Trembling Hand Equilibria of Plurality Voting

Svetlana Obraztsova
Hebrew University of Jerusalem
Israel

Zinovi Rabinovich
Mobileye Vision Technologies Ltd.
Israel

Edith Elkind
University of Oxford
United Kingdom

Maria Polukarov
University of Southampton
United Kingdom

Nicholas R. Jennings
Imperial College
United Kingdom

Abstract

Trembling hand (TH) equilibria were introduced by Selten in 1975. Intuitively, these are Nash equilibria that remain stable when players assume that there is a small probability that other players will choose off-equilibrium strategies. This concept is useful for equilibrium refinement, i.e., selecting the most plausible Nash equilibria when the set of all Nash equilibria can be very large, as is the case, for instance, for Plurality voting with strategic voters. In this paper, we analyze TH equilibria of Plurality voting. We provide an efficient algorithm for computing a TH best response and establish many useful properties of TH equilibria in Plurality voting games. On the negative side, we provide an example of a Plurality voting game with no TH equilibria, and show that it is NP-hard to check whether a given Plurality voting game admits a TH equilibrium where a specific candidate is among the election winners.

1 Introduction

Plurality voting is a simple and popular method for aggregating preferences of multiple stakeholders over a set of available choices. It involves each stakeholder, or *voter*, declaring his most preferred alternative; the winner is then selected among the most popular alternatives using a tie-breaking rule. Due to its simplicity, Plurality voting is particularly susceptible to manipulation by dishonest voters: indeed, one can easily figure out if he may benefit from voting for an alternative other than his true top choice [Bartholdi, III *et al.*, 1989]. Thus, when analyzing the outcomes of Plurality voting, it is imperative to take strategic considerations into account. This can be accomplished by viewing voting as a strategic game, and focusing on outcomes that are identified by classic *solution concepts* (e.g., Nash equilibria) or arise as a result of a natural iterative process (e.g., best response dynamics).

Both approaches have been considered in the literature, going back as far as the pioneering work of Farquharson [1969] (other important contributions include [Moulin, 1979; Dhillon and Lockwood., 2004; Myerson and Weber, 1993]), and have recently received a lot of attention from the computational social choice community (see [Desmedt and Elkind, 2010; Xia and Conitzer, 2010; Thompson *et al.*, 2013; Obraztsova *et al.*, 2013; Meir *et al.*, 2014; Elkind *et al.*, 2015] for the analysis of Nash equilibria and their refinements and [Meir *et al.*, 2010; Reijngoud and Endriss, 2012; Reyhani and Wilson, 2012; Grandi *et al.*, 2013; Rabinovich *et al.*, 2015] for an investigation of best response dynamics). In particular, it has been observed early on that Plurality voting games have many undesirable Nash equilibria: e.g., in elections with at least 3 voters it is a Nash equilibrium for all voters to vote for the same candidate, even if none of them likes him. For this reason, much of the subsequent work focused on equilibrium *refinements*, i.e., sets of Nash equilibria that satisfy additional constraints, or, more broadly, outcomes that are stable when we place additional assumptions on voters' utilities. Two examples of the latter approach are Plurality voting games with *lazy voters* [Borgers, 2004; Desmedt and Elkind, 2010; Elkind *et al.*, 2015], where voters prefer to abstain when they are not pivotal (see e.g. [Kolovos and Harris, 2005] for motivation), or with *truth-biased voters*, where voters prefer to vote truthfully when they are not pivotal [Dutta and Laslier, 2010; Dutta and Sen, 2012; Thompson *et al.*, 2013; Obraztsova *et al.*, 2013; Elkind *et al.*, 2015]. Both approaches eliminate many of the undesirable Nash equilibria; however, it is not clear if either assumption (or their combination) can be used to capture all real-life Plurality voting scenarios where voters may strategize.

The goal of our paper is to extend this line of work to a classic equilibrium refinement known as *Trembling Hand Equilibrium (THE)*. This concept was introduced by Selten [1975], and has received a considerable amount of attention in the game-theoretic literature (see [Mas-Colell *et al.*, 1995] for basic definitions and [van Damme, 1991] for a comprehensive treatment). It is based on the idea that, when choosing his strategy, a player assumes that other players' hands may tremble, and with a small probability they would choose a strategy at random instead of playing their best response. In the context of voting, this corresponds to voters being confused by the design of the ballot or to errors during the vote counting procedure (such mistakes are well documented in the popular press, see the discussion in [Xia, 2012]). However, to the best of our knowledge, very little is known about trembling hand equilibria of Plurality voting.

Specifically, Messner and Polborn [2002; 2011] were the first to investigate the impact of small errors on voters' equilibrium behavior. However, they combine the trembling hand approach with another tweak of the model, which enables the voters to coordinate their behavior. Thus, their results (which include a characterization of the associated class of equilibria for three candidates, and a proof that a variant of Duverger's law holds in this setting) are not directly applicable if one aims to understand the impact of the trembling hand refinement per se. Other authors who consider THEs in voting focus on dynamic games [Acemoglu *et al.*, 2009], incomplete information scenarios [Jackson and Tan, 2013] or multiwinner elections [Obraztsova *et al.*, 2015], and only the latter paper explicitly addresses the associated algorithmic issues.

Against this background, in this paper we analyze the *classic* trembling hand equilibria of Plurality voting games. We establish many useful properties of players' trembling hand (TH) best responses and THE. In particular, we characterize games that admit an outcome where some candidate is unanimously supported in a THE. Our analysis enables us to develop a polynomial-time algorithm for computing one's TH best response, and construct THE for a wide range of voters' preferences. On the negative side, we provide an example of a Plurality election that admits no THE, and show that it is NP-hard to check if a given game admits a THE that assigns a positive probability of winning to a given candidate.

2 Model and Preliminaries

We begin by introducing the notation that will be used throughout the paper. We then define Plurality voting games and trembling hand equilibria of these games.

Plurality voting games. There is a set $\mathbf{V} = \{v_1, v_2, \dots, v_n\}$ of n *voters*, who aim to elect a winner from a set $\mathbf{C} = \{c_1, c_2, \dots, c_m\}$ of m *alternatives*, or *candidates*. Each voter $v \in \mathbf{V}$ has a *preference order* $>_v$ over $\mathbf{C}$, which is a strict total order (a complete, transitive and anti-symmetric binary relation). We write $x >_v y$ to express that voter v prefers candidate x to candidate y; we extend this notation to subsets of candidates $\mathbf{C}' \subseteq \mathbf{C}$ by writing $x >_v \mathbf{C}'$ whenever $x >_v y$ for all $y \in \mathbf{C}'$.

In Plurality voting, each voter $v \in \mathbf{V}$ submits a *vote*, or *ballot*, $b_v \in \mathbf{C}$. A *ballot profile* of a subset $S \subseteq \mathbf{V}$ of voters is a vector of ballots $\mathbf{b}_S = (b_v)_{v \in S}$, one for each voter in S. When $S = \mathbf{V}$, we drop the subscript S, and simply write $\mathbf{b}$. For each $v \in \mathbf{V}$ we write $\mathbf{b}_{-v}$ to denote the ballot profile obtained from $\mathbf{b}$ by removing the ballot of voter v, and we write $(b'_v, \mathbf{b}_{-v})$ to denote the ballot profile obtained by adding a ballot b'_v of voter v to the profile $\mathbf{b}_{-v}$ of the remaining voters.

Consider a subset of voters $S \subseteq \mathbf{V}$. The *plurality score* of a candidate $c \in \mathbf{C}$ in a ballot profile $\mathbf{b}_S$ is given by $sc(c; \mathbf{b}_S) = |\{v \in S \mid b_v = c\}|$. Let $s^w(\mathbf{b}_S) = \max_{c \in \mathbf{C}} sc(c; \mathbf{b}_S)$ be the highest score observed in $\mathbf{b}_S$. The *winning set under* $\mathbf{b}_S$ is the set of alternatives with the highest score:

$$W(\mathbf{b}_S) = \{c \in \mathbf{C} \mid sc(c; \mathbf{b}_S) = s^w(\mathbf{b}_S)\}.$$

The unique Plurality winner under $\mathbf{b}$ is chosen from $W(\mathbf{b})$ using a tie-breaking rule. In this paper, we consider *randomized tie-breaking*: the winner is chosen uniformly at random from $W(\mathbf{b})$. It will also be useful to set $s^*(\mathbf{b}_S) = \max_{c \in \mathbf{C} \setminus W(\mathbf{b}_S)} sc(c; \mathbf{b}_S)$ and consider the set of *runners-up*

$$\mathbf{C}^*(\mathbf{b}_S) = \{c \in \mathbf{C} \mid sc(c; \mathbf{b}_S) = s^*(\mathbf{b}_S)\}.$$

Further, we define the following sets of candidates:

$$\begin{aligned} \mathbf{C}^+(\mathbf{b}_S) &= \{c \in \mathbf{C} \mid sc(c; \mathbf{b}_S) > 0\}, \\ \mathbf{C}'(\mathbf{b}_S) &= \{c \in \mathbf{C} \mid sc(c; \mathbf{b}_S) = s^*(\mathbf{b}_S) - 1\}, \\ \mathbf{C}''(\mathbf{b}_S) &= \{c \in \mathbf{C} \mid sc(c; \mathbf{b}_S) = s^*(\mathbf{b}_S) - 2\}. \end{aligned}$$

Let $c^*(v; \mathbf{b}_{-v})$ (respectively, $c'(v; \mathbf{b}_{-v})$ and $c''(v; \mathbf{b}_{-v})$) be voter v's most preferred candidate in $\mathbf{C}^*(\mathbf{b}_{-v})$ (respectively, in $\mathbf{C}'(\mathbf{b}_{-v})$ and in $\mathbf{C}''(\mathbf{b}_{-v})$), i.e.,

$$\begin{aligned} c^*(v; \mathbf{b}_{-v}) &>_v \mathbf{C}^*(\mathbf{b}_{-v}) \setminus \{c^*(v; \mathbf{b}_{-v})\}, \\ c'(v; \mathbf{b}_{-v}) &>_v \mathbf{C}'(\mathbf{b}_{-v}) \setminus \{c'(v; \mathbf{b}_{-v})\}, \\ c''(v; \mathbf{b}_{-v}) &>_v \mathbf{C}''(\mathbf{b}_{-v}) \setminus \{c''(v; \mathbf{b}_{-v})\}. \end{aligned}$$

Since the winner is chosen from $W(\mathbf{b})$ uniformly at random, to discuss voters' strategic behavior, we have to be able to reason about voters' preferences over *lotteries*. To deal with this issue, we follow the approach used in a number of recent papers [Desmedt and Elkind, 2010; Boutilier *et al.*, 2012; Filos-Ratsikas and Miltersen, 2014; Obraztsova *et al.*, 2011] and assume that each voter is endowed with a *utility function* $\delta_v : \mathbf{C} \to \mathbb{N}$, which assigns a utility to each candidate in $\mathbf{C}$; we require that $\delta_v(c) > \delta_v(c')$ whenever $c >_v c'$ (this means that a voter's preference order is fully described by his utility function). Now, for each voter $v \in \mathbf{V}$ the utility that he derives from a winning set $W(\mathbf{b}_S)$ is his expected utility when the winner is chosen from $W(\mathbf{b}_S)$ uniformly at random, i.e.,

$$\frac{1}{|W(\mathbf{b}_S)|} \sum_{c \in W(\mathbf{b}_S)} \delta_v(c);$$

we denote this quantity by $u_v(\mathbf{b}_S)$.

The setting described above induces a normal-form game $\Gamma = \langle \mathbf{V}, \mathbf{C}, (u_v)_{v \in \mathbf{V}} \rangle$, where $\mathbf{V}$ is the set of players, the set of strategies available to each player is given by $\mathbf{C}$, and the utility function of each player at a strategy profile $\mathbf{b}$ is given by $u_v(\mathbf{b})$. Games of this form are called *Plurality voting games*. A strategy profile $\mathbf{b}$ is a *Nash Equilibrium (NE)* of a game Γ if for each player $v \in \mathbf{V}$ and each ballot $b'_v \in \mathbf{C}$ it holds that $u_v(\mathbf{b}) \ge u_v(b'_v, \mathbf{b}_{-v})$.

Trembling hand equilibrium (THE) The notion of trembling hand equilibrium goes back to Selten [1975]. Intuitively, it allows for a small probability that, instead of playing his chosen strategy, a player errs and chooses a random one. We consider a popular implementation of this idea, where, for a given small value $\varepsilon > 0$, a player with m strategies available to him plays strategically with probability $1 - m\varepsilon$, and with probability $m\varepsilon$ he chooses a strategy uniformly at random.

More formally, fix a *trembling hand (TH) probability* $\varepsilon > 0$. For each voter v, we distinguish between his *intended ballot* b_v and his *actual ballot* $\widetilde{b}_v$. A voter selects his intended ballot b_v. Then, the probability that v's actual ballot is some $\widetilde{b}_v \in \mathbf{C}$, termed the *ballot TH probability*, is given by

$$P^\varepsilon(\widetilde{b}_v \mid b_v) = \begin{cases} 1 - (m-1)\varepsilon & \text{if } \widetilde{b}_v = b_v, \\ \varepsilon & \text{otherwise.} \end{cases} \qquad (1)$$

This definition can be extended to subsets of voters: for a subset $S \subseteq \mathbf{V}$ of voters, the *joint ballot TH probability* given the intended ballot profile $\mathbf{b}_S$ is given by

$$P^{\varepsilon}(\widetilde{\mathbf{b}}_S \mid \mathbf{b}_S) = \prod_{v \in S} P^{\varepsilon}(\widetilde{b}_v \mid b_v). \qquad (2)$$

Notice that voter v's ballot TH probability defines a fully mixed strategy for v. We will denote this mixed strategy by b^{ε}_v. Moreover, given a set of voters $S \subseteq \mathbf{V}$, we will denote their mixed strategy profile $(b^{\varepsilon}_v)_{v \in S}$ by $\mathbf{b}^{\varepsilon}_S$; we will refer to b^{ε}_v and $\mathbf{b}^{\varepsilon}_S$ as the *TH ballot* and the *joint TH ballot*, respectively.

Now, suppose that the intended joint ballot of players in $\mathbf{V} \setminus \{v\}$ is $\mathbf{b}_{-v}$. When deciding on his intended ballot, a voter realizes that other voters' hands may tremble and assumes that these trembles are independent from each other, but ignores the possibility that his own hand may tremble as well. Thus, v calculates his utility from submitting a ballot b_v against $\mathbf{b}_{-v}$ as follows:

$$\widetilde{U}^{\varepsilon}_v(b_v, \mathbf{b}_{-v}) = \sum_{\widetilde{\mathbf{b}}_{-v} \in \mathbf{C}^{n-1}} u_v(b_v, \widetilde{\mathbf{b}}_{-v}) P^{\varepsilon}(\widetilde{\mathbf{b}}_{-v} \mid \mathbf{b}_{-v}). \qquad (3)$$

We refer to this quantity as the *ε-TH utility* of voter v. We are now ready to define the notions of *ε-TH best response*, *TH best response* and *TH equilibrium*.

Definition 1. *Given an ε with $0 < \varepsilon < 1$, we say that a ballot b is an* ε-TH best response (ε-TH-BR) *of voter v to the intended ballot $\mathbf{b}_{-v}$ of the other voters if $b \in \arg\max_{b'} \widetilde{U}^{\varepsilon}_v(b', \mathbf{b}_{-v})$. We say that b is a* TH best response (TH-BR) *of voter v to $\mathbf{b}_{-v}$ if there exists an ε_0 with $0 < \varepsilon_0 < 1$ such that b is an ε-TH best response to $\mathbf{b}_{-v}$ for all ε with $0 < \varepsilon < \varepsilon_0$. A joint ballot $\mathbf{b}$ is a* TH equilibrium (THE) *if for each $v \in \mathbf{V}$ the ballot b_v is a TH-BR to $\mathbf{b}_{-v}$.*

Example 1. Let $\mathbf{C} = \{a, b\}$, $\mathbf{V} = \{u, v, w\}$ and assume that each voter $z \in \mathbf{V}$ has $\delta_z(a) = 1$, $\delta_z(b) = 0$. Then both (a, a, a) and (b, b, b) are Nash equilibria of the respective Plurality voting game, simply because no individual player can change the winning set by modifying his vote. However, (b, b, b) is not a THE of this game. Indeed, consider some ε with $0 < \varepsilon < 1$ and a voter v. The joint TH ballot of the remaining voters assigns probability $(1-\varepsilon)^2$ to (a, a); the ballot profiles (a, b) and (b, a) occur with probability $\varepsilon(1-\varepsilon)$ each, and the ballot profile (b, b) occurs with probability ε^2. Thus, if voter v changes his strategy from b to a, he increases his ε-TH utility by $2\varepsilon(1-\varepsilon) > 0$: whenever $W(\mathbf{b}_{-v}) = \{a, b\}$ (which happens with probability $2\varepsilon(1-\varepsilon)$), switching from b to a increases his utility by 1, and if $W(\mathbf{b}_{-v}) = \{a\}$ or $W(\mathbf{b}_{-v}) = \{b\}$, the outcome does not depend on his vote.

It will be convenient to rewrite the expression for ε-TH utility by grouping the terms according to how many voters' hands have trembled. Formally, for every possible value of the parameter η we define the set of *η-tremblings* as follows.

Definition 2. *Given a set of voters $S \subseteq \mathbf{V}$ and two ballot profiles $\mathbf{b}'_S$, $\mathbf{b}''_S$, let $d(\mathbf{b}'_S, \mathbf{b}''_S) = |\{v \in S \mid b'_v \neq b''_v\}|$ denote the Hamming distance between these two ballot profiles. Given an intended ballot profile $\mathbf{b}_S$ for S and a parameter η with $0 \leq \eta \leq |S|$, we say that a ballot $\widetilde{\mathbf{b}}_S$ is an* η-trembling *for $\mathbf{b}_S$ if $d(\widetilde{\mathbf{b}}_S, \mathbf{b}_S) = \eta$. We denote the set of all η-tremblings for $\mathbf{b}_S$ by $\mathbf{B}(\mathbf{b}_S, \eta)$.*

Every η-trembling for $\mathbf{V} \setminus \{v\}$ occurs with the same probability, namely, $\zeta(\varepsilon, \eta) = (1 - (m-1)\varepsilon)^{n-1-\eta}\, \varepsilon^{\eta}$. Consequently, we can rewrite the expression for ε-TH utility as a sum of n terms, where each term corresponds to a fixed value of η. Specifically, we define

$$\widetilde{U}^{\varepsilon}_v(b_v, \mathbf{b}_{-v}; \eta) = \sum_{\widetilde{\mathbf{b}}_{-v} \in \mathbf{B}(\mathbf{b}_{-v}, \eta)} u_v(b_v, \widetilde{\mathbf{b}}_{-v}) \zeta(\varepsilon, \eta), \qquad (4)$$

and obtain $$\widetilde{U}^{\varepsilon}_v(b_v, \mathbf{b}_{-v}) = \sum_{\eta=0}^{n-1} \widetilde{U}^{\varepsilon}_v(b_v, \mathbf{b}_{-v}; \eta). \qquad (5)$$

We will now define a class of utility functions that will play an important role in our analysis.

Definition 3. *A utility function $\delta_v : \mathbf{C} \to \mathbb{N}$ of a voter $v \in \mathbf{V}$ is* ordinary *if the following two conditions hold:*

(1) for each triple of distinct candidates $a, b, c \in \mathbf{C}$ and each k with $0 \leq k \leq n + m$ we have
$$(m + n - k)(\delta_v(a) - \delta_v(b)) \neq \delta_v(b) + \delta_v(c);$$

(2) for each pair of distinct candidates $a, b \in \mathbf{C}$, each subset $\mathbf{C}' \subseteq \mathbf{C} \setminus \{a, b\}$ and each k with $0 \leq k \leq n$ we have
$$\textstyle\sum_{c \in \mathbf{C}'} \delta_v(c) - \delta_v(a) \neq (m + n - k)(\delta_v(b) - \delta_v(a)).$$

Some of the results in this paper require that all voters have ordinary utilities. This restriction is needed for technical reasons: it ensures that, when comparing two available strategies, a voter can make his decision by considering the smallest number of trembles that change the outcome (see, e.g., the proof of Theorem 1). Note that almost all utility functions are ordinary, in the following sense: if we were to draw voters' utilities for each candidate uniformly at random from a real interval $[a, b]$, then all utility functions would be ordinary with probability 1. Moreover, for every utility function $\delta : \mathbf{C} \to \mathbb{N}$ we can find a constant $\gamma \in \mathbb{N}$ such that the utility function $\delta' : \mathbf{C} \to \mathbb{N}$ given by $\delta'(c) = \delta(c) + \gamma$ for each $c \in \mathbf{C}$ is ordinary. Thus, we believe that results for voters with ordinary utilities provide useful intuition for the general case; nevertheless, getting rid of this restriction is an important direction for future work.

3 Best Response Calculation

In this section we will explain how voter v can compute his TH best response given the intended ballot profile $\mathbf{b}_{-v}$ of other voters. We start by observing that if v can identify a vote that improves the outcome from his perspective (compared to the outcome at $\mathbf{b}_{-v}$), then he should simply ignore the possibility of trembles and submit his best response to the intended ballot profile.

Proposition 1. *Consider a voter $v \in \mathbf{V}$ and a ballot profile $\mathbf{b}_{-v}$ of the remaining voters. If there exists a ballot $b'_v \in \mathbf{C}$ such that $u_v(b'_v, \mathbf{b}_{-v}) > u_v(\mathbf{b}_{-v})$ then a ballot b is a TH best response of voter v to $\mathbf{b}_{-v}$ if and only if $u_v(b, \mathbf{b}_{-v}) \geq u_v(b', \mathbf{b}_{-v})$ for all $b' \in \mathbf{C}$.*

We omit the formal proof of this result due to space constraints; intuitively, Proposition 1 holds because for small ε the intended ballot profile is much more likely than all other ballot profiles taken together, so v can ignore the 'second-order effects' of his vote. Note that we can easily check whether v and $\mathbf{b}_{-v}$ satisfy the conditions of Proposition 1, by trying all $m = |\mathbf{C}|$ possible votes. Moreover, these conditions are satisfied whenever $|W(\mathbf{b}_{-v})| > 1$, as v can change the outcome in his favor by voting for his most preferred candidate in $W(\mathbf{b}_{-v})$. Thus, from now on we will focus on the case where $W(\mathbf{b}_{-v}) = \{w\}$ for some $w \in \mathbf{C}$.

Proposition 2. *Suppose that for some voter v and for some ballot profile $\mathbf{b}_{-v}$ of the remaining voters we have $|W(\mathbf{b}_{-v})| = 1$, and let c be a TH best response of v to $\mathbf{b}_{-v}$. Then $s(c) \geq s^*(\mathbf{b}_{-v}) - 2$.*

Proof. Fix a voter $v \in \mathbf{V}$, a candidate $w \in \mathbf{C}$ and a ballot $\mathbf{b}_{-v}$ such that $W(\mathbf{b}_{-v}) = \{w\}$. Let $s^w = s^w(\mathbf{b}_{-v})$, $s^* = s^*(\mathbf{b}_{-v})$, $c^* = c^*(v; \mathbf{b}_{-v})$, $c' = c'(v; \mathbf{b}_{-v})$. Fix a candidate c with $s(c) < s^* - 2$. To show that c is not among v's TH best responses to $\mathbf{b}_{-v}$, we consider two cases.

Case I: $c^* >_v w$ In this case, we show that for all sufficiently small $\varepsilon > 0$ the quantity $\Delta = \widetilde{U}_v^\varepsilon(c^*, \mathbf{b}_{-v}) - \widetilde{U}_v^\varepsilon(c, \mathbf{b}_{-v})$ is positive, i.e., the ε-TH utility that v derives from voting for c^* is higher than his ε-TH utility from voting for c. The argument depends on the parity of $s^w - s^*$.

$s^w - s^*$ is odd.
Let $\eta^* = (s^w - s^* - 1)/2$. Then $\Delta = T_1 + T_2 + T_3$, where

$$\begin{aligned} T_1 &= \sum_{\eta=0}^{\eta^*-1} \widetilde{U}_v^\varepsilon(c^*, \mathbf{b}_{-v}; \eta) - \widetilde{U}_v^\varepsilon(c, \mathbf{b}_{-v}; \eta), \\ T_2 &= \widetilde{U}_v^\varepsilon(c^*, \mathbf{b}_{-v}; \eta^*) - \widetilde{U}_v^\varepsilon(c, \mathbf{b}_{-v}; \eta^*), \\ T_3 &= \sum_{\eta=\eta^*+1}^{n-1} \widetilde{U}_v^\varepsilon(c^*, \mathbf{b}_{-v}; \eta) - \widetilde{U}_v^\varepsilon(c, \mathbf{b}_{-v}; \eta). \end{aligned}$$

First, let us analyze T_1. Note that

$$s^w - \eta^* = \frac{s^w + s^* + 1}{2}, \quad s^* + \eta^* = \frac{s^w + s^* - 1}{2}.$$

For every $\eta \leq \eta^* - 1$ and every ballot profile $\widetilde{\mathbf{b}}_{-v} \in \mathbf{B}(\mathbf{b}_{-v}, \eta)$ we have $sc(w; (c^*, \widetilde{\mathbf{b}}_{-v})) > s^w - \eta^*$, $sc(w; (c, \widetilde{\mathbf{b}}_{-v})) > s^w - \eta^*$, whereas for every $x \in \mathbf{C} \setminus \{w\}$ we have $sc(x; (c^*, \widetilde{\mathbf{b}}_{-v})) < s^* + \eta^*$, $sc(x; (c, \widetilde{\mathbf{b}}_{-v})) < s^* + \eta^*$. Thus, $W(c^*, \widetilde{\mathbf{b}}_{-v}) = W(c, \widetilde{\mathbf{b}}_{-v}) = \{w\}$, and hence $T_1 = 0$.

Next, consider T_3. Since for each $x \in \mathbf{C}$ we have $\widetilde{U}_v^\varepsilon(x, \mathbf{b}_{-v}; \eta) = g_x \zeta(\varepsilon, \eta)$, where g_x does not depend on ε, we can bound the absolute value of T_3 as $|T_3| \leq \alpha\varepsilon^{\eta^*+1}$, where $\alpha > 0$ does not depend on ε.

Now, consider T_2. There is a ballot $\widetilde{\mathbf{b}}_{-v} \in \mathbf{B}(\mathbf{b}_{-v}, \eta^*)$ where exactly η^* voters err by voting for c^* instead of w. We have $sc(w; \widetilde{\mathbf{b}}_{-v}) = (s^w + s^* + 1)/2$, $sc(c^*; \widetilde{\mathbf{b}}_{-v}) = (s^w + s^* - 1)/2$, so $W(c, \widetilde{\mathbf{b}}_{-v}) = \{w\}$, $W(c^*, \widetilde{\mathbf{b}}_{-v}) = \{w, c^*\}$. For every other ballot $\widehat{\mathbf{b}}_{-v} \in \mathbf{B}(\mathbf{b}_{-v}, \eta^*)$ it holds that $W(c, \widehat{\mathbf{b}}_{-v}) = W(c^*, \widehat{\mathbf{b}}_{-v}) = \{w\}$. The probability that $\widetilde{\mathbf{b}}_{-v}$ occurs under the mixed strategy profile $\mathbf{b}^\varepsilon_{-v}$ is $\zeta(\varepsilon, \eta^*)$. For $\varepsilon < 1/(2m-2)$ we have $\zeta(\varepsilon, \eta^*) > \varepsilon^{\eta^*}/2^n$. Thus, $T_2 \geq \beta\varepsilon^{\eta^*}$ for some $\beta > 0$ that does not depend on ε. If ε is sufficiently small, we have $\alpha\varepsilon^{\eta^*+1} < \beta\varepsilon^{\eta^*}$, meaning that for such values of ε the quantity $T_1 + T_2 + T_3$ is positive, and hence c cannot be v's ε-TH best response to $\mathbf{b}_{-v}$.

$s^w - s^*$ is even.
The argument for this case is similar. We set $\eta^* = (s^w - s^*)/2$, and represent Δ as a sum of three terms, corresponding to $\eta = 0, \dots, \eta^* - 1$, $\eta = \eta^*$, and $\eta = \eta^* + 1, \dots, n-1$. We show that the first term is 0, since in every η-trembling with $\eta < \eta^*$ candidate w is the unique winner no matter how v votes. Also, the absolute value of the third term can be bounded by $\alpha\varepsilon^{\eta^*+1}$ for some positive constant α. Thus, we can focus on analyzing $\widetilde{U}_v^\varepsilon(c^*, \mathbf{b}_{-v}; \eta^*) - \widetilde{U}_v^\varepsilon(c, \mathbf{b}_{-v}; \eta^*)$.

Just as in the previous case, we observe that $\mathbf{B}(\mathbf{b}_{-v}, \eta^*)$ contains a ballot $\widetilde{\mathbf{b}}_{-v}$ where exactly η^* voters err by voting for c^* instead of w. For this ballot we have $W(c, \widetilde{\mathbf{b}}_{-v}) = \{w, c^*\}$, $W(c^*, \widetilde{\mathbf{b}}_{-v}) = \{c^*\}$. For all other ballots $\widehat{\mathbf{b}}_{-v} \in \mathbf{B}(\mathbf{b}_{-v}, \eta^*)$ we have $W(c, \widehat{\mathbf{b}}_{-v}) \subseteq \{w, x\}$ for some $x \in \mathbf{C}^* \setminus \{c^*\}$ and either $W(c^*, \widehat{\mathbf{b}}_{-v}) = W(c, \widehat{\mathbf{b}}_{-v})$ or $W(c^*, \widehat{\mathbf{b}}_{-v}) = W(c, \widehat{\mathbf{b}}_{-v}) \cup \{c^*\}$; as $c^* >_v C^* \setminus \{c^*\}$ and $c^* >_v w$, this means that voting for c^* is never worse than voting for c. The probability of $\widetilde{\mathbf{b}}_{-v}$ under the mixed strategy profile $\mathbf{b}^\varepsilon_{-v}$ is at least $\varepsilon^{\eta^*}/2^n$ for sufficiently small ε, so the gain in expected utility that v can obtain by switching his vote from c to c^* is at least $\beta\varepsilon^{\eta^*}$ for some $\beta > 0$ that does not depend on ε. As in the previous case, this establishes our claim.

Case II: $w >_v c^*$ The argument for this case is somewhat more complicated. We establish that c is not an ε-TH best response to $\mathbf{b}_{-v}$ for all sufficiently small ε by proving that either w or c' is a more profitable strategy for small ε. If $s^w - s^*$ is even, we focus on η-tremblings with $\eta = (s^w - s^*)/2$ and show that w is more profitable than c, and for odd $s^w - s^*$, we look at η-tremblings with $\eta = (s^w - s^* + 1)/2$ and show that v's most preferred candidate in $\{w, c'\}$ is more profitable than c. We omit the details due to space constraints. ☐

By inspecting the proof of Proposition 2, one can observe that if $x >_v y$ and $sc(x; \mathbf{b}_{-v}) = sc(y; \mathbf{b}_{-v})$ then y cannot be v's TH best response to $\mathbf{b}_{-v}$. Combining this observation with Propositions 1 and 2, we discover the following feature of TH best responses.

Corollary 1. *A ballot b_v can be a TH best response to an intended ballot profile $\mathbf{b}_{-v}$ of the other voters only if for some set $Z \in \{W(\mathbf{b}_{-v}), \mathbf{C}^*(\mathbf{b}_{-v}), \mathbf{C}'(\mathbf{b}_{-v}), \mathbf{C}'(\mathbf{b}_{-v})\}$ we have $b_v \in Z$ and $b_v >_v Z \setminus \{b_v\}$.*

We will now leverage Corollary 1 to derive a polynomial-time algorithm for computing a TH best response of a given voter v. Importantly, the input to this algorithm is v's utility function and the other voters' ballots; in particular, to compute his best response, v does not need to know the utility functions of other voters.

Theorem 1. *Given a voter $v \in \mathbf{V}$ with an ordinary utility function δ_v and a ballot profile $\mathbf{b}_{-v}$ of the remaining voters, we can find in polynomial time a TH-BR of v to $\mathbf{b}_{-v}$.*

Proof. Fix a voter v and a ballot profile $\mathbf{b}_{-v}$. By Proposition 1 we can efficiently compute a TH best response if $|W(\mathbf{b}_{-v})| > 1$, so from now we will assume that $W(\mathbf{b}_{-v}) = \{w\}$ for some $w \in \mathbf{C}$. Let $c^* = c^*(v; \mathbf{b}_{-v})$, $c' = c'(v; \mathbf{b}_{-v})$, $c'' = c''(v; \mathbf{b}_{-v})$. We know that the set of v's TH best responses to $\mathbf{b}_{-v}$ is a subset of $\{w, c^*, c', c''\}$. Thus, it suffices to compare four values—$\widetilde{U}_v^\varepsilon(w, \mathbf{b}_{-v})$, $\widetilde{U}_v^\varepsilon(c^*, \mathbf{b}_{-v})$, $\widetilde{U}_v^\varepsilon(c', \mathbf{b}_{-v})$ and $\widetilde{U}_v^\varepsilon(c'', \mathbf{b}_{-v})$—for sufficiently small ε.

Suppose that $s^w - s^*$ is even and let $\eta^* = (s^w - s^*)/2$ (when $s^w - s^*$ is odd the analysis is similar, and we omit it due to space constraints). For $\eta < \eta^*$ we have $W(x, \widetilde{\mathbf{b}}_{-v}) = \{w\}$ for all $\widetilde{\mathbf{b}}_{-v} \in \mathbf{B}(\mathbf{b}_{-v}, \eta)$ and all $x \in \{w, c^*, c', c''\}$. Further, there is a value $\alpha > 0$ that does not depend on ε such that $|\widetilde{U}_v^\varepsilon(x, \mathbf{b}_{-v}; \eta)| < \alpha\varepsilon^{\eta^*+1}$ for each $\eta \geq \eta^* + 1$ and each $x \in \{w, c^*, c', c''\}$. Thus, just as in the proof of Proposition 2, we will focus on $\widetilde{U}_v^\varepsilon(x, \mathbf{b}_{-v}; \eta^*)$ for $x \in \{w, c^*, c', c''\}$.

Suppose first that $b_v = w$. Then for every ballot profile $\widetilde{\mathbf{b}}_{-v} \in \mathbf{B}(\mathbf{b}_{-v}, \eta^*)$ we have $W(w, \widetilde{\mathbf{b}}_{-v}) = \{w\}$.

Now, suppose that $b_v = c^*$. We have $W(c^*, \widetilde{\mathbf{b}}_{-v}) \neq \{w\}$ for a ballot profile $\widetilde{\mathbf{b}}_{-v} \in \mathbf{B}(\mathbf{b}_{-v}, \eta^*)$ if and only if one of the following four conditions holds:

(a) exactly η^* voters in $\widetilde{\mathbf{b}}_{-v}$ err by voting for c^* instead of w, in which case $W(c^*, \widetilde{\mathbf{b}}_{-v}) = \{c^*\}$;
(b) exactly η^* voters in $\widetilde{\mathbf{b}}_{-v}$ err by voting for some $x \in \mathbf{C}^* \setminus \{c^*\}$ instead of w, in which case $W(c^*, \widetilde{\mathbf{b}}_{-v}) = \{x, w\}$;
(c) exactly $\eta^* - 1$ voters in $\widetilde{\mathbf{b}}_{-v}$ err by voting for c^* instead of w, and there is one voter $u \in \mathbf{V} \setminus \{v\}$ with $b_u = w$, $\widetilde{b}_u = y$ for some $y \notin \{w, c^*\}$, in which case $W(c^*, \widetilde{\mathbf{b}}_{-v}) = \{c^*, w\}$;
(d) exactly $\eta^* - 1$ voters in $\widetilde{\mathbf{b}}_{-v}$ err by voting for c^* instead of w, and there is one voter $u \in \mathbf{V} \setminus \{v\}$ with $b_u = z$, $\widetilde{b}_u = c^*$ for some $z \notin \{w, c^*\}$, in which case $W(c^*, \widetilde{\mathbf{b}}_{-v}) = \{c^*, w\}$.

Let P_a, P_b, P_c, and P_d be the probability that a ballot profile selected according to $\mathbf{b}_{-v}^\varepsilon$ satisfies condition (a), (b), (c), or (d), respectively, and let $P_t = P_a + P_b + P_c + P_d$. We have

$$
\begin{aligned}
P_a &= \binom{s^w}{\eta^*}\zeta(\varepsilon, \eta^*); \\
P_b &= (|\mathbf{C}^*| - 1)\binom{s^w}{\eta^*}\zeta(\varepsilon, \eta^*); \\
P_c &= \binom{s^w}{\eta^*} \cdot \eta^* \cdot (m-2)\zeta(\varepsilon, \eta^*); \\
P_d &= \binom{s^w}{\eta^* - 1} \cdot (n - 1 - s^w - s^*) \cdot \zeta(\varepsilon, \eta^*).
\end{aligned}
$$

$$
\begin{aligned}
\text{Let } X &= P_a\delta_v(c^*) + \frac{P_b}{|\mathbf{C}^*| - 1}\sum_{x \in \mathbf{C}^* \setminus \{c^*\}} \frac{\delta_v(x) + \delta_v(w)}{2} \\
&+ (P_c + P_d)\frac{\delta_v(c^*) + \delta_v(w)}{2},
\end{aligned}
$$

with the convention that the second term in the left-hand side is 0 when $|\mathbf{C}^*| = 1$. We conclude that $\widetilde{U}_v^\varepsilon(w, \mathbf{b}_{-v}; \eta^*) > \widetilde{U}_v^\varepsilon(w, \mathbf{b}_{-v}; \eta^*)$ if $X > P_t\delta_v(w)$ and $\widetilde{U}_v^\varepsilon(w, \mathbf{b}_{-v}; \eta^*) < \widetilde{U}_v^\varepsilon(w, \mathbf{b}_{-v}; \eta^*)$ if $X < P_t\delta_v(w)$. Further, since δ_v is an ordinary utility function, we know that $X \neq P_t\delta_v(w)$, and hence $|X - P_t\delta_v(w)| \geq \beta\varepsilon^{\eta^*}$ for some constant β and all sufficiently small values of ϵ. Since the quantities $X/\zeta(\varepsilon, \eta^*)$ and $P_t\delta_v(w)/\zeta(\varepsilon, \eta^*)$ can be computed and compared in polynomial time, this shows that v can decide between w and c^* in polynomial time. Note that v may prefer voting for w over voting for c^* even if $c^* >_v w$; this may happen, for instance, if he finds some candidate in $\mathbf{C}^* \setminus \{c^*\}$ highly undesirable. In a similar fashion, v can compare all other pairs of candidates from $\{w, c^*, c'c''\}$ and select his best response. □

Theorem 1 provides evidence that voters may be able to apply TH-based reasoning when choosing their strategies; indeed, if computing a TH best response was intractable, it would be hard to defend TH equilibria as a reasonable solution concept for Plurality voting games.

4 Properties of Trembling Hand Equilibria

We have shown that a voter can efficiently compute his TH best response to other voters' intended ballot profile. We will now establish several useful properties of trembling hand equilibria of Plurality voting. Throughout this section, we assume that all voters have ordinary utilities, and leave simpler claims without proof due to space constraints.

We first revisit Example 1 and characterize the conditions under which a candidate can be a unanimous winner in a THE. Interestingly, the answer strongly depends on whether the number of voters is even or odd.

Theorem 2. *A profile $\mathbf{b} = (w, \ldots, w)$ is a THE if and only if*

- *$n = |\mathbf{V}|$ is even and each voter $v \in \mathbf{V}$ prefers w to all other candidates, or*
- *n is odd and for each voter $v \in \mathbf{V}$ and each candidate $a \in \mathbf{C}$ we have $m\delta_v(w) - \sum_{c \in \mathbf{C}} \delta_v(c) \geq \delta_v(a) - \delta_v(w)$.*

Proof. Suppose that we have $|\mathbf{V}| = 2k$ for some $k \in \mathbb{N}$. For $k = 1$ our claim follows from Proposition 1, so assume $k > 1$. Suppose that a ballot profile $\mathbf{b} = (w, \ldots, w)$ is a THE, and assume for the sake of contradiction that some voter $v \in \mathbf{V}$ has $a >_v w$ for some $a \in \mathbf{C}$. In all η-tremblings with $\eta < k - 1$ candidate w wins irrespective of v's vote, so we focus on $(k-1)$-tremblings. If v votes for w, then w is the unique winner in all such tremblings as well. However, there is a $(k-1)$-trembling $\widetilde{\mathbf{b}}_{-v}$ where all $k-1$ erring voters vote for a, and we have $W(a, \widetilde{\mathbf{b}}_{-v}) = \{a, w\}$, $W(a, \widehat{\mathbf{b}}_{-v}) = \{w\}$ for any other $(k-1)$-trembling $\widehat{\mathbf{b}}_{-v}$. As $(\delta_v(a) + \delta_v(w))/2 > \delta_v(w)$, voting for w is not a TH best response for v, a contradiction.

Now, suppose that $|\mathbf{V}| = 2k+1$ for some $k \in \mathbb{N}$. Consider a voter v, and let a be his most preferred candidate. From symmetry considerations, it is clear that v's TH best response to $\mathbf{b}_{-v}$ is either w or a. To compare the associated TH utilities, it suffices to focus on k-tremblings where all k erring voters vote for the same candidate $c_i \neq w$; if v responds with w, then w remains the unique winner, and if v responds with a, the winning set is $\{w, c_i\}$ when $a \neq c_i$ and $\{a\}$ otherwise. Since all such tremblings are equally likely and voters' utilities are ordinary, w is a TH best response to $\mathbf{b}_{-v}$ if and only if $\sum_{c \in \mathbf{C} \setminus \{w,a\}} \frac{\delta_v(w)+\delta_v(c)}{2} + \delta_v(a) \leq (m-1)\delta_v(w)$. By rearranging the terms, we get the required result. $\square$

Theorem 2 says that for an even number of voters a THE outcome can be a consensus only if all voters agree on the best candidate. For an odd number of voters, the situation is more complicated: a candidate w can be a consensus winner if every voter v believes that w provides a 'substantial' advantage (namely, $\frac{1}{m}(\delta_v(a) - \delta_v(w))$, where a is v's top choice) over a candidate sampled uniformly at random from $\mathbf{C}$.

Example 2. Let $\mathbf{C} = \{c_1, \dots, c_5\}$, $\mathbf{V} = \{u, v, w\}$. If we have $\delta_z(c_i) = i - 1$ for each $z \in \mathbf{V}$ and each $c_i \in \mathbf{C}$, then (c_5, c_5, c_5) and (c_4, c_4, c_4) are THE (since $5 \cdot 4 - 10 \geq 4 - 4$ and $5 \cdot 3 - 10 \geq 4 - 3$), but (c_3, c_3, c_3) is not (since $5 \cdot 2 - 10 < 4 - 2$). In contrast, if we have $\delta_z(c_i) = i - 1$ for $i = 1, 2, 3, 4$, $\delta_z(c_5) = 10$, for each $z \in \mathbf{V}$, then (c_4, c_4, c_4) is not a THE.

Our next result describes THE where the winning set contains at least two candidates. Interestingly, it is reminiscent of existing characterizations of NE with non-singleton winning sets for lazy voters (Theorem 2 in [Desmedt and Elkind, 2010]) and truth-biased voters (Theorem 4 in [Elkind *et al.*, 2015]), under randomized tie-breaking.

Theorem 3. *For every trembling hand equilibrium* $\mathbf{b}$ *with* $|W(\mathbf{b})| > 1$ *we have* $sc(c; \mathbf{b}) > 0$ *if and only if* $c \in W(\mathbf{b})$.

Importantly, Theorem 3 generally implies that contrary to Messner and Polborn [2002; 2011], in our setting the Duverger's law does not hold, i.e., there may exist THE with large winning sets. Even more interestingly, the example below demonstrates that there may exist THE with a single winner and several runners-up.

Example 3. Let $\mathbf{C} = \{a, b, c\}$ and let $|\mathbf{V}| = 14$. There are six voters with preference order abc (where we omit the '$>$' sign for readability), four voters with preference order bac, and another four voters with preference order cab. Each voter assigns utility 4 to his top choice, 2 to his second choice and 1 to his third choice. In this case, the truthful profile is a trembling hand equilibrium, where candidate a is a unique winner, and both b and c are runners-up.

We now focus on THE where the winning set is a singleton, and not all voters vote for the same candidate, i.e., the set of runners-up is non-empty. In what follows, we consider a THE $\mathbf{b}$ with $W(\mathbf{b}) = \{w\}$ and set $s^w = s^w(\mathbf{b})$, $s^* = s^*(\mathbf{b})$.

When the difference between the score of the winner and the score of the runners-up is even, the voters can be partitioned into three groups: the (unique) winner, the runner-up, and the losers (candidates with no votes). Moreover, each voter votes for his most preferred candidate among those with positive scores.

Theorem 4. *If* $s^w - s^*$ *is even then* $sc(c; \mathbf{b}) \in \{0, s^*\}$ *for all* $c \in \mathbf{C} \setminus \{w\}$. *Furthermore,* $b_v >_v \mathbf{C}^+(\mathbf{b}) \setminus \{b_v\}$, $\forall v \in \mathbf{V}$.

When $s^w - s^*$ is odd, there may also exist some candidates whose score is $s^* - 1$.

Theorem 5. *If* $s^w - s^*$ *is odd then* $sc(c; \mathbf{b}) \in \{0, s^*, s^* - 1\}$ *for all* $c \in \mathbf{C} \setminus \{w\}$.

Note, however, that in the latter case a voter may fail to vote for his most preferred candidate in $\mathbf{C}^+(\mathbf{b})$: for instance, v may prefer a candidate c with $sc(c; \mathbf{b}) = s^* - 1$ to w, yet vote for w, because it is a 'safer' choice.

We will now describe a large class of preference profiles that guarantee the existence of a trembling hand equilibrium.

Theorem 6. *Suppose that for a pair of candidates* $a, c \in \mathbf{C}$ *it holds that a strict majority of voters prefer* a *to* c*, yet at least three voters prefer* c *to* a*. Then there exists a THE* $\mathbf{b}$ *such that* $W(\mathbf{b}) = \{a\}$*,* $\mathbf{C}^*(\mathbf{b}) = \{c\}$*, and* $C^+(\mathbf{b}) = \{a, c\}$.

This result indicates that in large-scale elections we are very likely to have at least one THE; indeed, it is likely that *any* candidate can be the unique winner in some THE. However, it is possible to construct a small profile with no THE.

Example 4. Let $\mathbf{C} = \{a, b, c, d, e, f\}$ and let $|\mathbf{V}| = 9$. There are six voters with preference order $abcdef$, one voter with preferences $bcdaef$, one voter with preferences $feadcb$, and one voter with preferences $cbdefa$. Each voter assigns utility 100 to his top choice and utilities in the set $\{1, 2, 3, 4, 5\}$ to the remaining candidates. Note that no pair of candidates in this election satisfies the conditions of Theorem 6. An extensive case analysis shows that this profile admits no THE.

We conclude this section by showing that, even though we now know quite a bit about the structure of trembling hand equilibria, it may still be difficult to decide whether there is a THE that contains a specific candidate w in its winning set.

Theorem 7. *Given a candidate set* $\mathbf{C}$*, a voter set* $\mathbf{V}$*, a utility function* δ_v *for each voter* $v \in \mathbf{V}$ *and a distinguished candidate* $a \in \mathbf{C}$*, it is NP-hard to decide whether the associated Plurality voting game admits a THE* $\mathbf{b}$ *with* $a \in W(\mathbf{b})$.

5 Discussion and Conclusions

Our results show that trembling hand equilibria of Plurality voting have many good properties. In particular, 'silly' Nash equilibria where all voters vote for an undesirable candidate are eliminated by this equilibrium refinement. Further, THE encourage voters to coordinate on reasonable candidates even when these candidates do not win: all candidates other than the winners are either (almost) runners-up, or receive no votes at all. Also, the problem of computing one's TH best response, while seemingly complicated, turns out to be tractable. On the negative side, Theorem 6 shows that essentially any candidate that is not a Condorcet loser may still become a winner in a THE, i.e., the power of this equilibrium refinement as a tool to eliminate bad equilibria is somewhat limited. Moreover, it can 'overshoot', in that some games have no THE at all (however, such cases are very rare compared to other voting equilibrium refinements studied so far).

There are several open problems suggested by our work. First, it would be desirable to remove or weaken the technical requirement that voters' utilities are ordinary. However, this seems to require looking at η-tremblings for several values of η simultaneously, which complicates the analysis significantly. Further, it would be interesting to know whether there exists a polynomial-time algorithm for deciding whether a given preference profile admits a THE (note that this is not ruled out by Theorem 7). Another exciting direction, which is enabled by the tractability result of Theorem 1, is exploring the properties of TH best response dynamics, both theoretically and empirically.

Acknowledgments

Svetlana Obraztsova was supported by the I-CORE-ALGO. Edith Elkind was supported by ERC StG 639945. This research has also received financial support from Israel Science Foundation grant #1227/12, UK Research Council for project ORCHID grant EP/I011587/1, COST Action IC1205 on Computational Social Choice, and RFFI grant 14-01-00156-a.

References

[Acemoglu *et al.*, 2009] D. Acemoglu, G. Egorov, and K. Sonin. Equilibrium refinement in dynamic voting games. Technical report, MIT, Dept. of Economics, 2009.

[Bartholdi, III *et al.*, 1989] J. J. Bartholdi, III, C. A. Tovey, and M. Trick. The computational difficulty of manipulating an election. *Social Choice and Welfare*, pages 227–241, 1989.

[Borgers, 2004] T. Borgers. Costly voting. *American Economic Review*, 94(1):57–66, 2004.

[Boutilier *et al.*, 2012] C. Boutilier, I. Caragiannis, S. Haber, T. Lu, A. D. Procaccia, and O. Sheffet. Optimal social choice functions: a utilitarian view. In *ACM EC'12*, pages 197–214, 2012.

[Desmedt and Elkind, 2010] Y. Desmedt and E. Elkind. Equilibria of plurality voting with abstentions. In *ACM EC'10*, pages 347–356, 2010.

[Dhillon and Lockwood., 2004] A. Dhillon and B. Lockwood. When are plurality rule voting games dominance-solvable? *GEB*, 46:55–75, 2004.

[Dutta and Laslier, 2010] B. Dutta and J.-F. Laslier. Costless honesty in voting. 10th International Meeting of the Society for Social Choice and Welfare, Moscow, 2010.

[Dutta and Sen, 2012] B. Dutta and A. Sen. Nash implementation with partially honest individuals. *Games and Economic Behavior*, 74(1):154–169, 2012.

[Elkind *et al.*, 2015] E. Elkind, S. Obraztsova, E. Markakis, and P. Skowron. Equilibria of plurality voting: Lazy and truth-biased voters. In *SAGT'15*, pages 110–122, 2015.

[Farquharson, 1969] R. Farquharson. *Theory of Voting*. Yale University Press, 1969.

[Filos-Ratsikas and Miltersen, 2014] A. Filos-Ratsikas and P. B. Miltersen. Truthful approximations to range voting. In *WINE'14*, pages 175–188, 2014.

[Grandi *et al.*, 2013] U. Grandi, A. Loreggia, F. Rossi, K. B. Venable, and T. Walsh. Restricted manipulation in iterative voting: Condorcet efficiency and Borda score. In *ADT'13*, pages 181–192, 2013.

[Jackson and Tan, 2013] M. O. Jackson and X. Tan. Deliberation, disclosure of information, and voting. *Journal of Economic Theory*, 148(1):2–30, 2013.

[Kolovos and Harris, 2005] I. Kolovos and P. Harris. Voter apathy in British elections: causes and remedies. Technical report, University of Otago, 2005.

[Mas-Colell *et al.*, 1995] A. Mas-Colell, M. D. Whinston, and J. R. Green. *Microeconomic theory*. Oxford University Press, New York, 1st edition, 1995.

[Meir *et al.*, 2010] R. Meir, M. Polukarov, J. S. Rosenschein, and N. R. Jennings. Convergence to equilibria of plurality voting. In *AAAI'10*, pages 823–828, 2010.

[Meir *et al.*, 2014] R. Meir, O. Lev, and J. S. Rosenschein. A local dominance theory of voting equilibria. In *ACM EC'14*, pages 313–330, 2014.

[Messner and Polborn, 2002] M. Messner and M. K. Polborn. Robust political equilibria under plurality and runoff rule. Mimeo, Bocconi University, 2002.

[Messner and Polborn, 2011] M. Messner and M. K. Polborn. Miscounts, Duverger's law and Duverger's hypothesis. Technical report, Innocenzo Gasparini Institute for Economic Research, 2011.

[Moulin, 1979] H. Moulin. Dominance solvable voting schemes. *Econometrica*, 47:1337–1351, 1979.

[Myerson and Weber, 1993] R. Myerson and R. Weber. A theory of voting equilibria. *American Political Science Review*, 87(1):102–114, 1993.

[Obraztsova *et al.*, 2011] S. Obraztsova, E. Elkind, and N. Hazon. Ties matter: complexity of voting manipulation revisited. In *AAMAS'11*, pages 71–78, 2011.

[Obraztsova *et al.*, 2013] S. Obraztsova, E. Markakis, and D. R. M. Thompson. Plurality voting with truth-biased agents. In *SAGT'13*, pages 26–37, 2013.

[Obraztsova *et al.*, 2015] S. Obraztsova, E. Elkind, M. Polukarov, and Z. Rabinovich. Doodle poll games. In *AGT Workshop at IJCAI*, 2015.

[Rabinovich *et al.*, 2015] Z. Rabinovich, S. Obraztsova, O. Lev, E. Markakis, and J. S. Rosenschein. Analysis of equilibria in iterative voting schemes. In *AAAI'15*, 2015.

[Reijngoud and Endriss, 2012] A. Reijngoud and U. Endriss. Voter response to iterated poll information. In *AAMAS'12*, pages 635–644, 2012.

[Reyhani and Wilson, 2012] R. Reyhani and M. C. Wilson. Best reply dynamics for scoring rules. In *ECAI'2012*, pages 672–677, 2012.

[Selten, 1975] R. Selten. A reexamination of the perfectness concept for equilibrium points in extensive game. *International Journal of Game Theory*, 4:25–55, 1975.

[Thompson *et al.*, 2013] D. R. M. Thompson, O. Lev, K. Leyton-Brown, and J. S. Rosenschein. Empirical analysis of plurality election equilibria. In *AAMAS'13*, pages 391–398, 2013.

[van Damme, 1991] E. van Damme. *Stability and Perfection of Nash Equlibria*. New York: Springer, 2nd edition, 1991.

[Xia and Conitzer, 2010] L. Xia and V. Conitzer. Stackelberg voting games: Computational aspects and paradoxes. In *AAAI*, pages 805–810, 2010.

[Xia, 2012] L. Xia. Computing the margin of victory for various voting rules. In *ACM EC'12*, pages 982–999, 2012.