Representative Proxy Voting

Elliot Anshelevich,[1] **Zack Fitzsimmons,**[2] **Rohit Vaish,**[3] **Lirong Xia**[1]

[1] Rensselaer Polytechnic Institute, Troy, New York, USA 12180
[2] College of the Holy Cross, Worcester, Massachusetts, USA 01610
[3] Tata Institute of Fundamental Research, Mumbai, India 400005
{eanshel,xial}@cs.rpi.edu, zfitzsim@holycross.edu, rohit.vaish@tifr.res.in

Abstract

We study a model of proxy voting where the candidates, voters, and proxies are all located on the real line, and instead of voting directly, each voter delegates its vote to the closest proxy. The goal is to find a set of proxies that is θ-*representative*, which entails that for any voter located anywhere on the line, its favorite candidate is within a distance θ of the favorite candidate of its closest proxy. This property guarantees a strong form of representation as the set of voters is not required to be fixed in advance, or even be finite. We show that for candidates located on a line, an optimal proxy arrangement can be computed in polynomial time. Moreover, we provide upper and lower bounds on the number of proxies required to form a θ-representative set, thus showing that a relatively small number of proxies is enough to capture the preferences of any set of voters. An additional beneficial property of a θ-representative proxy arrangement is that for strict-Condorcet voting rules, the outcome of proxy voting is similarly close to the outcome of direct voting.

Introduction

It is natural to consider settings where voters either may not be able to or may not be willing to directly cast their vote, but instead decide to delegate their votes to a proxy. In much of the related work on proxy voting, the proxies are chosen from the set of eligible voters to then represent the electorate (see, e.g., Cohensius et al. 2017). In this paper, we consider a model for proxy voting that introduces proxies to the election using only the arrangement of the candidates and a given distance θ for *any* collection of voters.

We call our arrangement θ-*representative* since each voter's closest proxy is guaranteed to have a top preference that is within θ of the voter's. This can be interpreted as providing a set of allowed votes identified by a set of proxies placed in the metric space, and naturally the voters cast the vote of the proxy closest to them with the guarantee that this preference is close to the voter's. Another way to think about our model is that we form a set of representatives (proxies) whose choice does not depend on the locations of the voters (since these locations are difficult to determine exactly, or perhaps they are not static and change over time). The goal for this set of representatives is that it captures well the opinions of the voters, even as the voters change over time, if the voters are able to express their opinions by delegating their vote to their closest representative (proxy).

We consider elections where the voters, the proxies, and the candidates are all located in a metric space, and each voter and each proxy has spatial preferences determined by its distance to each candidate (see, e.g., Schofield 2008 for a survey on spatial voting). We focus our results on the one-dimensional case where all candidates and voters lie in the interval $[0, 1]$ (as in Cohensius et al. 2017). While not as general as an arbitrary metric space, it is an important step in understanding the behavior of the problem in more general domains. The one-dimensional assumption encompasses the well-studied domains of single-peaked (Black 1948) and single-crossing preferences (Mirrlees 1971).

What sets our contributions apart from the related work on proxy voting is that we determine a proxy arrangement *without* knowing the locations of the voters; that is, a θ-representative proxy arrangement is representative of *all* possible sets of voters simultaneously. In contrast, related work on proxy voting generally selects proxies or representatives from among a given set of voters (see, e.g., Cohensius et al. 2017; Meir, Sandomirskiy, and Tennenholtz 2020 and the references therein).

Our Contributions Our main contributions are as follows (see also Table 1):

- We introduce a new model of proxy voting and measure for the representation of voter preferences. We consider both the *unrestricted* case where the proxies can be placed anywhere, and the *restricted* case where the proxies can only be placed at the candidate locations.[1]

- We first provide algorithms for computing an optimal θ-representative proxy arrangement (i.e., one that uses a *minimum* number of proxies) in polynomial time (Theorems 1 and 3). However, these algorithms do not provide much insight into *how many* proxies are actually necessary in order to be θ-representative, for any possible set of candidates and voters. Because of this, we also prove upper and lower bounds on the number of proxies

---

[1]Our results also hold if the possible proxy locations are more general, as long as they include all candidate locations as a subset.

Restrictions on proxies ↓	Bounds on the number of proxies		Computational Results
	Upper bound	Lower bound	
On top of candidates (Restricted)	$2\lfloor\frac{1}{\theta}\rfloor$ (Theorem 2)	$2\lfloor\frac{1}{\theta}\rfloor$ (Proposition 1)	opt in poly time (Theorem 1)
Anywhere within $\mathbb{R}$ (Unrestricted)	$\frac{3}{2}\lceil\frac{1}{\theta}\rceil$ (Theorem 4)	$\lceil\frac{1}{\theta}\rceil$ (Proposition 2)	opt in poly time (Theorem 3)

Table 1: Summary of results. For each assumption on the positioning of the proxies (left column), the second and the third columns provide upper and lower bounds, respectively, on the number of proxies as a function of θ for a θ-representative proxy arrangement. The rightmost column contains the computational results for optimal proxy arrangements.

needed to have a θ-representative proxy arrangement (see Table 1). These bounds show that a relatively small number of proxies is enough to capture the preferences of any set of voters, even in the worst case.

- Our results also address the *dual* problem of minimizing the distance threshold θ for a given number of proxies. For example, we prove that for the unrestricted case, one can use a given budget of k proxies to compute a θ-representative arrangement with $\theta \leq \frac{3}{2}\lceil\frac{1}{k}\rceil$ (Corollary 1).

- We observe that the proxy arrangements determined by our algorithms are not only θ-representative with respect to the voters, but for elections using strict-Condorcet voting rules, the direct election outcome (i.e., without proxies) is within θ of the proxy-voting outcome (Proposition 3).

Related Work

Alger (2006) introduce a general proxy voting model where there is a fixed set of proxies and the voters can choose a proxy to represent them. Cohensius et al. (2017) consider a proxy voting setting closely related to ours where the voters, the proxies, and the candidates are on an interval, and the voters delegate their vote to their nearest proxy. However, in their work, the proxies are selected at random and the focus is on how the outcome is affected.

Green-Armytage (2015) consider proxy voting with spatial preferences where proxies are elected from among the voters, but additionally explore settings where proxies can further delegate their vote, which is often referred to as delegative democracy (see, e.g., Kahng, Mackenzie, and Procaccia 2018; Gölz et al. 2018). More general models of delegative democracy have also been studied (see, e.g., Brill and Talmon 2018; Abramowitz and Mattei 2019). Another line of research considers electing a set of representatives from the voters, and generally comparing the outcome from the vote of the committee and the outcome from direct voting (Skowron 2015; Pivato and Soh 2020; Meir, Sandomirskiy, and Tennenholtz 2020; Magdon-Ismail and Xia 2018).

Besides having a θ-representative proxy arrangement, we observe that for Condorcet-consistent voting rules, our results guarantee that the proxy-voting outcome is within θ of the direct-voting outcome. This notion of a θ-representative outcome is related to, but quite different than, the notion

of distortion Procaccia and Rosenschein (2006). A voting rule's distortion is the worst-case ratio of the distance from the voters to the winner and the candidate that minimizes this distance. The study of distortion is an active area of research in voting (see e.g., (Anshelevich et al. 2018; Goel, Krishnaswamy, and Munagala 2017; Abramowitz, Anshelevich, and Zhu 2019; Kempe 2020)).

Feldman et al. (2020) study a similar model to ours for a problem motivated by fairness in academic hiring. They consider a setup where a set of applicants (analogous to candidates in our model) are chosen by a set of experts (analogous to proxies). Each applicant is associated with a known quality score, and the applicants as well as the experts are arranged on a line. Each expert votes for its closest candidate, where closeness is defined as "distance minus quality." As with the related work on distortion, the aim of Feldman et al. (2020) is related to the outcome, while our paper is focused on finding a representative set of proxies.

Preliminaries

For any natural number $s \in \mathbb{N}$, let $[s] := \{1, \ldots, s\}$. Our model involves three types of entities—*candidates*, *voters*, and *proxies*—which are described below.

Candidates: Let $\mathcal{C} = \{c_1, \ldots, c_m\}$ denote a set of $m \in \mathbb{N}$ candidates that are arranged on a line segment $[0, 1]$. We will overload notation to denote the position of the ith candidate also by $c_i \in [0, 1]$. We will assume throughout that the extreme candidates are located at the endpoints of $[0, 1]$ (i.e., $c_1 = 0$ and $c_m = 1$), and that all candidates have distinct locations (i.e., for any distinct $i, j \in [m]$, $c_i \neq c_j$).

For any pair of adjacent candidates c_i and c_{i+1}, their *candidate bisector* is the vertical line at $(c_i + c_{i+1})/2$. Notice that with m candidates, there can, in general, be $\binom{m}{2}$ different bisectors. However, unless stated otherwise, the term 'candidate bisector' will refer to a bisector between *adjacent* candidates.

Voters: A voter can be located anywhere in $[0, 1]$ and is identified by its location. For any voter $v \in [0, 1]$, its favorite candidate, denoted by $\mathtt{top}(v)$, is the candidate that is closest to it. That is, $\mathtt{top}(v) := \arg\min_{c_i \in \mathcal{C}} |v - c_i|$, where ties are broken according to any fixed *directionally consistent* tie-breaking rule (see Definition 1). Note that we do not assume the set of voters to be fixed in advance. This is because our

results apply to any arbitrary collection of voters that could be located anywhere within $[0, 1]$.

Definition 1 (Directionally consistent tie-breaking rule).
A tie-breaking rule is a function $\tau : [0, 1]^3 \rightarrow [0, 1]$ *that maps any triple* $v, x, y \in [0, 1]$ *as follows:*

$$
\tau(v, x, y) = \begin{cases} x & \text{if } |v - x| < |v - y| \\ y & \text{if } |v - x| > |v - y| \\ \text{either } x \text{ or } y & \text{otherwise.} \end{cases}
$$

That is, a tie-breaking rule always maps to the point that is closer to v*, and in case of a tie, picks exactly one of the two points. We say that a tie-breaking rule* τ *is* directionally consistent *if, for any fixed* $v \in [0, 1]$*, either* τ *always tie-breaks to the left of* v *or always to the right of* v *(the choice of direction could depend on* v*). That is, for any fixed* $v \in [0, 1]$*, either* $\tau(v, x, y) = x$ *for all* $x, y \in [0, 1]$ *such that* $x \leq v \leq y$ *and* $|v - x| = |v - y|$*, or* $\tau(v, x, y) = y$ *for all* $x, y \in [0, 1]$ *such that* $x \leq v \leq y$ *and* $|v - x| = |v - y|$*.*

For any candidate $c_i \in \mathcal{C}$, its *Voronoi cell* V_i denotes the set of all voter locations $v \in [0, 1]$ whose favorite candidate is c_i, i.e., $V_i := \{v \in [0, 1] : \texttt{top}(v) = c_i\}$. Notice that the Voronoi cells $V_1, \ldots, V_m$ induce a partition of the line segment $[0, 1]$.

Proxies: Our model also includes a finite set $\mathcal{P}$ of proxies whose role is to vote on behalf of the voters. Specifically, each voter v delegates its vote to its *nearest proxy*, denoted by $p^v := \arg\min_{p \in \mathcal{P}} |v - p|$. Each proxy $p \in \mathbb{R}$ then votes for its favorite candidate, denoted by $\texttt{top}(p)$, which is defined as the candidate closest to it, i.e., $\texttt{top}(p) := \arg\min_{c_i \in \mathcal{C}} |p - c_i|$. As before, ties are broken according to a directionally consistent tie-breaking rule (see Definition 1). We will often use the term *proxy arrangement* to refer to a set of proxies placed on the real line.

Representative proxy arrangements: We will now formally define what it means for a proxy arrangement to be *representative*. The definition is stated in terms of a parameter $\theta \in [0, 1]$ that corresponds to a distance threshold. We will find it convenient to call two points $x, y \in [0, 1]$ to be *θ-close* if $|x - y| \leq \theta$, and call them *θ-far* otherwise.

Given any $\theta \in [0, 1]$, we say that a proxy arrangement is *θ-representative* if for *every* voter location, the favorite candidate of the voter is θ-close to the favorite candidate of its closest proxy.

Definition 2 (θ-representative proxy arrangement). *An arrangement of proxies is θ-representative if for any voter* $v \in [0, 1]$*, its favorite candidate is θ-close to the favorite candidate of its nearest proxy. That is, for every* $v \in [0, 1]$*,* $|\texttt{top}(v) - \texttt{top}(p^v)| \leq \theta$*.*

This property essentially says that for every voter, no matter where they may be located, it must be that their preference is not too different from the preference of their closest proxy. Therefore, the voter should feel reasonably satisfied with the proxy arrangement, as their closest proxy (to whom they yield the power of their vote) will be somewhat representative of their interests. On the other hand, if an arrangement is *not* θ-representative for some large θ, this means that

there are collections of voters which would be unhappy with this proxy arrangement, as the votes of the proxies would heavily disagree with the preferences of the voters.

We will now define the central problem studied in this paper.

Definition 3 (PROXY VOTING). *An instance of the* PROXY VOTING *problem* $\mathcal{I} = \langle \mathcal{C}, \theta \rangle$ *is specified by a set of candidates* $\mathcal{C}$ *in* $[0, 1]$ *and a parameter* $\theta \in (0, 1)$*. The goal is to compute an arrangement of proxies that is θ-representative for every location* $v \in [0, 1]$*.*[2]

We once again stress that a θ-representative proxy arrangement should be representative of *all possible sets of voters*: When choosing appropriate proxies to represent the populace, we do not have information about the voter locations; we only know the candidate locations. No matter where the voters are located, or how they change in the future, the proxies will still be representative of their views.

We will study two variants of PROXY VOTING which we call the *restricted* and *unrestricted* versions. Under the *restricted* version of the problem, we require that the set of proxies must be a subset of candidate locations, that is, the proxies must lie on top of the candidates. In the *unrestricted* version, the proxies can lie anywhere on the real line.

Note that the computational problem pertaining to Definition 3 can be formalized as a decision as well as an optimization problem. The decision version asks whether, given an instance $\mathcal{I}$ and a natural number $k \in \mathbb{N}$, there exists a θ-representative proxy arrangement for $\mathcal{I}$ consisting of at most k proxies. The optimization version asks whether, given an instance $\mathcal{I}$, the *optimal* θ-representative proxy arrangement for $\mathcal{I}$ can be computed in polynomial time. An arrangement of k proxies is *optimal* for instance $\mathcal{I}$ if there is no other arrangement of $k - 1$ or fewer proxies that is θ-representative. We will write $\texttt{opt}(\mathcal{I})$ to denote the number of proxies in any optimal θ-representative arrangement for the instance $\mathcal{I}$. To make the computational problem meaningful, we will assume throughout that all candidate locations in $\mathcal{C}$ and the parameter θ are rational.

Basic Properties and Examples

Let us start by discussing some of the issues that arise when constructing θ-representative sets of proxies, and attempt to build intuition about our techniques. First, consider the following lemma, which points out a precise relationship between candidate bisectors and proxy bisectors.

Lemma 1. *Let* $\mathcal{I} = \langle C, \theta \rangle$ *denote an instance with a pair of adjacent candidates* c_i *and* c_{i+1} *that are θ-far. Let* $\mathcal{P}$ *be any proxy arrangement such that none of the bisectors between adjacent proxies in* $\mathcal{P}$ *coincides with the bisector between* c_i *and* c_{i+1}*. Then,* $\mathcal{P}$ *is not θ-representative for* $\mathcal{I}$*.*

Proof. Let $p_1, \ldots, p_k$ denote the proxies in $\mathcal{P}$. Among all proxy bisectors between adjacent proxies, let the one be-

---

[2]We exclude the degenerate corner cases of $\theta = 0$ and $\theta = 1$ from the definition. Indeed, if $\theta = 0$, then it is easy to see that an optimal proxy arrangement requires as many proxies as candidates (realized by placing a proxy on each candidate). On the other hand, $\theta = 1$ involves placing a single proxy anywhere in $[0, 1]$.

tween p_j and p_{j+1} be closest to the candidate bisector between c_i and c_{i+1} (see Figure 1). Without loss of generality, we can assume that this proxy bisector is to the right of the candidate bisector, i.e., $\frac{c_i+c_{i+1}}{2} < \frac{p_j+p_{j+1}}{2}$. Let $d := \min\{\frac{p_j+p_{j+1}}{2}, c_{i+1}\} - \frac{c_i+c_{i+1}}{2}$ and observe that $d > 0$.
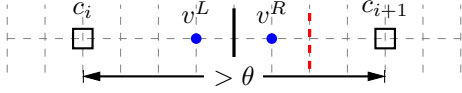

Figure 1: Failure of θ-representation when none of the proxy bisectors between adjacent proxies coincides with the candidate bisector between θ-far candidates. The candidate bisector is shown as solid vertical line, and its closest proxy bisector is shown as a dashed vertical line in red.

Consider a pair of voters $v^L \in (\frac{c_i+c_{i+1}}{2} - d, \frac{c_i+c_{i+1}}{2})$ and $v^R \in (\frac{c_i+c_{i+1}}{2}, \frac{c_i+c_{i+1}}{2} + d)$. Notice that the favorite candidates of v^L and v^R are c_i and c_{i+1}, respectively. Furthermore, since the closest proxy bisector is at least a distance d away from the candidate bisector, the two voters must have the same closest proxy, i.e., $p^{v^L} = p^{v^R} = p_j$. Then, regardless of the favorite candidate of p_j, θ-representation must be violated for at least one of v^L or v^R. ❑

It is relevant to note that the implication in Lemma 1 holds under both restricted and unrestricted positioning assumptions. Because of Lemma 1, it is easy to see that attempting something trivial like placing proxies at equal distances will not result in a θ-representative arrangement, as shown by Example 1 below. This further motivates the need for sophisticated algorithms for computing proxy arrangements, such as those discussed in upcoming sections.

Example 1 (Evenly spaced proxies may not be θ-representative). *Suppose we are given some $\theta \in (0,1)$ and a budget of $k \in \mathbb{N}$ proxies. We will construct an instance where evenly spacing these k proxies, i.e., placing the proxies at $\frac{\ell}{k-1}$ for every $\ell \in \{0,1,\ldots,k-1\}$, fails to be θ-representative (we will assume, without loss of generality, that $k \geq 2$). Notice that the bisectors between adjacent proxies are located at $\frac{2\ell-1}{2(k-1)}$ for every $\ell \in [k-1]$.*

Consider a sufficiently small $\varepsilon > 0$ such that $[\theta, \theta + \varepsilon] \subseteq [\frac{\ell-1}{k-1}, \frac{\ell}{k-1}]$ for some $\ell \in [k-1]$. Note that such a choice of ε must exist since $\theta \in (0,1)$. Pick a rational point $x \in (\theta, \theta + \varepsilon) \setminus \cup_{i \in [\ell]}\{\frac{2i-1}{k-1}\}$.

Consider an instance with three candidates c_1, c_2, and c_3 that are placed at 0, x, and 1, respectively. Notice that c_1 and c_2 are θ-far, and that the bisector between c_1 and c_2, which is located at $x/2$, does not coincide with any proxy bisector between adjacent proxies. Therefore, by Lemma 1, the evenly spaced proxy arrangement fails to be θ-representative.

We will now describe our results for restricted and unrestricted positioning of proxies.

Restricted Positioning of Proxies

This section provides an algorithm for computing an optimal proxy arrangement, followed by upper and lower bounds on the number of proxies needed to be θ-representative.

Algorithm for Computing Optimal Number of Restricted Proxies

Our first result for restricted positioning (Theorem 1) shows that a θ-representative arrangement with the smallest number of proxies can be computed in polynomial time. While the details are somewhat complex, the proof is via an essentially straightforward dynamic programming algorithm, and is presented in the full version of the paper (Anshelevich et al. 2020).

Theorem 1 (Optimal proxy arrangement under restricted positioning). *There is a polynomial-time algorithm that, given any instance $\langle \mathcal{C}, \theta \rangle$ of PROXY VOTING as input, terminates in polynomial time and returns an optimal θ-representative proxy arrangement satisfying restricted positioning.*

### Upper and Lower Bounds for Restricted Positioning of Proxies

Although the optimum number of proxies can be computed efficiently, these algorithms do not provide any insight into *how many* proxies are actually necessary in order to be θ-representative for any given set of candidates and any set of voters. Because of this, we now prove upper and lower bounds on the number of proxies needed to achieve θ-representation.

Theorem 2 (Upper bound under restricted positioning). *Given any instance $\langle \mathcal{C}, \theta \rangle$ of PROXY VOTING, there exists a θ-representative proxy arrangement satisfying restricted positioning that consists of at most $2(\frac{1}{\theta}-1)$ proxies if $\frac{1}{\theta} \in \mathbb{N}$, and at most $2\lfloor \frac{1}{\theta} \rfloor$ proxies otherwise. Furthermore, such an arrangement can be computed in polynomial time.*

Proof Sketch. The detailed proof of Theorem 2 is presented in the full version of the paper (Anshelevich et al. 2020). Here, we will describe the main steps in our algorithm by means of the example shown in Figure 2a.
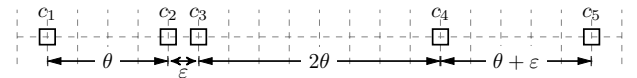
(a) An instance of PROXY VOTING.
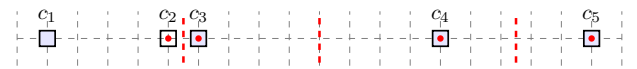
(b) The proxy bisectors computed by the algorithm are shown as dashed red lines, and the reference candidates are highlighted in blue. The locations of proxies are shown as solid red circles.

Figure 2: Illustrating the execution of the algorithm in Theorem 2 on a toy example.

To compute the desired proxy arrangement, our algorithm (see Anshelevich et al. 2020) computes a set of *proxy bisec-*

tors between adjacent proxies, according to the following strategy. Starting from the leftmost candidate as the *reference*, the bisector between the furthest θ-close and the closest θ-far candidates from the reference on the right is chosen as the first proxy bisector. (In Figure 2b, this is the bisector between c_2 and c_3.) In the next iteration, the candidate immediately to the right of the previous proxy bisector is chosen as the new reference in order to compute the next proxy bisector. (Thus, when c_3 is the reference in the next iteration, the furthest θ-close candidate from c_3 on the right is c_3 itself, and c_4 is the closest θ-far candidate. Therefore, the proxy bisector is chosen as the candidate bisector between c_3 and c_4.) This process repeats until the rightmost candidate c_m either becomes a reference or is θ-close to one.

Since, by construction, all proxy bisectors coincide with candidate bisectors, the desired proxy arrangement is realized by placing proxies on the equidistant candidates next to each proxy bisector. It is easy to show that the number of proxy bisectors is at most $\lfloor \frac{1}{\theta} \rfloor$, and therefore the number of proxies is at most $2\lfloor \frac{1}{\theta} \rfloor$. The θ-representation of this proxy arrangement follows from the fact that all candidates between consecutive proxy bisectors are θ-close. ❑

Proposition 1 shows that the upper bound derived in Theorem 2 is tight.

Proposition 1 (Lower bound under restricted positioning). *Given any $\theta \in (0,1)$, there exists an instance for which any θ-representative proxy arrangement under restricted positioning requires at least $2(\frac{1}{\theta} - 1)$ proxies if $\frac{1}{\theta} \in \mathbb{N}$, and $2\lfloor \frac{1}{\theta} \rfloor$ proxies otherwise.*

Proof. Let $p \in \mathbb{N}$ denote the unique positive integer such that $\frac{1}{p} \leq \theta < \frac{1}{p-1}$. Observe that $p-1$ equals $\frac{1}{\theta} - 1$ when $\frac{1}{\theta} \in \mathbb{N}$, and equals $\lfloor \frac{1}{\theta} \rfloor$ otherwise. Therefore, it suffices to show that any θ-representative arrangement requires $2p - 2$ proxies.

Our construction of the lower bound instance will depend on whether p is even or odd. Specifically, let $\varepsilon := \frac{1-(p-1)\theta}{5(p-1)/2}$ (if p is odd) or $\varepsilon := \frac{1-(p-1)\theta}{(5p/2-3)}$ (if p is even), and observe that $\varepsilon > 0$ in both cases. The distinction between even and odd cases is made in order to ensure that the distance between the extreme candidates is equal to 1.
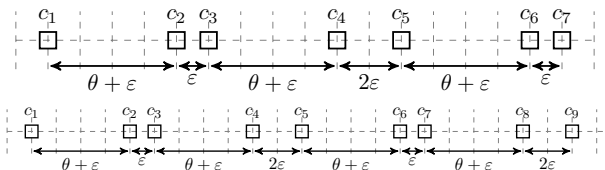

Figure 3: Lower bound instance in Proposition 1 when $\frac{1}{4} \leq \theta < \frac{1}{3}$ (top) and $\frac{1}{5} \leq \theta < \frac{1}{4}$ (bottom).

The instance consists of the candidate c_1 on the left, followed by a sequence of pairs of candidates $\{c_{2i}, c_{2i+1}\}$ for $i \in \mathbb{N}$ such that each pair is at a distance of $\theta + \varepsilon$ from the preceding pair, and the gap between the candidates in each

pair alternates between ε and 2ε (see Figure 3 for an example). In other words, the sequence of distances between adjacent pair of candidates from left to right is given by $\theta + \varepsilon, \varepsilon, \theta + \varepsilon, 2\varepsilon, \theta + \varepsilon, \varepsilon$ and so on. The rightmost candidate is c_{2p-1}. It is easy to see that all candidate locations are rational, and that the distance between the extreme candidates is equal to 1.

The fact that the candidate pairs $\{c_1, c_2\}$, $\{c_3, c_4\}$, $\{c_5, c_6\}$, and so on are each θ-far necessitates that for every candidate bisector between these pairs, there must exist a proxy bisector between adjacent proxies that coincides with it (contrapositive of Lemma 1). Furthermore, due to the alternating gaps property, any θ-representative proxy arrangement is required to place proxies on every candidate except for the last candidate c_{2p-1} (see Anshelevich et al. 2020 for a detailed argument). This implies that any θ-representative proxy arrangement for the above instance requires at least $2p - 2$ proxies. ❑

Unrestricted Positioning of Proxies

Let us now turn our attention to the *unrestricted* setting wherein the proxies can be placed anywhere on the real line. Clearly, any feasible proxy arrangement in the restricted model is also feasible under the unrestricted model. Therefore, the optimal number of proxies under the latter setting is at most that under the former; in fact, Example 2 shows that the separation between the two models can be strict.

Example 2 (Unrestricted positioning uses fewer proxies than restricted). *Fix $\theta = \frac{1}{3}$ and let $\varepsilon > 0$ be sufficiently small. Consider the instance shown in Figure 4 consisting of four candidates $c_1 = 0$, $c_2 = \frac{1}{3} + \varepsilon$, $c_3 = \frac{2}{3} - \varepsilon$, and $c_4 = 1$. Notice that the adjacent candidate pairs $\{c_1, c_2\}$ and $\{c_3, c_4\}$ are θ-far. Thus, by the contrapositive of Lemma 1, any θ-representative proxy arrangement must have bisectors between adjacent proxies that coincide with the candidate bisectors between these pairs.*
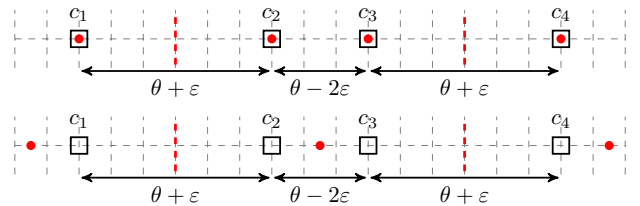

Figure 4: Unrestricted positioning (bottom) requires strictly fewer proxies than restricted positioning (top). The proxies in each case are shown as solid red circles.

Under restricted positioning, a θ-representative proxy arrangement must place proxies on all candidates, using four proxies in total (see top figure in Figure 4). By contrast, under unrestricted positioning, there exists a feasible proxy arrangement that only uses three proxies, namely $p_1 = -\frac{1}{6} + \varepsilon$, $p_2 = 0.5$, and $p_3 = \frac{7}{6} + \varepsilon$ (see bottom figure in Figure 4).

Notice that the saving in the number of proxies in the unrestricted case was achieved by "merging" the second and the third proxies. This, in turn, forces the first and the fourth

proxies to fall outside of $[0, 1]$, thus highlighting the importance of allowing the proxies to be anywhere on the real line.

Algorithm for Computing Optimal Number of Unrestricted Proxies

Let us now state our main result for unrestricted proxies.

Theorem 3 (**Optimal proxy arrangement under unrestricted positioning**). *There is a polynomial-time algorithm that, given any instance $\langle \mathcal{C}, \theta \rangle$ of* PROXY VOTING *as input, terminates in polynomial time and returns an optimal θ-representative proxy arrangement.*

The proof of Theorem 3 is technically the most involved part of the paper and is presented in (Anshelevich et al. 2020). Here we will outline a brief sketch of the proof.

Proof Sketch. For any fixed $k \in [m]$, our algorithm decides whether there exists a feasible (i.e., θ-representative) arrangement of k proxies for the given instance. The smallest k with a positive answer is returned as the output.

Given a proxy arrangement $\mathcal{P} = \{p_1, \ldots, p_k\}$, let us define the Voronoi cell W_j of proxy p_j as the set of all locations of voters whose closest proxy is p_j, i.e., $W_j := \{v \in [0, 1] : |v - p_j| \leq |v - p_\ell| \text{ for any } \ell \neq j\}$, where ties are broken according to a directionally-consistent tie-breaking rule (Definition 1). Notice that $v \in W_j$ if and only if $p_j = p^v$.

At a high level, the algorithm uses dynamic programming to compute the set of feasible locations (or the *feasibility set*) of the proxy p_{j+1} using the feasibility set of the preceding proxy p_j. It maintains the property that for each point s in the feasibility set of proxy p_{j+1}, there exists some point $s' < s$ in the feasibility set of proxy p_j such that θ-representation is satisfied for all voters in W_j (see Anshelevich et al. 2020 for a formal characterization result). Such a pair of locations $\{s', s\}$ is said to be *mutually feasible*.

To begin with, the feasibility set F^1 of the leftmost proxy p_1 is initialized as the union of the Voronoi cells of all candidates that are θ-close to the leftmost candidate c_1, along with the region $(-\infty, 0)$.[3] This corresponds to the set of all positions of the leftmost proxy such that θ-representation is satisfied for the leftmost voter.

The feasibility set F^{j+1} of the proxy p_{j+1} is computed using F^j as follows: Consider the restriction of F^j to the Voronoi cell of candidate c_h (denoted by $F^{j,h}$), as illustrated by the interval $[x, y]$ in Figure 5 (thus, $\texttt{top}(p_j) = c_h$). We will use $F^{j,h}$ to compute $F^{j+1,i}$, which is the restriction of F^{j+1} to the Voronoi cell of candidate c_i.

Let $\bar{b}_h$ denote the candidate bisector between the furthest θ-close and closest θ-far candidates to the right of c_h. Similarly, let $\underline{b}_i$ denote the candidate bisector between the furthest θ-close and closest θ-far candidates to the left of c_i.

By our characterization result, it follows that any point $z \in F^{j+1,i}$ that is mutually feasible with some point in $F^{j,h}$ must be (weakly) to the left of the mirror image of the point x about the bisector $\bar{b}_h$ (namely, x'). This is because the proxy bisector between p_j and p_{j+1} has to be (weakly) to

---

[3]Recall from Example 2 that proxies can lie outside $[0, 1]$ in an optimal arrangement under unrestricted positioning.

the left of $\bar{b}_h$. Similarly, z must be (weakly) to the right of the mirror image of point y about the bisector $\underline{b}_i$ (namely, y'). The contribution of the set $[x, y]$ to $F^{j+1,i}$ is given by $V_i \cap [y', x']$. In general, the restriction $F^{j,h}$ could comprise of several disjoint intervals. In the full version of the paper (Anshelevich et al. 2020), we describe how the feasibility set F^{j+1} can nevertheless be efficiently computed.
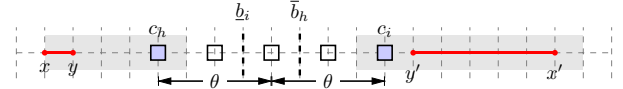

Figure 5

Having computed the feasibility set F^k of the rightmost proxy, the algorithm now checks whether it overlaps with the union of the Voronoi cells of candidates that are θ-close to the rightmost candidate c_m along with the region $(1, \infty)$. If yes, then there exists a feasible location of the proxy p_k that is θ-representative for the rightmost voter. By the aforementioned property, there must exist a feasible location p_{k-1} of the previous proxy such that p_{k-1} and p_k are mutually feasible, thus implying θ-representation for the voters in W_{k-1}. Similarly, there must exist a feasible location of the proxy p_{k-2} that is mutually feasible with p_{k-1}, and so on. Continuing backwards in this manner, we obtain a set of proxy locations $p_1, \ldots, p_k$ wherein the adjacent pairs are mutually feasible, which immediately implies θ-representation. On the other hand, the absence of an overlap certifies that with k proxies, there is no θ-representative proxy arrangement in the given instance. □

Upper and Lower Bounds for Unrestricted Positioning of Proxies

To compute an upper bound on the number of proxies, we will show that an algorithm similar to that for the restricted setting turns out to be useful. We will defer the detailed description of our algorithm and its formal analysis to the full version of the paper (Anshelevich et al. 2020), and instead revisit the example from Figure 2 considered previously in the restricted case.

Our algorithm proceeds in two phases. The first phase is identical to that of the algorithm for the restricted case (Theorem 2), and returns a set of proxy bisectors. In the second phase, the algorithm starts with a proxy arrangement that is consistent with the proxy bisectors computed in Phase 1 by placing a pair of equidistant proxies on either side of each bisector. This results in twice as many proxies as there are bisectors. To shrink this number down to $3/2$ times the number of bisectors, the algorithm utilizes the additional flexibility of the unrestricted setting via an "expand and merge" step. Specifically, the algorithm pulls all the proxies away from their bisectors at equal speeds until there is a "collision" event in some interval (recall that an interval is the area between adjacent proxy bisectors). At this point, the two proxies in that interval can be merged into a single proxy, and the locations of their 'partner' proxies are frozen (see Figure 6). This process is repeated until all proxies are frozen.
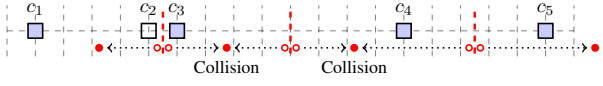
Figure 6: Illustrating the execution of the algorithm in Theorem 4 on a toy example. The initial and the final proxy locations are shown as empty and solid red circles, respectively.

By construction, all candidates within each interval are θ-close. Furthermore, since each voter is in the same interval as its closest proxy, θ-representation is satisfied for all voters. By an analysis similar to that in Theorem 2, it follows that the number of proxy bisectors is at most $\lceil \frac{1}{\theta} \rceil$. The expand-and-merge step ensures that no two consecutive intervals can have two proxies each. This readily implies the desired bound of $\frac{3}{2} \lceil \frac{1}{\theta} \rceil$ for the number of proxies.

Theorem 4 (**Upper bound under unrestricted positioning**). *Given any instance $\langle \mathcal{C}, \theta \rangle$ of* PROXY VOTING *as input, there exists a θ-representative arrangement consisting of at most $\frac{3}{2} \lceil \frac{1}{\theta} \rceil$ proxies. Furthermore, such an arrangement can be computed in polynomial time.*

Notice that Theorem 4 can be used to obtain upper bounds for the dual problem to PROXY VOTING wherein the input consists of a proxy budget k and the goal is to provide an upper bound on θ. Indeed, given as input a budget of $k > 1$ proxies, we can invoke our algorithm with $\theta = 1/\lfloor \frac{2k}{3} \rfloor$. Then, by Theorem 4, we know that the proxy arrangement returned by the algorithm is $\left(1/\lfloor \frac{2k}{3} \rfloor \right)$-representative and uses at most $\frac{3}{2} \lceil \frac{1}{\theta} \rceil \leq k$ proxies, which is within the given budget. Corollary 1 formalizes this observation.

Corollary 1 (**Upper bound on θ**). *There is a polynomial-time algorithm that, given any positive integer $k \in \mathbb{N}$ as input, returns an arrangement of at most k proxies that is $\left(1/\lfloor \frac{2k}{3} \rfloor \right)$-representative.*

Next, we will show that the dependence on θ in the upper bound cannot be improved.

Proposition 2 (**Lower bound under unrestricted positioning**). *Given any $\theta \in (0, 1)$, there exists an instance such that any θ-representative proxy arrangement requires at least $\lceil \frac{1}{\theta} \rceil$ proxies.*

Proof. Define $\theta' \in (0, 1)$ as follows:

$$\theta' := \begin{cases} \frac{1}{p-1}, & \text{if } \theta = \frac{1}{p} \text{ for some positive integer } p \\ \frac{1}{p}, & \text{if } \frac{1}{p+1} < \theta < \frac{1}{p} \text{ for some positive integer } p. \end{cases}$$

Observe that $\theta' > \theta$. Consider an instance where the candidates are evenly spaced at a distance of θ', as shown in Figure 7. Notice that the total number of candidates is $m = \frac{1}{\theta'} + 1 = \lceil \frac{1}{\theta} \rceil$.

Since any pair of adjacent candidates are θ-far, it must be that each candidate is the favorite candidate of some proxy. Indeed, in any arrangement with fewer than m proxies, there must exist a candidate, say c, that is not the favorite candidate of any proxy. Consider a voter v that is located at c. The distance between the favorite candidate of v and that of v's
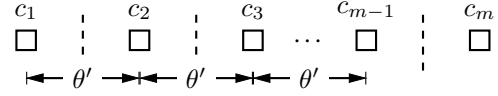
nearest proxy must then be strictly greater than θ, which violates θ-representation. Thus, m proxies are necessary, which gives the desired bound. ❑

Representative Election Outcomes

So far, we have focused on achieving θ-representation for each individual voter. A natural question is whether the voting outcome by proxies is a good representation of the direct voting outcome. This is formally defined as follows.

Definition 4. *An arrangement of proxies is said to be θ-representative under a voting rule r, if for any preference profile $P \in [0, 1]^n$, $r(P)$ is θ-close to $r(\{p^v : v \in P\})$.*

In Definition 4, $r(P)$ is the outcome of direct voting and $\{p^v : v \in P\}$ is the multiset of proxy votes. It turns out that θ-representation of an arrangement implies θ-representation under any strict-Condorcet rule with *consistent* tie-breaking (i.e., any strict-Condorcet rule which must choose a weak Condorcet winner as the winner, and in case there is more than one weak Condorcet winner, the leftmost (or, the rightmost) one must be chosen).

Proposition 3. *Let r be a strict-Condorcet rule with consistent tie-breaking. If an arrangement of proxies is θ-representative, then it is θ-representative under the rule r.*

In light of Proposition 3, all θ-representation results proved in this paper (see Table 1) naturally extend to θ-representation under strict-Condorcet rules. The proof of Proposition 3 is presented in (Anshelevich et al. 2020).

Concluding Remarks

In this paper, we gave efficient algorithms for computing optimal sets of θ-representative proxies, as well as proved upper and lower bounds on the number of proxies needed to achieve this property. Unlike in most related work, our proxy arrangements do not depend on the voter locations, and will remain representative for any set of voters. In fact, all our results hold (although some of the proofs become far more complex) even with additional requirements on the proxy arrangement, such as requiring that for every voter $v \in [0, 1]$, its closest proxy is within a distance θ, in addition to them being θ-representative.

Many interesting open problems remain, however, beginning with closing the gap between the upper and lower bounds for the number of proxies under unrestricted positioning, and analysing the ability to form θ-representative proxies in more general metric spaces. More generally, it would be interesting to expand the scope of θ-representation from "representing the top candidate well" to "representing the top-k candidates well", or to more general fairness properties.

Figure 7: Lower bound on the number of proxies under unrestricted positioning (Proposition 2)

References

Abramowitz, B.; Anshelevich, E.; and Zhu, W. 2019. Awareness of Voter Passion Greatly Improves the Distortion of Metric Social Choice. In *Proceedings of the 15th International Conference on Web and Internet Economics*, 3–16.

Abramowitz, B.; and Mattei, N. 2019. Flexible Representative Democracy: An Introduction with Binary Issues. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, 3–10.

Alger, D. 2006. Voting by Proxy. *Public Choice* 126(1-2): 1–26.

Anshelevich, E.; Bhardwaj, O.; Elkind, E.; Postl, J.; and Skowron, P. 2018. Approximating Optimal Social Choice under Metric Preferences. *Artificial Intelligence* 264: 27–51.

Anshelevich, E.; Fitzsimmons, Z.; Vaish, R.; and Xia, L. 2020. Representative Proxy Voting. *arXiv preprint arXiv:2012.06747* .

Black, D. 1948. On the Rationale of Group Decision-Making. *Journal of Political Economy* 56(1): 23–34.

Brill, M.; and Talmon, N. 2018. Pairwise Liquid Democracy. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, 137–143.

Cohensius, G.; Mannor, S.; Meir, R.; Meirom, E.; and Orda, A. 2017. Proxy Voting for Better Outcomes. In *Proceedings of the 16th Conference on Autonomous Agents and Multiagent Systems*, 858–866.

Feldman, M.; Mansour, Y.; Nisan, N.; Oren, S.; and Tennenholtz, M. 2020. Designing Committees for Mitigating Biases. In *Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence*, 1942–1949.

Goel, A.; Krishnaswamy, A. K.; and Munagala, K. 2017. Metric Distortion of Social Choice Rules: Lower Bounds and Fairness Properties. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, 287–304.

Gölz, P.; Kahng, A.; Mackenzie, S.; and Procaccia, A. D. 2018. The Fluid Mechanics of Liquid Democracy. In *Proceedings of the 14th International Conference on Web and Internet Economics*, 188–202.

Green-Armytage, J. 2015. Direct Voting and Proxy Voting. *Constitutional Political Economy* 26(2): 190–220.

Kahng, A.; Mackenzie, S.; and Procaccia, A. 2018. Liquid Democracy: An Algorithmic Perspective. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence*, 1095–1102.

Kempe, D. 2020. Communication, Distortion, and Randomness in Metric Voting. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence*, 2087–2094.

Magdon-Ismail, M.; and Xia, L. 2018. A Mathematical Model For Optimal Decisions In A Representative Democracy. In *Advances in Neural Information Processing Systems*, 4702–4711.

Meir, R.; Sandomirskiy, F.; and Tennenholtz, M. 2020. Representatitve Committees of Peers. Technical Report arXiv:2006.07837 [cs.GT], arXiv.org.

Mirrlees, J. 1971. An Exploration in the Theory of Optimum Income Taxation. *The Review of Economic Studies* 38(2): 175–208.

Pivato, M.; and Soh, A. 2020. Weighted Representative Democracy. *Journal of Mathematical Economics* 88: 52–63.

Procaccia, A.; and Rosenschein, J. 2006. The Distortion of Cardinal Preferences in Voting. In *Proceedings of the 10th International Workshop on Cooperative Information Agents*, 317–331.

Schofield, N. 2008. *The Spatial Model of Politics*. Number 95 in Routledge Frontiers of Political Economy. New York: Routledge.

Skowron, P. 2015. What Do We Elect Committees For? A Voting Committee Model for Multi-Winner Rules. *Proceedings of the 24th International Joint Conference on Artificial Intelligence* 1141—1147.