Iterative Voting and Acyclic Games[☆]

Reshef Meir[*]

Technion—Israel Institute of Technology

Maria Polukarov

King's College London, United Kingdom

Jeffrey S. Rosenschein

The Hebrew University of Jerusalem, Israel

Nicholas R. Jennings

Imperial College London, United Kingdom
King Abdulaziz University, Saudi Arabia

Abstract

Multi-agent decision problems, in which independent agents have to agree on a joint plan of action or allocation of resources, are central to artificial intelligence. In such situations, agents' individual preferences over available alternatives may vary, and they may try to reconcile these differences by voting.

We consider scenarios where voters cannot coordinate their actions, but are allowed to change their vote after observing the current outcome, as is often the case both in offline committees and in online voting. Specifically, we are interested in identifying conditions under which such iterative voting processes are guaranteed to converge to a Nash equilibrium state—that is, under which this process is acyclic. We classify convergence results based on the underlying assumptions about the agent scheduler (the order in which the agents take their actions) and the action scheduler (the actions available to the agents at each step). By so doing, we

[☆]Preliminary versions of this paper were presented at AAAI-2010 [1] and at SAGT-2016 [2].

[*]Corresponding author

Email addresses: `reshefm@ie.technion.ac.il` (Reshef Meir), `maria.polukarov@kcl.ac.uk` (Maria Polukarov), `jeff@cs.huji.ac.il` (Jeffrey S. Rosenschein), `n.jennings@imperial.ac.uk` (Nicholas R. Jennings)

position iterative voting models within the general framework of acyclic games and game forms.

In more detail, our main technical results provide a complete picture of conditions for acyclicity in several variations of Plurality voting. In particular, we show that (a) under the traditional lexicographic tie-breaking, the game converges from any state and for any order of agents, under a weak restriction on voters' actions; and that (b) Plurality with randomized tie-breaking is not guaranteed to converge under arbitrary agent schedulers, but there is always *some* path of better replies from any initial state of the game to a Nash equilibrium. We thus show a first separation between order-free acyclicity and weak acyclicity of game forms, thereby settling an open question from [Kukushkin 2011]. In addition, we refute another conjecture of Kukushkin regarding strongly acyclic voting rules, by proving the existence of strongly acyclic separable game forms.

Keywords: Iterative voting, Acyclicity, Convergence, Nash equilibrium

---

1. Introduction

Voting mechanisms are a popular tool for preference aggregation and collective decision making in multi-agent systems. One major concern when applying such mechanisms is that voters may misreport their real preferences in order to affect the outcome in their favor. Indeed, most voting rules are known to be susceptible to such strategic behavior [3, 4]. It is therefore natural to employ game-theoretic tools in order to model voting behavior and assess the outcome of a voting process. Specifically, in this work we are interested in identifying conditions under which the process will converge to an equilibrium state where no voter has an incentive to change his vote—that is, conditions such that the game induced by the voting process is acyclic.

Researchers in economics and game theory since Cournot [5] have been developing a formal framework to study questions about acyclicity (see, for example, [6, 7, 8]). Acyclic games have several attractive features: not only do they possess an equilibrium in pure strategies, they also guarantee that local improvement dynamics will always lead to one. These properties are highly desirable both from an economic and a computational perspective, as they imply that a system has a stable state which is reachable in a decentralized way, often with little information and communication.

The analysis of acyclicity of voting games is of particular interest to AI as it tackles the fundamental problem of multi-agent decision making, where autonomous agents (that may be distant, self-interested and/or unknown to one another) have to choose a joint plan of action or allocate resources or goods. Agents may vote strategically based on their current information, and keep updating their vote as the current state changes, and thus it is the local dynamics (in addition to preferences) that determines the outcome.

Now, there are several degrees of acyclicity, depending on the initial state of the process, the type of improvement steps that agents may take and the order in which they may act (see the classification scheme of Kukushkin [9, 10, 7]). In particular, a better reply denotes any change of strategy that strictly improves the utility of the agent, and a game that admits no cycles of better replies whatsoever is called strongly acyclic. In contrast, weak acyclicity means that while cycles may generally occur, there is at least one path of better replies that leads to an equilibrium from any initial state. Order-free acyclicity is a middle ground,[1] requiring convergence for any order of agents (agent scheduler), but allowing the action scheduler to restrict the way they choose among several available replies (e.g., only allowing best replies that maximize the agent's utility among all better replies). In this work, we apply this classification to games that arise in the context of voting.

Specifically, we consider the model of iterative voting. In this model, voters have fixed private preferences and start from some announcement (e.g., sincerely report their preferences or submit random votes). Votes are then aggregated via some predefined rule (e.g., Plurality, Veto or Borda), but the agents can change their votes after observing the current announcement and the outcome, but not other voters' true preferences. The game proceeds in turns, where a single voter changes his vote at each turn, until no voter has objections and the final outcome is announced. Note that voters remain ignorant regarding the true preferences of the other voters. This process is similar to online polls via Doodle or Facebook, where users can log-in at any time and change their ballot. Similarly, some offline committees (e.g., for recruitment processes or art competitions) often hold straw votes, or an informal process where a member can ask to change his vote when she sees the current outcome or receives additional information about the applicants. Another voter may react by changing his vote as well, and so on. Interestingly,

[1] Order-free acyclicity is sometimes referred to as *restricted acyclicity* [7], but we find this term needlessly ambiguous.

even in presidential elections, in several states of the United States people can change their votes if they cast their ballots early [11].

The formal study of iterative voting rules was initiated about seven years ago in our AAAI paper that was a preliminary version of this work [1]. Subsequent papers on iterative voting typically focused on common voting rules such as Plurality, Veto and Borda, and studied the conditions under which convergence of the iterative process to an equilibrium is guaranteed [12, 13, 14]. However, despite the fact that both fields ask similar questions, the iterative voting literature has remained largely detached from the more general literature on acyclicity in games. Bridging this gap is the main conceptual contribution of this work, and is important for two reasons. First, the analysis of conditions that entail acyclicity of games and game forms is crucial to the understanding of iterative voting scenarios and the ability to properly compare their convergence properties (e.g., convergence of any best reply dynamics is a special case of order-free acyclicity and convergence under a particular order of voters implies weak acyclicity). Second, convergence results for specific voting rules under best/better reply dynamics may shed light on more general questions regarding acyclicity of voting processes. To this end, we apply the formalism of Kukushkin [7] for strong/order-free/weak acyclicity of game forms, which allows us to re-interpret both known and new results on convergence of better and best reply dynamics in voting games, and to answer some open questions.

1.1. Related work

Kukushkin [7] provided several partial characterizations for game forms with strong acyclicity. In particular, he showed that if we further strengthen the acyclicity requirement to demand an ordinal potential [6], then this is attained if and only if the game form is dictatorial—i.e., there is at most one voter that can affect the outcome. He further characterized game forms that are strongly acyclic under coalitional improvements, and provided broad classes of game forms that are "almost strongly acyclic"—i.e., order-free acyclic with only mild restrictions on voters' actions. Other partial characterizations have been provided for acyclicity in complete information extensive form games [15, 16]. The most relevant aspects of this work are explained in more detail in the following sections.

The study of classes of games with specific utility structures that are guaranteed to be acyclic or weakly acyclic has attracted much attention, in particular regarding the existence and properties of potential functions [6, 17, 18, 8]. We discuss these below.

Strategic voting

The notion of strategic voting has been highlighted in research on social choice as crucial to understanding the relationship between preferences of a population and the final outcome of elections (see, for example, [19, 20, 21]). In various applications, ranging from political domains to AI, the most widely used voting rule is Plurality, in which each voter has one vote and the winner is the candidate who receives the highest number of votes. While it is known that no reasonable voting rule is completely immune to strategic behavior [3, 4], Plurality has been shown to be particularly susceptible, both in theory [21, 22] and in practice [23]. This makes the analysis of any election campaign—even one where the simple Plurality rule is used—a challenging task. As voters may speculate and counter-speculate, it would be beneficial to have formal tools that help us understand (and perhaps predict) the final outcome.

In particular, natural tools for this task include the well-studied solution concepts developed for normal form games, such as better/best replies, dominant strategies or different variants of equilibrium. Now, while voting settings are commonly presented in other forms, several natural normal form formulations have been proposed in the past [24, 25, 26, 27, 28]. These formulations are extremely simple for Plurality voting games, where voters have only few available ways to vote. Specifically, some of this previous work has been devoted to the analysis of solution concepts such as elimination of dominated strategies [24] and strong equilibria [26]. Other multi-step voting procedures that have been proposed in the literature are the iterated majority vote [29] and extensive form games, where voters vote one by one [30]. However, in contrast to iterative voting, these models are inconsistent with the better reply dynamics in the framework of normal form games, and are analyzed via different techniques. A model somewhat more similar to ours was recently studied in [31], where voters can choose between voting truthfully or manipulating under the assumption that everyone else is truthful. That is, in this model each voter has exactly two available actions, whereas in ours all valid votes are allowed.

Convergence of better reply dynamics in iterative voting for particular voting rules has been studied extensively in the computational social choice literature. We summarize and compare these findings with our results in the concluding section, particularly in Table 3.

An important question in the context of strategic voting, including the iterative voting model, is whether an obtained equilibrium state is good for the society according to various metrics. To this end, Branzei et al. [32] showed bounds on what they term the dynamic price of anarchy, that evaluates how far the final outcome can be from the initial truthful outcome. Other work in this line used simulations to show that iterative voting may improve the social welfare or Condorcet efficiency [33, 34, 35].

A similar variant of iterative voting in the context of multi-issue voting was studied by [36] via simulations.

Biased and sophisticated voting

Some recent work on iterative voting deals with voters who are uncertain, truth-biased, lazy-biased, bounded-rational, non-myopic, or apply some other restrictions and/or heuristics that diverge from the standard notion of better reply in games [37, 38, 33, 39, 34, 40, 13, 41, 42]. The outcomes of such dynamics are not necessarily Nash equilibria, which means that some voters could still benefit from changing their votes in such states, should the limitations on their available actions be removed. In this work, we exclude the possibility of this potential instability, and deal exclusively with myopic better and best reply dynamics that (if they converge) lead to a Nash equilibrium state.[2]

1.2. Contribution

On the conceptual level we introduce a model to handle dynamics of strategic behavior in voting settings and position variants of this iterative voting model within the general framework of acyclic games and game forms.

In Section 3 we consider strong acyclicity, and settle an open question regarding the existence of acyclic non-separable game forms, by explicitly constructing one, thereby refuting a conjecture by Kukushkin [7].

Section 4 focuses on order-free acyclicity of the Plurality rule. Our main result in this section shows that to guarantee convergence, it is necessary and sufficient that voters restrict their actions in a natural way that we term *direct reply*—meaning that a voter will only reassign his vote to a candidate that will become a winner as a result. Importantly, it is not sufficient to restrict the set of actions to best (but possibly indirect) replies: best reply dynamics, as we demonstrate, may contain cycles. However, best reply dynamics is guaranteed to converge *from the truthful state*, under either lexicographic or randomized tie-breaking.

In Section 5, we use variations of Plurality to show a strict separation between order-free acyclicity and weak acyclicity, thereby settling another open question

[2]However, we do consider two standard ways to handle ties that slightly relax the better-reply definition (see Section 4.4).

from [7]. In particular, we show that if we add either weights (plus some restriction on the votes) or random tie-breaking to the Plurality rule, we get a game form that is weakly acyclic, but not order-free acyclic, since the order of voters is crucial for convergence.

We conclude in Section 6, where we also classify all known convergence results in iterative voting according to the standard taxonomy of acyclicity in games.

2. Model and preliminaries

We usually denote sets by uppercase letters (e.g., $A, B, \ldots$), and vectors by bold letters (e.g., $\boldsymbol{a} = (a_1, \ldots, a_n)$). In some cases the i'th entry of $\boldsymbol{a}$ is referred to as $a(i)$. For a set X we denote by $\mathcal{L}(X)$ the set of permutations over X.

2.1. *Voting rules and game forms*

There is a set C of m alternatives (or, *candidates*), and a set N of n strategic agents (*voters*). A game form (also called a *voting rule*) f allows each agent $i \in N$ to select an action a_i from a set A_i (we also refer to a_i as the *vote* of agent i). The input to f is therefore a vector $\boldsymbol{a} = (a_1, \ldots, a_n)$ called an *action profile* (or, a *voting profile*). Then, f chooses a winning alternative—i.e., it is a function $f : \mathcal{A} \to C$, where $\mathcal{A} = \times_{i \in N} A_i$ (see Figure 1 for examples).

A voting rule f is *standard* if $A_i = A$ for all i, and A is either $\mathcal{L}(C)$ or a coarsening of $\mathcal{L}(C)$.[3] For example, in *Plurality*—one of the most prominent voting rules—we have that $A = C$, and the winner is the candidate with the most votes; thus, all permutations with the same leading candidate are considered to be the same action. For a permutation $L \in \mathcal{L}(C)$, we denote the first element—i.e., the leading candidate—in L by $\text{top}(L)$. Like Plurality, most common voting rules except Approval are standard.

We allow for a broader set of "Plurality game forms" by considering *weighted* and *fixed* voters, and varying the *tie-breaking* method. Specifically, each of the strategic voters $i \in N$ has an integer weight $w_i \in \mathbb{N}$, and there are also $\hat{n}$ fixed voters who do not play strategically or change their vote. The value $\hat{s}(c)$ (called the *initial score*) specifies the number of fixed votes for each candidate c. The vector $\boldsymbol{w} \in \mathbb{N}^n$ of weights and the vector $\hat{\boldsymbol{s}} \in \mathbb{N}^m$ of initial scores are part of the game form.[4]

[3]Standard voting rules are also called *resolute social choice functions*, or *SCF* (see [43]).

[4]All of our results still hold if there are no fixed voters, but allowing fixed voters enables the

f_1	a	b	c
a	a	a	a
b	b	b	b
c	c	c	c

f_2	a	b	c
a	a	a	a
b	a	b	b
c	a	b	c

f_3	x	y
a	a	b
b	b	c
c	c	a

f_4	x	y	z	w
a	ax	ay	az	aw
b	bx	by	bz	bw
c	cx	cy	cz	cw

Figure 1: Four examples of game forms with two agents. f_1 is a dictatorial game form with 3 candidates (the row agent is the dictator). f_2 is the Plurality voting rule with 3 candidates and lexicographic tie-breaking. f_3 and f_4 are non-standard game forms. In f_3, $A_1 = C = \{a, b, c\}, A_2 = \{x, y\}$. Note that f_4 is completely general (there are 3×4 possible outcomes in C, one for each voting profile) and can represent any 3-by-4 game.

$f^{PL}_{\boldsymbol{w},\hat{\boldsymbol{s}}}$	a	b	c
a	$(14, 9, 3)$ $\{a\}$	$(10, 13, 3)$ $\{b\}$	$(10, 9, 7)$ $\{a\}$
b	$(11, 12, 3)$ $\{b\}$	$(7, 16, 3)$ $\{b\}$	$(7, 12, 7)$ $\{b\}$
c	$(11, 9, 6)$ $\{a\}$	$(7, 13, 6)$ $\{b\}$	$(7, 9, 10)$ $\{c\}$

Figure 2: A game form $f^{PL}_{\boldsymbol{w},\hat{\boldsymbol{s}}}$, where $N = \{1, 2\}$, $A_1 = A_2 = C = \{a, b, c\}$, $\hat{\boldsymbol{s}} = (7, 9, 3)$ and $\boldsymbol{w} = (3, 4)$ (i.e., voter 1 has weight 3 and voter 2 has weight 4). The table shows the final score vector $\boldsymbol{s}_{(a_1,a_2)}$ for every joint action of the two voters, and the respective winning candidate $f^{PL}_{\boldsymbol{w},\hat{\boldsymbol{s}}}(a_1, a_2)$ in curly brackets.

The *final score* of a candidate c for a given profile $\boldsymbol{a} \in A^n$ in the Plurality game form $f_{\hat{\boldsymbol{s}},\boldsymbol{w}}$ is the sum of the initial score and the total weight of strategic voters that vote for c. We denote the final score vector by $\boldsymbol{s}_{\hat{\boldsymbol{s}},\boldsymbol{w},\boldsymbol{a}}$ (often just $\boldsymbol{s}_{\boldsymbol{a}}$ or $\boldsymbol{s}$ when the other parameters are clear from the context), where $s(c) = \hat{s}(c) + \sum_{i \in N : a_i = c} w_i$.

The Plurality rule selects a candidate from $W = \text{argmax}_{c \in C}\, s_{\hat{\boldsymbol{s}},\boldsymbol{w},\boldsymbol{a}}(c)$, breaking ties according to some specified method. As our results show, acyclicity properties may strongly depend on the tie-breaking method. The two primary variations we consider are $f^{PL}_{\hat{\boldsymbol{s}},\boldsymbol{w}}$ which breaks ties lexicographically, and $f^{PR}_{\hat{\boldsymbol{s}},\boldsymbol{w}}$ which selects a winner from W uniformly at random. As with $\boldsymbol{s}$, we omit the subscripts $\boldsymbol{w}$ and $\hat{\boldsymbol{s}}$ when they are clear from the context.

For illustration, consider an example in Figure 2, demonstrating a specific weighted Plurality game form with two agents.

---

introduction of simpler examples, and facilitates some of the proofs, see Remark 4.1. For further discussion on fixed voters see [31].

2.2. Incentives

Games are attained by adding either cardinal or ordinal utility to a game form. The linear order relation $L_i \in \mathcal{L}(C)$ reflects the preferences of agent i. That is, i prefers c over c' (denoted $c \succ_i c'$) if $(c, c') \in L_i$. The vector containing the preferences of all n agents is called a *preference profile*, and is denoted by $\boldsymbol{L} = (L_1, \ldots, L_n)$. The game form f, coupled with a preference profile $\boldsymbol{L}$, defines an ordinal utility normal form game $G = \langle f, \boldsymbol{L} \rangle$ with n agents, where agent i prefers outcome $f(\boldsymbol{a})$ over outcome $f(\boldsymbol{a}')$ if $f(\boldsymbol{a}) \succ_i f(\boldsymbol{a}')$. In standard game forms, the action a_i may indicate the agent's preferences, hence their common identification with voting rules.

Improvement steps and equilibria

Having defined a normal form game, we can now apply standard solution concepts. Let $G = \langle f, \boldsymbol{L} \rangle$ be a game, and let $\boldsymbol{a} = (\boldsymbol{a}_{-i}, a_i)$ be a joint action in G.

We denote by $\boldsymbol{a} \xrightarrow{i} \boldsymbol{a}'$ an *individual improvement step* (or, *better reply*), if (1) $\boldsymbol{a}, \boldsymbol{a}'$ differ only by the action of agent i; and (2) $f(a_{-i}, a'_i) \succ_i f(a_{-i}, a_i)$. We sometimes omit the actions of the other voters $\boldsymbol{a}_{-i}$ when they are clear from the context, only writing $a_i \xrightarrow{i} a'_i$. We denote by $I_i(\boldsymbol{a}) \subseteq A_i$ the set of actions a'_i s.t. $a_i \xrightarrow{i} a'_i$ is an improvement step of agent i in $\boldsymbol{a}$. Similarly, $I(\boldsymbol{a}) = \bigcup_{i \in N} \bigcup_{a'_i \in I_i(\boldsymbol{a})} (\boldsymbol{a}_{-i}, a'_i)$ contains all states accessible from $\boldsymbol{a}$ by some better reply. $\boldsymbol{a} \xrightarrow{i} a'_i$ is called a *best reply* if i weakly prefers $f(\boldsymbol{a}_{-i}, a'_i)$ over any candidate $f(\boldsymbol{a}_{-i}, b_i)$ s.t. $b \in I_i(\boldsymbol{a})$.

A joint action $\boldsymbol{a}$ is a (pure strategy) *Nash equilibrium* (NE) in G if $I(\boldsymbol{a}) = \emptyset$. That is, no agent can gain by changing his vote, provided that others keep their strategies unchanged. *A priori*, a game may not admit any NE in pure strategies.

Now, observe that when f is a standard voting rule the preference profile $\boldsymbol{L}$ induces a special joint action $\boldsymbol{a}^* = \boldsymbol{a}^*(\boldsymbol{L})$, termed the *truthful state*, where a^*_i equals (the coarsening of) L_i. For example, in Plurality $a^*_i = \text{top}(L_i)$. We refer to $f(\boldsymbol{a}^*)$ as the *truthful outcome* of the game $\langle f, \boldsymbol{L} \rangle$.

The truthful state may or may not be included in the NE points of the game, as can be seen from Tables 3 and 4 that demonstrate games that are induced by adding incentives to the game form shown in Figure 2, and indicate the truthful states and the NE points in these games.

2.3. Iterative Games

We consider natural *dynamics* in iterative games. Assume that agents start by announcing some initial profile $\boldsymbol{a}^0$, and then proceed as follows: at each step t a

$\langle f, \boldsymbol{L}^1 \rangle$	a	b	$* \, c$
$* \, a$	$\boldsymbol{\{a\} \, 3, 2}$	$\{b\} \, 2, 1$	$* \, \boldsymbol{\{a\} \, 3, 2}$
b	$\{b\} \, 2, 1$	$\boldsymbol{\{b\} \, 2, 1}$	$\{b\} \, 2, 1$
c	$\{a\} \, 3, 2$	$\{b\} \, 2, 1$	$\{c\} \, 1, 3$

Figure 3: A game $G = \langle f, \boldsymbol{L}^1 \rangle$, where $f = f^{PL}_{\boldsymbol{w}, \hat{\boldsymbol{s}}}$ is as in Fig. 2, and $\boldsymbol{L}^1$ is defined by $a \succ_1 b \succ_1 c$ and $c \succ_2 a \succ_2 b$. The table shows the ordinal utility of the outcome to each agent, where 3 means the best candidate. **Bold** outcomes are the NE points. Here the truthful vote (marked with *) is also a NE.

$\langle f, \boldsymbol{L}^2 \rangle$	a	b	$* \, c$
$* \, a$	$\{a\} \, 3, 1$	$\boldsymbol{\{b\} \, 1, 2}$	$* \, \{a\} \, 3, 1$
b	$\{b\} \, 1, 2$	$\boldsymbol{\{b\} \, 1, 2}$	$\{b\} \, 1, 2$
c	$\{a\} \, 3, 1$	$\{b\} \, 1, 2$	$\{c\} \, 2, 3$

Figure 4: This game has the same game form as in Fig. 2, and the preference profile $\boldsymbol{L}^2$ is $a \succ_1 c \succ_1 b$ and $c \succ_2 b \succ_2 a$. In this case, the truthful vote $\boldsymbol{a}^*(\boldsymbol{L}^2)$ is not a NE.

single agent i may change his vote to $a'_i \in I_i(\boldsymbol{a}^{t-1})$, resulting in a new state (joint action) $\boldsymbol{a}^t = (\boldsymbol{a}^{t-1}_{-i}, a'_i)$. The process ends when no agent has objections, and the outcome is set by the last state.

Local improvement graphs and schedulers. Any game G induces a directed graph whose vertices are all action profiles (states) $\mathcal{A}$, and edges are all local improvement steps [44, 16]. The pure Nash equilibria of G are all states with no outgoing edges.

Since generally a state may have multiple outgoing edges ($|I(\boldsymbol{a})| > 1$), we need to specify which one is selected in a given play. An *agent scheduler* is a function ϕ^N that given a finite sequence of states $\boldsymbol{a}^0, \dots, \boldsymbol{a}^t$ where $\boldsymbol{a}^t$ is not a PNE, selects a player i such that $I_i(\boldsymbol{a}^t) \neq \emptyset$ [8].[5] Since we also need to decide which action is played by i, we define an *action scheduler* as a function ϕ^A that selects a single action from $I_i(\boldsymbol{a}^t)$. Thus a scheduler $\phi = (\phi^N, \phi^A)$ is a function mapping any sequence not ending in a PNE to a better reply of some agent.

[5] Apt and Simon [8] make a finer distinction to state-based and other subclasses of schedulers, that is not important for our results.

Convergence and acyclicity. Given a game G, an initial action profile $\boldsymbol{a}^0$ and a scheduler ϕ, we get a unique (possibly infinite) path of steps.[6] Also, it is readily apparent to see that the path is finite if and only if it reaches a Nash equilibrium (which is the last state in the path). We say that the triple $\langle G, \boldsymbol{a}^0, \phi\rangle$ *converges* if the induced path is finite.

Following [6, 17], a game G has the *finite improvement property* (we say that G *is* FIP), if $\langle G, \boldsymbol{a}^0, \phi\rangle$ converges for *any* $\boldsymbol{a}^0$ and scheduler ϕ. Games that are FIP are also known as (strongly) *acyclic games* and as *generalized ordinal potential games* [6]. It is quite easy to see that not all Plurality games are FIP (see examples in Section 4).

However, there are alternative, weaker notions of acyclicity and convergence.

- A game G is *weakly-FIP* if there is *some* scheduler ϕ such that $\langle G, \boldsymbol{a}^0, \phi\rangle$ converges for any $\boldsymbol{a}^0$. Such games are known as *weakly acyclic* [8].
- A game G is *order-free-FIP* (or, *order-free acyclic*) if there is *some action scheduler* ϕ^A such that $\left\langle G, \boldsymbol{a}^0, (\phi^N, \phi^A)\right\rangle$ converges for any $\boldsymbol{a}^0$ and ϕ^N [7].

Intuitively, order-free-FIP means that there is some restriction agents can adopt such that convergence is guaranteed regardless of the order in which they play. Kukushkin identifies a particular restriction of interest, namely restriction to best reply improvements. Formally, an action scheduler ϕ^A is a *BR action scheduler* if it always selects from $I_i(\boldsymbol{a}^t)$ a best reply of i. $\phi = \left\langle \phi^N, \phi^A\right\rangle$ is BR if ϕ^A is BR. We get the following definitions for a game G, where FBRP stands for *finite best reply property*:

- G is *FBRP* if $\langle G, \boldsymbol{a}^0, \phi\rangle$ converges for any $\boldsymbol{a}^0$ and any BR ϕ.
- G is *weakly-FBRP* if there is a BR ϕ such that $\langle G, \boldsymbol{a}^0, \phi\rangle$ converges for any $\boldsymbol{a}^0$.
- G is *order-free-FBRP* if there is a BR ϕ^A such that $\left\langle G, \boldsymbol{a}^0, (\phi^N, \phi^A)\right\rangle$ converges for any $\boldsymbol{a}^0$ and ϕ^N.

In this paper, we identify a different restriction, namely *direct reply*, that is well defined under the Plurality rule. Formally, a step $\boldsymbol{a} \xrightarrow{i} \boldsymbol{a}'$ is a direct reply if $f(\boldsymbol{a}') = a'_i$, i.e., if i votes for the new winner (see labeled examples in Section 4). Another rule where a natural direct reply exists is Veto, where a voter can veto the

[6]This is in contrast to some definitions of schedulers that allow multiple possible paths [8, 45].

$$\begin{array}{ccc|ccc|ccccc}
 & & \text{FBRP} & & & & \text{order-free-FBRP} & \Rightarrow & \text{weak-FBRP} & & \\
 & & \Uparrow & & & & \Downarrow & & \Downarrow & & \\
\text{ordinal potential} & \Rightarrow & \text{FIP} & \Rightarrow & \text{FDBRP} & \Rightarrow & \text{order-free-FIP} & \Rightarrow & \text{weak-FIP} & \Rightarrow & \text{pure Nash} \\
\text{exists} & & \Downarrow & & & & \Uparrow & & \Uparrow & & \text{exists} \\
 & & \text{FDRP} & & & & \text{order-free-FDRP} & \Rightarrow & \text{weak-FDRP} & &
\end{array}$$

Figure 5: A double arrow X $\Rightarrow$ Y means that any game or game form with the X property also has the Y property. A triple arrow means that any property on the premise side entails all properties on the conclusion side. The third row is only relevant for Plurality/Veto, where direct-reply is well defined.

current winner [46]. [7] Formally, ϕ^A is a ***DR action scheduler*** if it always selects a direct reply from $I_i(\boldsymbol{a}^t)$, and $\phi = \left\langle \phi^N, \phi^A \right\rangle$ is DR if ϕ^A is DR. We get the following definitions for a Plurality game G, where FDRP stands for *finite direct reply property*:

- G is *FDRP* if $\langle G, \boldsymbol{a}^0, \phi \rangle$ converges for any $\boldsymbol{a}^0$ and any DR ϕ.
- G is *weakly-FDRP* if there is a DR ϕ such that $\langle G, \boldsymbol{a}^0, \phi \rangle$ converges for any $\boldsymbol{a}^0$.
- G is *order-free-FDRP* if there is a DR ϕ^A such that $\left\langle G, \boldsymbol{a}^0, (\phi^N, \phi^A) \right\rangle$ converges for any $\boldsymbol{a}^0$ and ϕ^N.
- FDBRP means that replies are both best and direct. Note that it is unique and thus cannot be further restricted. In Veto there is only one direct reply and thus FDBRP and FDRP coincide.

Restricted dynamics that may disallow all better replies (as in [38, 33]) do not select an action from $I_i(\boldsymbol{a}^t)$ and thus do not fall under the definition of order-free-FIP (but can be considered as a separate dynamics).

Finally, a game form f has the X property (where X is any of the above versions of finite improvement) if $\langle f, \boldsymbol{L} \rangle$ is X for all preference profiles $\boldsymbol{L} \in (\mathcal{L}(C))^n$. Since some convergence properties entail others, we describe these entailments in Figure 5.

[7] It is not obvious how to define direct replies in other voting rules. A good candidate for "direct strategy" in positional scoring rules is to rank the new winner first, and then all other candidates by increasing order of their current score. Note that the direct replies for Plurality and Veto are special cases.

Kukushkin notes that there are no known examples of game forms that are weak-FIP, but not order-free-FIP. In this paper, we settle this question by constructing such examples explicitly (see Section 5.2).

Convergence from the truth. We say that a game G is *FIP from state* $\boldsymbol{a}$ if $\langle G, \boldsymbol{a}, \phi \rangle$ converges for any ϕ. Clearly, a game is FIP iff it is FIP from $\boldsymbol{a}$ for any $\boldsymbol{a} \in A^n$. The definitions for all other notions of finite improvement are analogous.

We are particularly interested in convergence from the truthful state $\boldsymbol{a}^*$. This is because it is reasonable to assume that agents will start by voting truthfully, especially when not sure about others' preferences. Even with complete information, voters may be inclined to start truthfully, as they can always later change their vote.

Heuristic voting

Much work on iterative voting deals with heuristics that apply different limitations on the sets of available moves at each state. The properties of strong, weak and order-free convergence can be defined in the same way as for best and better replies, where the only difference is in the way of defining the set $I_i(\boldsymbol{a})$ (i.e., the set of all steps that are allowed for agent i at state $\boldsymbol{a}$ by the considered heuristics). For example, the *truth-bias* heuristic assumes that if a voter does not have any local improvement step, she reverts to her truthful vote [1, 39]. Some heuristics only allow an agent a single action at a given state (for example, "k-pragmatist" [37], "second chance" or "best upgrade" [33]). In such cases, the only meaningful distinction is between FIP and weak-FIP.

However, as we mentioned earlier, the outcomes of such heuristics are not necessarily Nash equilibria, meaning that voters could further improve by changing their votes, should the restrictions on their actions be removed. Therefore, in this paper we focus on better and best reply dynamics that (if they converge) reach a Nash equilibrium state. We refer the reader to [13] for a more thorough discussion of heuristics.

3. Properties of Strongly Acyclic Rules

Recall that a voting rule (or game-form) is strongly acyclic if any sequence of better-replies by voters must converge to an equilibrium, regardless of voters' preferences. An *ordinal potential* is a function that strictly increases if and only if some agent plays a better reply [6]. A *generalized ordinal potential* is a function that strictly increases with every better reply, but may also increase with other steps. Clearly, a game is FIP if and only if it has a generalized potential (by a topological sort of the better reply graph) [6]. In the context of voting, even most common rules may not admit a generalized ordinal potential or an ordinal potential function; in fact, the latter only exists for dictatorships.

Theorem 3.1 (Kukushkin [7])**.** *A game form f guarantees an ordinal potential (i.e., every derived game has an ordinal potential function) if and only if f is a dictatorship.*

Let us stress that this result *does not* preclude the existence of other game forms with FIP, as an ordinal potential is a stronger condition than a generalized ordinal potential. Indeed, Kukushkin provides a partial characterization of FIP game forms. For example, the *Lexicographic rule* where there is a linear order L over C and the winner is the first candidate according to L that is top-ranked by at least one voter, has this property.

A game form f is called "separable" [7] if there are mappings $g_i : A_i \to C$ for $i \in N$ s.t. for all $\boldsymbol{a} \in \mathcal{A}$, $f(\boldsymbol{a}) \in \{g_1(a_1), g_2(a_2), \dots, g_n(a_n)\}$. That is, the vote of each voter is mapped to a single candidate via some function g_i, and the outcome is always one of the candidates in the range. Examples of separable rules include Plurality, the Lexicographic rule, and dictatorial rules, in all of which g_i are the identity functions.

We now demonstrate another separable game form which (unlike the examples above) is non-standard. In the *direct kingmaker* voting rule [47] all voters $i \in N \setminus \{1\}$ specify a single candidate $a \in C$, whereas voter 1 selects $i \in N \setminus \{1\}$ to be a "dictator of the day." The direct kingmaker is a separable game form, since $f(\boldsymbol{a}) \in \{a_2, \dots, a_n\}$. However, since A_1 is not a coarsening of $\mathcal{L}(C)$, it is non-standard. We prove the following.

Proposition 3.2. *The direct kingmaker is FIP.*

Proof. Denote $d^t = a^t_1$ as the dictator in $\boldsymbol{a}^t$. In every state $\boldsymbol{a}^t$, only agents 1 and d^t may have a better reply. Further, any better reply of d^t is selecting a more preferred candidate, i.e., $a^{t+1}_{d^t} \succ_{d^t} a^t_{d^t}$. Thus any agent except agent 1 may move at most $m-1$ times. Since any cycle implies an unlimited number of steps by at least 2 agents, there can be no cycles. □

It was conjectured by Kukushkin that separability is a necessary condition for a game form to be FIP.

Conjecture 3.3 (Kukushkin [7])**.** *Any FIP game form is separable.*

Some weaker variations of this conjecture have been proved. In particular, the statement is true for game forms with finite *coalitional improvement* property [7], and for FIP game forms with $n = 2$ voters [48] (separable game forms are called "assignable" there). We believe that Conjecture 3.3 holds for other small values of n. Yet, next we show that for sufficiently large n, there are non-separable FIP game forms, thereby refuting the conjecture.

Theorem 3.4. *For any $n \geq 7$, there is a non-separable game form f_n^* with n agents s.t. f_n^* is FIP.*

Proof. The proof outline is as follows. Suppose that each voter only has two possible actions. Clearly any separable game form may have at most $2n$ possible outcomes in C, otherwise it is impossible to find mappings g_i from actions to outcomes as required. However the number of possible action profiles is exponential in n. We will define a voting rule/game form with $2n + 1$ outcomes, such that almost all 2^n profiles yield the same outcome z, with "special" profiles that are few (only $2n$) and far between (several steps are required to get from one special profile to another). Note that for sufficiently high n, such game forms are very easy to construct, but with small n construction has to be more careful. Such game forms must be FIP, and for a very "boring" reason: any cycle must contain a several consecutive steps where the outcome remains z. Such steps cannot be a better-reply for any agent.

We begin by constructing f_n^*. The following construction works for $n \geq 8$. For $n = 7$ we use a somewhat different construction, see appendix. Let $X = \{x^1, \dots, x^n\}$, $Y = \{y^1, \dots, y^n\}$ and $C = X \cup Y \cup \{z\}$. Let $A_i = \{0, 1\}$ for each voter. Every voting function f_n is a function from the n dimensional binary cube $\mathcal{B} = \{0, 1\}^n$ to C.

The challenge is to "pack" our $2n$ special profiles corresponding to $X \cup Y$ in this cube such that they are "far between." Fortunately, coding theory deals with exactly this type of challenges, albeit for very different reasons. What we will do is write our $2n$ special outcomes as (short) binary vectors, then use an error-correcting code to map them back into n-bits (i.e., to profiles), but with the guarantee that they are sufficiently far from one another (a distance of 3 'bits' will suffice).

Denote by $\overline{k}$ the total number of bits required for an optimal single error-correcting code (Hamming code) with k data bits. For example, for $k = 4$ we need $\overline{4} = 7$ bits. In particular, there is a mapping $q : \{0, 1\}^k \rightarrow \{0, 1\}^{\overline{k}}$ such that the Hamming distance between any two words is at least 3 [49]. Formally, for all $w, w' \in \{0, 1\}^k$, it holds that $|\{j : q(w)_j \neq q(w')_j\}| \geq 3$.

For reasons that will become apparent later, we want to have $\overline{k} \leq n-1$, and $n = |X| \leq 2^k$.

Set $r = \lfloor \log_2(n) \rfloor$, and $k = \lceil \log_2(n) \rceil$. Thus for $n \in [8, 15]$,

$$2^r - r - 1 = 2^{\lfloor \log_2(n) \rfloor} - \lfloor \log_2(n) \rfloor - 1 = 2^3 - 3 - 1 = 4 \geq \lceil \log_2(15) \rceil \geq \lceil \log_2(n) \rceil = k;$$

for $n \in [16, 31]$,

$$2^r - r - 1 = 2^{\lfloor \log_2(n) \rfloor} - \lfloor \log_2(n) \rfloor - 1 = 2^4 - 4 - 1 = 11 \geq \lceil \log_2(31) \rceil \geq \lceil \log_2(n) \rceil = k;$$

and for $n \geq 32$,

$$2^r - r - 1 = 2^{\lfloor \log_2(n) \rfloor} - \lfloor \log_2(n) \rfloor - 1 \geq 2^{\log_2(n)-1} - \log_2(n) - 1 = n/2 - \log_2(n) - 1 \geq \lceil \log_2(n) \rceil = k.$$

So for all $n \geq 8$ we get $k \leq 2^r - r - 1$. From coding theory [49], for all $r \geq 2$, if $k \leq 2^r - r - 1$, then $\overline{k} = 2^r - 1$ bits are sufficient to code all k-bit strings. That is, there is a valid code q from $\{0,1\}^k$ to $\{0,1\}^{\overline{k}}$.

It thus holds that:

$$\overline{k} = 2^r - 1 = 2^{\lfloor \log_2(n) \rfloor} - 1 \leq 2^{\log_2(n)} - 1 = n - 1 \quad (1)$$

$$n = 2^{\log_2(n)} \leq 2^{\lceil \log_2(n) \rceil} = 2^k \quad (2)$$

Let $bin(t, k) \in \{0,1\}^k$ be the k-bit binary representation of $t \in \{1, 2, \ldots, n\}$ (e.g., $bin(5, 4) = 0101$). Since $n \leq 2^k$, all of $t \leq n$ are represented with k bits. Using the Hamming code q, we map each outcome $x^t \in X$ to a specific profile $\boldsymbol{a}^t \in \{0,1\}^n$ by setting the first $\overline{k}$ bits of $\boldsymbol{a}^t$ to $q(bin(t,k)) \in \{0,1\}^{\overline{k}}$, the n'th bit to 0. Since $n \geq \overline{k} + 1$ we pad any remaining bits with 0's. Similarly, we map each $y^t \in Y$ to the profile $\boldsymbol{b}^t = (q(bin(t,k)), 0, \ldots, 0, 1) \in \{0,1\}^n$.

This provides us with a mapping from candidates to profiles. We now define f^*_n (which is a reverse mapping from profiles to candidates) as follows:

- For all $t \leq n$, set $f^*_n(\boldsymbol{a}^t) = x^t$;
- For all $t \leq n$, set $f^*_n(\boldsymbol{b}^t) = y^t$;
- For all other $2^n - 2n$ profiles, set $f^*_n(\boldsymbol{a}) = z$.

For any two profiles $\boldsymbol{a}, \boldsymbol{a}'$, let $d(\boldsymbol{a}, \boldsymbol{a}')$ be the number of voters that disagree in $\boldsymbol{a}, \boldsymbol{a}'$ (the Hamming distance on the cube). By the construction above, we have that for all $t \neq t'$, $d(\boldsymbol{a}^t, \boldsymbol{b}^t) = 1$, $d(\boldsymbol{a}^t, \boldsymbol{a}^{t'}) \geq 3$, and $d(\boldsymbol{a}^t, \boldsymbol{b}^{t'}) \geq 3$. Let $C^t = \{x^t, y^t\}$.

Assume towards a contradiction that there is some cycle of better replies in f_n^*. Since each "special outcome" in $X \cup Y$ appears in only one profile, there must be a path containing at least three distinct outcomes. This is possible in one of two ways:

1. There are two distinct outcomes in the cycle from the same set $C^t = \{x^t, y^t\}$. In this case the cycle contains a path from $\boldsymbol{a}^t$ to $\boldsymbol{b}^t$ *and* a different path from $\boldsymbol{b}^t$ to $\boldsymbol{a}^t$. At least one of these paths must be of length 3, and contain at least two consequent profiles whose outcome is z.

2. There are two outcomes in the cycle from distinct sets $C^t, C^{t'}$. In this case the cycle must contain a path from $\{\boldsymbol{a}^t, \boldsymbol{b}^t\}$ to $\{\boldsymbol{a}^{t'}, \boldsymbol{b}^{t'}\}$ *or* vice versa. In either case this path must be of length 3, and contain at least two consequent profiles whose outcome is z.

A path that contains two consequent profiles with outcome z cannot be a better reply path, since a better reply must in particular change the outcome. Hence we get a contradiction, and f_n^* is FIP.

Finally, note that $\boldsymbol{a}^t, \boldsymbol{b}^t$ are all distinct profiles, and f_n^* has $2n+1 > \sum_{i \leq n} |A_i|$ possible outcomes. In contrast, for any *separable* rule f_n the size of the range of f_n is at most $\sum_{i \leq n} |A_i|$, since $f_n(\boldsymbol{a}) = g_i(b)$ for some voter $i \in N$ and action $b \in A_i$. This means that f_n^* is non-separable. $\square$

For $n \leq 6$ such a construction is impossible, since any function with $2n + 1$ distinct outcomes will necessarily contain a cycle. It may still be possible to construct non-separable rules with fewer outcomes, but we believe that Conjecture 3.3 holds for all $n \leq 6$.

For most common voting rules, separable or not, it is easy to find examples where some cycles occur. Thus the importance of Conjecture 3.3 (or its refutation) is mainly to game theory rather than social choice. In the remainder of this paper we focus on the weaker notions of convergence as discussed in Section 1, which are more relevant to social choice.

4. Plurality is Order-Free Acyclic

In this section, we analyze convergence of order-free dynamics in Plurality—that is, there is no scheduler prescribing in which order the agents should take their improvement steps. We look at sequences of better, best and direct best replies, initiated at an arbitrary or a specific (namely, truthful) profile. We consider variations of the game where the agents may have different or equal weights, and the ties are broken lexicographically or arbitrarily.

4.1. Lexicographic Tie-Breaking

We start with the case of lexicographic tie-breaking. Given some score vector $\boldsymbol{s}$, we denote by $\ddot{s}(c) \in \mathbb{R}$ the score of $c \in C$ that includes the lexicographic tie-breaking component. One way to formally define it is by setting $\ddot{s}(c) = s(c) + \frac{1}{m+1}(m - L(c))$, where $L(c)$ is the lexicographic index of candidate c. However the only important property of $\ddot{\boldsymbol{s}}$ is that $\ddot{s}(c) > \ddot{s}(c')$ if either $s(c) > s(c')$ or the score is equal and c has a higher priority (lower index) than c'.

Thus, for Plurality with lexicographic tie-breaking, a given weight vector $\boldsymbol{w}$ and a given initial score vector $\hat{\boldsymbol{s}}$, we denote the outcome by

$$f^{PL}_{\hat{\boldsymbol{s}},\boldsymbol{w}}(\boldsymbol{a}) = \text{argmax}_{c \in C}\, \ddot{s}_{\hat{\boldsymbol{s}},\boldsymbol{w},\boldsymbol{a}}(c).$$

As with $\boldsymbol{s}$, we omit the scripts $\boldsymbol{w}, \hat{\boldsymbol{s}}$ and *PL* when they are clear from the context.

Lemma 4.1. *Consider a game $\left\langle f^{PL}_{\boldsymbol{w},\hat{\boldsymbol{s}}}, \boldsymbol{L}\right\rangle$. If there exists a better reply for a given agent i at state $\boldsymbol{a}^{t-1}$, then i has a direct best reply at state $\boldsymbol{a}^{t-1}$.*

The proof is trivial under lexicographic tie-breaking, by letting i vote for her most preferred candidate among all better replies. In this case the direct best reply is also unique.

One implication of the lemma is that it is justified and natural to restrict our discussion to direct replies and focus on FDRP, as w.l.o.g. a voter always has a direct reply that is at least as good as any other reply.

We next introduce some additional notation that we can use to classify all possible improvement steps under the lexicographic Plurality rule into three types. This classification will be useful in the proofs and examples throughout the section.

Types of improvement steps in Plurality

Recall that along a given path, $\boldsymbol{a}^t \in A^n = C^n$ denotes the voting profile at time t. We next denote by $\boldsymbol{s}^t = \boldsymbol{s}_{\boldsymbol{a}^t}$ the score vector at time t; by $cw^t = f^{PL}(\boldsymbol{a}^t)$ the candidate that wins at time t; and by $sw^t = \ddot{s}^t(cw^t)$ the highest score at time t. Note that the score of a losing candidate is always strictly lower than sw^t, even if she has the same number of votes as the winner and loses only by tie-breaking.

Suppose that agent i has an improvement step (i.e., better reply) $a_i^{t-1} \xrightarrow{i} a_i^t$ at time t. We classify all possible steps into the following types (an example of such a step appears in parentheses):

Type 1. from $a_i^{t-1} \neq cw^{t-1}$ to $a_i^t = cw^t$; (step 1 in Ex. 4.3)

Type 2. from $a_i^{t-1} = cw^{t-1}$ to $a_i^t = cw^t$; (step 1 in Ex. 4.4),

Type 3. from $a_i^{t-1} = cw^{t-1}$ to $a_i^t \neq cw^t$; (step 2 in Ex. 4.3)

Note that steps of type 1 and 2 are direct, whereas type 3 steps are indirect.

We first show our main result: when voters all have unit weight, any sequence of direct replies converges to equilibrium. Then, we show that this no longer holds if voters may have different weights.

Unweighted voters

Theorem 4.2. *$f^{PL}_{\hat{\mathbf{s}}}$ is FDRP. Moreover, any path of direct replies will converge after at most m^2n^2 steps. In particular, Plurality is order-free acyclic.*

This extends a weaker version of the theorem that appeared in the preliminary version of this paper [1], which only showed FDBRP. The bound on the number of direct best reply steps was improved to $O(mn)$ in [12, Theorem 5.4].

Proof. By our restriction to direct replies, there can only be moves of types 1 and 2. We first consider moves of type 1, and inductively prove two invariants that yield a bound on the total number of such moves. Next, we bound the number of moves of type 2 by a given voter between any of his moves of type 1, which completes the proof.

Consider time $t-1$ and denote the score of the current winner (including tie-breaking) by $\bar{s} = sw^{t-1}$. Suppose that a move $a \xrightarrow{i} b$ of type 1 occurs at time t: that is, $a \neq cw^{t-1}$ and $b = cw^t$. We then have (see Figure 6):

$$\ddot{s}^t(b) = sw^t \geq sw^{t-1} = \bar{s} > \ddot{s}^{t-1}(a) = \ddot{s}^t(a) + 1. \tag{3}$$

The strict inequality is since cw^{t-1} beats a. We claim that at any later time $t' \geq t$ the following two invariants hold:

I. Either there is a candidate $c \neq a$ whose score is at least $\bar{s}+1$, or there are at least two candidates $c, c' \neq a$ whose score is at least $\bar{s}$. In particular it holds in either case that $sw^{t'} \geq \bar{s}$.

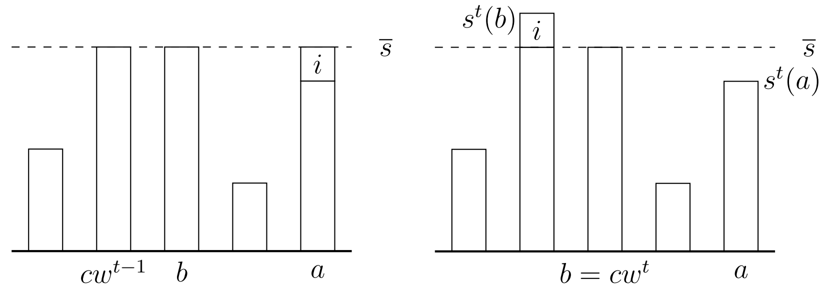


Figure 6: An illustration of a type 1 move. Tie-breaking is in favor of the left-most candidate.

II. The score of a does not increase: $\ddot{s}^{t'}(a) \leq \ddot{s}^t(a)$.

Note that this, coupled with Eq. (3), implies that candidate a will never win again, as its score will stay strictly below $\bar{s}$, and there will always be a candidate with a score of at least $\bar{s}$.

We now prove both invariants by induction on time t'. In the base case $t' = t$, (I) holds since both cw^{t-1} and b have a score of at least $\bar{s}$, and (II) holds trivially.

Assume by induction that both invariants hold until time $t' - 1$, and consider step t' by voter j. Due to (I), we either have at least two candidates whose score is at least $\bar{s}$, or a candidate with a score of at least $\bar{s} + 1$. Due to (II) and Eq. (3) we have that $\ddot{s}^{t'-1}(a) \leq \ddot{s}^t(a) < \bar{s} - 1$.

Let $d \xrightarrow{j} d'$ be the step at time t' by voter j (that is, $d = a_j^{t'-1}, d' = a_j^{t'}$). We first argue that $d' \neq a$: by adding the vote of j to a its score will still be strictly less than $\bar{s}$, whereas by removing a vote from any other candidate d, we still have at least one candidate c with score at least $\bar{s}$. Thus a cannot be a direct reply for any voter j, and (II) still holds after step t'.

It remains to show that (I) holds. If d is *not* one of the candidates in (I) with the score of at least $\bar{s}$ at time $t' - 1$, then their score does not decrease after step t', and we are done. Otherwise, we divide into the following cases:

1. At $t' - 1$, $d = c$ is the (only) candidate with a score of at least $\bar{s} + 1$.

2. At $t' - 1$, candidates c, c' have scores of at least $\bar{s}$, and d is one of them (w.l.o.g. $d = c$).

In the first case, $\ddot{s}^{t'}(d) = \ddot{s}^{t-1}(d) - 1 \geq \bar{s} + 1 - 1 = \bar{s}$, whereas $\ddot{s}^{t'}(d') > \ddot{s}^{t'}(d) \geq \bar{s}$. Thus, both d, d' have scores of at least $\bar{s}$ at time t', as required. In the second case, since only $c = d$ can lose votes, then if $d' \neq c'$,

$$\ddot{s}^{t'}(d') = sw^{t'} \geq \ddot{s}^{t'}(c') = \ddot{s}^{t'-1}(c') \geq \bar{s},$$

and thus both c', d' have scores of at least $\bar{s}$ at time t, as required. If $d' = c'$, then

$$\ddot{s}^{t'}(d') = \ddot{s}^{t'-1}(d') + 1 = \ddot{s}^{t'-1}(c') + 1 \geq \bar{s} + 1,$$

that is, d' has a score of at least $\bar{s} + 1$, as required.

Next, we demonstrate that invariants (I) and (II) supply us with a polynomial bound on the rate of convergence. Indeed, as we mentioned before, at every step of type 1, at least one candidate is ruled out permanently, and there are at most n times that a vote can be withdrawn from a given candidate. Thus in total there can be at most mn steps of type 1. Also note that, since a type 2 move by a given voter i implies that he prefers a_i^t to a_i^{t-1}, each voter can make at most $m-1$ type 2 moves before making a move of type 1, and the total number of steps between any two type 1 steps is $(m-1)n$. Hence, there are in total at most m^2n^2 steps until convergence. □

Next, we show that the restriction to direct replies is necessary to guarantee convergence, whereas a restriction to best replies is insufficient.

Proposition 4.3. *f^{PL} is not FBRP.*

Remark 4.1. *In the example below and in other examples throughout the paper we use an initial score vector $\hat{\mathbf{s}}$. However, this is w.l.o.g. since we could replace $\hat{\mathbf{s}}$ with additional voters that do not participate in the cycle. Initial scores are only useful to construct examples that are simpler and/or with fewer strategic agents. This holds for all negative results in the paper.*[8] *For positive results, we have to show convergence for every initial score vector $\hat{\mathbf{s}}$. Clearly, any such positive result also holds for the case of $\hat{\mathbf{s}} = \mathbf{0}$.*

Example 4.3. Let $C = \{a, b, c\}$ and $N = \{1, 2\}$. We have a single fixed voter voting for a, thus $\hat{\mathbf{s}} = (1, 0, 0)$. The preference profile is defined as $a \succ_1 b \succ_1 c$, $c \succ_2 b \succ_2 a$. The following cycle consists of individual improvement steps of the strategic voters (the vector denotes the votes (a_1, a_2) at time t, and the winner appears in curly brackets):

$$(b, c)\{a\} \xrightarrow{2} (b, b)\{b\} \xrightarrow{1} (c, b)\{a\} \xrightarrow{2} (c, c)\{c\} \xrightarrow{1} (b, c).$$

[8] Note that the remark no longer holds if $\hat{\mathbf{s}}$ is used to construct a counter-example for weak-FIP. However we do not use such examples in this paper.

Note that all steps are best replies, but the steps of agent 1 are indirect. $\diamondsuit$

Proposition 4.3 in particular implies that f^{PL} is not FIP. As we show below, cycles of better replies may occur even from the truthful state.

Proposition 4.4. *f^{PL} is not FIP even from the truthful state.*

Example 4.4. Let $C = \{a, b, c, d\}$. Candidates a, b, and c have 2 fixed voters each, thus $\hat{\boldsymbol{s}} = (2, 2, 2, 0)$. We use 3 strategic agents $N = \{1, 2, 3\}$ with the following preferences: $d \succ_1 a \succ_1 b \succ_1 c$, $c \succ_2 b \succ_2 a \succ_2 d$ and $d \succ_3 a \succ_3 b \succ_3 c$. Starting from the truthful state (d, c, d) the agents can make the following two improvement steps, which are direct replies (showing only the outcome $\boldsymbol{s}$ and the winner): $(2, 2, 3, 2)\{c\} \xrightarrow{1} (2, 3, 3, 1)\{b\} \xrightarrow{3} (3, 3, 3, 0)\{a\}$, after which agents 1 and 2 repeat the cycle shown in Example 4.3. $\diamondsuit$

There is still the question of convergence from the truthful state, when voters are restricted to best replies that are not necessarily direct. This was shown to be positive as well.

Proposition 4.5 (Reijngoud [50])**.** *f^{PL} is FBRP from the truth. Moreover, any path of best replies will converge after at most mn steps.*

The reason is that under best-response dynamics from the truthful state, only type 1 steps can occur. This fact was also shown independently by Branzei et al. [32].

Weighted voters

Next, we show that if the voters may have non-identical weights, then convergence to equilibrium is not guaranteed even if they start from the truthful state and use direct best replies.

Proposition 4.6. *$f^{PL}_{\boldsymbol{w}}$ is not order-free-FDRP for some $\boldsymbol{w}$, even from the truthful state.*

Example 4.6. Let the initial fixed score of candidates $C = \{a, b, c, d\}$ be $\hat{\boldsymbol{s}} = (0, 1, 2, 3)$. The weight of each voter $i \in N = \{1, 2, 3\}$ is i. The preference profile is as follows: $c \succ_1 d \succ_1 b \succ_1 a$, $b \succ_2 c \succ_2 a \succ_2 d$, and $a \succ_3 b \succ_3 c \succ_3 d$. The initial truthful profile is thus $\boldsymbol{a}^0 = (c, b, a)$, which results in the score vector

$\boldsymbol{s}^0 = (3,3,3,3)$ where a is the winner.

$$
\begin{array}{lccccc}
\text{votes:} & (c,b,a) & \xrightarrow{1} & (d,b,a) & \xrightarrow{2} & (d,c,a) \\
\text{scores:} & (3,3,3,3)\{a\} & & (3,3,2,4)\{d\} & & (3,1,4,4)\{c\} \\
 & \uparrow_{3} & & & & \downarrow_{3} \\
 & (c,b,b) & \xleftarrow{2} & (c,c,b) & \xleftarrow{1} & (d,c,b) \\
 & (0,6,3,3)\{b\} & & (0,4,5,3)\{c\} & & (0,4,4,4)\{b\}
\end{array}
$$

Our example shows a cycle of direct replies. Note that at every step there is only one direct reply available to the agent, thus it is not possible to eliminate the cycle by further restricting the action scheduler. $\diamondsuit$

However, if there are *only two* strategic weighted voters (and possibly other fixed voters), either restriction to direct replies or to a truthful initial state is sufficient to guarantee convergence.

Lemma 4.7. *Suppose that there are $n = 2$ weighted voters, and that at some state t' both voters vote for distinct candidates, and one of these candidates is the winner. Then the score of the winner is strictly higher at any later state $t > t'$.*

Proof. W.l.o.g, voter 1 plays in all the odd steps, and voter 2 in all the even steps (if a voter plays several consecutive steps we consider them as a single step). Denote the current votes by $a_1^{t'} = x$ and $a_2^{t'} = y \neq x$. In order to change the outcome, voter 1 must vote in step $t' + 1$ for some $z \notin \{x, y\}$ s.t. $cw^{t'+1} = z$. Since $x \neq y$, we have $sw^{t'+1} = \ddot{s}^{t'+1}(z) > \ddot{s}^{t'+1}(y) = \ddot{s}^{t'}(y) = sw^t$. We can now apply the same argument inductively, as in state $t' + 1$ voters vote for distinct candidates, one of whom is the winner. $\square$

Theorem 4.8. *$f^{PL}_{\hat{\boldsymbol{s}},\boldsymbol{w}}$ is FDRP for $n = 2$.*

Proof. Let state $t' > 2$ be any state where voters vote for two distinct candidates. W.l.o.g. t' is even, meaning voter 2 just moved to y from some other candidate. By our restriction to direct replies, $cw^{t'} = y$. By applying Lemma 4.7, the score of the winner must increase in every step until convergence. $\square$

Theorem 4.9. *$f^{PL}_{\hat{\boldsymbol{s}},\boldsymbol{w}}$ is FIP from the truth for $n = 2$.*

Proof. The first step must be of type 1 since initial votes are truthful. Thus at state $\boldsymbol{a}^1$ at least one voter (w.l.o.g. voter 2) is voting for the winner $a_2^1 = cw^1 = x$. If $a_1^1 = x$ as well then we are done since neither voter will want to move. Otherwise, the voters are voting for distinct candidates and thus by Lemma 4.7, the score of the winner must increase in every step until convergence. $\square$

Thus, in either case convergence is guaranteed after at most $2m$ steps.

It remains an open question whether there is any restriction on better replies that guarantees order-free acyclicity in weighted games, i.e., if $f^{PL}_{\boldsymbol{w}}$ is order-free-FIP for any $n > 2$ and weights $\boldsymbol{w}$. However, Prop. 4.6 shows that if such a restricted order-free dynamic exists, it must make use of indirect replies, which is rather unnatural. We thus conjecture that such order-free dynamics do not exist.

We next consider how the tie-breaking method affects the convergence properties of the (unweighted) Plurality rule.

4.2. Arbitrary tie-breaking

Lev and Rosenschein [46, 14] showed that for any positional scoring rule (including Plurality), we can assign some deterministic tie-breaking rule, so that the resulting voting rule may contain cycles. For any positional scoring rule f_α with score vector α, denote by f^{LR}_α the same rule with the Lev-Rosenschein tie-breaking.

Proposition 4.10 (Theorem 1 in [46])**.** *f^{LR}_α is not FBRP for any α, even for $n = 2$, and even from the truth. In particular, Plurality with the Lev-Rosenschein tie-breaking (f^{PLR}) is not FBRP.*

In fact, a slight modification of their example yields the following:

Proposition 4.11. *f^{PLR} is not order-free-FIP, even for $n = 2$, and even from the truth.*

Example 4.11. The original example used in [46] for Plurality has four candidates $\{a, b, c, d\}$ and two voters with preferences $a \succ_1 b \succ_1 c \succ_1 d$ and $c \succ_2 d \succ_2 b \succ_2 a$. The (non linear) tie-breaking rule is defined such that: a beats d; b beats a and c; c beats a and d; and d beats b. We modify the preferences by switching a and b in voter 2's preferences, so that $c \succ_2 d \succ_2 a \succ_2 b$.

We get the following cycle from the truthful state $\boldsymbol{a}^0 = (a, c)$:

$$\begin{array}{ccc} (a,c)\{c\} & \xrightarrow{1} & (b,c)\{b\} \\ \uparrow_2 & & \downarrow_2 \\ (a,d)\{a\} & \xleftarrow{1} & (b,d)\{d\} \end{array}$$

Since each voter has only one available better reply in every step, no restriction on the action scheduler would break the cycle. $\Diamond$

4.3. Randomized tie-breaking

Let $F^P_{\hat{\boldsymbol{s}},\boldsymbol{w}}$ denote the game form that maps any state $\boldsymbol{a} \in A^n$ to the set $\text{argmax}_{c \in C}\, s_{\hat{\boldsymbol{s}},\boldsymbol{w},\boldsymbol{a}}(c)$ (all candidates with maximal Plurality score). Let $W^t = F^P(\boldsymbol{a}^t) \subseteq C$ denote the set of winners at time t. We thus define a direct reply $a_i^{t-1} \xrightarrow{i} a_i^t$ as one where $a_i^t \in W^t$.

It is easy to see that a preference order L_i does *not* induce a complete order over outcomes of F^P. For instance, the order $a \succ_i b \succ_i c$ does not determine if i prefers $\{b\}$ over $\{a, c\}$. However, we can naturally extend L_i to a *partial preference order* over subsets, using the following axioms:

Axiom	Name and reference	Definition
K1	Kelly [20]	$(\forall a \in X, b \in Y, a \succ_i b) \;\Rightarrow\; X \succ_i Y$
K2	Kelly [20]	$(\forall a \in X, b \in Y, a \succeq_i b) \;\Rightarrow\; X \succeq_i Y$
G	Gärdenfors [19]	$(\forall b \in X, a \succ_i b) \;\Rightarrow\; \{a\} \succ_i (\{a\} \cup X) \succ_i X$
R	Responsiveness [51]	$a \succ_i b$ if and only if $\forall X \subseteq C \setminus \{a, b\},\; (\{a\} \cup X) \succ_i (\{b\} \cup X)$

These axioms reflect various beliefs a rational voter may have on the tie-breaking procedure: the K axiom reflects no assumptions whatsoever; the K+G axioms are consistent with tie-breaking according to a fixed and unknown order [52]; and K+G+R axioms are consistent with random tie-breaking with equal probabilities (see Lemma 4.15 and Prop. 4.23). In this section, we assume all axioms hold, however our results do not depend on the interpretation of this preference extension as a result of random tie-breaking, and we do not specify the voters' preferences in cases not covered by the above axioms. Under strict preferences, it also holds that G entails K [53]. We can also define "weak" variants G2 and R2 for axioms G and R, by replacing all strict relations with weak ones; however, as long as we restrict attention to strict preferences over the alternatives, the weak variants are not required.

We define the *Plurality with random tie-breaking* rule $f^{PR}_{\hat{\boldsymbol{s}},\boldsymbol{w}}$ by selecting the winner uniformly at random from $F^P_{\hat{\boldsymbol{s}},\boldsymbol{w}}$. We emphasize that the ties are resolved only *after* the voting process ends, and thus voters' decisions cannot take the realized outcome into consideration (*ex-ante* better reply).[9] In order to be able to identify all better replies in a game $\left\langle f^{PR}_{\hat{\boldsymbol{s}},\boldsymbol{w}}, \boldsymbol{L} \right\rangle$, we need to extend each L_i using

[9] With ex-post better replies, any convergence proof for lexicographic tie-breaking entails convergence of random tie-breaking, since eventually there will be a long enough sequence of steps where ties are broken (by chance) lexicographically.

Axioms K+G+R. The extended relation L_i is just a partial order over subsets of C, but nonetheless this is sufficient to prove our main result in this section.

For the following lemma we only need Axiom K, i.e., it does not depend on the tie-breaking assumptions.

Lemma 4.12. *If there exists a better reply in $f^{PR}_{\hat{\mathbf{s}},\boldsymbol{w}}$ for agent i at state $\boldsymbol{a}^{t-1}$, then i has a direct best reply.*

Proof. Suppose there is a better reply $a_i^{t-1} \xrightarrow{i} b$ at time $t-1$. As some best reply always exists, denote by b' an arbitrary best reply. Let $W = F^P_{\hat{\mathbf{s}},\boldsymbol{w}}(\boldsymbol{a}^{t-1}_{-i}, b')$, and let a' be the most preferred candidate of i in W. Then we argue that $a_i^{t-1} \xrightarrow{i} a'$ is a direct best reply of i. Since a' is a direct reply by definition, it is left to show that a' is a best reply (for the lexicographic case this follows immediately from $W = \{a'\}$ and $f^{PL}(\boldsymbol{a}^{t-1}_{-i}, a') = W = \{a'\}$).

If b' is a direct reply then $b' = a'$ and we are done. Thus assume that b' is not a direct reply from a_i^{t-1}. Then $b' \notin W$. By voting for $a' \in W$, we get that $F^P_{\hat{\mathbf{s}},\boldsymbol{w}}(\boldsymbol{a}^{t-1}_{-i}, a') = \{a'\}$, i.e., a' remains the unique winner. If $|W| = 1$ then we are done as in the lexicographic case. Otherwise we apply Axiom K2 with $X = \{a'\}, Y = W$, and get that $a' \succeq_i W$. That is,

$$F^P_{\hat{\mathbf{s}},\boldsymbol{w}}(\boldsymbol{a}^{t-1}_{-i}, a') = \{a'\} \succeq_i W = F^P_{\hat{\mathbf{s}},\boldsymbol{w}}(\boldsymbol{a}^{t-1}_{-i}, b'),$$

which means that a' is also a best reply in $f^{PR}_{\hat{\mathbf{s}},\boldsymbol{w}}$. □

Weighted voters

In contrast with the lexicographic case, the weighted randomized case does not always converge to equilibrium, even with (only) two strategic agents. Moreover, a pure strategy Nash equilibrium may not exist at all [1]. We therefore restrict attention in the rest of this section to unweighted voters.

Unweighted voters

In this case, we show that while better replies may not converge, best replies converge from the truthful profile.

Proposition 4.13. *f^{PR} is not FIP.*

Example 4.13. $C = \{a, b, c\}$ with initial score $\hat{\boldsymbol{s}} = (0, 1, 0)$. The initial state is $\boldsymbol{a}_0 = (a, a, b)$—that is, $\boldsymbol{s}(\boldsymbol{a}_0) = (2, 2, 0)$ and the outcome is the winner set

$F^P_{\hat{\mathbf{s}}}(\boldsymbol{a}_0) = \{a, b\}$. The preferences are $a \succ_1 c \succ_1 b$, $b \succ_2 a \succ_2 c$ and $c \succ_3 b \succ_3 a$. We get the following cyclic sequence:

$$\begin{array}{ccccc}
(2,2,0)\{a,b\} & \xrightarrow{2} & (1,2,1)\{b\} & \xrightarrow{1} & (0,2,2)\{b,c\} \\
\uparrow{\scriptstyle 1} & & & & \downarrow{\scriptstyle 3} \\
(1,2,1)\{b\} & \xleftarrow{3} & (2,1,1)\{a\} & \xleftarrow{2} & (1,1,2)\{c\}
\end{array}$$

We emphasize that each step is justified as a better reply by either Axiom K or Axiom G. For example, in the step of agent 2 in the top row, agent 2 prefers $b \succ_2 a$, and thus $b \succ_2 \{a, b\}$ by Axiom G. This will be used later in Section 4.4. $\diamondsuit$

Theorem 4.14. *$f^{PR}_{\hat{\mathbf{s}}}$ is FBRP from the truth.*

Proof. We denote the sets of winners and runner-ups at time t as $W^t = F^P_{\hat{\mathbf{s}}}(\boldsymbol{a}^t)$ (i.e., we omit the scripts); and $R^t = \{c : s^t(c) = sw^t - 1\}$. We will show by induction that at any step $\boldsymbol{a}^{t-1} \xrightarrow{i} \boldsymbol{a}^t$:

1. $W^t \cup R^t \subseteq W^{t-1} \cup R^{t-1}$ (i.e., the set of potential winners is never expanding).

2. $a^t_i \in W^t$ (i.e., step t is a direct reply).

3. a^t_i is the most preferred candidate for i in $W^t \cup R^t$.

4. $a^{t-1}_i \succ_i a^t_i$ (in the terminology of [34], this is a *compromise step*).

5. Either $sw^t > sw^{t-1}$ (the score of the winner strictly increases), or $sw^t = sw^{t-1}$ and $|W^t| > |W^{t-1}|$.

Since each voter can make at most $m - 1$ compromise steps, convergence is guaranteed within nm steps.

Assume that for some $t \geq 1$, all of the above holds for any $t' < t$ (so we prove the base case together with the other cases). Since $\boldsymbol{a}^0$ is truthful, the first step of any voter is always a compromise move. Also, if i had already moved at some previous time $t' < t$, then $a^{t'}_i$ is most preferred in $W^{t'} \cup R^{t'}$.

By induction, $a = a^{t-1}_i$ is the most preferred candidate in some C' that contains $W^{t-1} \cup R^{t-1}$ ($C' = C$ in i's first step, and $C' = W^{t'} \cup R^{t'}$ at any other step t'). Let x and y be i's most preferred candidates in W^{t-1} and in R^{t-1}, respectively, and denote the best reply by $a' = a^t_i$. Each of a or a' may belong to W^{t-1}, to R^{t-1}, or to neither set. This means there are $3 \cdot 3 = 9$ cases to check. Fortunately, we can show that some of these cases immediately lead to a contradiction, and in the other cases all invariants 1–5 will hold after step t.

Consider first the case $a \in W^{t-1}$. Since a is most preferred in C', it is strictly more preferred than any other candidate in W^{t-1} or in R^{t-1} (in particular, $a = x$).

- If $a' \in W^{t-1}$, then we show that all invariants hold: (1) $W^t \cup R^t = W^{t-1} \setminus \{a\} \subseteq W^{t-1} \cup R^{t-1}$; (2) holds since a' becomes the unique winner; (3) if there was another candidate $z \in W^t \cup R^t = W^{t-1} \setminus \{a\}$ such that $z \succ_i a'$, then z would be i's best reply; (4) follows since $a' \in C'$ and a is the most preferred in C'; and (5) follows since the score of the winner a' increases by 1.

- If $a' \in R^{t-1}$ we get $W^t = (W^{t-1} \setminus \{a\}) \cup \{a'\} \prec_i W^{t-1}$ by Axiom R. Thus $a \xrightarrow{i} a'$ is not an improvement step.

- If $a' \notin W^{t-1} \cup R^{t-1}$ then we further split into two cases:

  - If $|W^{t-1}| > 1$, then $W^t = W^{t-1} \setminus \{a\} = W^{t-1} \setminus \{x\}$. Then since x is the most preferred in W^{t-1} we have by Axiom G that $W^{t-1} \succ_i W^t$, meaning that this cannot be an improvement step.
  - If $|W^{t-1}| = 1$, i.e., $W^{t-1} = \{a\}$, then let $t' < t$ be the previous step by agent i. We argue that step t' must be an earlier violation of some of the five invariants, thus contradicting our inductive hypothesis:
    * $W^t = R^{t-1} \cup \{a'\}$. Otherwise (i.e., if a' does not become a winner) we have $\{a\} = W^{t-1} \succ_i W^t = R^{t-1}$ by Axiom K.
    * The score of a' at time $t-1$ is exactly $s = s^{t-1}(a) - 2$, otherwise a' cannot become a winner. Denote by RR^{t-1} all candidates with score s at time $t-1$ (second runner-ups), then a' is i's most preferred in RR^{t-1}.
    * a' must be more preferred than all of $R^{t-1} \cup \{a\}$. Otherwise, let z be the most preferred candidate in $R^{t-1} \cup \{a\}$, and by Axiom G, $\{z\} \succ_i W^t$. Since i can make z the unique winner by voting for it, a' cannot be a best reply.
    * Since step $t-1$ was a direct reply (by invariant (2)), it was by some agent j moving from some candidate $b = a_j^{t-2} \notin W^{t-2}$ to a, meaning that the winning score at time $t-2$ is $s+1$. Thus

$W^{t-2} = R^{t-1} \cup \{a\}$, and $R^{t-2} = RR^{t-1} \setminus \{b\}$. In particular $a \in W^{t-2}$ and $a' \in R^{t-2}$, and both are in the union.

* By invariant (1) again, and since $t' \leq t-1$, $a' \in R^{t-2} \cup W^{t-2}$.
* We get that at step t', voter i votes for a even though there is a more preferred candidate $a' \in R^{t'} \cup W^{t'}$. This is a violation of invariant (3).

This shows that a step that strictly decreases the score of the winner at time t implies an earlier violation, which is a contradiction to our induction assumption.

Next, suppose $a \notin W^{t-1}$. We further split to subcases based on a'.

- If $a' \in W^{t-1}$ then $W^t = F^P_{\hat{\boldsymbol{s}}}(\boldsymbol{a}_{-i}, a') = \{a'\}$. Then $a' = x$, as otherwise $F^P_{\hat{\boldsymbol{s}}}(\boldsymbol{a}_{-i}, x) = \{x\} \succ_i \{a'\}$, and i is strictly better off by voting for x. This entails $W^t = \{x\}, R^t = W^{t-1} \setminus \{x\}$ so all invariants 1–5 hold: (1) since $W^{t-1} = W^t \cup R^t$; (2) since $W^t = \{a'\}$; (3) follows from (1) since $a' = x$ is the most preferred in W^{t-1}; (4) follows from (1) since $a = a_i^{t-1}$ is the most preferred in C', and $a' \in C'$; (5) follows since the score of the unique winner a' increases by 1.

- If $a' \in R^{t-1}$ then $W^t = F^P_{\hat{\boldsymbol{s}}}(\boldsymbol{a}_{-i}, a') = \{a'\} \cup W^{t-1}$. Then $a' = y$, as otherwise $F(\boldsymbol{a}_{-i}, y) = \{y\} \cup W^{t-1} \succ_i \{a'\} \cup W^{t-1}$ by Axiom R, which means i is strictly better off by voting for y. This entails $W^t = \{y\} \cup W^{t-1}$, $R^t = R^{t-1} \setminus \{y\}$. We also get that $a' = y \succ_i x$ or else x would have been a strictly better reply since $F^P_{\hat{\boldsymbol{s}}}(\boldsymbol{a}_-i, x) = \{x\} \succ_i \{y\} \cup W^{t-1}$ by Axiom K. Thus all invariants 1–5 hold: (1) $W^t = W^{t-1} \cup \{y\} \subseteq W^{t-1} \cup R^{t-1}$ and $R^t = R^{t-1} \setminus \{y\}$; (2) since $a' \in W^t$; (3) follows from (1) since $a' = a_i^t = y$ is most preferred in R^{t-1} and strictly preferred to x; (4) follows from (1) as in the previous case; and (5) follows since $W^t = W^{t-1} \cup \{a'\}$.

- If $a' \notin W^{t-1} \cup R^{t-1}$, then $W^t = F(\boldsymbol{a}_{-i}, a') = W^{t-1}$. The outcome does not change so this cannot be an improvement step for i.

□

The proofs and examples above make use only of the axioms, without specifying an explicit cardinal scale. To show that the result is tight, we next consider games with specific cardinal utility scales.

Cardinal utilities

A (cardinal) utility function is a mapping of candidates to real numbers $u : C \to \mathbb{R}$, where $u_i(c) \in \mathbb{R}$ denotes the utility of candidate c to agent i. We say that u_i is *consistent* with a preference relation L_i if $u_i(c) > u_i(c') \Leftrightarrow c \succ_i c'$. The definition of cardinal utility naturally extends to multiple winners by setting $u_i(W) = \frac{1}{|W|}\sum_{c \in W} u_i(c)$ for any subset $W \subseteq C$. As explained above, one interpretation of this is that the winner is selected uniformly at random from the set W.

Lemma 4.15. *Consider any cardinal utility function u and the partial preference order L it induces on subsets by random tie-breaking. Then, L satisfies Axioms K+G+R.*

The proof is rather straightforward, and is deferred to the appendix. Note that a cardinal scale u_i may induce additional preference relations that are not implied by (nor contradict) the axioms.

Proposition 4.16. *f^{PR} is not FIP from the truth, even for $n = 2$.*

Example 4.16. We use 5 candidates with initial score $\hat{\boldsymbol{s}} = (1, 1, 2, 0, 0)$, and 2 agents with preferences that imply $\{b, c\} \succ_1 c$, $\{a, c\} \succ_1 \{a, b, c\}$ (with $d =$ top(L_1)), and $\{a, b, c\} \succ_2 \{b, c\}$, $c \succ_2 \{a, c\}$ (with $e =$ top(L_2)). The following cycle occurs: $(1, 1, 2, 1, 1)\{c\} \xrightarrow{1} (1, 2, 2, 0, 1)\{b, c\} \xrightarrow{2} (2, 2, 2, 0, 0)\{a, b, c\} \xrightarrow{1} (2, 1, 2, 1, 0)\{a, c\} \xrightarrow{2} (1, 1, 2, 1, 1)\{c\}$.

To see why there must exist valid preferences as above, note that L_1 is consistent with (for example) $u_1 = (5, 3, 2, 8, 0)$ and likewise L_2 is consistent with $u_2 = (4, 2, 5, 0, 8)$. Then by Lemma 4.15 the preferences hold all axioms. $\diamondsuit$

Finally, in contrast to the lexicographic case, convergence is no longer guaranteed if agents start from an arbitrary profile of votes, or are allowed to use direct replies that are not best replies. In the next propositions we define voters' preferences directly as cardinal utilities. The following example shows that in the randomized tie-breaking setting even direct best reply dynamics may have cycles, albeit for specific utility scales.

Proposition 4.17. *f^{PR} is not order-free-FIP.*

Example 4.17. There are 4 candidates $\{a, b, c, x\}$ and 3 agents with utilities $u_1 = (7, 3, 0, 4)$, $u_2 = (0, 7, 3, 4)$ and $u_3 = (3, 0, 7, 4)$. In particular, the following preference relations hold: $a \succ_1 \{a, b\} \succ_1 x \succ_1 \{a, c\}$; $b \succ_2 \{b, c\} \succ_2 x \succ_2 \{a, b\}$; and $c \succ_3 \{a, c\} \succ_3 x \succ_3 \{b, c\}$.

Consider the initial state $\boldsymbol{a}_0 = (a, b, x)$ with $\boldsymbol{s}(\boldsymbol{a}_0) = (1, 1, 0, 1)$ and the outcome $\{a, b, x\}$. We have the following cycle where every step is the unique reply of the playing agent.

$$\begin{array}{ccccc}
(1,1,0,1)\{a,b,x\} & \xrightarrow{2} & (1,0,0,2)\{x\} & \xrightarrow{3} & (1,0,1,1)\{a,x,c\} \\
\uparrow_1 & & & & \downarrow_1 \\
(0,1,0,2)\{x\} & \xleftarrow{3} & (0,1,1,1)\{x,b,c\} & \xleftarrow{2} & (0,0,1,2)\{x\}
\end{array}$$

$\diamondsuit$

Proposition 4.18. *f^{PR} is not FDRP even from the truth.*

Example 4.18. We take the game from Ex. 4.17, and add for each voter $i \in \{1, 2, 3\}$ a candidate d_i, s.t. $u_i(d_i) = 8, u_i(d_j) = j$ for $j \neq i$. Thus initially each voter i votes for d_i. We also add an initial score of 3 to each of the candidates $\{a, b, c, x\}$. Voter 3 moves first to $a_3^1 = x$, which is a direct reply. Then voters 1 and 2 move to their best replies a, b, respectively. Now the cycle continues as per Ex. 4.17. $\diamondsuit$

4.4. Stochastic Dominance and Local Dominance

Extending L_i to a complete preference over subsets (that is consistent with the axioms) is one way to define a better reply dynamics for F^P. Another way to derive a well-defined dynamics from any *partial order* L_i over subsets of candidates, is by assuming that a voter performs a better reply step if she strictly prefers the new outcome according to L_i, and otherwise (if the new outcome is same, worse, or incomparable) she does not move.

One example of such a partial order is *stochastic dominance* (SD), which was applied to tie-breaking by [12]. A different partial order is implied by *local dominance* (LD) which was defined for voting with uncertainty about the outcome [54, 34]. We show how convergence results for LD/SD dynamics fit with other results.

Stochastic dominance

Reyhani and Wilson assume that ties are broken uniformly at random, and that a voter will only perform a step that stochastically dominates the current winner(s), if such exists.

Theorem 4.19 (Theorem 5.7 in [12])**.** *Plurality with stochastic dominance tie-breaking is FDBRP.*

A natural question is how this result related to the results in Section 4.3, where voters have cardinal utilities.

Since any SD step is also a better reply under any cardinal utility scale [12], any strong or order-free convergence result for the latter applies to the former, but not vice-versa. In particular, if we restrict attention to convergence from the truth, we have the following immediate corollary from Theorem 4.14:

Corollary 4.20. *Plurality with stochastic dominance tie-breaking is FBRP from the truth.*

Local dominance

Suppose that there are several candidates with maximal score. A voter may consider all of them as "possible winners," without specifying how the actual winner is selected. If the voter is concerned about making a move that will leave him worse off, he will only make moves that will improve his utility with certainty, i.e., that dominates his current action (where possible worlds are all strict tie-breaking orders) [54, 34, 41].[10]

Theorem 4.21 (Theorem 11 in the full version of [41])**.** *Plurality with Local-Dominance tie-breaking is FDRP.*

Which tie-breaking axioms are required for convergence?

We can use our axioms to characterize all better replies. Note that by our assumption voters that cannot conclude that a step will myopically benefit them prefer to keep their current vote. Thus adding more axioms lets the voter make a more informed decision and therefore only increases the number of improvement steps and may only add cycles. We emphasize that in either case convergence may be to a state that is not a Nash equilibrium.

If we assume that voters only follow steps that are better replies by Axiom K (an extreme risk-averse behavior), then it is easy to see that only moves to a more-preferred candidate can be better replies (any move to or from a tie cannot follow from Axiom K and is thus forbidden), which means that there are trivially no cycles, and Plurality becomes FIP.

[10] These papers consider a setting where the voter is uncertain about the outcome in general. [34, 41] consider uncertainty over the candidates' score, and [54] considers arbitrary information sets. We simply apply the same decision criterion to the case where uncertainty is regarding the tie breaking.

Proposition 4.22 (See appendix for the proof)**.** *A step $\boldsymbol{a} \xrightarrow{i} \boldsymbol{a}'$ is a better reply under unknown tie-breaking and local dominance, if and only if $F^P(\boldsymbol{a}') \succ_i F^P(\boldsymbol{a})$ is entailed by L_i, Axioms K+G, and transitivity.*

Proposition 4.23 (Meir [55])**.** *A step $\boldsymbol{a} \xrightarrow{i} \boldsymbol{a}'$ is a better reply under random tie-breaking and stochastic dominance, if and only if $F^P(\boldsymbol{a}') \succ_i F^P(\boldsymbol{a})$ is entailed by L_i, Axioms K+G+R, and transitivity.*

Thus by Propositions 4.21 and 4.22, Plurality with voters that follow Axioms K+G is FDRP (i.e., guaranteed to converge if voters stick to direct replies). In contrast, the analysis of Ex. 4.13 shows that all steps are entailed by Axioms K+G, meaning that imposing Axioms K+G make it possible to have cycles of arbitrary better replies.

Similarly, by Propositions 4.19, 4.20 and 4.23, Plurality with voters that follow all Axioms K+G+R is both FDBRP and FBRP from the truth. That is, convergence is guaranteed if voters both stick to best replies, and either keep them direct or start from the truthful state. It is left as an open question if cycles of arbitrary direct replies are possible, i.e., whether FDRP holds even with Axiom R.

5. Separating Order-Free and Weak Acyclicity

Except for Plurality and Veto, convergence is not guaranteed even under restrictions on the action scheduler and the initial state. In contrast, simulations [33, 34, 35] show that iterative voting almost always converges even when this is not guaranteed by theory. We believe that weak acyclicity is an important part of the explanation to this gap. Recall that a voting rule is weakly acyclic if from any initial profile, there is some sequence of better-replies that reaches a pure strategy Nash equilibrium. Indeed, we show that there are voting rules, including a common variation of the Plurality rule, that are not order-free acyclic, but are weakly acyclic. This explains convergence in practice, since in a weakly acyclic game with a random scheduler, every cycle will only be repeated a finite number of times before convergence must occur.

5.1. Plurality with Random tie-breaking

We have seen in Section 4 that while f^{PR} is FBRP from the truthful initial state, this is no longer true from arbitrary states, and in fact f^{PR} is not order-free-FIP under any action scheduler. Our main theorem in this section shows that under a certain scheduler (of both agents and actions), convergence is guaranteed from *any* state. Further, this still holds if actions are restricted to direct replies. We use the following lemma, whose proof is in the appendix.

Lemma 5.1. *Consider any game $G = \langle f_{\hat{\boldsymbol{s}}}^{PR}, \boldsymbol{L}\rangle$. Consider some candidate z, and suppose that in $\boldsymbol{a}^0$, there are x, y s.t. $s^0(x) \geq s^0(y) \geq s^0(z) + 2$. Then, for any sequence of direct replies, $z \notin f(\boldsymbol{a}^t)$.*

Theorem 5.2. *$f_{\hat{\boldsymbol{s}}}^{PR}$ is weak-FDRP.*

Proof. Consider a game $G = \langle f_{\hat{\boldsymbol{s}}}^{PR}, \boldsymbol{L}\rangle$, and an initial state $\boldsymbol{a}^0$. For a state $\boldsymbol{a}$, denote by $B(\boldsymbol{a}) \subseteq A^n$ all states reachable from $\boldsymbol{a}$ via paths of direct replies. Let $B = B(\boldsymbol{a}^0)$, and assume towards a contradiction that B does not contain a Nash equilibrium. For every $\boldsymbol{b} \in B$, let $C(\boldsymbol{b}) = \{c \in C : \exists \boldsymbol{a} \in B(\boldsymbol{b}) \wedge c \in f(\boldsymbol{a})\}$, i.e., all candidates that are winners in some state reachable from $\boldsymbol{b}$.

For any $\boldsymbol{b} \in B(\boldsymbol{a}^0)$, define a game $G_{\boldsymbol{b}}$ by taking G and eliminating all candidates *not in* $C(\boldsymbol{b})$. Since we only consider direct replies, for any $\boldsymbol{a} \in B(\boldsymbol{b})$, the set of outgoing edges $I(\boldsymbol{a})$ is the same in G and in $G_{\boldsymbol{b}}$ (as any direct reply must be to a candidate in $C(\boldsymbol{b})$). Thus by our assumption, the set $B(\boldsymbol{b})$ in game $G_{\boldsymbol{b}}$ does not contain a NE.

For any $\boldsymbol{b} \in B(\boldsymbol{a}^0)$, let $\boldsymbol{b}^*$ be the truthful state of game $G_{\boldsymbol{b}}$, and let $T(\boldsymbol{b}) \subseteq N$ be the set of agents who are truthful in $\boldsymbol{b}$. That is, $i \in T(\boldsymbol{b})$ if $b_i = b^*_i$.

Let $\boldsymbol{b}^0$ be some state $\boldsymbol{b} \in B(\boldsymbol{a}^0)$ s.t. $|T(\boldsymbol{b})|$ is maximal, and let $T^0 = T(\boldsymbol{b}^0)$. If $|T^0| = n$ then $\boldsymbol{b}^0$ is the truthful state of $G_{\boldsymbol{b}^0}$, and thus by Theorem 4.14 all best reply paths from $\boldsymbol{b}^0$ in $G_{\boldsymbol{b}^0}$ lead to a NE, in contradiction to $B(\boldsymbol{b}^0)$ not containing any NE. Thus $T^0 < n$. We will prove that there is a path from $\boldsymbol{b}^0$ to a state $\boldsymbol{b}'$ s.t. $|T(\boldsymbol{b}')| > |T^0|$.

Let $i \notin T(\boldsymbol{b}^0)$ (must exist by the previous paragraph). Consider the score of candidate b^*_i at state $\boldsymbol{b}^0$. We divide into 6 cases. All scores specified below are in the game $G_{\boldsymbol{b}^0}$.

Case 1. $|f(\boldsymbol{b}^0)| > 1$ and $b^*_i \in f(\boldsymbol{b}^0)$ (i.e b^*_i is one of several winners). Then consider the step $\boldsymbol{b}^0 \xrightarrow{i} b^*_i$. This make b^*_i the unique winner, and thus it is a direct best reply for i. In the new state $\boldsymbol{b}' = (\boldsymbol{b}^0_{-i}, b^*_i)$, and we have $T(\boldsymbol{b}') = T(\boldsymbol{b}^0) \cup \{i\}$.

Case 2. $s^0(b^*_i) = sw^0 - 1$ (i.e., b^*_i needs one more vote to become a winner). By Axioms G and R, i prefers $f(\boldsymbol{b}^0_{-i}, b^*_i)$ over $f(\boldsymbol{b}^0)$. Then similarly to case 1, i has a direct step $\boldsymbol{b}^0 \xrightarrow{i} b^*_i$, which results in a state $\boldsymbol{b}'$ that is closer to being truthful.[11]

[11] Note that this is not necessarily a best reply.

Case 3. $b_i^* = f(\boldsymbol{b}^0)$ (i.e b_i^* is the unique winner). Then the next step $\boldsymbol{b}^0 \xrightarrow{j} \boldsymbol{b}^1$ will bring us to one of the two previous cases. Moreover, it must hold that $j \notin T(\boldsymbol{b}^0)$ since otherwise $b_j^0 = b_j^* = f(\boldsymbol{b}^0)$ which means $I_j(\boldsymbol{b}^0) = \emptyset$. Thus $|T(\boldsymbol{b}')| = |T(\boldsymbol{b}^1)| + 1 \geq |T(\boldsymbol{b}^0)| + 1$.

Case 4. There are x, y s.t. $s^0(x) \geq s^0(y) \geq s^0(b_i^*) + 2$. Then by Lemma 5.1 b_i^* can never be selected, in contradiction to $b_i^* \in C(\boldsymbol{b}^0)$.

Case 5. There is a unique winner $x = f(\boldsymbol{b}^0)$, and $s^0(x) \geq s^0(y) + 2$ for all other candidates (note that there has to be an equality for at least one candidate, or else there is no better reply). Then the next step (by some voter j) must be from x, which brings us to one of the Cases 1,2 (if $s^0(b_i^*) = sw^0 - 2$) or in Case 4 (if $s^0(b_i^*) < sw^0 - 2$).

Case 6. The last case is when there is a unique winner $x = f(\boldsymbol{b}^0)$, some other candidate y with $s^0(y) = sw^0 - 1$, and $s^0(b_i^*) = sw^0 - 2$. Then there are two types of steps: a type I step is from x to y (by a voter that prefers y over x), and a type II step is any other step. A type I step puts us back in Case 6, but there can be at most 1 such step by each voter. Thus eventually there will be a type II step. Finally, any type II move puts us either in Case 1 or 2 (if x loses a vote) or in Case 4 (if y gains a vote).

Therefore we either construct a path of direct replies to $\boldsymbol{b}' \in B(\boldsymbol{b}^0)$ with $|T(\boldsymbol{b}')| > |T(\boldsymbol{b}^0)|$ in contradiction to our maximality assumption, or we reach another contradiction. Thus $B(\boldsymbol{b}^0)$ must contain some NE (both in $G_{\boldsymbol{b}^0}$ and in G), which means by construction that G is weakly-FDRP from $\boldsymbol{b}^0$. However since $\boldsymbol{b}^0 \in B(\boldsymbol{a}^0)$, we get that G is weakly-FDRP from $\boldsymbol{a}^0$ as well. $\square$

Remark 5.1. *Theorem 5.2 and Ex. 4.17 provide a partial answer to an open question regarding whether there are game forms that admit weak-FIP but not order-free-FIP [7]. Indeed, the game form f^{PR} for $m = 4, n = 3$ is such an example, but one that uses randomization. However, if we think of f^{PR} as a deterministic game form with $2^m - 1$ possible outcomes (all nonempty subsets of candidates), where agents are restricted to m actions each, then the allowed utility profiles are constrained (by Axioms G and R) and thus this result does not settle Kukushkin's question completely.*

5.2. Weighted Plurality

When voters are weighted, cycles of direct replies can emerge [1, 56]. We conjecture that such cycles must depend on the order of agents, and that certain orders will break such cycles and reach an equilibrium, at least from the truthful state.

Conjecture 5.3. *$f^{PL}_{\hat{\boldsymbol{s}},\boldsymbol{w}}$ is weak-FDRP (in particular weak-FIP).*

Similar techniques to those used so far (where some invariants are maintained throughout the better reply path) appear to be insufficient to prove the conjecture. For example, in contrast to the unweighted case, a voter might return to a candidate she deserted in *any scheduler*, even if only two weight levels are present. We thus leave the proof of the general conjecture for future work, possibly with the aid of "non-constructive" convergence proofs, such as the ones in [41].

Yet, we want to demonstrate the power of weak acyclicity over order-free acyclicity, even when there are no randomness or restrictions on the utility space. That is, we intend to provide a definite (negative) answer to Kukushkin's question of whether weak acyclicity entails order-free acyclicity. To this end, we use a slight variation of Plurality with weighted voters and lexicographic tie-breaking.

Theorem 5.4. *There exists a game form f^* s.t. f^* is weak-FIP but not order-free-FIP.*

Proof. Consider the game $\left\langle f^{PL}_{\hat{\boldsymbol{s}},\boldsymbol{w}}, \boldsymbol{L} \right\rangle$ from Example 4.6: The initial fixed score of candidates $\{a, b, c, d\}$ is $\hat{\boldsymbol{s}} = (0, 1, 2, 3)$. The weight of each voter $i \in \{1, 2, 3\}$ is i. The preference profile is as follows: $c \succ_1 d \succ_1 b \succ_1 a$, $b \succ_2 c \succ_2 a \succ_2 d$, and $a \succ_3 b \succ_3 c \succ_3 d$. This game was used in Section 4.1 to demonstrate that Plurality with weighted voters is not FDRP, however it can be verified that this game is order-free-FIP so it is not good enough for our use.

If we ignore the agents' preferences, we get a particular game form $f^{PL}_{\hat{\boldsymbol{s}},\boldsymbol{w}}$ where $N = \{1, 2, 3\}$, $M = \{a, b, c, d\}$, $\hat{\boldsymbol{s}} = (0, 1, 2, 3)$ and $\boldsymbol{w} = (1, 2, 3)$.

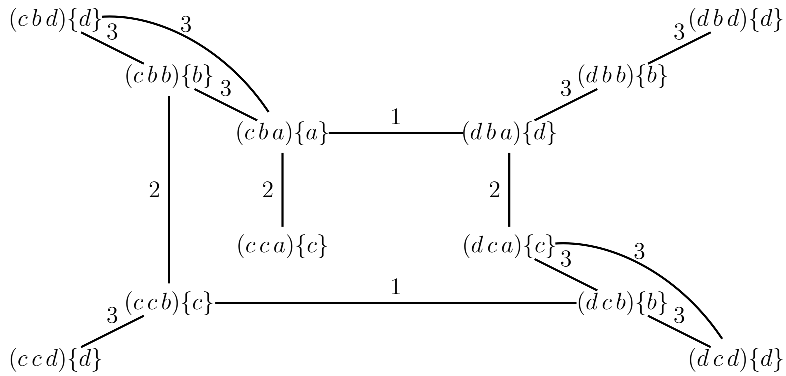
We define f^* by modifying $f^{PL}_{\hat{\boldsymbol{s}},\boldsymbol{w}}$ with the following restrictions on agents' actions: $A_1 = \{c, d\}, A_2 = \{b, c\}$, and $A_3 = \{a, b, d\}$. Thus f^* is a $2 \times 2 \times 3$ game form, presented in Figure 7(a).

We first show that f^* is not order-free-FIP. Indeed, consider the game G^* accepted from f^* with the same preferences from game G (Figure 7(b)). We can see that there is a cycle of length 6 (in bold). An agent scheduler that always selects the agent with the bold reply guarantees that convergence does not occur, since in all 6 relevant states the selected agent has no alternative replies.

Next, we show that f^* is weak-FIP. That is, for any preference profile there is some scheduler that guarantees convergence. We thus divide into cases according to the preferences of agent 3. In each case, we specify a state where the scheduler selects agent 3, the action of the agent, and the new state.

We note that since all thick edges must be oriented in the same direction, $a \succ_3 b$ if and only if $b \succ_3 c$. Thus the following three cases are exhaustive.

(a) The game form f^*



(b) The game G^*

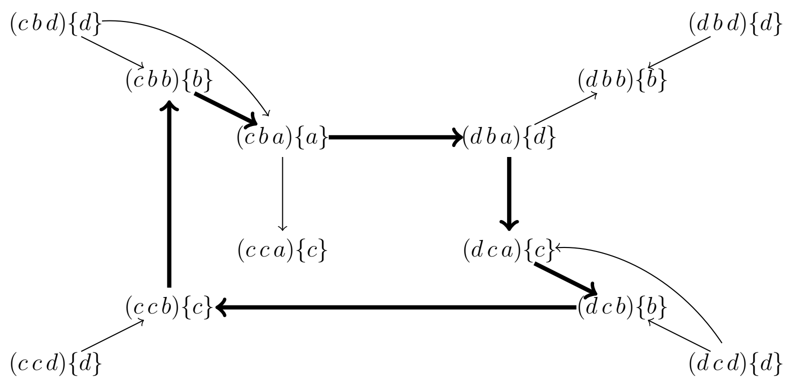


Figure 7: In each state we specify the actions of all 3 agents, and the outcome in curly brackets. Agent 1 controls the horizontal axis, agent 2 the vertical axis, and agent 3 the in/out axis. We omit edges between states with identical outcomes, since such moves are impossible for any preferences. A directed edge in (b) is a better reply in G^*.

	L_3	state	action	new state
1	$b \succ d$	(d, b, a)	b	(d, b, b)
2	$d \succ b \;\&\; d \succ a$	(c, b, b)	d	(c, b, d)
3	$a \succ d \succ b \succ c$	(d, c, b)	d	(d, c, d)

In either case, agent 3 moves from a state on the cycle to a Nash equilibrium.

□

6. Conclusions

Acyclicity is a highly desirable property of games, as it means that the agents will reach a deterministic stable state (equilibrium), even if they act myopically with little knowledge and coordination. It is particularly useful in the analysis and design of voting mechanisms: allowing voters to freely modify their votes will let them reach an equilibrium regardless of their initial votes and the information about other voters' preferences. Importantly, in the context of voting such myopic dynamics define a natural and straightforward process, which, coupled with the convergence properties demonstrated in this paper, makes it an attractive candidate method for predicting human voter behavior in elections and designing artificial agents with strategic voting capabilities—two of the most important, and also the hardest, goals of social choice research.

The outcome of such voting processes can be thought of as *compromise* reached by means of an iterative game—the agents might start voting for their favorites, but looking at the current scores, might like to compromise and support less preferred candidates if they (unlike the top choice candidates) have more chances to win the election. Importantly, recent theoretical and empirical work demonstrates that these compromise outcomes are never much worse than the truthful Plurality outcome, and are often significantly better according to Condorcet efficiency, voters' social welfare and other metrics [32, 33, 34]. These findings, together with our convergence results, thus suggest that online voting platforms such as Doodle and Facebook should enable and maybe even encourage their users to look at current candidates' scores and update their votes dynamically, as this would guarantee both stable and socially desirable outcome.

Beyond the direct implication of various acyclicity properties on interactive settings where agents vote one by one, strong/weak acyclicity is tightly linked to convergence properties of more sophisticated learning strategies in repeated games [57, 58], which gives another reason to understand them. However, while there exist some broad classes of acyclic games (with congestion/potential games being probably the best known representative), there are not many natural game forms that demonstrate acyclicity [7]. To this end, Fabrikant et al. [18] provide a sufficient condition for weak acyclicity, namely that any subgame contains a unique Nash equilibrium; another sufficient condition due to Apt and Simon [8] is by eliminating never-best reply strategies. Unfortunately, both these criteria are not applicable to most voting rules, where typically (at least) all unanimous votes form equilibria, and every strategy is a best reply to some joint vote of the others.

In this work, we focus on the common Plurality voting rule and show that it

code	author(s)	citation
$*1$	Koolyk, Strangway, Lev and Rosenschein'17	[35]
$*2$	Lev'15	[59]
$*3$	Lev and Rosenschein'12	[46]
$*4$	Lev and Rosenschein'16	[14]
$*5$	Meir'15	[41]
$*6$	Meir'16	[56]
$*7$	Reyhani and Wilson'12	[12]

Table 1: Reference codes for Tables 3 and 4.

has the desired acyclicity property: natural better reply dynamics guarantees that voters will converge to an equilibrium, while the exact conditions under which this will occur may depend on the tie-breaking method. A key insight for these results is the identification of *direct replies*—a natural restriction on voters' actions that might be generalized to other voting rules. Moreover, we provide a joint rigorous framework for the study of iterative voting, as part of the broader literature on acyclicity of games and game forms, which allows us to compare all known convergence results from the literature, and derive some new entailments. In particular, we demonstrate variations of Plurality that are weakly acyclic but not order-free acyclic, thereby settling an open question on whether such game forms exist [7].

We summarize all known results on iterative voting that we are aware of in Tables 3 and 4. Note that in some cases we get positive results if we restrict the initial state, the number of voters, or some other parameter (not shown in the table). For Plurality we provide a more detailed picture in Figs. 8 and 9. The tables also visualize which questions are still open (mainly for rules other than Plurality). Note that previous papers whose results are covered in the tables often use different terminology and thus theorems and examples need to be rephrased (and sometimes slightly modified) to be directly comparable. These rephrasing and necessary modifications are explained in detail in [56].

Based on the progress made in this paper and the other results published since the introduction of iterative voting in [1], we believe that research on iterative voting should focus on three primary directions:

1. Weak acyclicity seems more indicative than order-free acyclicity to determine convergence in practice. Thus, theorists should study which voting rules are weak-FIP, perhaps under reasonable restrictions (as we demon-

Property	Game converges for:
FIP	any selection of players and replies
FBRP	any selection of players and best replies
FDRP	any selection of players and direct replies
FDBRP	any selection of players, and the unique direct best reply
Order-free-FIP	any selection of players, for some reply
Weak-FDRP	some selection of players and direct replies
Weak-FIP	some selection of players and replies

Table 2: Reminder of the different notions of convergence properties and what they mean. The entailment relations among them are detailed in Figure 5.

Voting rule	FIP	FDRP	FDBRP	Order-free-FIP	Weak-FIP
Plurality (Lex.)	✗(4.3)	✓(4.2)	✓	✓	✓
Plurality (LD)	✗(4.13)	✓^{*5}	✓	✓	✓
Plurality (SD)	✗(4.13)	?	✓^{*7}	✓	✓
Plurality (Rand.)	✗	✗	✗	✗(4.17)	✓(5.2)
Weighted Plurality (Lex.)	✗	✗	✗(4.6)	?	?
Veto (Lex.)	✗^{*6}	✓^{*7*3}		✓	✓
Veto (SD)	?	✓^{*7}		✓	✓

Table 3: Positive results (in light green) carry to the right side, negative results (in dark gray) to the left side. Tie-breaking methods: Lex.- Lexicographic, LD - Local Dominance, SD - Stochastic Dominance, Rand. - Randomized with cardinal utilities. The number in brackets points to the Theorem or Proposition proving this result. The numbered superscripts are references to results shown in other papers, see Table 1.

strated, this property is distinct from order-free-FIP). We highlight that even in rules where there are counter-examples for weak acyclicity (k-approval, Borda), these examples use only two voters, and games with more voters may well be weakly acyclic.

2. It is important to experimentally study how people really vote in iterative settings (both in and out of the lab), so that this behavior can be formalized and behavioral models can be improved. The work of [60] is a preliminary step in this direction, but there is much more to learn. Ideally, we would like to identify a few types of voters, such that for each type we can relatively accurately predict the next action in a particular state. It would be even better if these types are not specific to a particular voting rule or contextual details.

Voting rule	FIP	FBRP	order-free-FIP	Weak-FIP
Direct Kingmaker	✓(3.2)	✓	✓	✓
Plurality	✗	✗(4.3)	✓(4.2)	✓
Veto	✗	✗^{*6}	✓^{*7*3}	✓
k-approval ($k \geq 2$)	✗	✗^{*3*2}	✗	✗^{*6}
Borda	✗	✗^{*7*3}	✗	✗^{*7}
PSRs (except k-approval)	✗	✗^{*3*2}	?	?
Approval	✗	✗^{*6}	✓^{*6}	✓
Other common rules	✗	✗^{*1}	?	?

Table 4: Positive results carry to the right side, negative to the left side. All rules in the table use lexicographic tie-breaking.

(from truth)						
FBRP ✓(4.5)	⇐	FBRP ✗(4.3)		Order-free-FBRP ✓	⇒	Weak-FBRP ✓
⇑		⇑		⇓		⇓
FIP ✗	⇐	FIP ✗	⇒	Order-free-FIP ✓	⇒	Weak-FIP ✓
⇓		⇓		⇑		⇑
FDRP ✓	⇐	FDRP ✓(4.2)		Order-free-FDRP ✓	⇒	Weak-FDRP ✓

Figure 8: Convergence results for Plurality under lexicographic tie-breaking. Positive results (in light green) carry with the direction of the arrows ad in Fig. 5, whereas negative results (dark gray) carry in the opposite direction. For example non-convergence from the truthful state (leftmost column) implies non-convergence from an arbitrary initial state.

3. We would like to know not only if a voting rule converges under a particular dynamics (always or often), but also what are the properties of the attained outcome—in particular, whether the iterative process improves welfare or fairness, avoids "voting paradoxes" [61], and so on. Towards this end, several researchers (e.g., [37, 32, 34, 36, 35]) have started to explore these questions via theory and simulations. However, a good understanding of how iterative voting shapes the outcome, whether the population of voters consists of humans or artificial agents, is still under way.

We hope that this paper will help in focusing and classifying future work.

(from truth)						
FBRP ✓(4.14)	$\Leftarrow$	FBRP ✗		Order-free-FBRP ✗	$\Rightarrow$	Weak-FBRP ?
$\Uparrow$		$\Uparrow$		$\Downarrow$		$\Downarrow$
FIP ✗	$\Leftarrow$	FIP ✗	$\Rightarrow$	order-free-FIP ✗(4.17)	$\Rightarrow$	weak-FIP ✓
$\Downarrow$		$\Downarrow$		$\Uparrow$		$\Uparrow$
FDRP ✗(4.18)	$\Leftarrow$	FDRP ✗		order-free-FDRP ✗	$\Rightarrow$	weak-FDRP ✓(5.2)

Figure 9: Convergence results for Plurality under random tie-breaking.

Acknowledgments

The authors thank Ulle Endriss, Omer Lev, and Svetlana Obraztsova for their helpful comments. This research was partly supported by the Israel Science Foundation (ISF) under Grant #773/16 and Grant #1227/12.

[1] R. Meir, M. Polukarov, J. S. Rosenschein, N. Jennings, Convergence to equilibria of plurality voting, in: Proc. of 24th AAAI, 2010, pp. 823–828.

[2] R. Meir, Strong and weak acyclicity in iterative voting, in: Proc. of 9th SAGT, Springer, 2016, pp. 182–194.

[3] A. Gibbard, Manipulation of voting schemes, Econometrica 41 (1973) 587–602.

[4] M. Satterthwaite, Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions, Journal of Economic Theory 10 (1975) 187–217.

[5] A.-A. Cournot, Recherches sur les principes mathématiques de la théorie des richesses par Augustin Cournot, chez L. Hachette, 1838.

[6] D. Monderer, L. S. Shapley, Potential games, Games and Economic Behavior 14 (1) (1996) 124–143.

[7] N. S. Kukushkin, Acyclicity of improvements in finite game forms, International Journal of Game Theory 40 (1) (2011) 147–177.

[8] K. R. Apt, S. Simon, A classification of weakly acyclic games, in: Proc. of 5th SAGT, 2012, pp. 1–12.

[9] N. S. Kukushkin, Congestion games: a purely ordinal approach, Economics Letters 64 (1999) 279–283.

[10] N. S. Kukushkin, Perfect information and congestion games, Games and Economic Behavior 38 (2002) 306–317.

[11] D. Sherfinski, Buyers remorse? some states allow early voters to change ballots, The Washington Times (2016).
URL `http://www.washingtontimes.com/news/2016/nov/1/early-voting-buyers-remorse-some-states-allow-ball/`

[12] R. Reyhani, M. C. Wilson, Best-reply dynamics for scoring rules, in: Proc. of 20th ECAI, 2012.

[13] S. Obraztsova, E. Markakis, M. Polukarov, Z. Rabinovich, N. R. Jennings, On the convergence of iterative voting: How restrictive should restricted dynamics be?, in: Proc. of 29th AAAI, 2015.

[14] O. Lev, J. S. Rosenschein, Convergence of iterative scoring rules, Journal of Artificial Intelligence Research 57 (2016) 573–591.

[15] E. Boros, V. Gurvich, K. Makino, W. Shao, Nash-solvable bidirected cyclic two-person game forms, Tech. rep., Rutcor Research Report 26-2007 and DIMACS Technical Report 2008-13, Rutgers University (2008).

[16] D. Andersson, V. Gurvich, T. D. Hansen, On acyclicity of games with cycles, Discrete Applied Mathematics 158 (10) (2010) 1049–1063.

[17] I. Milchtaich, Congestion games with player-specific payoff functions, Games and economic behavior 13 (1) (1996) 111–124.

[18] A. Fabrikant, A. D. Jaggard, M. Schapira, On the structure of weakly acyclic games, in: Proc. of 3rd SAGT, 2010, pp. 126–137.

[19] P. Gärdenfors, Manipulation of social choice functions, Journal of Economic Theory 13 (2) (1976) 217–228.

[20] J. S. Kelly, Strategy-proofness and social choice functions without single-valuedness, Econometrica: Journal of the Econometric Society (1977) 439–446.

[21] D. G. Saari, Susceptibility to manipulation, Public Choice 64 (1990) 21–41.

[22] E. Friedgut, G. Kalai, N. Keller, N. Nisan, A quantitative version of the gibbard-satterthwaite theorem for three alternatives, SIAM Journal on Computing 40 (3) (2011) 934–952.

[23] R. Forsythe, T. Rietz, R. Myerson, R. Weber, An experimental study of voting rules and polls in three-candidate elections, International Journal of Game Theory 25 (3) (1996) 355–83.

[24] A. Dhillon, B. Lockwood, When are plurality rule voting games dominance-solvable?, Games and Economic Behavior 46 (2004) 55–75.

[25] S. Chopra, E. Pacuit, R. Parikh, Knowledge-theoretic properties of strategic voting, in: Proc. of 9th JELIA, 2004, pp. 18–30.

[26] M. R. Sertel, M. R. Sanver, Strong equilibrium outcomes of voting games are the generalized condorcet winners, Social Choice and Welfare 22 (2004) 331–347.

[27] D. Falik, R. Meir, M. Tennenholtz, On coalitions and stable winners in plurality, in: WINE'12, 2012, pp. 256–269.

[28] M. Messner, M. K. Polborn, Robust political equilibria under plurality and runoff rule, mimeo, Bocconi University (2002).

[29] S. Airiau, U. Endriss, Iterated majority voting, in: Proc. of 1st ADT, 2009, pp. 38–49.

[30] Y. Desmedt, E. Elkind, Equilibria of plurality voting with abstentions, in: Proc. of 11th ACM-EC, 2010, pp. 347–356.

[31] E. Elkind, U. Grandi, F. Rossi, A. Slinko, Gibbard–satterthwaite games, in: Proc. of 24th IJCAI, 2015.

[32] S. Brânzei, I. Caragiannis, J. Morgenstern, A. D. Procaccia, How bad is selfish voting?, in: Proc. of 27th AAAI, 2013.

[33] U. Grandi, A. Loreggia, F. Rossi, K. B. Venable, T. Walsh, Restricted manipulation in iterative voting: Condorcet efficiency and Borda score, in: Proc. of 3rd ADT, 2013, pp. 181–192.

[34] R. Meir, O. Lev, J. S. Rosenschein, A local-dominance theory of voting equilibria, in: Proc. of 15th ACM-EC, 2014.

[35] A. Koolyk, T. Strangway, O. Lev, J. S. Rosenschein, Convergence and quality of iterative voting under non-scoring rules, in: Proc. of 26th IJCAI, 2017, to appear.

[36] C. Bowman, J. K. Hodge, A. Yu, The potential of iterative voting to solve the separability problem in referendum elections, Theory and decision 77 (1) (2014) 111–124.

[37] A. Reijngoud, U. Endriss, Voter response to iterated poll information, in: Proc. of 11th AAMAS, 2012, pp. 635–644.

[38] N. Gohar, Manipulative voting dynamics, Ph.D. thesis, University of Liverpool (2012).

[39] S. Obraztsova, E. Markakis, D. R. M. Thompson, Plurality voting with truth-biased agents, in: Proc. of 6th SAGT, 2013, pp. 26–37.

[40] Z. Rabinovich, S. Obraztsova, O. Lev, E. Markakis, J. S. Rosenschein, Analysis of equilibria in iterative voting schemes, in: Proc. of 29th AAAI, 2015.

[41] R. Meir, Plurality voting under uncertainty, in: Proc. of 29th AAAI, 2015, pp. 2103–2109.

[42] U. Endriss, S. Obraztsova, M. Polukarov, J. S. Rosenschein, Strategic voting with incomplete information, in: Proc. of 25th IJCAI, 2016.

[43] W. S. Zwicker, Introduction to the theory of voting, in: F. Brandt, V. Conitzer, U. Endriss, J. Lang, A. D. Procaccia (Eds.), Handbook of Computational Social Choice, Cambridge University Press, 2016.

[44] H. P. Young, The evolution of conventions, Econometrica: Journal of the Econometric Society (1993) 57–84.

[45] I. Milchtaich, Schedulers, potentials and weak potentials in weakly acyclic games, Tech. rep., Working Papers, Bar-Ilan University, Department of Economics (2013).

[46] O. Lev, J. S. Rosenschein, Convergence of iterative voting, in: Proc. of 11th AAMAS, 2012, pp. 611–618.

[47] B. Dutta, Effectivity functions and acceptable game forms, Econometrica: Journal of the Econometric Society (1984) 1151–1166.

[48] E. Boros, V. Gurvich, K. Makino, D. Papp, Acyclic, or totally tight, two-person game forms: Characterization and main properties, Discrete Mathematics 310 (6) (2010) 1135–1151.

[49] R. W. Hamming, Error detecting and error correcting codes, Bell System technical journal 29 (2) (1950) 147–160.

[50] A. Reijngoud, Voter response to iterated poll information, Ph.D. thesis, Universiteit van Amsterdam (2011).

[51] A. E. Roth, The college admissions problem is not equivalent to the marriage problem, Journal of economic Theory 36 (2) (1985) 277–288.

[52] C. Geist, U. Endriss, Automated search for impossibility theorems in social choice theory: Ranking sets of objects, Journal of Artificial Intelligence Research 40 (1) (2011) 143–174.

[53] U. Endriss, Sincerity and manipulation under approval voting, Theory and Decision 74 (3) (2013) 335–355.

[54] V. Conitzer, T. Walsh, L. Xia, Dominating manipulations in voting with partial information., in: Proc. of 25th AAAI, Vol. 11, 2011, pp. 638–643.

[55] R. Meir, Random tie-breaking with stochastic dominance, CoRR abs/1609.01682.

[56] R. Meir, Strong and weak acyclicity in iterative voting, in: COMSOC'16, 2016.

[57] M. Bowling, Convergence and no-regret in multiagent learning, Advances in neural information processing systems 17 (2005) 209–216.

[58] J. R. Marden, G. Arslan, J. S. Shamma, Regret based dynamics: convergence in weakly acyclic games, in: Proc. of 6th AAMAS, 2007.

[59] O. Lev, Agent modeling of human interaction: Stability, dynamics and cooperation, Ph.D. thesis, The Hebrew University of Jerusalem (2015).

[60] M. Tal, R. Meir, Y. Gal, A study of human behavior in voting systems, in: Proc. of 14th AAMAS, 2015, pp. 665–673.

[61] L. Xia, J. Lang, M. Ying, Sequential voting rules and multiple elections paradoxes, in: Proc. of 13thTARK, 2007, pp. 279–288.

Appendix A. Proofs

Theorem 3.4. *For any $n \geq 7$, there is a non-separable game form f_n^* with n agents s.t. f_n^* is FIP.*

Proof for $n = 7$. Let $C' = \{c^1, \ldots, c^{2n}\}$ and $C = C' \cup \{z\}$. Let $A_i = \{0, 1\}$ for each voter. Every voting function f_n is a function from the n dimensional binary cube $\mathcal{B} = \{0, 1\}^n$ to C.

Denote by $\overline{k}$ the total number of bits required for an optimal single error-correcting code (Hamming code) with k data bits. For example, for $k = 4$ we need $\overline{4} = 7$ bits. In particular, there is a mapping $q : \{0, 1\}^k \to \{0, 1\}^{\overline{k}}$ such that the Hamming distance between any two words is at least 3 [49]. Formally, for all $w, w' \in \{0, 1\}^k$, it holds that $|\{j : q(w)_j \neq q(w')_j\}| \geq 3$.

For reasons that will become apparent later, we want to have $\overline{k} \leq n$, and $2n = |C'| \leq 2^k$.

For any $n = 7$, set $r = \lceil \log_2(n + 1) \rceil = 3$, and $k = \lceil \log_2(2n) \rceil = 4$. Thus $2^r - r - 1 = 2^3 - 3 - 1 = 4 = k$. From coding theory [49], for all $r \geq 2$, if $k \leq 2^r - r - 1$, then $\overline{k} = 2^r - 1 = 7$ bits are sufficient to code all $(k = 4)$-bit strings. That is, there is a valid code q from $\{0, 1\}^4$ to $\{0, 1\}^7$.

Note that $\overline{k} = 7 = n$, and that $|C'| = 2n = 14 < 2^4 = 2^k$.

Let $bin(t, k) \in \{0, 1\}^k$ be the k-bit binary representation of $t \leq 2n$ (e.g., $bin(5, 4) = 0101$). Since $2n < 16 = 2^4$, all of $t \leq 2n = 14$ have a 4-bit representation. Using the Hamming code q, we map each outcome $c^t \in C'$ to a specific profile $\boldsymbol{a}^t = q(bin(t, 4)) \in \{0, 1\}^7 = \{0, 1\}^n$.

We now define f_n^* by setting $f_n^*(\boldsymbol{a}^t) = c^t$ for all $t \leq 2n$, and $f_n^*(\boldsymbol{a}) = z$ for all $2^n - 2n = 114$ other profiles.

For any two profiles $\boldsymbol{a}, \boldsymbol{a}'$, let $d(\boldsymbol{a}, \boldsymbol{a}')$ be the number of voters that disagree in $\boldsymbol{a}, \boldsymbol{a}'$ (the Hamming distance on the cube). Let $B \subseteq \mathcal{B}$ be all 14 profiles whose outcome is not z. By the construction above, we have that $d(\boldsymbol{a}, \boldsymbol{a}') \geq 3$ for all $\boldsymbol{a}, \boldsymbol{a}' \in B$.

Assume towards a contradiction that there is some cycle of better replies in f_n^*. Then there must be a path containing at least three distinct outcomes, and thus at least two profiles from B. Denote these profiles by $\boldsymbol{a}, \boldsymbol{b} \in B$. Since any path between $\boldsymbol{a}$ and $\boldsymbol{b}$ is of length at least 3, the path must contain at least two consequent profiles whose outcome is z. This path cannot be a better reply path, since a better reply must in particular change the outcome. Hence we get a contradiction, and f_n^* is FIP.

Finally, f_n^* cannot be separable, exactly as in the proof for $n \geq 8$. □

Lemma 4.15. *Consider any cardinal utility function u and the partial preference order L it induces on subsets by random tie-breaking. L holds Axioms K+G+R.*

Proof. Let u be any utility scale, we will show that all axioms hold. Let $a, b \in C$ and $W \subseteq C \setminus \{a, b\}$.

$$u(\{a\} \cup W) = \frac{1}{|W|+1|}\left(u(a) + \sum_{c \in W} u(c)\right),$$

$$u(\{b\} \cup W) = \frac{1}{|W|+1|}\left(u(b) + \sum_{c \in W} u(c)\right) = u(\{b\} \cup W),$$

thus $\{a\} \cup W \succ_L \{b\} \cup W$, and Axiom R holds.

Let $a \in C, W \subseteq C$ s.t. $\forall b \in W, a \succ_b$. Then

$$u(a) = \frac{1}{|W|+1}\left(u(a) + \sum_{b \in W} u(a)\right) > \frac{1}{|W|+1}\left(u(a) + \sum_{b \in W} u(b)\right) = u(\{a\} \cup W)$$
$$> \frac{1}{|W|+1}\left(u(W) + \sum_{b \in W} u(a)\right) = \frac{1}{|W|+1} u(W) + \frac{|W|}{|W|+1} u(W) = u(W),$$

thus $a \succ_L \{a\} \cup W \succ_L W$ and Axiom G holds.

Axiom K1 follows immediately from G. K2 also follows if preferences are strict. Even if there are ties, and $a \succeq w$ for all $a \in A, w \in W$ then:

$$u(A) \geq \min_{a \in A} u(a) \geq \max_{w \in W} u(w) \geq u(W),$$

i.e., $A \succeq_L W$. □

Proposition 4.22. *A step $\boldsymbol{a} \xrightarrow{i} \boldsymbol{a}'$ is a better-response under unknown tie-breaking and local dominance, if and only if $f(\boldsymbol{a}') \succ_i f(\boldsymbol{a})$ is entailed by L_i, Axioms K+G, and transitivity.*

Proof. Suppose that $X = f(\boldsymbol{a}')$ locally-dominates $Y = f(\boldsymbol{a})$. Let $Z = X \cap Y$, and $X' = X \setminus Z$, $Y' = Y \setminus Z$. We must have $x \succ_i y$ for any $x \in X, y \in Y'$, otherwise, a tie-breaking order that selects y first and x second would make i strictly lose when moving from Y to X. Similarly, $x \succ_i y$ for any $x \in X', y \in Y$. If $Z = \emptyset$ then $X = X' \succ_i Y' = Y$ follows from Axiom K. Otherwise, by repeatedly applying Axiom G we get $X \succeq_i Z \succeq_i Y$ with at least one relation being strict.

In the other direction, since Axiom G can only be used to add elements lower (or higher) than all existing elements, it may only induce relations of the form $Z \succ Z \cup Y'$ where $z \succ y$ for all $z \in Z, y \in Y'$; or relations of the form $Z \cup X' \succ Z$ where $x \succ z$ for all $z \in Z, x \in X'$. Thus if $X \succ Y$ follows from Axiom G, they must have the form $X = Z \cup X', Y = Z \cup Y'$ where $x \succ z \succ y$ for all $x \in X', z \in Z, y \in Y'$. To see that this entails local dominance, let $x_L = L(X)$ be the first element in X according to order $L \in \pi(C)$, and likewise for Y. For any L, $x_L \succeq y_L$ (with equality iff $L(X) = L(Y) \in Z$). Further, either X' or Y' are non-empty (w.l.o.g. X'). Consider an order L' such that $L'(X) \in X'$, then $x_{L'} \succ y$ for all $y \in Y$ and in particular $x_{L'} \succ y_{L'}$. $\square$

Lemma 5.1. *Consider any game $G = \langle f^{PR}_{\hat{\boldsymbol{s}}}, \boldsymbol{L} \rangle$. Consider some candidate z, and suppose that in $\boldsymbol{a}^0$, there are x, y s.t. $s^0(x) \geq s^0(y) \geq s^0(z) + 2$. Then, for any sequence of direct replies, $z \notin f(\boldsymbol{a}^t)$.*

Proof. We show that at any time $t \geq 0$ there are x^t, y^t s.t. $s^0(x), s^0(y) \geq s^0(z) + 2$. For $t = 0$ this holds for $x^t = x, y^t = y$. Assume by induction that the premise holds for $\boldsymbol{a}^{t-1}$. Then there are two cases:

1. $|f(\boldsymbol{a}^{t-1})| \geq 2$. Then since step t must be a direct reply, it must be to some candidate z with $s^{t-1}(z) \geq sw^{t-1} - 1$. Also, either x^{t-1} or y^{t-1} did not lose votes (w.l.o.g. x^{t-1}). Thus $s^t(x), s^t(z) \geq sw^{t-1} \geq s^{t-1}(z) + 2 \geq s^t(z) + 2$.

2. $|f(\boldsymbol{a}^{t-1})| = 1$. Then suppose $f(\boldsymbol{a}^{t-1}) = \{x^{t-1}\}$, and we have that $sw^{t-1} \geq s^{t-1}(z) + 3$. The next step is z where either $s^{t-1}(z) = sw^{t-1} - 1$ (and then we conclude as in case 1), or $s^{t-1}(z) = sw^{t-1} - 2$ and x^{t-1} loses 1 vote. In the latter case, $s^t(x^{t-1}) = s^t(z) = sw^{t-1} - 1 \geq s^{t-1}(z) + 2 \geq s^t(z) + 2$. $\square$