Do we Care about Poll Manipulation in Political Elections?

Vincent Mousseau[a], **Henri Surugue**[a,*] **and Anaëlle Wilczynski**[a]

[a]MICS, CentraleSupélec, Université Paris-Saclay

Abstract. We consider the problem of poll manipulation in political elections. In the context of strategic voting, we are interested in whether a polling institute can manipulate the information it communicates to voters in order to influence the outcome of the election. We start with a version of the problem where the polling institute is allowed to send any score to voters. Then, for realistic reasons, we investigate a restricted version in which the polling institute cannot announce scores which are too far from the truthful ones. While we show that both decision problems are computationally hard, we go beyond this worst-case complexity analysis by using probabilistic tools to address the possibility of successful and efficient manipulation in practice, w.r.t. several natural preference distributions.

1 Introduction

Strategic voting [23] is a major issue in political elections. Ideally, one would like to avoid such a strategic behavior. However, by the Gibbard-Satterthwaite theorem [16, 29], no reasonable voting rule is immune to voter manipulation. Since we cannot escape strategic voting, one way to tackle manipulation is to study it via a game-theoretical analysis. This approach has been followed by several works in computational social choice [6]. Notably, in *iterative voting* [22], the idea is to analyze the convergence and the quality of a sequential process where voters are allowed to make successive strategic deviations. In the classical iterative voting framework [24], complete knowledge is assumed, in the sense that voters are aware of all others' current ballot, sometimes even of their full preferences. However, this assumption is highly unrealistic and does not capture real scenarios with large electorates, such as political elections.

The question of the information available to the voters is key and has a strong impact on the manipulability of voting processes [11, 28]. To deal with partial information in voting, one can naturally follow a Bayesian approach by considering a probability distribution over a set of possible preference orders for other voters [26, 19]. Alternatively, a set of possible preference profiles can be derived from partial votes [8, 9] or from a given maximum distance to the voters' actual preferences [1]. Another possibility is to assume local information for the voters, which is captured by a social network [18]. Finally, an aggregated global information coming from opinion polls can be communicated to voters [3, 11, 28, 33].

Following this latter line of research, in this article, inspired by political elections, we assume that voters receive only a global information about the voting intentions within the population, which is communicated through *opinion polls*. Voters trust the information communicated in the polls and compute their best response ballot on the basis of this information. This confidence in the polls grants an important power to the polling institute which disseminates it, raising the natural question of *poll manipulation*. Indeed, a polling institute might have its own interests in the election and try to orient votes toward them. This problem is close to the question of election control [13], where an external agent aims to alter the outcome of the election, but here no structural change is made on the election.

In the line of seminal works analyzing the complexity of voter manipulation [2], one can analyze the complexity of the poll manipulation problem. However, computational intractability may not constitute a relevant barrier to manipulation, as it relies on worst-case analysis [12]. Therefore, to complement complexity results, an average-case study using a probabilistic approach is relevant, as it has been widely investigated for voter manipulation (see, e.g., [14, 21, 27, 35]). In particular, the asymptotic study is meaningful since political elections are characterized by a large number of voters. Considering election control problems, as far as we know, this approach has been surprisingly neglected. A notable exception is a recent work by Xia [34] which investigates the likelihood of manipulability for several coalition influence problems, including control by adding or deleting votes. Up to our best knowledge, no such study has been conducted so far for the poll manipulation problem.

In this article, we study the constructive poll manipulation problem where the polling institute wishes to favor a specific candidate by broadcasting manipulated candidates' scores. This problem has been introduced by Wilczynski [33] and further extended by Baumeister et al. [3], who also consider the destructive variant where the polling institute aims to prevent the election of a given candidate. While both works consider a framework where voters are embedded in a social network and analyze the complexity of the problem with respect to the structure of the graph, we consider a simpler model with no social network, which clarifies the role of the opinion polls. In particular, we analyze the following two versions of the problem. In the unrestricted problem, the polling institute is free to send any score information. The restricted problem considers a more realistic context where only score information that would be close enough to truthful scores are allowed. The idea for this second problem is for the polling institute to lie in a reasonable manner, by submitting realistic scores, not too far from a ground truth that may correspond to the results of a past election, or another poll. Such restrictions help to gain the trust and confidence from the voters. We prove that both versions of the problem are computationally hard, answering an open question from Baumeister et al. [3], but also analyze the probability of existence of a successful and efficiently computable poll manip-

* Corresponding Author. Email: henri.surugue@centralesupelec.fr.

ulation. For this latter purpose, we introduce a natural condition on statistical cultures which is satisfied by most natural preference distributions [31]. In fact, we exhibit a simple heuristic and prove its success for the unrestricted problem, which means that, without restriction, the polling institute can almost always efficiently influence large elections. For the restricted manipulation problem, we prove that if the allowed distance is negligible with respect to the number of voters, then no manipulation is possible. However, when this distance becomes significant, e.g., when it is a fixed proportion of the number of voters, easy manipulation is almost always successful in large elections. Finally, we show that most results still hold when assuming a more general strategic behavior for voters [33].

2 The Model

We first present the poll manipulation problem in the context of strategic voting.

2.1 *A voting system*

For any positive integer k, let $[k]$ denote the set $\{1, \ldots, k\}$. Let N be a set of voters where $N = [n]$, and M be a set of candidates where $M = [m]$. Since the goal is to study strategic voting in large political elections, we naturally assume that $n > m > 2$ (by the Gibbard-Satterthwaite theorem, voting rules are susceptible to manipulation only when there are more than two candidates). Each voter $i \in N$ has preferences over candidates represented by a linear order $\succ_i$ over candidates. Let $top_{\succ_i}$ and $worst_{\succ_i}$ denote the most preferred and worst candidates, respectively, of voter $i \in N$, i.e., $top_{\succ_i} \succeq_i x$ and $x \succeq_i worst_{\succ_i}$ for every candidate $x \in M$. The rank of a given candidate x in a preference order $\succ_i$ is denoted by $r_{\succ_i}(x)$, i.e., $r_{\succ_i}(x) := |\{y \in M : y \succeq_i x\}|$. We use the Kendall tau distance to evaluate the similarity between two preference orders, by counting the number of pairwise comparisons on which the two orders disagree, i.e., $dist_{KT}(\succ_i, \succ_j) = |\{(x, y) \in M^2 : x \succ_i y \text{ and } y \succ_j x\}|$. The set of linear orders $\succ_i$ for all voters $i \in N$ is called a preference profile and is denoted by $\succ$. Let us denote by $N^{x \succ y}$ the set of voters who prefer x over y, i.e., $N^{x \succ y} := \{i \in N : x \succ_i y\}$.

The winner of an election is determined by the Plurality voting rule where ties are broken lexicographically. Let $b_i \in M$ denote the ballot of voter i and $b \in M^n$ denote the ballot profile. The ballot profile b from which b_i is excluded is denoted by b_{-i}. The winner under Plurality of the ballot profile b is $\mathcal{W}_P(b) \in \arg\max_{x \in M} s_x(b)$, where $s_x(b) := |\{i \in N : b_i = x\}|$ and a lexicographic tie-breaking, denoted by $\rhd$, is used if necessary. By abuse of notation, we sometimes directly write $\mathcal{W}_P(s)$ to refer to the winner of a score vector s. Let b^T denote the truthful ballot profile, i.e., $b_i^T \succeq_i x$ for every candidate $x \in M$ and voter $i \in N$, and s^T denote the candidates' scores in b^T.

An election is given by the tuple $(N, M, \succ, \rhd)$.

2.2 *A strategic voting framework*

In this model, we consider an iterative voting process where voters are strategic with respect to the information they get, which only consists in the score broadcast by the polling institute. Like some previous works in iterative voting [3, 33], we make the assumption that voters trust the announced result by the polling institute but, in contrast, we assume that voters have no other local information on which to rely. Let us describe more in details the voters' strategic behavior.

Initially, all voters vote sincerely since they have no information yet, therefore the initial ballot profile b^0 is exactly the truthful ballot profile b^T. Then, the polling institute sends the results of this initial election by announcing a score vector s of size m such that $\sum_{j \in M} s_j = n$, where s_j is supposed to stand for the score of candidate $j \in M$ in the initial election. Let s^{-i} denote the score vector s where the truthful ballot of voter i has been removed, i.e., $s_j^{-i} = s_j$, for every $j \in M \setminus \{b_i^T\}$, and $s_{b_i^T}^{-i} = s_{b_i^T} - 1$. After the score is announced, every voter considers possible moves from her initial truthful ballot with respect to that information. Possible deviations are captured by the notion of potential winners: A candidate y is a *potential winner* for voter i w.r.t. announced score vector s if i believes that voting for y will make candidate y the new winner, i.e., $s_{\mathcal{W}_P(s^{-i})}^{-i} - s_y^{-i} + \mathbb{1}_{\mathcal{W}_P(s^{-i}) \rhd y} \leqslant 1$. Let PW_i^s denote the set of potential winners for voter i w.r.t. announced scores s.

We say that a candidate y is a potential winner if there exists a voter i such that $y \in PW_i^s$. By definition, the announced winner $\mathcal{W}_P(s)$ is also a potential winner. A voter is said to be *pivotal* if her considered set of potential winners contains more candidates than the winner exclusively. We consider the following best response for each voter i w.r.t. announced score s: i deviates from her current ballot b_i^T to another ballot supporting candidate $y \in PW_i^s \setminus \{\mathcal{W}_P(s)\}$ if y is her most preferred candidate within PW_i^s. Each voter can deviate at most once since she only gets the information about the scores s provided by the polling institute, and cannot see the deviations from other voters (thus the order of voters' deviations does not matter, they could even be simultaneous). Hence, the deviation process ends after at most n steps and converges to a final ballot profile denoted by b^s.

2.3 *A decision problem of manipulation*

In this model, a polling institute sends out a score and then each voter votes strategically w.r.t. that information, and finally the winner of the election is computed. We want to describe the behavior of the polling institute who may have its own interest in the election. Let x^* be the target candidate of the polling institute, i.e., it wants x^* to be elected. Let I be the space of all possible scores that the polling institute can announce, i.e., $I := \{s \in \mathbb{N}^m \mid \sum_{j=1}^{m} s_j = n\}$. We consider the following poll manipulation problem:

UNRESTRICTED MANIPULATION PROBLEM	
Instance:	Election $(N, M, \succ, \rhd)$, target candidate $x^* \in M$
Question:	Does there exist a score $s \in I$ to announce such that $\mathcal{W}_P(b^s) = x^*$?

However, the fact that the polling institute is allowed to send any score is not very realistic. We use a restricted version of the decision problem where the distance between the truthful poll and the one sent by the polling institute is bounded. We use the number of vote changes to evaluate the distance between possible scores, i.e., $d(s, s') = \frac{1}{2} \sum_{j \in M} |s_j - s'_j|$, for every scores $s, s' \in I$. Note that this distance is equivalent to the restriction of the ℓ_1 distance on I divided by 2 and sometimes called the "earth mover distance" in the literature [23]. We let $I_k := \{s \in \mathbb{N}^m \mid d(s, s^T) \leq k \text{ and } \sum_{j \in M} s_j = n\}$ be the restricted space of action of the polling institute. We then analyze the following poll manipulation problem:

RESTRICTED MANIPULATION PROBLEM	
Instance:	Election $(N, M, \succ, \rhd)$, target candidate $x^* \in M$, integer k
Question:	Does there exist a score $s \in I_k$ to announce such that $\mathcal{W}_P(b^s) = x^*$?

A poll manipulation is illustrated in the next example.

Example 1. *Let us consider an election* $(N, M, \succ, \rhd)$ *where* $N = \{1, \ldots, 8\}$, $M = \{a, b, c, d, e\}$, *the tie-breaking* $\rhd$ *follows the alphabetical order and the preferences* $\succ$ *are as follows:*

1:	$a \succ c \succ d \succ e \succ b$	*5, 6:*	$b \succ d \succ e \succ c \succ a$
2:	$e \succ b \succ d \succ c \succ a$	*7:*	$c \succ d \succ a \succ b \succ e$
3, 4:	$a \succ d \succ b \succ c \succ e$	*8:*	$d \succ e \succ a \succ b \succ c$

The initial truthful scores are given by $s^0 = (3, 2, 1, 1, 1)$ *(candidates are indexed w.r.t. their alphabetical position). Suppose that the polling institute communicates the following score vector* $s^M = (0, 2, 2, 3, 1)$*, at distance 3 to the truthful one. The set of potential winners w.r.t.* s^M *is equal to* $PW_i^{s^M} = \{b, c, d\}$ *for every voter* $i \in \{1, 2, 3, 4, 8\}$*, while* $PW_5^{s^M} = PW_6^{s^M} = \{c, d\}$ *and* $PW_7^{s^M} = \{b, d\}$*. Voters 3, 4, 5, 6, 7, and 8 do not have an incentive to deviate since the announced winner d is their most preferred candidate among the potential winners. However, voters 1 and 2 have an incentive to deviate to a ballot supporting c and b, respectively. After their deviations, we reach the final scores* $s^2 = (2, 3, 2, 1, 0)$ *where b is the winner. Hence, the polling institute can enforce the election of b, whereas a would remain the winner without poll manipulation.*

3 Preliminaries on Voting Cultures

In this section, we present voting cultures that are commonly used in the literature [31] to represent the distributions of preferences in elections. We will show that most of them satisfy a general condition on cultures, which is very useful for the purpose of our paper. Let Π^m be the set of all preference orders for m candidates, and $\succ_i \in \Pi^m$ be voter i's preference order for an arbitrary $i \in N$. We denote as $C(n, \Pi^m_{sub})$ the probability distribution of drawing n preference orders from $\Pi^m_{sub} \subseteq \Pi^m$ to constitute our preference profile. Such a probability distribution $C(n, \Pi^m_{sub})$ is called a *culture* and simply denoted by C when the context is clear, and its associated probability is denoted by $\mathbb{P}_C$. We use $\mathbb{P}_C(a \succ_i b)$ when it is clear from the context instead of $\mathbb{P}_C(\succ_i |\ a \succ_i b)$. In the following of the paper, we consider independent and identical drawings of voters' preferences such that we can either look at the distribution $C(n, \Pi^m_{sub})$ as a whole object or n drawings of preferences $\succ_i$. For technical reasons, we suppose that there are more than two different candidates which are ranked first by a preference order with a positive probability to be drawn under the considered culture. Note that this assumption is also natural since we focus on strategic voting and manipulation only occurs with at least three candidates [16, 29].

Definition 1 (Impartial culture)**.** *The impartial culture, called* IC*, draws every preference order* $\succ_i$ *from* Π^m *with uniform probability.*

One can also define variants of impartial cultures which are uniform but only on a given subset of Π^m, e.g., on single-peaked orders. A preference profile $\succ$ is said to be *single-peaked* [4] if there exists an axis $>$ on M such that, for every voter $i \in N$, and each triple of candidates $x > y > z$, we have $y \succ_i x$ or $y \succ_i z$. Let $\Pi^m_>$ be the set of single-peaked preference orders w.r.t. a given axis $>$ on M.

Definition 2 (Single-peaked culture)**.** *For a given axis* $>$ *over* M*, a culture* $C(n, \Pi^m_{sub})$ *is said to be single-peaked if* $C(n, \Pi^m_{sub}) = C(n, \Pi^m_>)$.

We might sample single-peaked preference orders by drawing them uniformly on the restricted space of single-peaked preference orders $\Pi^m_>$. The associated culture then refers to Walsh's model [32]. Another way to impartially sample single-peaked preference orders is to use Conitzer's model [7] which draws preference orders in $\Pi^m_>$ so that each candidate has the same chance to be ranked first.

Definition 3 (Mallows culture)**.** *For given* $\sigma \in \Pi^m$ *and* $\phi \in [0, 1]$*, the Mallows culture, called* $\mathcal{M}^{\phi,\sigma}$*, draws every preference order with a probability related to its distance to the reference ranking* σ*, more precisely,* $\mathbb{P}_{\mathcal{M}^{\phi,\sigma}}(\succ_i) = \frac{1}{Z}\phi^{dist_{KT}(\succ_i, \sigma)}$ *where* $Z = \sum_{\succ_i \in \Pi^m} \phi^{dist_{KT}(\succ_i, \sigma)}$.

Note that culture $\mathcal{M}^{1,\sigma}$ corresponds to the impartial culture.

We introduce below a simple property on cultures which will be key in the poll manipulation analysis.

Definition 4 (Balanced culture)**.** *A distribution* $C(n, \Pi^m)$ *is said to be balanced for a given candidate* $c \in M$ *if there exists another sufficiently worst candidate* $\ell \in M \setminus \{c\}$*, in the sense that* $\mathbb{P}_C(c \succ_i \ell) \geq \frac{1}{2}$*. The set of such candidates* ℓ *for* x *is denoted by* $B_C(x)$*. In general, a distribution* C *is said to be balanced if it is balanced for every candidate* $c \in M$.

It turns out that all cultures that we consider are balanced.

Proposition 1. *The impartial culture is balanced.*

Proof. Let $i \in N$ be a voter. The impartial culture is balanced for every candidate because for any pair of candidates x and y, we have $\mathbb{P}_{IC}(x \succ_i y) = \mathbb{P}_{IC}(y \succ_i x) = \frac{1}{2}$ since each preference order in Π^m has the same probability to be drawn. □

For a given axis $>$ on M, let $e_1^>$ and $e_2^>$ denote the two extreme candidates of $>$.

Proposition 2. *If* $x \in M \setminus \{e_1^>, e_2^>\}$*, then every single-peaked culture* $C(n, \Pi^m_>)$ *is balanced for* x*. If* $x \in \{e_1^>, e_2^>\}$*, then every single-peaked culture* $C(n, \Pi^m_>)$ *which also satisfies* $\mathbb{P}_C(\{\succ_i |worst_{\succ_i} = x\}) \leqslant \frac{1}{2}$*, is balanced for* x.

Sketch of proof. We mostly use the following fact: $\mathbb{P}_C(\{\succ_i \mid worst_{\succ_i} = e_1^>\} \cup \{\succ_i \mid worst_{\succ_i} = e_2^>\}) = 1$. □

In particular, the previous proposition shows that both Walsh's [32] and Conitzer's [7] cultures are balanced.

Proposition 3. *Any Mallows culture* $\mathcal{M}^{\phi,\sigma}$ *is balanced for every candidate* $x \in M \setminus \{worst_\sigma\}$.

Sketch of proof. Consider any candidate $x \in M \setminus \{worst_\sigma\}$ and the candidate $\ell := worst_\sigma$. Let $\Pi^m_{y \succ z}$ denote the set of all preference orders where y is ranked before z, i.e., $\Pi^m_{y \succ z} := \{\succ_i \in \Pi^m : y \succ_i z\}$. Consider the bijection $\tau : \Pi^m_{\ell \succ x} \to \Pi^m_{x \succ \ell}$, where for every $\succ_i \in \Pi^m_{\ell \succ x}$, we construct the preference order $\tau(\succ_i) \in \Pi^m_{x \succ \ell}$ which is the same as $\succ_i$ except that the positions of x and ℓ are swapped. One can show that $dist_{KT}(\sigma, \succ_i) \geq dist_{KT}(\sigma, \tau(\succ_i))$, for every $\succ_i \in \Pi^m_{\ell \succ x}$, by analyzing the differences between $\succ_i$ and $\tau(\succ_i)$ in terms of agreement on pairwise comparisons with σ. By definition of the Mallows culture $\mathcal{M}^{\phi,\sigma}$, we thus have $\mathbb{P}_{\mathcal{M}^{\phi,\sigma}}(\tau(\succ_i)) \geq \mathbb{P}_{\mathcal{M}^{\phi,\sigma}}(\succ_i)$. Hence, we conclude that $\mathbb{P}_{\mathcal{M}^{\phi,\sigma}}(\{\succ'_i : x \succ'_i \ell\}) \geq \mathbb{P}_{\mathcal{M}^{\phi,\sigma}}(\{\succ'_i : \ell \succ'_i x\})$, and thus $\mathbb{P}_{\mathcal{M}^{\phi,\sigma}}(\{\succ'_i : x \succ'_i \ell\}) \geq \frac{1}{2}$, implying that $\mathcal{M}^{\phi,\sigma}$ is balanced for x. □

4 The Unrestricted Poll Manipulation Problem

This section is devoted to the study of the unrestricted manipulation problem where the polling institute can send any score in I. We first give some results on the computational complexity of the problem then we continue our work with a probabilistic approach of the problem to capture what can happen in practice.

We first prove that, even in the unrestricted case, the poll manipulation problem is NP-complete. Our result answers an open question from Baumeister et al. [3].

Theorem 4. *The unrestricted manipulation problem is NP-complete.*

Sketch of proof. Membership to NP is straightforward: given communicated scores, we can efficiently derive the possible unique deviation of each voter and compute the winner in the deviating profile.

For hardness, we perform a reduction from a variant of EXACT COVER BY 3-SETS (X3C) known to be NP-complete [17]: Given a set $X = \{x_1, x_2, \ldots, x_{3q}\}$ and a set $S = \{S_1, S_2, \ldots, S_{3q}\}$ of 3-element subsets of X, where each element x_i occurs in exactly three subsets of S, we ask whether there exists an exact cover, i.e., a subset $S' \subseteq S$ such that every element of X occurs in exactly one member of S'. From an instance (X, S) of X3C, we construct an instance of our unrestricted manipulation problem as follows.

For each element x_i, for $i \in [3q]$, we create a candidate y_i, and for each subset S_j where $j \in [3q]$, we create a candidate c_j. We add three candidates w, z, and t where t is our target candidate.

There are $12q + 7$ voters: for each element x_i, for $i \in [3q]$, we create one voter Y_i, for each subset S_j, for $j \in [3q]$, we create three voters C_j^r where $r \in [3]$, and we finally add two voters T^ℓ, two voters Z^ℓ, two voters W^ℓ for $\ell \in [2]$, and one voter D.

Their preferences are defined below, where $y(s_j^r)$ denotes the candidate y_i associated with the r^{th} element of subset S_j, and when a subset of candidates is mentioned, the candidates are ranked according to the increasing order of their indices.

Y_i:	$w \succ y_i \succ z \succ \{y_{i'}\}_{i' \neq i} \succ \{c_j\}_j \succ t$	for $i \in [3q]$
C_j^r:	$y(s_j^r) \succ c_j \succ z \succ w \succ \{y_{i'}\}_{i' \neq i} \succ \{c_j\}_j \succ t$	for $j \in [3q], r \in [3]$
T^ℓ:	$t \succ z \succ w \succ \{y_i\}_i \succ \{c_j\}_j$	for $\ell \in [2]$
Z^ℓ:	$z \succ w \succ \{y_i\}_i \succ \{c_j\}_j \succ t$	for $\ell \in [2]$
W^ℓ:	$w \succ z \succ \{y_i\}_i \succ \{c_j\}_j \succ t$	for $\ell \in [2]$
D:	$w \succ t \succ z \succ \{y_i\}_i \succ \{c_j\}_j$	

Finally, the tie-breaking rule is as follows: $w \rhd t \rhd z \rhd y_1 \rhd \cdots \rhd y_{3q} \rhd c_1 \rhd \cdots \rhd c_{3q}$.

The winner of the election with the truthful ballot profile is candidate w. The details of the scores for this truthful ballot profile are given in the second column of Table 1.

Table 1. Candidates' scores in the complexity proof of Theorem 4

candidate	initial score	announced score	score after manipulation
y_i ($i \in [3q]$)	3	3	3
c_j ($j \in [3q]$)	0	3 if $S_j \in S'$ 0 otherwise	3 if $S_j \in S'$ 0 otherwise
w	$3q + 3$	2	2
t	2	2	3
z	2	3	2
winner	w	z	t

We claim that there exists an exact cover in (X, S) iff we can force the election of candidate t in the constructed instance.

Suppose first that there exists a subset $S' \subseteq S$ such that every element of X occurs in exactly one element of S'. Let us consider manipulated communicated scores which differ from the sincere ones by taking $3q + 1$ votes initially given to w to give one additional vote to z and three votes to c_j for each $S_j \in S'$. These scores are summarized in the third column of Table 1. One can prove that these communicated scores trigger deviations which lead to the final scores presented in the fourth column of Table 1, where t is the winner.

Suppose now that there exist communicated scores such that the target candidate t becomes the winner after deviations from the voters. One can show that the only possibility for communicated scores to lead to the victory of t is to announce candidate z the winner and, as potential winners, the target candidate t and exactly q candidates c_j which correspond to subsets S_j forming an exact cover of X. □

Note that even though we have proved that the problem is NP-complete, we know from Baumeister et al. [3] that it is FPT when parameterized by the number of candidates m. Another way to go beyond the NP-hardness result, which focuses on worst-case complexity, is to analyze the actual possibility of poll manipulation using a probabilistic approach which works even when m is large. We will see that the poll manipulation problem is often easy to tackle in a probabilistic point of view, following natural statistical cultures as defined in Section 3. We will start by considering a balanced culture.

For a given target candidate x^* the polling institute wants to elect, we say its poll manipulation is successful if after all strategic moves from voters, the desired candidate x^* is elected. Let us denote by S the associated event of success, which corresponds to the yes-instances of the unrestricted poll manipulation problem.

Let 2PW-H(x^*, ℓ) be the heuristic which announces a score with exactly two potential winners x^* and ℓ, with x^* the target candidate and ℓ the announced winner. For realistic conditions, one point is given to candidates with a positive score in the truthful ballot profile. Assuming $n > m+5$ is sufficient to guarantee the possibility of making any pair of candidates the only potential winners (this hypothesis is rather weak since we focus on large elections in terms of voters). It then suffices to check whether the associated communicated polling score leads to the victory of x^*. This heuristic can be called by a global heuristic, which tests it with different candidates ℓ. Our heuristics are computable in polynomial time and are inspired from the heuristics of Wilczynski [33] and Baumeister et al. [3], where the idea is to find a candidate ℓ, which is a threatening winner, i.e., enough voters prefer x^* to ℓ, while x^* is the only credible alternative to ℓ, in order to incentivize voters to deviate to x^*.

Let $S_{\text{2PW-H}(x^*,\, \ell)}$ denote the event of success for heuristic 2PW-H(x^*, ℓ). Let $X_{-\ell}$ be the random variable which counts the number of voters who prefer x^* over ℓ, i.e., $X_{-\ell} = |N^{x^* \succ \ell}|$.

Similarly, let $Y_{\ell,j}$ be the random variable which counts the number of voters who prefer ℓ over x^* while their most preferred candidate is j, i.e., $Y_{\ell,j} = |\{i \in N^{\ell \succ x^*} : top_{\succ_i} = j\}|$. If our heuristic 2PW-H(x^*, ℓ) indeed succeeds to announce exactly two potential winners x^* and ℓ with ℓ as a winner, then only voters who prefer x^* over ℓ and currently vote for another candidate, will deviate and they will do so in favor of x^*. Note that voters already having x^* as their top choice would keep this vote because there is no other potential winner. Therefore, in total, after deviations, x^* obtains a number of votes which is equal to the numbers of voters who prefer x^* over ℓ. It follows that x^* would win only if the number of voters preferring x^* over ℓ is greater than the number of voters who keep their vote for another candidate, implying that for a given culture $C(n, \Pi^m)$, $\mathbb{P}_C(S_{\text{2PW-H}(x^*,\, \ell)}) = \mathbb{P}_C(\forall j \in M \setminus \{x^*\}, Y_{\ell,j} \leqslant X_{-\ell})$.

Our first theorem provides a high lower bound on the probability of success of the poll manipulation heuristic.

Theorem 5. *For a balanced culture $C(n, \Pi^m)$, there exists $\ell \in B_C(x^*)$ such that the probability of success of the sub-heuristic 2PW-H(x^*, ℓ), is as follows: $\mathbb{P}_C(S_{\text{2PW-H}(x^*,\, \ell)}) \geqslant 1 - 2(m-2)(e^{-2n(p_{x^*,\ell} - q_{x^*,\ell,j^*})^2})$ where:*

- $p_{x^*,\ell} := \mathbb{P}_C(x^* \succ_i \ell)$,
- $r_{x^*,\ell,j} := \mathbb{P}_C(\{\ell \succ_i x^*\} \cap \{j = top_{\succ_i}\})$, *for* $j \neq \ell$,
- $q_{x^*,\ell,j} := \frac{p_{x^*,\ell} + r_{x^*,\ell,j}}{2}$,
- $j^* := \arg\min_{j \in M \setminus \{x^*,\ell\}} \mathbb{P}_C(Y_{\ell,j} \leqslant X_{-\ell})$.

In particular, the probability of success of the global heuristic satisfies the same lower bound.

Sketch of proof. For our target candidate x^* and a balanced culture $C(n, \Pi^m)$, let us consider a candidate $\ell \in B_C(x^*)$. Since $\ell \in B_C(x^*)$, we always have $Y_{\ell,\ell} \leqslant X_{-\ell}$ and thus $\mathbb{P}_C(S_{\text{2PW-H}(x^*,\,\ell)}) = \mathbb{P}_C(\forall j \in M \setminus \{x^*, \ell\}, Y_{\ell,j} \leqslant X_{-\ell})$. We will show a lower bound to this latter quantity. Using Bonferroni's inequality, we have: $\mathbb{P}_C(\forall j \in M \setminus \{x^*, \ell\}, Y_{\ell,j} \leqslant X_{-\ell}) \geqslant \sum_{j \in M\setminus\{x^*,\ell\}} \mathbb{P}_C(Y_{\ell,j} \leqslant X_{-\ell}) - (m-3) \geqslant (m-2) \cdot \min_{j\in M\setminus\{x^*,\ell\}} \mathbb{P}_C(Y_{\ell,j} \leqslant X_{-\ell}) - (m-3)$. By considering $j^* := \arg\min_{j\in M\setminus\{x^*,\ell\}} \mathbb{P}_C(Y_{\ell,j} \leqslant X_{-\ell})$, we then have that $\mathbb{P}_C(\forall j \in M \setminus \{x^*, \ell\}, Y_{\ell,j} \leqslant X_{-\ell}) \geqslant (m-2) \cdot (\mathbb{P}_C(Y_{\ell,j^*} \leqslant X_{-\ell}) - 1) + 1$. Let us now treat the term $\mathbb{P}_C(Y_{\ell,j^*} \leqslant X_{-\ell})$. We remark that $X_{-\ell}$ follows a binomial distribution of parameters n and $p_{x^*,\ell}$, and Y_{ℓ,j^*} follows a binomial distribution of parameters n and r_{x^*,ℓ,j^*}, where $p_{x^*,\ell}$ and $r_{x^*,\ell,j}$ are defined as $p_{x^*,\ell} = \mathbb{P}_C(x^* \succ_i \ell)$ and $r_{x^*,\ell,j} = \mathbb{P}_C(\{\ell \succ_i x^*\} \cap \{j = top_{\succ_i}\})$, for every $j \neq \ell$. We introduce $q_{x^*,\ell,j^*} := \frac{p_{x^*,\ell} + r_{x^*,\ell,j^*}}{2}$ to lower bound our probability as follows: $\mathbb{P}_C(Y_{\ell,j^*} \leqslant X_{-\ell}) \geqslant \mathbb{P}_C(\{Y_{\ell,j^*} < q_{x^*,\ell,j^*} \cdot n\} \bigcap \{X_{-\ell} > q_{x^*,\ell,j^*} \cdot n\})$. We use again Bonferroni's inequality and an inclusion of events to get: $\mathbb{P}_C(Y_{\ell,j^*} \leqslant X_{-\ell}) \geqslant \mathbb{P}_C(\{Y_{\ell,j^*} < q_{x^*,\ell,j^*} \cdot n\}) + \mathbb{P}_C(\{X_{-\ell} > q_{x^*,\ell,j^*} \cdot n\}) - 1$. We then use Hoeffding's inequality to find these lower bounds: $\mathbb{P}_C(Y_{\ell,j^*} < q_{x^*,\ell,j^*} \cdot n) \geqslant 1 - e^{-2n(q_{x^*,\ell,j^*} - r_{x^*,\ell,j^*})^2}$ and $\mathbb{P}_C(X_{-\ell} > q_{x^*,\ell,j^*} \cdot n) \geqslant 1 - e^{-2n(p_{x^*,\ell} - q_{x^*,\ell,j^*})^2}$. Putting the last two inequalities together we get: $\mathbb{P}_C(Y_{\ell,j^*} \leqslant X_{-\ell}) \geqslant 1 - e^{-2n(p_{x^*,\ell} - q_{x^*,\ell,j^*})^2} - e^{-2n(q_{x^*,\ell,j^*} - r_{x^*,\ell,j^*})^2}$. Coming back to the first work of the proof we have: $\mathbb{P}_C(S_{\text{2PW-H}(x^*,\,\ell)}) \geqslant 1 - (m-2)(e^{-2n(p_{x^*,\ell} - q_{x^*,\ell,j^*})^2} + e^{-2n(q_{x^*,\ell,j^*} - r_{x^*,\ell,j^*})^2})$. Finally, since q_{x^*,ℓ,j^*} is defined as the middle between $p_{x^*,\ell}$ and r_{x^*,ℓ,j^*}, we can simplify the inequality: $\mathbb{P}_C(S_{\text{2PW-H}(x^*,\,\ell)}) \geqslant 1 - 2(m-2)e^{-2n(p_{x^*,\ell} - q_{x^*,\ell,j^*})^2}$. □

We can thus deduce the same lower bound for the probability of existence of a successful unrestricted poll manipulation.

Corollary 6. *For a balanced culture $C(n, \Pi^m)$, the probability of success of an unrestricted poll manipulation is as follows: $\mathbb{P}_C(S) \geqslant 1 - 2(m-2)(e^{-2n(p_{x^*,\ell} - q_{x^*,\ell,j^*})^2})$.*

Our next theorem considers the asymptotic case and shows the convergence of the lower bound probability toward 1 when n becomes large. Since the number of voters is typically large in political elections, this shows an important susceptibility to poll manipulation.

Theorem 7. *For a balanced culture $C(n, \Pi^m)$, there exists $\ell \in B_C(x^*)$ such that the probability of success of the sub-heuristic 2PW-H(x^*, ℓ), and thus of the Global Heuristic, tends toward 1, i.e., $\lim_{n\to\infty} \mathbb{P}_C(S_{\text{2PW-H}(x^*,\,\ell)}) = 1$ and thus $\lim_{n\to\infty} \mathbb{P}_C(S) = 1$.*

Proof. We use the lower bound from Theorem 5 to deduce the convergence toward 1 of this probability. In fact it is enough to pass to the limit on both sides in n the number of voters. The only tricky point might be when $p_{x^*,\ell} = q_{x^*,\ell,j^*}$. However, this situation can happen only when the culture puts positive probability only on preference orders whose top can only be x^* or j and in an equal manner, which is not possible by natural assumption on the culture. □

Observe that the quantities $p_{x^*,\ell}$ and $q_{x^*,\ell,j}$ from Theorem 5 are constants and different, we thus have exponentially fast convergence toward 1 for the probability of success of 2PW-H(x^*, ℓ) w.r.t. the number of voters. To give a quick intuition, observe that for $m = 5$ and $n = 50$, we get a lower bound of 0.82 and for $m = 5$ and $n = 100$, we already have a lower bound of 0.99 which is very fast!

Beyond this general result on balanced cultures, the goal would be to capture realistic cultures regarding real elections [31]. From Propositions 1–3, we can derive the following corollary which shows that our general result covers very natural concrete cultures.

Corollary 8. *For a culture $C(n, \Pi^m)$, there exists $\ell \in B_C(x^*)$ such that the probability of success of 2PW-H(x^*, ℓ), and thus of the Global Heuristic, tends toward 1, i.e., $\lim_{n\to\infty} \mathbb{P}_C(S_{\text{2PW-H}(x^*,\,\ell)}) = 1$ and thus $\lim_{n\to\infty} \mathbb{P}_C(S) = 1$, when:*

- *C corresponds to the impartial culture, or*
- *C is a single-peaked culture and x^* is not an extreme candidate or x^* is extreme but $\mathbb{P}_C(\succ_i \mid worst_{\succ_i} = x^*) \leqslant \frac{1}{2}$, which includes Walsh's and Conitzer's cultures, or*
- *C corresponds to a Mallows culture $\mathcal{M}^{\phi,\sigma}$ where $x^* \neq worst_\sigma$.*

Our results show that even if the poll manipulation problem is hard, it is very likely for the polling institute to efficiently and successfully control the election, under natural preference distributions. However, the hypothesis that allows to send any score is questionable since the polling institute might be forced to meet some legal quality standards or to maintain voter trust by sending a reasonable score.

5 The Restricted Poll Manipulation Problem

This section is devoted to the study of the manipulation problem in its restricted version i.e., the polling institute is restricted in its ability to lie about the scores and can only send a score vector from I_k.

The restricted poll manipulation problem is known to be NP-hard [3]. However, one could hope to get a fixed-parameter tractable algorithm w.r.t. the maximum allowed distance k to the truthful scores. We show below that such an efficient algorithm is unlikely to exist since we prove that the problem is W[1]-hard.

Theorem 9. *The restricted manipulation problem is W[1]-hard.*

Sketch of proof. From an instance $(G = (V, E), k)$ of k-Clique, known to be W[1]-complete [10], where $n := |V|$ and, w.l.o.g., $2 < k < n - 1$, we construct an instance of the restricted manipulation problem as follows.

For each vertex $v_i \in V$, we create a candidate v_i, and for each edge $\{v_i, v_j\} \in E$, we create a candidate e_{ij} (we suppose $i < j$ for this notation). We add three other candidates w, t, and z.

Let $K := (n-k)k$. For each vertex $v_i \in V$, we create k voters U_i^ℓ for $\ell \in [k]$, and $K - 1 - \delta(v_i)$ voters D_i^ℓ for $\ell \in [K - 1 - \delta(v_i)]$ (by our assumption on k, this quantity cannot be negative), where $\delta(v_i)$ denotes the degree of vertex v_i in G.

For each edge $\{v_i, v_j\} \in E$, we create two voters F_{ij}^i and F_{ij}^j, and $K - 2$ voters E_{ij}^ℓ for $\ell \in [K - 2]$. Finally, we add K voters T^ℓ for $\ell \in [K]$ and $K - 1$ voters Z^ℓ for $\ell \in [K - 1]$.

The preferences of the voters over the candidates are described below, for each $i \in [n]$, and each $\{v_p, v_q\} \in E$:

U_i^ℓ:	$w \succ v_i \succ z \succ \{v_j\}_{j\neq i} \succ \{e_{r,s}\}_{\{r,s\}} \succ t$	for $\ell \in [k]$
F_{pq}^p:	$v_p \succ e_{pq} \succ z \succ w \succ \{v_j\}_{j\neq p} \succ \{e_{r,s}\}_{\{r,s\}\neq\{p,q\}} \succ t$	
F_{pq}^q:	$v_q \succ e_{pq} \succ z \succ w \succ \{v_j\}_{j\neq q} \succ \{e_{r,s}\}_{\{r,s\}\neq\{p,q\}} \succ t$	
D_i^ℓ:	$v_i \succ z \succ w \succ \{v_j\}_{j\neq i} \succ \{e_{r,s}\}_{\{r,s\}} \succ t$	for $\ell \in [K - 1 - \delta(v_i)]$
T^ℓ:	$t \succ z \succ w \succ \{v_j\}_j \succ \{e_{r,s}\}_{\{r,s\}}$	for $\ell \in [K]$
E_{pq}^ℓ:	$e_{pq} \succ z \succ w \succ \{v_j\}_j \succ \{e_{r,s}\}_{\{r,s\}\neq\{p,q\}} \succ t$	for $\ell \in [K - 2]$
Z^ℓ:	$z \succ w \succ \{v_j\}_j \succ \{e_{r,s}\}_{\{r,s\}} \succ t$	for $\ell \in [K - 1]$

Finally, the tie-breaking rule is as follows: $z \triangleright t \triangleright \cdots \triangleright w$.

The winner of the election with the truthful ballot profile is candidate w. The details of the scores for this truthful ballot profile are given in the second column of Table 2.

Table 2. Candidates' scores in the complexity proof of Theorem 9

candidate	initial score	announced score	score after manipulation
v_i $(i \in [n])$	$K-1$	K if $v_i \in S$ $K-1$ otherwise	K if $v_i \in S$ $K-1$ otherwise
e_{ij} $(\{v_i, v_j\} \in E)$	$K-2$	K if $v_p, v_q \in S$ $K-2$ otherwise	K if $v_p, v_q \in S$ $K-2$ otherwise
w	kn	K	K
t	K	$K-1$	K
z	$K-1$	K	$K-1$
winner	w	z	t

We claim that G admits a clique of size k iff we can force the election of candidate t by announcing scores which differ from the truthful ones by at most k^2+1 vote changes.

Suppose first that there exists a subset of vertices $S \subseteq V$ such that S is a k-Clique of G, i.e., $|S| = k$ and $\{v_i, v_j\} \in E$ for every $v_i, v_j \in S$. Consider manipulated communicated scores which differ from the sincere ones by taking k^2 votes initially given to w in order to give one additional vote to each $v_i \in S$ (there are k such candidates) and two additional votes to each e_{ij} such that $v_i, v_j \in S$ (there are $\frac{k(k-1)}{2}$ such candidates), and finally by taking one vote initially given to t in order to give it to z. In total, these scores differ from the sincere ones by exactly k^2+1 vote changes, they are summarized in the third column of Table 2. One can prove that these communicated scores trigger deviations which lead to the final scores presented in the fourth column of Table 2, where t is the winner.

Suppose now that there exist communicated scores such that candidate t becomes the winner after all voters' deviations. One can show that the only possibility for communicated scores to lead to the victory of the target candidate t is to announce candidate z the winner and, as potential winners, k candidates v_i, as well as $k(k-1)/2$ candidates e_{pq}, such that for each potential winner v_i, there are $k-1$ potential winners e_{ij} (or e_{ji}) corresponding to edges incident to v_i. □

Nevertheless, we prove below that the restricted poll manipulation problem can be efficiently solved if the parameter k of the maximum distance to the truthful scores is a constant.

Proposition 10. *The restricted manipulation problem is in XP w.r.t. the maximum distance k to the truthful scores. More precisely, it can be solved by an algorithm which runs in time $\Theta(m^{2k+1} \cdot n)$.*

Sketch of proof. We provide an upper bound of m^{2k} to $|I_k|$. It is then enough to visit every score vector of I_k to check whether it leads to the victory of the target candidate. □

However, the previous result cannot be used if k is large and does not tell whether there actually exists a successful manipulation. We thus use a probabilistic approach to analyze the possibility of poll manipulation. Let S_k denote the event of success for the restricted poll manipulation where k denotes the maximum allowed distance to the truthful scores. We first prove that when k is small compared to $\sqrt{n}$, the restricted poll manipulation tends to be impossible.

Theorem 11. *For any culture $C(n, \Pi^m)$, if the maximum distance k to the truthful scores is such that $k = o(\sqrt{n})$ and the target candidate x^* is not winning in the initial score, then the probability of existence of a successful poll manipulation to elect x^* tends toward zero, i.e.,* $\lim_{n\to\infty} \mathbb{P}_C(S_k) = 0$.

Sketch of proof. We first identify the law of the truthful scores.

Observation 12. *For a culture $C(n, \Pi^m)$, the truthful scores s^T follow a multinomial law Multi(q, n) where $q = (q_1, \dots, q_m)$ and $q_j := \mathbb{P}_C(\{\mathcal{W}_P(s^T) = j\})$, for every $j \in M$.*

The truthful scores follow a multinomial law because there are n voters' preferences drawn independently at random with the same law, and we have m possibilities for the most preferred candidate of each voter, and these are the only necessary elements to compute scores s^T. Let c^* be the truthful winner, i.e., $c^* := \mathcal{W}_P(b^T)$. Informally, a necessary condition for the existence of a successful manipulation with the two-candidate heuristic is that there is at least one candidate that is sufficiently close to the winner. The pair of candidates would then be this candidate and the current winner. Of course, this is not necessarily sufficient, as the pair may not be the right one. However, we will see that this necessary condition occurs with probability 0, and that is enough for us to conclude. We then write $S_k \subset \{\bigcup_{z \neq c^*} \{|s^T_{c^*} - s^T_z + \mathbb{1}_{c^* \rhd z}| \leq k\}\}$.

We will analyze the probability of the second event to get an upper bound on the probability of success of the restricted poll manipulation problem. By using Observation 12 and a central limit theorem for multinomial law s^T, we get: $\frac{1}{\sqrt{n}}(s^T - nq) \underset{n\to+\infty}{\longrightarrow} \mathcal{N}(0; K)$, where $K_{i,j} = q_i\delta_{i,j} - q_i q_j$, for every $1 \leq i, j \leq m$, with $\delta_{i,j} = 1$ if $i = j$ and $\delta_{i,j} = 0$ otherwise. We denote $\mathcal{N}(0; K) = (\mathcal{N}_1, \dots, \mathcal{N}_m)$ and remark that each $\mathcal{N}_j$ follows a Gaussian law. Using the previous point we show that for any $z \in M \setminus \{x^*\}$, we have: $\lim_{n\to+\infty} \mathbb{P}_C(|s^T_{c^*} - s^T_z + \mathbb{1}_{c^* \rhd z}| \leq k) = 0$. It follows for the probability of the success event that $\lim_{n\to+\infty} \mathbb{P}_C(S_k) \leqslant \lim_{n\to+\infty} \mathbb{P}_C(\bigcup_{z \neq c^*} |s^T_{c^*} - s^T_z + \mathbb{1}_{c^* \rhd z}| \leq k) \leq \lim_{n\to+\infty} \sum_{z \neq c^*} \mathbb{P}_C(|s^T_{c^*} - s^T_z + \mathbb{1}_{c^* \rhd z}| \leq k) = 0$. We then get: $\lim_{n\to+\infty} \mathbb{P}_C(S_k) = 0$, which concludes the proof. □

Then, we get immediately the following corollary if we include the case where x^* might win in the initial poll, because it is always possible to communicate scores that keep the same winner.

Corollary 13. *For any culture $C(n, \Pi^m)$, if the maximum distance k to the truthful scores is such that $k = o(\sqrt{n})$, then* $\lim_{n\to\infty} \mathbb{P}_C(S_k) = \mathbb{P}_C(\{\mathcal{W}_P(s^T) = x^*\})$.

We might note that, e.g., $\mathbb{P}_C(\{\mathcal{W}_P(s^T) = x^*\}) \approx \frac{1}{m}$ when considering the impartial culture.

We now focus on a case where poll manipulation can be successful, and prove that we can even efficiently compute it, thanks to an adaptation of the global heuristic where the sub-heuristic to call is Restricted 2PW-H(x^*, ℓ) which, starting from s^T, tries to announce ℓ as the winner and x^* as the only other potential winner, while taking into account the maximum allowed distance k. Let $S_{\text{Restr-2PW-H}(x^*, \ell)}$ denote the event of success of this sub-heuristic.

Theorem 14. *For a balanced culture C, if the maximum distance k to the truthful scores is such that $n = o(k)$ where and $c^* := \mathcal{W}_P(b^T)$, then there exists $\ell \in B_C(x^*)$ such that the probability of success of Restricted 2PW-H(x^*, ℓ) tends toward 1, i.e.,* $\lim_{n\to\infty} \mathbb{P}_C(S_{\text{Restr-2PW-H}(x^*, \ell)}) = 1$ *and thus* $\lim_{n\to\infty} \mathbb{P}_C(S_k) = 1$.

Proof. We can first observe that each score the polling institute may send can be summarized by its set of potential winners and its winner, since two announced scores with the same potential winners and winner produce the same voters' deviations. A type $\mathcal{T}(s)$ for a score vector s is thus defined as a pair $(PW, w) \in 2^M \times M$ where $w \in PW$, representing its potential winners and its winner. The set of all possible score types is denoted by $\mathcal{T}$. We will then show that: $\mathbb{P}_C(\{\bigcup_{s \in I_k} \mathcal{T}(s) = \mathcal{T}\}) = 1$. Let c^* be the truthful winner, i.e., $c^* := \mathcal{W}_P(b^T)$. Informally, a sufficient condition for the existence of a strategy of each type is that all candidates are sufficiently close to the winner. More precisely, we would like them all to be closer than $\frac{k}{m}$, so that the cost of making potential winners any pair of candidates never exceeds k. We then get: $\{\bigcap_{z\neq c^*}\{|s_{c^*}^T - s_z^T + \mathbb{1}_{c^*\triangleright z}| < \frac{k}{m}\}\} \subset \{\bigcup_{s\in I_k} \mathcal{T}(s) = \mathcal{T}\}$. We again use the same technique adding and subtracting $n \cdot q_{c^*}$ and $n \cdot q_z$ and a central limit theorem on the truthful scores s^T following a multinomial law (Observation 12). However, we have this time a remaining term $\sqrt{n}(q_{c^*} - q_z)$ that is bounded by assumption ($n = o(k)$). We then get: $\lim_{n\to+\infty} \mathbb{P}_C(|s_{c^*}^T - s_z^T + \mathbb{1}_{c^*\triangleright z}| < \frac{k}{m}) = 1$. Since a countable intersection of events of probability 1 is of probability 1, we have: $\lim_{n\to+\infty} \mathbb{P}_C(\bigcap_{z\neq c^*}\{|s_{c^*}^T - s_z^T + \mathbb{1}_{c^*\triangleright z}| < \frac{k}{m}\}) = 1$. Then, we have: $\mathbb{P}_C(\bigcup_{s\in I_k} \mathcal{T}(s) = \mathcal{T}) = 1$. Using the fact that a successful strategy exists in the unrestricted case and since all strategies are accessible, we get: $\lim_{n\to\infty} \mathbb{P}_C(S_{\text{Restr-2PW-H}(x^*,\, \ell)}) = 1$. Therefore, we have also: $\lim_{n\to\infty} \mathbb{P}_C(S_k) = 1$. □

The case $k = \alpha \cdot n$ with $\alpha \in]0, 1]$ is included in Theorem 14. This has a clear interpretation: if the polling institute is allowed to lie by a fraction α on scores then we will fall in the manipulation regime for a sufficiently large number of voters.

Like for the unrestricted problem, the general result of Theorem 14 holds for the concrete cultures mentioned in Section 3.

Corollary 15. *For a culture C and $n = o(k)$ and $c^* := \mathcal{W}_P(b^T)$, there exists $\ell \in B_C(x^*)$ such that the probability of success of Restricted 2PW-H(x^*, ℓ) tends toward 1, i.e., $\lim_{n\to\infty} \mathbb{P}_C(S_{\textit{Restr-2PW-H}(x^*,\, \ell)}) = 1$ and thus $\lim_{n\to\infty} \mathbb{P}_C(S_k) = 1$, when:*

- *C corresponds to the impartial culture, or*
- *C is a single-peaked culture and x^* is not an extreme candidate or x^* is extreme but $\mathbb{P}_C(\succ_i | \; worst_{\succ_i} = x^*) \leqslant \frac{1}{2}$, which includes Walsh's and Conitzer's cultures, or*
- *C corresponds to a Mallows culture $\mathcal{M}^{\phi,\sigma}$ where $x^* \neq worst_\sigma$.*

6 Toward a Generalization of Strategic Behavior

Until now, we only considered strategic moves from pivotal voters. However, one can argue that voters might want to deviate when they are close enough to be pivotal. Such a strategic behavior can be captured by considering pivotal thresholds $p_i \in \mathbb{N}$ for each voter i, as done by Wilczynski [33] in an idea close to local-dominance [25]. This slightly modifies the definition of potential winners: A candidate y is a *general potential winner* for voter i w.r.t. score s if i believes that adding p_i votes to y will make candidate y the new winner, i.e., $s^{-i}_{\mathcal{W}_P(s^{-i})} - s^{-i}_y + \mathbb{1}_{\mathcal{W}_P(s^{-i})\triangleright y} \leqslant p_i$. We denote PW_i^{s,p_i} the set of general potential winners for i w.r.t. score s. The definition of best response naturally follows by considering general potential winners. Our initial setting corresponds to the case where $p_i = 1$.

Let us first analyze the impact of pivotal thresholds on strategic voting. For this purpose, we suppose that the polling institute is sincere and sends truthful scores $s = s^T$, and that all thresholds are equal and denoted by p, i.e., $p_i = p$, for every voter i. Let us define the expected proportion of strategic voters $\mathcal{P}_{SV}$ w.r.t. culture $C(n, \Pi^m)$, n, m, and p. Let U_i^p denote the event where the top candidate of voter i is not a general potential winner for i, i.e., $U_i^p = \{top_{\succ_i} \notin PW_i^{s,p}\}$, and D_i^p the event where voter i could favor a potential winner other than the current winner, that she prefers to it, i.e., $D_i^p = \{\exists w \in M \setminus \{top_{\succ_i}\} : w \succ_i \mathcal{W}_P(s)$ and $w \in PW_i^{s,p}\}$. By definition, the proportion of strategic voters counts the voters for who the two events are true, i.e., $\mathcal{P}_{SV}(C, n, m, p) = \mathbb{E}_C[\frac{1}{n}\sum_{i=1}^n \mathbb{1}_{U_i^p \cap D_i^p}]$.

The following proposition provides several insights on the proportion of strategic voters at the limits, by showing that the variations of the dependent events U_i and D_i are opposed with respect to p.

Proposition 16. *1. $U = \mathbb{E}_C[\frac{1}{n}\sum_{i=1}^n \mathbb{1}_{U_i^p}]$ is decreasing w.r.t. p.*
2. $D = \mathbb{E}_C[\frac{1}{n}\sum_{i=1}^n \mathbb{1}_{D_i^p}]$ is increasing w.r.t. p.
3. $\mathcal{P}_{SV}(C, n, m, p) \leqslant \min(U, D)$.
4. $\lim_{p\to+\infty} \mathcal{P}_{SV}(C, n, m, p) = 0$ and $\mathcal{P}_{SV}(C, n, m, 0) = 0$.
5. $\lim_{n\to+\infty} \mathcal{P}_{SV}(C, n, m, p) = 0$ when p is fixed.

Although the previous proposition helps to better understand the proportion of strategic voters at the limits, it is still difficult to exactly determine the behavior for other values of p, in particular when $\mathcal{P}_{SV}$ is maximum, because of the dependency between U and D.

Let us now analyze the poll manipulation problems. Let S^G (resp., S_k^G) denote the associated event of success for the unrestricted (resp., restricted) problem with generalized strategic behavior.

Proposition 17. *For a balanced culture $C(n, \Pi^m)$, $p > 0$ and $p = o(n)$, we have $\mathbb{P}_C(S^G) \geqslant \mathbb{P}_C(S)$ and $\mathbb{P}_C(S_k^G) \geqslant \mathbb{P}_C(S_k)$.*

It follows from Proposition 17 that Theorems 5 and 7, for successful unrestricted poll manipulation, also hold under a generalized strategic behavior with the given weak hypotheses. Similarly, the convergence result toward 1 for the probability of a successful restricted poll manipulation (Theorem 14) and the generalization of Theorem 11 can also be extended to a generalized strategic behavior.

7 Conclusion

In the context of political elections where voters are assumed to be strategic, we have studied the poll manipulation problem: *Can a polling institute lie about candidates' scores it communicates to voters in order to influence the outcome of the election?* Two variants are investigated: an unrestricted one where any scores can be sent, and a restricted one, more realistic, where the polling institute cannot announce scores too far from the reality. We show that both problems are computationally hard and answer an open question from Baumeister et al. [3]. However, we go beyond this worst-case analysis by using probabilistic tools to balance computational hardness. Under a broad condition on cultures, satisfied by many concrete preference distributions, we prove a lower bound on the probability of success of an easily computable heuristic for the unrestricted problem. This enables us to obtain a rapid convergence toward 1 of the manipulation probability, meaning that large elections are highly manipulable when the polling institute can freely manipulate without altering the trust of voters. When it may not be the case, i.e., in a restricted context, our asymptotic results show that manipulability strongly relies on whether the allowed distance to truthful scores depends on the election size. Manipulation tends to fail when this distance is negligible w.r.t. the number of voters. However, when the distance is significant, e.g., is a given proportion of the election size, which appears as a very natural assumption, efficient and successful manipulation tends to be always possible, showing that political elections are highly susceptible to poll manipulation in practice.

Future work should be devoted to studying the generalization of these results in different directions. Considering other voting rules or other types of information communicated in the poll would be natural. Another avenue of work could be to examine different strategic voting behaviors, such as local dominance [23]. Finally, a challenging future direction would be to adapt our analysis to dependent cultures such as the Pólya-Eggenberger urn [31].

References

[1] A. Anand and P. Dey. Distance restricted manipulation in voting. *Theoretical Computer Science*, 891:149–165, 2021.

[2] J. J. Bartholdi, C. A. Tovey, and M. A. Trick. The computational difficulty of manipulating an election. *Social choice and welfare*, 6:227–241, 1989.

[3] D. Baumeister, A.-K. Selker, and A. Wilczynski. Manipulation of opinion polls to influence iterative elections. In *Proceeding of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020)*, pages 132–140, 2020.

[4] D. Black. On the rationale of group decision-making. *Journal of Political Economy*, 56(1):23–34, 1948.

[5] C. Bonferroni. Teoria statistica delle classi e calcolo delle probabilita. *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commericiali di Firenze*, 8:3–62, 1936.

[6] F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. D. Procaccia. *Handbook of computational social choice*. Cambridge University Press, 2016.

[7] V. Conitzer. Eliciting single-peaked preferences using comparison queries. In *Proceedings of the 6th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2007)*, pages 420–427, 2007.

[8] V. Conitzer, T. Walsh, and L. Xia. Dominating manipulations in voting with partial information. In *Proceedings of the 25th AAAI Conference on Artificial Intelligence (AAAI 2011)*, volume 25, pages 638–643, 2011.

[9] P. Dey, N. Misra, and Y. Narahari. Complexity of manipulation with partial information in voting. *Theoretical Computer Science*, 726:78–99, 2018.

[10] R. G. Downey and M. R. Fellows. Fixed-parameter tractability and completeness II: On completeness for W[1]. *Theoretical Computer Science*, 141(1-2):109–131, 1995.

[11] U. Endriss, S. Obraztsova, M. Polukarov, and J. S. Rosenschein. Strategic voting with incomplete information. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI 2016)*, pages 236–242, 2016.

[12] P. Faliszewski and A. D. Procaccia. AI's war on manipulation: Are we winning? *AI Magazine*, 31(4):53–64, 2010.

[13] P. Faliszewski and J. Rothe. Control and bribery in voting. In F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. D. Procaccia, editors, *Handbook of Computational Social Choice*, chapter 7, pages 146–168. Cambridge University Press, 2016.

[14] E. Friedgut, G. Kalai, and N. Nisan. Elections can be manipulated often. In *Proceeding of the 49th Annual IEEE symposium on Foundations of Computer Science (FOCS 2008)*, pages 243–249, 2008.

[15] M. R. Garey and D. S. Johnson. *Computers and intractability : A guide to the theory of NP-completeness*. W.H. Freeman, 1979.

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

[17] T. F. Gonzalez. Clustering to minimize the maximum intercluster distance. *Theoretical computer science*, 38:293–306, 1985.

[18] U. Grandi. Social choice and social networks. In U. Endriss, editor, *Trends in Computational Social Choice*, chapter 9, pages 169–184. AI Access, 2017.

[19] N. Hazon, Y. Aumann, S. Kraus, and M. Wooldridge. On the evaluation of election outcomes under uncertainty. *Artificial Intelligence*, 189:1–18, 2012.

[20] W. Hoeffding. Probability inequalities for sums of bounded random variables. *The collected works of Wassily Hoeffding*, pages 409–426, 1994.

[21] M. Isaksson, G. Kindler, and E. Mossel. The geometry of manipulation—a quantitative proof of the Gibbard-Satterthwaite theorem. *Combinatorica*, 32(2):221–250, 2012.

[22] R. Meir. Iterative voting. In U. Endriss, editor, *Trends in Computational Social Choice*, chapter 4, pages 69–86. AI Access, 2017.

[23] R. Meir. Strategic voting. *Synthesis lectures on artificial intelligence and machine learning*, 13(1):1–167, 2018.

[24] R. Meir, M. Polukarov, J. Rosenschein, and N. Jennings. Convergence to equilibria in plurality voting. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI 2010)*, volume 24, pages 823–828, 2010.

[25] R. Meir, O. Lev, and J. S. Rosenschein. A local-dominance theory of voting equilibria. In *Proceedings of the 15th ACM conference on Economics and computation (EC 2014)*, pages 313–330, 2014.

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

[27] A. D. Procaccia and J. S. Rosenschein. Average-case tractability of manipulation in voting via the fraction of manipulators. In *Proceedings of the 6th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2007)*, page 718–720, 2007.

[28] A. Reijngoud and U. Endriss. Voter response to iterated poll information. In *Proceeding of the 4th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2012)*, pages 635–644, 2012.

[29] M. A. Satterthwaite. Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of economic theory*, 10(2):187–217, 1975.

[30] T. A. Severini. *Elements of distribution theory*. Cambridge University Press, 2005.

[31] S. Szufa, P. Faliszewski, P. Skowron, A. Slinko, and N. Talmon. Drawing a map of elections in the space of statistical cultures. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019)*, pages 1341–1349, 2020.

[32] T. Walsh. Generating single peaked votes. *arXiv preprint arXiv:1503.02766*, 2015.

[33] A. Wilczynski. Poll-confident voters in iterative voting. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI 2019)*, volume 33, pages 2205–2212, 2019.

[34] L. Xia. The impact of a coalition: Assessing the likelihood of voter influence in large elections. In *Proceedings of the 24th ACM Conference on Economics and Computation (EC 2023)*, pages 1156–1156, 2023.

[35] L. Xia and V. Conitzer. Generalized scoring rules and the frequency of coalitional manipulability. In *Proceedings of the 9th ACM Conference on Electronic Commerce (EC 2008)*, pages 109–118, 2008.

Technical Appendix

A Missing Proofs

Proposition 2. *If $x \in M \setminus \{e_1^>, e_2^>\}$, then every single-peaked culture $C(n, \Pi_>^m)$ is balanced for x. If $x \in \{e_1^>, e_2^>\}$, then every single-peaked culture $C(n, \Pi_>^m)$ which also satisfies $\mathbb{P}_C(\{\succ_i | worst_{\succ_i} = x\}) \leqslant \frac{1}{2}$, is balanced for x.*

Proof. Let $>$ be an axis on M. Observe first that, by definition, every preference order $\succ_i$ which is single-peaked w.r.t. $>$, must rank last an extreme candidate of $>$. Therefore, for every single-peaked culture $C(n, \Pi_>^m)$, we must have $\mathbb{P}_C(\{\succ_i \mid worst_{\succ_i} = e_1^>\} \cup \{\succ_i \mid worst_{\succ_i} = e_2^>\}) = 1$. Moreover, since $\mathbb{P}_C(\{\succ_i \mid worst_{\succ_i} = e_1^>\} \cup \{\succ_i \mid worst_{\succ_i} = e_2^>\}) \leqslant \mathbb{P}_C(\{\succ_i \mid worst_{\succ_i} = e_1^>\}) + \mathbb{P}_C(\{\succ_i \mid worst_{\succ_i} = e_2^>\})$, this implies that there exists an extreme candidate $e_\ell^>$, for $\ell \in [2]$, such that $\mathbb{P}_C(\{\succ_i \mid worst_{\succ_i} = e_\ell^>\}) \geqslant \frac{1}{2}$. It follows that, for every candidate $x \in M \setminus \{e_1^>, e_2^>\}$, $\mathbb{P}_C(x \succ_i e_\ell^>) \geqslant \frac{1}{2}$, proving the first part of the statement. Consider now a candidate $e_\ell^>$ for $\ell \in [2]$. Assuming that $\mathbb{P}_C(\{\succ_i |worst_{\succ_i} = e_\ell^>\}) \leqslant \frac{1}{2}$, implies that $\mathbb{P}_C(\{\succ_i |worst_{\succ_i} = e_{3-\ell}^>\}) \geqslant \frac{1}{2}$ and thus $\mathbb{P}_C(e_\ell^> \succ_i e_{3-\ell}^>) \geqslant \frac{1}{2}$, proving the second part. □

Proposition 3. *Any Mallows culture $\mathcal{M}^{\phi,\sigma}$ is balanced for every candidate $x \in M \setminus \{worst_\sigma\}$.*

Proof. Consider any candidate $x \in M \setminus \{worst_\sigma\}$ and the candidate $\ell := worst_\sigma$. Let $\Pi_{y \succ z}^m$ denote the set of all preferences orders where y is ranked before z, i.e., $\Pi_{y\succ z}^m := \{\succ_i \in \Pi^m : y \succ_i z\}$. Consider the bijection $\tau : \Pi_{\ell \succ x}^m \to \Pi_{x \succ \ell}^m$, where for every preference order $\succ_i \in \Pi_{\ell \succ x}^m$, we construct the preference order $\tau(\succ_i) \in \Pi_{x\succ\ell}^m$ which is the same as $\succ_i$ except that the positions of x and ℓ are swapped. We will show that $\mathbb{P}_{\mathcal{M}^{\phi,\sigma}}(\tau(\succ_i)) \geq \mathbb{P}_{\mathcal{M}^{\phi,\sigma}}(\succ_i)$ for every $\succ_i \in \Pi_{\ell\succ x}^m$. For this purpose, we will show that $d_{KT}(\sigma, \succ_i) \geq d_{KT}(\sigma, \tau(\succ_i))$, by analyzing the differences between $\succ_i$ and $\tau(\succ_i)$ in terms of agreement on pairwise comparisons with σ.

By definition, for any arbitrary preference order $\succ_i'$, we have that $d_{KT}(\sigma, \succ_i') = d_{KT}([\sigma]_{|M\setminus\{x,\ell\}}, [\succ_i']_{|M\setminus\{x,\ell\}}) + |\{y \in M : \ell \succ_i' y\}| + |\{y \in M \setminus \{\ell\} : x \succ_i' y \text{ and } y\sigma x\}| + |\{y \in M \setminus \{\ell\} : y \succ_i' x \text{ and } x\sigma y\}|$, where $[\succ_i'']_{|Y}$ denotes the restriction of the preference order $\succ_i''$ on $Y \subseteq M$. Observe that, by construction, for any $\succ_i \in \Pi_{\ell\succ x}^m$, $\succ_i$ and $\tau(\succ_i)$ agree on all pairwise comparisons within $M \setminus \{\ell, x\}$. Therefore, we have $d_{KT}([\sigma]_{|M\setminus\{x,\ell\}}, [\succ_i]_{|M\setminus\{x,\ell\}}) = d_{KT}([\sigma]_{|M\setminus\{x,\ell\}}, [\tau(\succ_i)]_{|M\setminus\{x,\ell\}})$. Moreover, by construction, for all candidates y such that $\ell\ \tau(\succ_i)\ y$ it implies that $\ell \succ_i y$, for all candidates $y \in M \setminus \{\ell\}$ such that $x \succ_i y$ it implies that $x\ \tau(\succ_i)\ y$, and for all candidates $y \in M \setminus \{\ell\}$ such that $y\ \tau(\succ_i)\ x$ it implies that $y \succ_i x$. It follows that we have $d_{KT}(\sigma, \succ_i) - d_{KT}(\sigma, \tau(\succ_i)) = |\{y \in M : \ell \succ_i y\ \tau(\succ_i)\ \ell\}| - |\{y \in M \setminus \{\ell\} : y\sigma x \text{ and } x\ \tau(\succ_i)\ y \succ_i x\}| + |\{y \in M \setminus \{\ell\} : x\sigma y \text{ and } x\ \tau(\succ_i)\ y \succ_i x\}|$. By construction, it holds that $|\{y \in M : \ell \succ_i y\ \tau(\succ_i)\ \ell\}| = r_{\succ_i}(x) - r_{\succ_i}(\ell)$. Moreover, $|\{y \in M \setminus \{\ell\} : y\sigma x \text{ and } x\ \tau(\succ_i)\ y \succ_i x\}| + |\{y \in M \setminus \{\ell\} : x\sigma y \text{ and } x\ \tau(\succ_i)\ y \succ_i x\}| = r_{\succ_i}(x) - r_{\succ_i}(\ell) - 1$, which implies that $-(r_{\succ_i}(x) - r_{\succ_i}(\ell) - 1) \leqslant -|\{y \in M \setminus \{\ell\} : y\sigma x \text{ and } x\ \tau(\succ_i)\ y \succ_i x\}| + |\{y \in M \setminus \{\ell\} : x\sigma y \text{ and } x\ \tau(\succ_i)\ y \succ_i x\}| \leqslant r_{\succ_i}(x) - r_{\succ_i}(\ell) - 1$. Therefore, in total, we have $1 \leq d_{KT}(\sigma, \succ_i) - d_{KT}(\sigma, \tau(\succ_i)) \leq 2(r_{\succ_i}(x) - r_{\succ_i}(\ell)) - 1$, and thus $d_{KT}(\sigma, \succ_i) \geq d_{KT}(\sigma, \tau(\succ_i))$. By definition of the Mallows culture $\mathcal{M}^{\phi,\sigma}$, we thus have $\mathbb{P}_{\mathcal{M}^{\phi,\sigma}}(\tau(\succ_i)) \geq \mathbb{P}_{\mathcal{M}^{\phi,\sigma}}(\succ_i)$. Hence, we conclude that $\mathbb{P}_{\mathcal{M}^{\phi,\sigma}}(\{\succ_i' : x \succ_i' \ell\}) \geq \mathbb{P}_{\mathcal{M}^{\phi,\sigma}}(\{\succ_i' : \ell \succ_i' x\})$, and thus $\mathbb{P}_{\mathcal{M}^{\phi,\sigma}}(\{\succ_i' : x \succ_i' \ell\}) \geq \frac{1}{2}$, implying that $\mathcal{M}^{\phi,\sigma}$ is balanced for x. □

Theorem 4. *The unrestricted manipulation problem is NP-complete.*

Proof. Membership to NP is straightforward: given communicated scores, we can efficiently derive the possible unique deviation of each voter and compute the winner in the deviating profile.

For hardness, we perform a reduction from a variant of EXACT COVER BY 3-SETS (X3C) known to be NP-complete [15]: Given a set $X = \{x_1, x_2, \ldots, x_{3q}\}$ and a set $S = \{S_1, S_2, \ldots, S_{3q}\}$ of 3-element subsets of X, where each element x_i occurs in exactly three subsets of S, we ask whether there exists an exact cover, i.e.,., a subset $S' \subseteq S$ such that every element of X occurs in exactly one member of S'. From an instance (X, S) of X3C, we construct an instance of our unrestricted manipulation problem as follows. For each element x_i, for $i \in [3q]$, we create a candidate y_i, and for each subset S_j where $j \in [3q]$, we create a candidate c_j. We add three candidates w, z, and t where t is our target candidate.

There are $12q + 7$ voters: for each element x_i, for $i \in [3q]$, we create one voter Y_i, for each subset S_j, for $j \in [3q]$, we create three voters C_j^r where $r \in [3]$, and we finally add two voters T^ℓ, two voters Z^ℓ, two voters W^ℓ for $\ell \in [2]$, and one voter D.

Their preferences are defined below, where $y(s_j^r)$ denotes the candidate y_i associated with the r^{th} element of subset S_j, and when a subset of candidates is mentioned, the candidates are ranked according to the increasing order of their indices.

Y_i:	$w \succ y_i \succ z \succ \{y_{i'}\}_{i' \neq i} \succ \{c_j\}_j \succ t$	for $i \in [3q]$
C_j^r:	$y(s_j^r) \succ c_j \succ z \succ w \succ \{y_{i'}\}_{i' \neq i} \succ \{c_j\}_j \succ t$	for $j \in [3q], r \in [3]$
T^ℓ:	$t \succ z \succ w \succ \{y_i\}_i \succ \{c_j\}_j$	for $\ell \in [2]$
Z^ℓ:	$z \succ w \succ \{y_i\}_i \succ \{c_j\}_j \succ t$	for $\ell \in [2]$
W^ℓ:	$w \succ z \succ \{y_i\}_i \succ \{c_j\}_j \succ t$	for $\ell \in [2]$
D:	$w \succ t \succ z \succ \{y_i\}_i \succ \{c_j\}_j$	

Finally, the tie-breaking rule is as follows: $w \rhd t \rhd z \rhd y_1 \rhd \cdots \rhd y_{3q} \rhd c_1 \rhd \cdots \rhd c_{3q}$.

The winner of the election with the truthful ballot profile is candidate w. The details of the scores for this truthful ballot profile are given in the second column of Table 3.

Table 3. Candidates' scores in the complexity proof of Theorem 4

candidate	initial score	announced score	score after manipulation
y_i ($i \in [3q]$)	3	3	3
c_j ($j \in [3q]$)	0	3 if $S_j \in S'$ 0 otherwise	3 if $S_j \in S'$ 0 otherwise
w	$3q + 3$	2	2
t	2	2	3
z	2	3	2
winner	w	z	t

We claim that there exists an exact cover in (X, S) iff we can force the election of candidate t in the constructed instance.

$\Longrightarrow$**:** Suppose first that there exists a subset $S' \subseteq S$ such that every element of X occurs in exactly one subset of S'. Since $|X| = 3q$ and all elements of S are subsets of X of size 3, we have $|S'| = q$. Let us consider manipulated communicated scores which differ from the sincere ones by taking $3q + 1$ votes initially given to w to give one additional vote to z and three votes to c_j for each $S_j \in S'$. These scores are summarized in the third column of Table 3. By the tie-breaking rule, candidate z is the announced winner.

It follows from these communicated scores that all candidates are potential winners except the candidates c_j such that $S_j \notin S'$. Therefore, each voter Y_i will deviate from ballot w to ballot y_i, for $i \in [3q]$, all voters C_j^r such that $S_j \in S'$ will deviate from ballot $y(s_j^r)$ to ballot c_j, and voter D will deviate from ballot w to ballot t. Since S' is an exact cover, each additional vote for y_i by voter Y_i will be balanced by the removal of one vote for y_i by the voter C_j^r, such that $S_j \in S'$ and $y(s_j^r) = y_i$, who deviates from y_i to c_j. Therefore, in total, these deviations will remove $3q + 1$ votes from w, give three votes to q candidates c_j and add one vote to t, leading to the victory of t, as summarized in the fourth column of Table 3.

$\Longleftarrow$ **:** Suppose now that there exist communicated scores such that the target candidate t becomes the winner after deviations from the voters. The global idea of the proof is that the only possibility for communicated scores to lead to the victory of the target candidate t is to announce candidate z the winner and, as potential winners, the target candidate t and exactly q candidates c_j which correspond to subsets S_j forming an exact cover of X.

We will first show by disjunction case that the announced winner can only be candidate z.

Observe that t cannot win if it does not gain any additional vote. Indeed, for t to win with at most two votes, w cannot get more than one vote, and all the other candidates more than two votes, which sums to at most $12q + 3$ votes for other candidates, whereas there would be $12q + 5$ voters who do not vote for t, a contradiction. It follows that t cannot be announced as the winner, and must be a potential winner. However, by construction of the preferences, the only voter who can deviate to a ballot t is voter D. Therefore, in the deviating profile, t can get at most three votes.

If w is announced the winner, then the $3q$ voters Y_i will keep their vote for w, therefore t can never win with its maximum score of three, a contradiction.

Let us now analyze the case where the announced winner is a candidate y_i or c_j, by considering the candidates that can be announced potential winners:

- If candidate z is a potential winner, then at least voters T^ℓ and W^ℓ will deviate to it, which leads to at least four votes for z, whereas t can get at most three votes. Therefore, z cannot be a potential winner.
- Now, if candidate w is a potential winner, then at least voters T^ℓ and Z^ℓ will deviate to it, leading to at least four votes for w, whereas t can get at most three votes. Therefore, w cannot be a potential winner.
- Now, if a candidate $y_{i'}$ is a potential winner, for $i' < i$ or when c_j is the winner, then at least voters T^ℓ, Z^ℓ, and W^ℓ, for $\ell \in [2]$, will deviate to the candidate $y_{i'}$, that we call y^*, which is declared potential winner with the smallest index i', by construction of their preferences. Therefore, y^* would get at least six votes, whereas t can get at most three votes. Thus, such $y_{i'}$ cannot be a potential winner.
- Now, if a candidate $y_{i'}$ or $c_{j'}$ is a potential winner, for $i' > i$ and y_i winner, then voter Y_i will keep her vote for w as well as voters W^ℓ for $\ell \in [2]$, which leads to at least three votes for w whereas w is preferred to t in the tie-breaking rule. Therefore, such $y_{i'}$ or $c_{j'}$ cannot be a potential winner.
- Now, if a candidate $c_{j'}$ is a potential winner, for $j' < j$ and c_j winner, then at least voters T^ℓ, Z^ℓ, and W^ℓ, for $\ell \in [2]$, will deviate to the candidate $c_{j'}$, that we call c^*, which is declared potential winner with the smallest index j', by construction of their preferences. Therefore, c^* would get at least six votes, whereas t can get at most three votes. Thus, such $c_{j'}$ cannot be a potential winner.
- Now, finally, if a candidate $c_{j'}$ is a potential winner, for $j' > j$ and c_j winner, then all voters Y_i will keep their vote for w, which leads to at least $3q$ votes for w, whereas t can get at most three votes. Therefore, such $c_{j'}$ cannot be a potential winner.
- It follows that t is the only potential winner, and thus all voters Y_i keep their vote for w, which leads to at least $3q$ votes for w, and thus t cannot win, a contradiction.

Consequently, the announced winner must be candidate z. Since t can get at most three votes, and w initially gets $3q + 3$ votes, at least $3q + 1$ votes must be removed from w (w is preferred to t in the tie-breaking). Voters W^ℓ will not deviate from w if z is the announced winner, therefore all voters Y_i and D must deviate from w. It follows that each candidate y_i must be a potential winner as well as t. However, for each candidate y_i, we need that at least one of the three voters C_j^r such that $y(s_j^r) = y_i$ deviates from her initial vote to y_i, otherwise y_i would get four votes and t could not win. For such a voter C_j^r to deviate, the only solution is to make candidate c_j a potential winner. By construction, it follows that we need to find a subset of candidates c_j (to make them potential winners) such that the associated subsets S_j entirely cover the elements in X. Thus, we need to make at least q candidates c_j potential winners.

Let us now analyze the compatible scores that can be communicated. If z is the announced winner with at most two votes, then by the tie-breaking rule, candidates w and t can get at most one vote, and all the other candidates at most two votes, which sums to at most $12q + 2$ votes for other candidates, whereas there would be $12q + 5$ voters who do not vote for z, a contradiction. If z is the announced winner with at least four votes, then to be potential winners, t should get at least three votes, all candidates y_i at least four votes, and at least q candidates c_j at least four votes, which sums to at least $16q+3$ votes for other candidates, whereas there are $12q + 7$ voters in total, a contradiction. Consequently, z must be announced the winner with exactly three votes, and thus t must be announced with two votes, all candidates y_i with three votes, and at least q candidates c_j with three votes. The only possibility to announce such scores is to take $3q + 1$ votes from w and to distribute them to give three votes to q candidates c_j and one vote to candidate z. The only possible margin then is given by the two remaining votes for w, however they are not sufficient to make another candidate c_j a potential winner. Hence, there are exactly q candidates c_j which are potential winners such that the associated subsets of S_j entirely cover X, which means that the union of such subsets is an exact cover. $\square$

Theorem 5. *For a balanced culture $C(n, \Pi^m)$, there exists $\ell \in B_C(x^*)$ such that the probability of success of the sub-heuristic 2PW-H(x^*, ℓ), is as follows: $\mathbb{P}_C(S_{\text{2PW-H}(x^*,\ell)}) \geqslant 1 - 2(m - 2)(e^{-2n(p_{x^*,\ell} - q_{x^*,\ell,j^*})^2})$ where:*

- $p_{x^*,\ell} := \mathbb{P}_C(x^* \succ_i \ell)$,
- $r_{x^*,\ell,j} := \mathbb{P}_C(\{\ell \succ_i x^*\} \cap \{j = top_{\succ_i}\})$, *for* $j \neq \ell$,
- $q_{x^*,\ell,j} := \frac{p_{x^*,\ell} + r_{x^*,\ell,j}}{2}$,
- $j^* := \arg\min_{j \in M \setminus \{x^*,\ell\}} \mathbb{P}_C(Y_{\ell,j} \leqslant X_{-\ell})$.

In particular, the probability of success of the global heuristic satisfies the same lower bound.

Proof. For our target candidate x^* and a balanced culture $C(n, \Pi^m)$, let us consider a candidate $\ell \in B_C(x^*)$. Since $\ell \in B_C(x^*)$, ℓ will never be better and we can simplify the equality: $\mathbb{P}_C(S_{\text{2PW-H}(x^*,\ell)}) = \mathbb{P}_C(\forall j \in M \setminus \{x^*, \ell\}, Y_{\ell,j} \leqslant X_{-\ell})$. Our goal is to show a lower bound to $\mathbb{P}_C(\forall j \in M \setminus \{x^*, \ell\}, Y_{\ell,j} \leqslant X_{-\ell})$. For this purpose, the following lemma will be useful.

Lemma 18 (Bonferroni [5])**.** $\mathbb{P}(\bigcap_{i=1}^n A_i) \geqslant \sum_{i=1}^n \mathbb{P}(A_i) - (n - 1)$.

We deduce from Bonferroni's inequality (Lemma 18) that $\mathbb{P}_C(\forall j \in M \setminus \{x^*, \ell\}, Y_{\ell,j} \leqslant X_{-\ell}) \geqslant \sum_{j \in M\setminus\{x^*,l\}} \mathbb{P}_C(Y_{\ell,j} \leqslant X_{-\ell}) - (m-2-1) \geqslant \sum_{j\in M\setminus\{x^*,\ell\}} \min_{j\in M\setminus\{x^*,\ell\}} \mathbb{P}_C(Y_{\ell,j} \leqslant X_{-\ell}) - (m-3) \geqslant (m-2)\cdot \min_{j\in M\setminus\{x^*,\ell\}} \mathbb{P}_C(Y_{\ell,j} \leqslant X_{-\ell}) - (m-3)$. By considering $j^* := \arg\min_{j\in M\setminus\{x^*,\ell\}} \mathbb{P}_C(Y_{\ell,j} \leqslant X_{-\ell})$, we then have that $\mathbb{P}_C(\forall j \in M \setminus \{x^*, \ell\}, Y_{\ell,j} \leqslant X_{-\ell}) \geqslant (m-2)\cdot(\mathbb{P}_C(Y_{\ell,j^*} \leqslant X_{-\ell}) - 1) + 1$. Let us now treat the term $\mathbb{P}_C(Y_{\ell,j^*} \leqslant X_{-\ell})$. We remark that $X_{-\ell}$ follows a binomial distribution of parameters n and $p_{x^*,\ell}$ and Y_{ℓ,j^*} follows a binomial distribution of parameters n and r_{x^*,ℓ,j^*}, where $p_{x^*,\ell}$ and $r_{x^*,\ell,j}$ are defined as $p_{x^*,\ell} = \mathbb{P}_C(x^* \succ_i \ell)$ and $r_{x^*,\ell,j} = \mathbb{P}_C(\{\ell \succ_i x^*\} \cap \{j = top_{\succ_i}\})$, for every $j \neq \ell$. We introduce $q_{x^*,\ell,j^*} := \frac{p_{x^*,\ell}+r_{x^*,\ell,j^*}}{2}$ to lower bound our probability as follows: $\mathbb{P}_C(Y_{\ell,j^*} \leqslant X_{-\ell}) \geqslant \mathbb{P}_C(\{Y_{\ell,j^*} < q_{x^*,\ell,j^*}\cdot n\} \bigcap \{X_{-\ell} > q_{x^*,\ell,j^*}\cdot n\})$. We use again Bonferroni's inequality (Lemma 18) to get: $\mathbb{P}_C(Y_{\ell,j^*} \leqslant X_{-\ell}) \geqslant \mathbb{P}_C(\{Y_{\ell,j^*} < q_{x^*,\ell,j^*}\cdot n\}) + \mathbb{P}_C(\{X_{-\ell} > q_{x^*,\ell,j^*}\cdot n\}) - 1$. We then use the following lemma to treat each term of this inequality.

Lemma 19 (Hoeffding [20]). *Let X_k be some independent real random variables, and $(a_k)_{k\in[n]}$ and $(b_k)_{k\in[n]}$ two real sequences such that for every $k \in [n]$, we have $a_k < b_k$ and $\mathbb{P}(a_k \leqslant X_k \leqslant b_k) = 1$. Then, for every $t > 0$, $\mathbb{P}(S_n - \mathbb{E}(S_n) \geqslant t) \leqslant e^{\frac{-2t^2}{\sum_{k=1}^n (b_k-a_k)^2}}$, where $S_n = \sum_{k=1}^n X_k$.*

Applying the inequality from Lemma 19 on Bernoulli variables X_k with $a_k = 0$ and $b_k = 1$, for every $k \in [n]$, and $t = x\cdot\sqrt{n}$, we get: $\mathbb{P}(S_n - \mathbb{E}(S_n) \geqslant x\cdot\sqrt{n}) \leqslant e^{-2x^2}$, where $S_n = \sum_{i=1}^n X_k$. Now, by taking $x = \sqrt{n}\cdot(q_{x^*,\ell,j^*} - r_{x^*,\ell,j^*})$ and applying Lemma 19 to our sum of Bernoulli variables Y_{ℓ,j^*} (i.e., a binomial of parameters n and r_{x^*,ℓ,j^*}), we get: $\mathbb{P}_C(Y_{\ell,j^*} < q_{x^*,\ell,j^*}\cdot n) = 1 - \mathbb{P}_C(Y_{\ell,j^*} \geqslant q_{x^*,\ell,j^*}\cdot n) = 1 - \mathbb{P}_C(Y_{\ell,j^*} - r_{x^*,\ell,j^*}\cdot n \geqslant q_{x^*,\ell,j^*}\cdot n - r_{x^*,\ell,j^*}\cdot n) = 1 - \mathbb{P}_C(Y_{\ell,j^*} - r_{x^*,\ell,j^*}\cdot n \geqslant \sqrt{n}(\sqrt{n}(q_{x^*,\ell,j^*} - r_{x^*,\ell,j^*}))) \geqslant 1 - e^{-2n(q_{x^*,\ell,j^*}-r_{x^*,\ell,j^*})^2}$. Now, we want to apply a similar treatment to variables $X_{-\ell}$. Let us denote $X'_{-\ell} = n - X_{-\ell}$ the random variable following a binomial distribution of parameters n and $1 - p_{x^*,\ell}$. We have: $\mathbb{P}_C(X_{-\ell} > q_{x^*,\ell,j^*}\cdot n) = \mathbb{P}_C(n - X'_{-\ell} > q_{x^*,\ell,j^*}\cdot n) = \mathbb{P}_C(X'_{-\ell} < n - q_{x^*,\ell,j^*}\cdot n) = 1 - \mathbb{P}_C(X'_{-\ell} - (1-p_{x^*,\ell})\cdot n \geqslant (1-q_{x^*,\ell,j^*})\cdot n - (1-p_{x^*,\ell})\cdot n) = 1 - \mathbb{P}_C(X'_{-\ell} - (1-p_{x^*,\ell})\cdot n \geqslant \sqrt{n}\sqrt{n}(p_{x^*,\ell} - q_{x^*,\ell,j^*})) \geqslant 1 - e^{-2n(p_{x^*,\ell}-q_{x^*,\ell,j^*})^2}$. Putting the last two inequalities together we get: $\mathbb{P}_C(Y_{\ell,j^*} \leqslant X_{-\ell}) \geqslant 1 - e^{-2n(p_{x^*,\ell}-q_{x^*,\ell,j^*})^2} - e^{-2n(q_{x^*,\ell,j^*}-r_{x^*,\ell,j^*})^2}$. Coming back to the first work of the proof we have: $\mathbb{P}_C(S_{\text{2PW-H}(x^*,\ell)}) \geqslant 1 - (m-2)(e^{-2n(p_{x^*,\ell}-q_{x^*,\ell,j^*})^2} + e^{-2n(q_{x^*,\ell,j^*}-r_{x^*,\ell,j^*})^2})$. Finally, since q_{x^*,ℓ,j^*} is defined as the middle between $p_{x^*,\ell}$ and r_{x^*,ℓ,j^*}, we can simplify the inequality: $\mathbb{P}_C(S_{\text{2PW-H}(x^*,\ell)}) \geqslant 1 - 2(m-2)e^{-2n(p_{x^*,\ell}-q_{x^*,\ell,j^*})^2}$. □

Theorem 9. *The restricted manipulation problem is W[1]-hard.*

Proof. From an instance $(G = (V, E), k)$ of k-Clique where $n := |V|$, $m := |E|$ and, w.l.o.g., $2 < k < n-1$, we construct an instance of our restricted poll manipulation problem as follows.

For each vertex $v_i \in V$, we create a candidate v_i, and for each edge $\{v_i, v_j\} \in E$, we create a candidate e_{ij} (we suppose $i < j$ for this notation). We add three other candidates w, t, and z. In total, we thus have $n + m + 3$ candidates.

Let $K := (n-k)k$. For each vertex $v_i \in V$, we create k voters U_i^ℓ for $\ell \in [k]$, and $K - 1 - \delta(v_i)$ voters D_i^ℓ for $\ell \in [K - 1 - \delta(v_i)]$ (by our assumption on k, this quantity cannot be negative), where $\delta(v_i)$ denotes the degree of vertex v_i in G.

For each edge $\{v_i, v_j\} \in E$, we create two voters F_{ij}^i and F_{ij}^j, and $K - 2$ voters E_{ij}^ℓ for $\ell \in [K - 2]$. Finally, we add K voters T^ℓ for $\ell \in [K]$ and $K - 1$ voters Z^ℓ for $\ell \in [K - 1]$.

The preferences of the voters over the candidates are described below, for each $i \in [n]$, and each $\{v_p, v_q\} \in E$:

U_i^ℓ:	$w \succ v_i \succ z \succ \{v_j\}_{j\neq i} \succ \{e_{r,s}\}_{\{r,s\}} \succ t$	for $\ell \in [k]$
F_{pq}^p:	$v_p \succ e_{pq} \succ z \succ w \succ \{v_j\}_{j\neq p} \succ \{e_{r,s}\}_{\{r,s\}\neq\{p,q\}} \succ t$	
F_{pq}^q:	$v_q \succ e_{pq} \succ z \succ w \succ \{v_j\}_{j\neq q} \succ \{e_{r,s}\}_{\{r,s\}\neq\{p,q\}} \succ t$	
D_i^ℓ:	$v_i \succ z \succ w \succ \{v_j\}_{j\neq i} \succ \{e_{r,s}\}_{\{r,s\}} \succ t$	for $\ell \in [K-1-\delta(v_i)]$
T^ℓ:	$t \succ z \succ w \succ \{v_j\}_j \succ \{e_{r,s}\}_{\{r,s\}}$	for $\ell \in [K]$
E_{pq}^ℓ:	$e_{pq} \succ z \succ w \succ \{v_j\}_j \succ \{e_{r,s}\}_{\{r,s\}\neq\{p,q\}} \succ t$	for $\ell \in [K-2]$
Z^ℓ:	$z \succ w \succ \{v_j\}_j \succ \{e_{r,s}\}_{\{r,s\}} \succ t$	for $\ell \in [K-1]$

Finally, the tie-breaking rule is as follows: $z \rhd t \rhd \cdots \rhd w$.

The winner of the election with the truthful ballot profile is candidate w. The details of the scores for this truthful ballot profile are given in the second column of Table 4.

Table 4. Candidates' scores in the complexity proof of Theorem 9

candidate	initial score	announced score	score after manipulation
v_i ($i \in [n]$)	$K-1$	K if $v_i \in S$ $K-1$ otherwise	K if $v_i \in S$ $K-1$ otherwise
e_{ij} ($\{v_i, v_j\} \in E$)	$K-2$	K if $v_p, v_q \in S$ $K-2$ otherwise	K if $v_p, v_q \in S$ $K-2$ otherwise
w	kn	K	K
t	K	$K-1$	K
z	$K-1$	K	$K-1$
winner	w	z	t

We claim that G admits a clique of size k iff we can force the election of candidate t by announcing scores which differ from the truthful ones by at most $k^2 + 1$ vote changes.

$\Longrightarrow$ **:** Suppose first that there exists a subset of vertices $S \subseteq V$ such that S is a k-Clique of G, i.e., $|S| = k$ and $\{v_i, v_j\} \in E$ for every $v_i, v_j \in S$. Let us consider manipulated communicated scores which differ from the sincere ones by taking k^2 votes initially given to w in order to give one additional vote to each $v_i \in S$ (there are k such candidates) and two additional votes to each e_{ij} such that $v_i, v_j \in S$ (there are $\frac{k(k-1)}{2}$ such candidates), and finally by taking one vote initially given to t in order to give it to z. In total, the communicated scores differ from the sincere ones by exactly $k^2 + 1$ vote changes.

In the manipulated scores, z is winning with K votes thanks to the tie-breaking, while only the k candidates v_i corresponding to the vertices of the clique are announced as potential winners with K votes, as well as the $k(k-1)/2$ candidates corresponding to the edges of the clique, and candidate w. These manipulated scores are summarized in the third column of Table 4.

It follows from these communicated scores that all voters U_i^ℓ such that $v_i \in S$ deviate from w to v_i. By these deviations, candidate w loses k^2 votes, and thus obtains in total K votes, while each candidate $v_i \in S$ gains k votes. However, by definition of the clique, for each $v_i \in S$, there are exactly $k-1$ voters F_{ij}^i (or F_{ji}^i) who will deviate from v_i to the potential winner e_{ij} (or e_{ji}) corresponding to an edge incident to v_i. Therefore, each $v_i \in S$ also loses $k-1$ votes, and thus obtains in total K votes. Note that, by these deviations, each candidate e_{ij} such that $v_i, v_j \in S$ gains two additional votes and thus obtains in total K votes. No other deviation is possible because all remaining voters prefer z to all potential winners that are not at top of their preferences. The scores after all deviations are summarized in the fourth column of Table 4. The maximum score is K, which is obtained by w, k candidates v_i, and $\frac{k(k-1)}{2}$ candidates e_{ij}, and

t. Candidate t is favored by the tie-breaking among these candidates and thus wins the election.

$\Longleftarrow$ **:** Suppose now that there exist communicated scores such that the target candidate t becomes the winner after deviations from the voters. The global idea of the proof is that the only possibility for communicated scores to lead to the victory of the target candidate t is to announce candidate z the winner and, as potential winners, k candidates v_i, as well as $k(k-1)/2$ candidates e_{pq}, such that for each potential winner v_i, there are $k-1$ potential winners e_{ij} (or e_{ji}) corresponding to edges incident to v_i.

We will first prove that z must be announced as the winner. Observe that no voter can deviate to t because every voter, except all voters T^ℓ who already vote for t, ranks it last. It follows that we need that at least k^2 voters U_i^ℓ, who currently vote for w, deviate to another candidate, and thus w cannot be announced as the winner.

Let us analyze the case where the announced winner would be a candidate v_i, e_{pq} or candidate t, by considering the candidates that can be announced potential winners:

- If candidate z or w is a potential winner, then at least all voters D_i^ℓ and all voters E_{rs}^ℓ (except voters E_{pq}^ℓ if e_{pq} is announced as the winner) would deviate to z if z is a potential winner or to w otherwise, and thus z or w would gain too many votes compared to t and t would never win. Therefore, none of them is a potential winner.
- Now, if a candidate $v_{i'}$ is a potential winner, for $i' < i$ or when e_{pq} or t is the winner, then all voters D_i^ℓ and all voters E_{rs}^ℓ (except voters E_{pq}^ℓ if e_{pq} is announced as the winner) would deviate to such candidate $v_{i'}$, that we call v^*, which is declared potential winner with the smallest index i'. Thus, such v^* would gain too many votes compared to t and t would never win. Therefore, such $v_{i'}$ cannot be a potential winner.
- Now, if a candidate $v_{i'}$ or e_{rs} is a potential winner, for $i' > i$ and v_i winner, then all voters $U_{i''}^\ell$, for $i'' \neq i'$, would keep their vote for w and thus w would have too many votes compared to t and t would never win. Therefore, such $v_{i'}$ or e_{rs} cannot be a potential winner.
- Now, if a candidate e_{rs} is a potential winner, for $\{r,s\} < \{p,q\}$ when e_{pq} winner or for t winner, then at least all voters D_i^ℓ and all voters E_{rs}^ℓ (except voters E_{pq}^ℓ if e_{pq} is announced as the winner) would deviate to such candidate e_{rs}, that we call e^*, which is declared potential winner with the smallest index $\{r,s\}$. Therefore, e^* would get too many votes compared to t and t would never win. Thus, such e_{rs} cannot be a potential winner.
- Now, finally, if a candidate e_{rs} is a potential winner, for $\{r,s\} > \{p,q\}$ and e_{pq} winner, then all voters $U_{i'}^\ell$ would keep their vote for w and thus w would have too many votes compared to t and t would never win. Therefore, such e_{rs} cannot be a potential winner.
- It follows that t is the only potential winner, and thus all voters $U_{i'}^\ell$ keep their vote for w. Thus, w has too many votes compared to t and t cannot win, a contradiction.

Hence the communicated scores must announce z as the winner.

Since z is ranked among the first two most preferred candidates by all voters D_i^ℓ, T^ℓ, E_{pq}^ℓ and Z^ℓ, none of these voters will deviate. Recall that we need at least k^2 voters U_i^ℓ (for $i \in [n]$ and $\ell \in [k]$) who deviate to another candidate, and the only candidate other than their top candidate that voters U_i^ℓ prefer to z is v_i, for all $\ell \in [k]$. Therefore, we need to announce at least k candidates v_i as potential winners. In such a way, each chosen candidate v_i gains k additional votes, whereas it initially had $K-1$ votes from voters D_i^ℓ, who cannot deviate, and from voters F_{ij}^i (or F_{ji}^i) for each edge $\{v_i, v_j\} \in E$. Since t will have at most K votes, we need at least $k-1$ voters F_{ij}^i (or F_{ji}^i) who deviate from ballot v_i. The only other candidate that such voters prefer to z is candidate e_{ij} (or e_{ji}). Therefore, for each chosen v_i potential winner, we also need to announce as potential winners at least $k-1$ candidates e_{ij} (or e_{ji}) which correspond to edges incident to v_i.

Recall that we can only announce scores which differ from the truthful ones by at most k^2+1 vote changes. If we announce z the winner with at most $K-1$ votes, then we need to remove at least k^2+1 votes for w and one vote for t, therefore we have already exceeded our budget. If we announce z the winner with at least $K+1$ votes, then we need to add two votes to at least k candidates v_i, three votes to at least $\frac{k(k-1)}{2}$ candidates e_{ij} and one vote to z, therefore we have already exceeded our budget. It follows that we need to announce z the winner with exactly K votes. In this case, we need to add one vote to z, one vote to at least k candidates v_i and two votes to at least $\frac{k(k-1)}{2}$ candidates e_{ij}. Therefore, to meet our budget, we need to declare exactly k candidates v_i and exactly $\frac{k(k-1)}{2}$ candidates e_{ij} as potential winners, in such a way that for potential winner v_i there exist $k-1$ potential winners e_{ij} corresponding to incident edges. Hence, the chosen candidates v_i correspond to a k-clique in G. $\square$

Proposition 10. *The restricted manipulation problem is in XP w.r.t. the maximum distance k to the truthful scores. More precisely, it can be solved by an algorithm which runs in time* $\Theta(m^{2k+1} \cdot n)$.

Proof. We give an upper bound to $|I_k|$. We denote that any move of voters is characterized by the origin and the destination candidate. Since our distance counts the number of swaps, one swap is defined by choosing two candidates, we then get $\binom{m}{2} = \frac{m(m-1)}{2}$ and $|I_1| \leqslant \frac{m(m-1)}{2}$. We start from s^T and iterate the upper bound argument and we get: $|I_k| \leqslant (\frac{m(m-1)}{2})^k \leqslant m^{2k}$. It is then enough to visit every score of I_k and add the winner determination in $\Theta(m \cdot n)$. At the end, we get $\Theta(m^{2k+1} \cdot n)$. $\square$

Theorem 11. *For any culture* $C(n, \Pi^m)$, *if the maximum distance* k *to the truthful scores is such that* $k = o(\sqrt{n})$ *and the target candidate* x^* *is not winning in the initial score, then the probability of existence of a successful poll manipulation to elect* x^* *tends toward zero, i.e.,* $\lim_{n\to\infty} \mathbb{P}_C(S_k) = 0$.

Proof. Let us start by identifying the probability law of the truthful scores.

Observation 20. *For a culture* $C(n, \Pi^m)$, *the truthful scores* s^T *follow a multinomial law Multi*(p, n) *where* $q = (q_1, \ldots, q_m)$ *and* $q_j := \mathbb{P}_C(\{\mathcal{W}_P(s^T) = j\})$, *for every* $j \in M$.

The truthful scores follow a multinomial law because there are n voters' preferences drawn independently at random with the same law, and we have m possibilities for the most preferred candidate of each voter, and these are the only necessary elements to compute scores s^T. We will use the following result on multinomial laws.

Lemma 21 (Severini [30]). *If* $(N_n)_{n\geqslant 0}$ *is a multinomial law in* $\mathbb{R}^m$ *with parameters n and* $q = (q_1, \ldots, q_m)$ *and* $\mathcal{N}(0; K)$ *a multivariate normal distribution then* $\frac{1}{\sqrt{n}}(N_n - nq) \underset{n\to+\infty}{\longrightarrow} \mathcal{N}(0; K)$, *where* $K_{i,j} = q_i\delta_{i,j} - q_i q_j$, *for every* $1 \leq i, j \leq m$, *with* $\delta_{i,j} = 1$ *if* $i = j$ *and* $\delta_{i,j} = 0$ *otherwise.*

Let c^* be the truthful winner, i.e., $c^* := \mathcal{W}_P(b^T)$. Informally, a necessary condition for the existence of a successful manipulation with the two-candidate heuristic is that there is at least one candidate that is sufficiently close to the winner. The pair of candidates would then be this candidate and the current winner. Of course, this is not necessarily sufficient, as the pair may not be the right one. However, we will see that this necessary condition occurs with probability 0, and that's enough for us to conclude. We then write $S_k \subset \{\bigcup_{z\neq c^*}\{|s^T_{c^*} - s^T_z + \mathbb{1}_{c^*\rhd z}| \leq k\}\}$.

We will analyze the probability of the second event to get an upper bound on the probability of success of the restricted poll manipulation problem. By using Observation 20 and Lemma 21 with $N_n = s^T$, we get: $\frac{1}{\sqrt{n}}(s^T - nq) \underset{n\to+\infty}{\longrightarrow} \mathcal{N}(0; K)$, where $K_{i,j} = q_i\delta_{i,j} - q_iq_j$, for every $1 \leq i, j \leq m$. We denote $\mathcal{N}(0; K) = (\mathcal{N}_1, \dots, \mathcal{N}_m)$ and remark that each $\mathcal{N}_j$ follows a Gaussian law. For any $z \in M\setminus\{x^*\}$, we have $\lim_{n\to+\infty} \mathbb{P}_C(|s^T_{c^*} - s^T_z + \mathbb{1}_{c^*\rhd z}| \leq k) = \lim_{n\to+\infty} \mathbb{P}_C(|\frac{1}{\sqrt{n}}s^T_{c^*} - nq_{c^*} - \frac{1}{\sqrt{n}}s_z + nq_z + \frac{\mathbb{1}_{c^*\rhd z}}{\sqrt{n}} + n(q_{c^*} - q_z)| \leq \frac{k}{\sqrt{n}})$. Combining this equality with the previous convergence result using the test function $\Phi(s^T) = \mathbb{1}_{\{|\frac{1}{\sqrt{n}}s^T_{c^*} - nq_{c^*} - \frac{1}{\sqrt{n}}s^T_z + nq_z + \frac{\mathbb{1}_{c^*\rhd z}}{\sqrt{n}} + n(q_{c^*}-q_z)| \leq \frac{k}{\sqrt{n}}\}}$ and $\lim\limits_{n\to+\infty} \frac{k}{\sqrt{n}} = 0$ by assumption, we deduce that $\lim_{n\to+\infty} \mathbb{P}_C(|\frac{1}{\sqrt{n}}s^T_{c^*} - nq_{c^*} - \frac{1}{\sqrt{n}}s^T_z + nq_z + \frac{\mathbb{1}_{c^*\rhd z}}{\sqrt{n}} + n(q_{c^*} - q_z)| \leq \frac{k}{\sqrt{n}})) = \mathbb{P}_C(|\mathcal{N}_{c^*} - \mathcal{N}_z + \frac{\mathbb{1}_{c^*\rhd z}}{\sqrt{n}} + n(q_{c^*} - q_z)| \leq 0) = 0$. Therefore, $\lim_{n\to+\infty} \mathbb{P}_C(|s^T_{c^*} - s^T_z + \mathbb{1}_{c^*\rhd z}| \leq k) = 0$.

It follows for the probability of the success event that $\lim_{n\to+\infty} \mathbb{P}_C(S_k) \leqslant \lim_{n\to+\infty} \mathbb{P}_C(\bigcup_{z\neq c^*} |s^T_{c^*} - s^T_z + \mathbb{1}_{c^*\rhd z}| \leq k) \leq \lim_{n\to+\infty} \sum_{z\neq c^*} \mathbb{P}_C(|s^T_{c^*} - s^T_z + \mathbb{1}_{c^*\rhd z}| \leq k) = 0$. We then get: $\lim\limits_{n\to+\infty} \mathbb{P}_C(S_k) = 0$, which concludes the proof. □

Proposition 16. *1.* $U = \mathbb{E}_C[\frac{1}{n}\sum_{i=1}^n \mathbb{1}_{U_i^p}]$ *is decreasing w.r.t. p.*

2. $D = \mathbb{E}_C[\frac{1}{n}\sum_{i=1}^n \mathbb{1}_{D_i^p}]$ *is increasing w.r.t. p.*

3. $\mathcal{P}_{SV}(C, n, m, p) \leqslant \min(U, D)$*.*

4. $\lim\limits_{p\to+\infty} \mathcal{P}_{SV}(C, n, m, p) = 0$ *and* $\mathcal{P}_{SV}(C, n, m, 0) = 0$*.*

5. $\lim\limits_{n\to+\infty} \mathcal{P}_{SV}(C, n, m, p) = 0$ *when p is fixed.*

Proof.

1-2. The statements follow from the inclusions $U_i^{p'} \subseteq U_i^p$ and $D_i^p \subseteq D_i^{p'}$, for each $p' > p$.

3. Using the inclusions $U \cap D \subset U$ and $U \cap D \subset D$, we show that: $\mathcal{P}_{SV}(C, n, m, p) \leqslant U$ and $\mathcal{P}_{SV}(C, n, m, p) \leqslant D$.

4. If p is maximum, then all candidates are potential winners and thus each voter keeps her truthful vote, while when $p = 0$ there are no potential winners to deviate to.

5. Using Lemma 21, we know that the winner c^* and any other candidate z will be spread out at least of order $\sqrt{n}$ asymptotically. We then deduce that there are no potential winners other than the winner in that case, since p is fixed. □

Proposition 17. *For a balanced culture* $C(n, \Pi^m)$*,* $p > 0$ *and* $p = o(n)$*, we have* $\mathbb{P}_C(S^G) \geqslant \mathbb{P}_C(S)$ *and* $\mathbb{P}_C(S^G_k) \geqslant \mathbb{P}_C(S_k)$*.*

Proof. In an idea similar to the proof of Theorem 14, for each score of a given type in the initial setting, we can always choose a score of the same type which works for the generalized strategic behavior, since they would trigger the same deviations. Indeed, even if the polling institute is not sending the same score, it might construct a score with the same potential winners and the same winner since p is negligible against n. □

B Heuristics and Figures

Algorithm 1: Global Heuristic

Input: $(N, M, \succ, \rhd)$, Target candidate x^*
foreach $\ell \in M \setminus \{x^*\}$ **do**
  $(is\_successful, s) \leftarrow$ 2PW-H(x^*, ℓ);
  if $is\_successful$ **then return** $(True, s)$;
return $(False, None)$

Algorithm 2: 2PW-H(x^*, ℓ)

Input: $(N, M, \succ, \rhd)$, Target candidate x^*, Candidate ℓ
$s \leftarrow m$-vector with zeros; $R \leftarrow n$;
foreach $j \in M \setminus \{x^*, \ell\}$ **do**
  if $\exists i \in N$ *such that* $top_{\succ_i} = j$ **then** $s_j \leftarrow 1$; $R \leftarrow R - 1$;
$s_{x^*} \leftarrow \lfloor \frac{R}{2} \rfloor$; $s_\ell \leftarrow \lfloor \frac{R}{2} \rfloor$; $j^* \leftarrow \arg\min_{j\in M\setminus\{x^*,\ell\}} s_j$;
if $x^* \rhd \ell$ *and* R *is even* **then** $s_{x^*} \leftarrow s_{x^*} - 1$; $s_{j^*} \leftarrow s_{j^*} + 1$;
if $x^* \rhd \ell$ *and* R *is odd* **then** $s_\ell \leftarrow s_\ell + 1$;
if $\ell \rhd x^*$ *and* R *is odd* **then** $s_{j^*} \leftarrow s_{j^*} + 1$;
if $\mathcal{W}_P(b^s) = x^*$ **then return** $(True, s)$;
else return $(False, None)$;

Algorithm 3: Restricted 2PW-H(x^*, ℓ)

Input: $(N, M, \succ, \rhd, k)$, Target candidate x^*, Candidate ℓ
$s \leftarrow s^T$; $R \leftarrow 0$;
while $\exists c \in M \setminus \{\ell\}$ *s.t.* $s_c \geq s_\ell - \mathbb{1}_{c\neq x^*\rhd\ell} + \mathbb{1}_{\ell\rhd c=x^*}$ *and* $R < k$ **do**
  $y \leftarrow \arg\min_{x^*,\ell}\{s_{x^*}, s_\ell - 1\}$; $s_y \leftarrow s_y + 1$;
  $s_c \leftarrow s_c - 1$; $R \leftarrow R + 1$;
while $s_{x^*} < s_\ell - \mathbb{1}_{x^*\rhd\ell}$ *and* $R < k$ **do**
  $j^* \leftarrow \arg\max_{j\in M\setminus\{\ell\}} s_j$;
  if $s_\ell > \max_{j\in M\setminus\{\ell\}} s_j + 2$ **then** $j^* \leftarrow \ell$;
  $s_{x^*} \leftarrow s_{x^*} + 1$; $s_{j^*} \leftarrow s_{j^*} - 1$; $R \leftarrow R + 1$;
if $\mathcal{W}_P(b^s) = x^*$ **then return** $(True, s)$;
else return $(False, None)$;

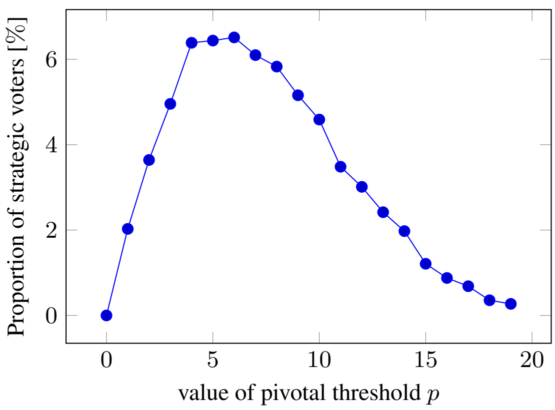

Figure 1. Proportion of strategic voters depending on the pivotal threshold p in an election with 100 voters and 4 candidates under impartial culture.