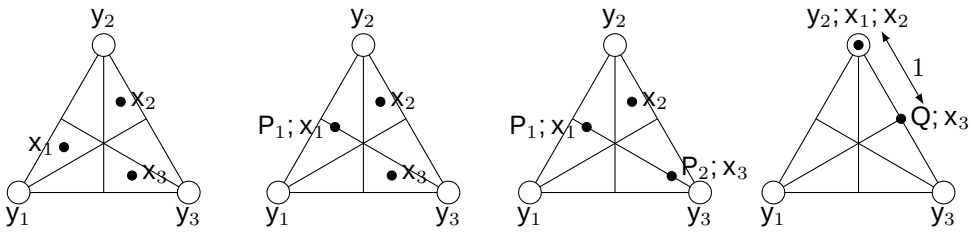
□

Proof. of Lemma 15: We prove the lower bounds for the case of two agents and two candidates. Let $\mathsf{y}_1 = -1; \mathsf{y}_2 = 1$. Let $\mathbf{a} = (\mathsf{a}_1; \mathsf{a}_2)$ be the ranking profile in which the two agents prefer different candidates, that is: $\mathsf{a}_1 = \mathsf{y}_1 \succeq \mathsf{y}_2$, $\mathsf{a}_2 = \mathsf{y}_2 \succeq \mathsf{y}_1$.

Examine the following two location profiles $\mathbf{x}; \mathbf{x}'$, (in both cases for which $\mathbf{a}$ is a truthful single candidate profile): $\mathbf{x} = (-1;\ )$, $\mathbf{x}' = (-\ ; 1)$.

We show that any decision of the mechanism makes will cause an approximation ratio of 2 either in $\mathbf{x}$ or in $\mathbf{x}'$.

Figure 19: An illustration of the proof of Theorem 14. The four figures, from left to right, show the dynamics of the proof (profiles $\mathbf{x}$, $\mathbf{x}'$, $\mathbf{x}''$ and $\mathbf{x}'''$ respectively).



It is easy to see that:

$$\begin{aligned}
\mathrm{SC}(\mathsf{y}_1;\mathbf{x}) &= 1+ \quad = \mathrm{SC}(\mathrm{OPT};\mathbf{x})\\
\mathrm{SC}(\mathsf{y}_2;\mathbf{x}) &= 3-\\
\mathrm{SC}(\mathsf{y}_1;\mathbf{x}') &= 3-\\
\mathrm{SC}(\mathsf{y}_2;\mathbf{x}') &= 1+ \quad = \mathrm{SC}(\mathrm{OPT};\mathbf{x}')
\end{aligned}$$

Denote $\mathsf{p} = \Pr[\mathsf{M}(\mathbf{a}) = \mathsf{y}_1]$. Therefore:

$$\begin{aligned}
\mathrm{SC}(\mathsf{M};\mathbf{x}) &= \mathsf{p}(1+\ \ ) + (1-\mathsf{p})(3-\ \ )\\
\mathrm{SC}(\mathsf{M};\mathbf{x}') &= \mathsf{p}(3-\ \ ) + (1-\mathsf{p})(1+\ \ )
\end{aligned}$$

The approximation ratio is therefore at least:

$$\begin{aligned}
&\min_{0\le p\le 1} \max \frac{\mathrm{SC}(\mathsf{M};\mathbf{x})}{\mathrm{SC}(\mathrm{OPT};\mathbf{x})}; \frac{\mathrm{SC}(\mathsf{M};\mathbf{x}')}{\mathrm{SC}(\mathrm{OPT};\mathbf{x}')} \quad =\\
&\min_{0\le p\le 1} \max \frac{\mathsf{p}(1+\ \ ) + (1-\mathsf{p})(3-\ \ )}{1+}; \frac{(1-\mathsf{p})(1+\ \ ) + \mathsf{p}(3-\ \ )}{1+} \quad =\\
&\min_{0\le p\le 1} \{\max\{1+2\mathsf{p}-\mathsf{p}\ \ ; 3-2\mathsf{p}+2\mathsf{p}\ \ -\ \ \}\}
\end{aligned}$$

The optimal value is reached at $\mathsf{p} = 0{:}5$, and it is $2-\frac{}{2}$, which tends to 2 as tends to 0. □

A.4.2 Upper Bound

Proof. of Lemma 16 (Random Dictator): Random dictator (RD) is a WPSC mechanism and so it is universally truthful according to Lemma 2.

We start by showing that the ratio can be arbitrarily close to 3 (up to a factor of $\frac{2}{n}$). Let $\mathsf{y}_1 = -1; \mathsf{y}_2 = 1$, and let $\mathsf{x}_1 = \cdots = \mathsf{x}_{n-1} = -1$ and $\mathsf{x}_n = 1$. Therefore the costs are: $\mathrm{SC}(\mathsf{y}_1;\mathbf{x}) = 1+ \quad = \mathrm{SC}(\mathrm{OPT};\mathbf{x})$ and $\mathrm{SC}(\mathsf{y}_2;\mathbf{x}) = 2(\mathsf{n}-1) + (1-\ \ )$.

Ergo: $\mathrm{SC}(\mathsf{RD};\mathbf{x}) = \frac{n-1}{n}\cdot\mathrm{SC}(\mathsf{y}_1;\mathbf{x}) + \frac{1}{n}\cdot\mathrm{SC}(\mathsf{y}_2;\mathbf{x}) = 3-\frac{2}{n}+\frac{2}{n}+\ \ $. The approximation ratio is therefore:

$$\frac{\mathrm{SC}(\mathsf{RD};\mathbf{x})}{\mathrm{SC}(\mathrm{OPT};\mathbf{x})} = \frac{3-\frac{2}{n}+\frac{2}{n}+}{1+} = 3 - \frac{2\ +\frac{2}{n}-\frac{2}{n}}{1+}$$

Clearly this ratio tends to 3 as $\mathsf{n}\to\infty$; $\to 0$.

We now show that the approximation ratio is bounded from above by 3. The social cost is:

$$
\begin{aligned}
\mathrm{SC}(\mathsf{RD};\mathbf{x}) &= \frac{1}{n}\sum_i \left(\sum_j |x_j - y(a_i)|\right) \\
&\le \frac{1}{n}\sum_i \left(\sum_j |x_j - y_{\mathrm{opt}}| + |y_{\mathrm{opt}} - y(a_i)|\right) \\
&\le \frac{1}{n}\sum_i \left(\sum_j |x_j - y_{\mathrm{opt}}| + |y_{\mathrm{opt}} - x_i| + |x_i - y(a_i)|\right) \\
&\le \frac{1}{n}\sum_i \left(\sum_j |x_j - y_{\mathrm{opt}}| + 2|y_{\mathrm{opt}} - x_i|\right) \\
&= \frac{1}{n}\sum_i \left(\sum_j |x_j - y_{\mathrm{opt}}|\right) + \frac{1}{n}\sum_i \left(\sum_j 2|y_{\mathrm{opt}} - x_i|\right) \\
&= \frac{1}{n}\sum_j \left(|x_j - y_{\mathrm{opt}}| \sum_i (1)\right) + \frac{2}{n}\sum_i \left(|y_{\mathrm{opt}} - x_i| \sum_j (1)\right) \\
&= \frac{1}{n}\sum_j (|x_j - y_{\mathrm{opt}}| \cdot n) + \frac{2}{n}\sum_i (|y_{\mathrm{opt}} - x_i| \cdot n) \\
&= \mathrm{SC}(\mathrm{OPT};\mathbf{x}) + 2\cdot \mathrm{SC}(\mathrm{OPT};\mathbf{x}) = 3\cdot \mathrm{SC}(\mathrm{OPT};\mathbf{x})
\end{aligned}
$$

The first two transitions hold due to the triangle inequality, and the third inequality holds due to fact that no candidate is closer to x_i than $y(a_i)$ is.

Notice that this holds in any metric space since we only used the triangle inequality, and did not use any notion which is specific to the line. □

A mechanism is *group-strategyproof (GSP)* if for any location profile and any coalition $S \subseteq N$, there is no joint deviation of the agents in S from the truthful reports such that they all gain. That is:

$$
\begin{aligned}
&\forall S \subseteq N; \forall a_S \in \mathcal{A}(\mathbf{x}_S); \forall a_{-S} \in \mathcal{A}^{n-|S|}; \forall a'_S \in \mathcal{A}^{|S|}; \exists i \in S : \\
&\mathrm{cost}_{x_i}(M;(a_S;a_{-S})) \le \mathrm{cost}_{x_i}(M;(a'_S;a_{-S}))
\end{aligned}
$$

In the continuous model, random dictator is GSP on the line (and even on the circle, see [2]). We show that in our candidate model random dictator is not GSP on the line. Notice that random dictator is in particular a WPSC mechanism, therefore a corollary of the lemma is that there exist WPSC mechanisms which are not GSP.

Lemma 30. *Random dictator is not group group-strategyproof*

Proof. Let there be three candidates at locations $y_1 = 1; y_2 = 0; y_3 = 1$ and let there be two agents at $x_1 = -0{:}51$ and $x_2 = 0{:}51$. When both agents report truthfully ($a_1 = y_1$, $a_2 = y_3$), the mechanism chooses $y_1; y_3$, each with probability 0.5. The cost of each of the agents in this case is: $\mathrm{cost}_{x_1}(\mathsf{RD};\mathbf{a}) = \mathrm{cost}_{x_2}(\mathsf{RD};\mathbf{a}) = 0{:}5 \cdot (0{:}51 + 1{:}49) = 1$.

However, if both agents misreport together to $\mathbf{a}' = (a'_1 = y_2; a'_2 = y_2)$, then y_2 will always be chosen. The costs in this case will be: $\mathrm{cost}_{x_1}(\mathsf{RD};\mathbf{a}') = \mathrm{cost}_{x_2}(\mathsf{RD};\mathbf{a}') = 0{:}51$. □

A.5 Missing Proofs from Section 6

As noted previously, in the continuous model on $\Re$, the mechanism which locates the facility on the report of the median agent is truthful and results in the optimal social cost. We define the median mechanism in the context of candidate constraints, and assess its social cost.

Definition 31 (Median mechanism)**.** *Median is a single candidate mechanism which chooses the median vote, that is* $\mathsf{a}_{(\lceil n=2 \rceil)}$.

Proof. of Theorem 18: The proofs of the two claims are:

1. Claim 1: Let there be 2 candidates such that $\mathsf{y}_1 = -1$ and $\mathsf{y}_2 = 1$. According to the sixth claim in Theorem 1 (Claim 26), any truthful location mechanism M is necessarily reducible to a single candidate mechanism. Let $\mathbf{x} = (-1;\ )$, $\mathbf{x}' = (-\ ;1)$ be two location profiles, and let B be the border between them. Clearly, both profiles correspond with the same votes ($\mathsf{x}_1; \mathsf{x}'_1 \in \mathcal{Z}_1 \setminus \mathsf{B}$ and $\mathsf{x}_2; \mathsf{x}'_2 \in \mathcal{Z}_2 \setminus \mathsf{B}$), therefore their outcome will be identical.

   If $\mathsf{M}(\mathbf{x}) = \mathsf{M}(\mathbf{x}') = \mathsf{y}_1$, then the ratio for M given $\mathbf{x}$ is $\frac{\mathrm{SC}(M;\mathbf{x})}{\mathrm{SC}(\mathrm{OPT};\mathbf{x})} = \frac{3-}{1+}$.

   If $\mathsf{M}(\mathbf{x}) = \mathsf{M}(\mathbf{x}') = \mathsf{y}_2$ then the ratio for M given $\mathbf{x}'$ is $\frac{\mathrm{SC}(M;\mathbf{x}')}{\mathrm{SC}(\mathrm{OPT};\mathbf{x}')} = \frac{3-}{1+}$.

   In either case, the approximation ratio tends to 3 as tends to 0.

2. Claim 2: "Median" is truthful - any agent located at the median location has no incentive to misreport since the only possible consequence is for the mechanism to select a different location. Similarly, other agents have no incentive to misreport, since misreporting either has no effect or moves the chosen location further away.

   We now move on to the approximation ratio - For a given single candidate profile $\mathbf{a}$, let $\mathsf{a}_{(\lceil n=2 \rceil)}$ and $\mathsf{y}_{\mathrm{opt}}$ be the median and optimal candidates respectively.

   Let A be the set of agents which not farther from $\mathsf{a}_{(\lceil n=2 \rceil)}$ than from $\mathsf{y}_{\mathrm{opt}}$, and let B be the set of all other candidates:

$$\mathsf{A} = \{\mathsf{i} : |\mathsf{x}_i - \mathsf{a}_{(\lceil n=2 \rceil)}| \leq |\mathsf{x}_i - \mathsf{y}_{\mathrm{opt}}|\}$$
$$\mathsf{B} = \{\mathsf{i} : |\mathsf{x}_i - \mathsf{a}_{(\lceil n=2 \rceil)}| > |\mathsf{x}_i - \mathsf{y}_{\mathrm{opt}}|\}$$

   The social cost of median is:

$$\mathrm{SC}(\mathsf{a}_{(\lceil n=2 \rceil)}; \mathbf{a}) \quad = \quad \sum_{i \in A} |\mathsf{x}_i - \mathsf{a}_{(\lceil n=2 \rceil)}| + \sum_{i \in B} |\mathsf{x}_i - \mathsf{a}_{(\lceil n=2 \rceil)}|$$

   Denote the first term as , and the second as .

   The social cost of the optimal candidate is:

$$\mathrm{SC}(\mathsf{y}_{\mathrm{opt}}; \mathbf{a}) \quad = \quad \sum_{i \in A} |\mathsf{x}_i - \mathsf{y}_{\mathrm{opt}}| + \sum_{i \in B} |\mathsf{x}_i - \mathsf{y}_{\mathrm{opt}}|$$

   Denote the first term by , and the second by .

   It is easy to see that $\leq$ since for any agent $\mathsf{i} \in \mathsf{A}$: $|\mathsf{x}_i - \mathsf{a}_{(\lceil n=2 \rceil)}| \leq |\mathsf{x}_i - \mathsf{y}_{\mathrm{opt}}|$, and this obviously holds when taking the sum.

We now show that $\leq \quad + \quad +$, due to the following inequalities (justifications for the transitions appear below):

$$
\begin{aligned}
&= \sum_{i \in B} |x_i - a_{(\lceil n=2 \rceil)}| \\
&\leq \sum_{i \in B} |a_{(\lceil n=2 \rceil)} - y_{\text{opt}}| + \sum_{i \in B} |y_{\text{opt}} - x_i| \\
&= \sum_{i \in B} |a_{(\lceil n=2 \rceil)} - y_{\text{opt}}| \quad + \\
&\leq \sum_{i \in A} |a_{(\lceil n=2 \rceil)} - y_{\text{opt}}| \quad + \\
&\leq \sum_{i \in A} |a_{(\lceil n=2 \rceil)} - x_i| + |x_i - y_{\text{opt}}| \quad + \quad = \quad + \quad +
\end{aligned}
$$

The inequalities in the second and fifth lines hold due to the triangle inequality, and the inequality in the fourth line holds because we are summing over a greater or equal amount of non-negative numbers (since $|A| \geq |B|$, by definition of the median – moving away from the median gets does not get us closer to at least half of the agents).

Putting this all together:

$$
\frac{\mathrm{SC}(a_{(\lceil n=2 \rceil)}; \mathbf{x})}{\mathrm{SC}(\mathrm{OPT}; \mathbf{x})} = \frac{\quad +}{\quad +} \leq \frac{\quad + \quad + \quad +}{\quad +} \leq \frac{3 \quad +}{\quad +} \leq \frac{3 \quad + 3}{\quad +} = 3
$$

□

Michal Feldman
The Blavatnik School of Computer Science
Tel Aviv University
Tel Aviv, Israel
Email: `michal.feldman@cs.tau.ac.il`

Amos Fiat
The Blavatnik School of Computer Science
Tel Aviv University
Tel Aviv, Israel
Email: `fiat@tau.ac.il`

Iddan Golomb
The Blavatnik School of Computer Science
Tel Aviv University
Tel Aviv, Israel
Email: `igolomb@gmail.com`