Probabilistic Aspects of Voting, Intransitivity and Manipulation

Elchanan Mossel*

Abstract

Marquis de Condorcet, a French philosopher, mathematician, and political scientist, studied mathematical aspects of voting in the eighteenth century. Condorcet was interested in studying voting rules as procedures for aggregating noisy signals and in the paradoxical nature of ranking 3 or more alternatives. These notes will survey some of the main mathematical models, tools, and results in a theory that studies probabilistic aspects of social choice. Our journey will take us through major results in mathematical economics from the second half of the 20th century, through the theory of boolean functions and their influences and through recent results in Gaussian geometry and functional inequalities.

*MIT, Cambridge MA, USA. E-mail: `elmos@mit.edu`. Partially supported by NSF Grant CCF 1665252 and DOD ONR grant N00014-17-1-2598

1 Introduction

Marquis de Condorcet, a French philosopher, mathematician and political scientist, studied mathematical aspects of voting in the eighteenth century. It is remarkable that already in the 18th century Condorcet was an advocate of equal rights for women and people of all races and of free and equal public education [64]. His applied interest in democratic processes led him to write an influential paper in 1785 [14], where in particular he was interested in voting as an aggregation procedure and where he pointed out the paradoxical nature of voting in the presence of 3 or more alternatives.

1.1 The law of large numbers and Condorcet Jury Theorem

In what is known as Condorcet Jury Theorem, Condorcet considered the following setup. There are n voters and two alternatives denoted $+$ (which stands for $+1$) and $-$ (which stands for -1). Each voter obtains a *signal* which indicates which of the alternatives is preferable. The assumption is that there is an a priori better alternative and that each voter independently obtains the correct information with probability $p > 1/2$ and incorrect information with probability $1-p$. The n voters then take a majority vote to decide the winner. Without loss of generality we may assume that the correct alternative is $+$ and therefore the individual signals are i.i.d RV x_i where $\mathbb{P}[x_i = +] = p$ and $\mathbb{P}[x_i = -] = 1-p$. Let m denote the Majority function, i.e the function that returns the most popular values among its inputs. Condorcet Jury Theorem is:

Theorem 1.1. *For every $p > 1/2$:*

- $\lim_{n \text{ odd} \to \infty} \mathbb{P}[m(x_1, \ldots, x_n) = +] = 1.$
- *If $n_1 < n_2$ are odd then* $\mathbb{P}[m(x_1, \ldots, x_{n_1}) = +] < \mathbb{P}[m(x_1, \ldots, x_{n_2}) = +]$.

The first part of the theorem is immediate from the law of large numbers (which was known at the time), so the novel contribution was the second part. In the early days of modern democracy, Condorcet used his model to argue that the more people participating in decision making, the more likely that the correct decision is arrived at. We leave the proof of the second part of the theorem as an exercise.

1.2 Condorcet Paradox and Arrow Theorem

As hinted earlier, things are more interesting when there are 3 or more alternatives. In the same 1785 paper, Condorcet proposed the following "paradox". Consider three voters named $1, 2$ and 3, and three alternatives named a, b and c. Each voter ranks the three alternatives in one of six linear orders. While it is tempting to represent the orders as elements of the permutation group S_3, it will be more useful for us to use the following representation. Voter i preference is given by (x_i, y_i, z_i) where $x_i = +$, if she prefers a

to b and $-$, otherwise, $y_i = +$, if she prefers b to c and $-$, otherwise and $z_i = +$ if she prefers c to a and $-$, otherwise. Each of the 6 rankings corresponds to one of the vectors in $\{-1,1\}^3 \setminus \{\pm(1,1,1)\}$.

Condorcet considered 3 voters, with rankings given by: $a > b > c, c > a > b, b > c > a$, or in our notation by the rows of the following matrix:

$$\begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{pmatrix} = \begin{pmatrix} + & + & - \\ + & - & + \\ - & + & + \end{pmatrix}$$

How should we decide how to aggregate the individual rankings? If we use the majority rule to decide between each pair of preferences, then we apply the majority rule on each of the columns of the matrix and conclude that overall preference is $(+,+,+)$. In other words the overall preference is that $a > b$, the overall preference is that $b > c$ and the overall preference is that $c > a$. This does not correspond to an order! This is what is known as Condorcet Paradox. It is part of a long and interesting discussion in political science, economics, mathematics and computer science on how to aggregate rankings.

Of particular interest to us is to follow Ken Arrow who asked if perhaps the paradox is the result of using the majority function to decide between every pair of alternatives? Can we avoid paradoxes if we aggregate pairwise preferences using a different function?

One function that never results in paradoxes is the dictator function $f(x) = x_1$ as the aggregate ranking is $(x_1, y_1, z_1) \neq \pm(1,1,1)$.

Arrow in his famous theorem proved [2, 3]:

Theorem 1.2. *Let $f : \{-1,1\}^n \to \{-1,1\}$ and suppose that f never results in a paradox, so for all $(x_i, y_i, z_i) \neq \pm(1,1,1)$ it holds that $(f(x), f(y), f(z)) \neq \pm(1,1,1)$. Then f is a dictator: there exists an i such that $f(x) = x_i$ for all x, or there exists an i such that $f(x) = -x_i$ for all x.*

With the right notation and formulation the proof of Arrow's Theorem is very short [4, 43]:

Proof. First note that if f is a constant function, then the outcome is always $\pm(1,1,1)$. Suppose that f is not a dictator and not a constant, then f depends on at least two coordinates. Without loss of generality, let these coordinates be 1 and 2. Therefore:

$$\exists x_2, \ldots, x_n : \quad f(+, x_2, \ldots, x_n) \neq f(-, x_2, \ldots, x_n)$$

$$\exists y_1, y_3, \ldots, y_n : \quad f(y_1, +, y_3, \ldots, y_n) \neq f(y_1, -, y_3, \ldots, y_n)$$

We now choose $z_1 = -y_1$ and $z_i = -x_i$ for $i \geq 2$. This guarantees that $(x_i, y_i, z_i) \neq \pm(1,1,1)$ for all i no matter what the values of x_1 and y_2 are. This can be also verified from the matrix in (1). Now choose x_1 and y_2 so that $f(x) = f(y) = f(z)$ to arrive at a contradiction.

$$\begin{pmatrix} x_1 & y_1 & -y_1 \\ x_2 & y_2 & -x_2 \\ \vdots & \vdots & \vdots \\ x_n & y_n & -x_n \end{pmatrix} \tag{1}$$

□

Arrow also considered a more general setting where f, g and h are allowed to be different functions, there are other functions that result in 0 probability of paradox. For example if $f = 1$ and $g = -1$ for all inputs then h can be arbitrary. This corresponds to the case where the second alternative b is ignored (always ranked last) and the choice between a and c is determined by h. Arrow Theorem in this setting can be stated as follows:

Theorem 1.3. *Let $f, g, h : \{-1,1\}^n \to \{-1,1\}$ and suppose that (f, g, h) never results in a paradox, so for all $(x_i, y_i, z_i) \neq \pm(1,1,1)$ it holds that $(f(x), g(y), h(z)) \neq \pm(1,1,1)$. Then either two of the functions are constant of opposite sign, or there exists an i such that f, g and h are dictator on voter i.*

Proof. If two of the functions, say f and g take the same constant value and the third function h is not constant, then clearly one can find x, y, z such that $f(x) = g(y) = h(z)$ and $(x, y, z) \neq \pm(1,1,1)$. So WLOG we may assume at least two of the functions, say f and g are not constant. Let $A(f)$ denote the set of non-zero influence variables of f and similarly $A(g)$ and $A(h)$. Since f and g are not constant $A(f)$ and $A(g)$ are not empty. If there exists a variable $i \in A(f)$ and a variable $i \neq j \in A(g)$, then by the same argument as in Theorem 1.2, there exist (x, y, z) resulting in a paradox. Thus, the only case remaining is where $A(f) = A(g) = i$ and $A(h) = i$ or $A(h) = \emptyset$. In either case, the functions f, g and h are all functions of variable i only. It is now easy to verify that it must be the case that $f = g = h$ is a dictator on voter i. □

1.3 Manipulation and the GS Theorem

A naturally desirable property of a voting system is *strategyproofness* (a.k.a. nonmanipulability): no voter should benefit from voting strategically, i.e., voting not according to her true preferences. However, Gibbard [27] and Satterthwaite [62] showed that no reasonable voting system can be strategyproof. Before stating their result, let us specify the problem more formally.

The setting here is different than the setup of Arrow's Theorem: We consider n voters electing a winner among k alternatives. The voters specify their opinion by ranking the alternatives, and the winner is determined according to some predefined *social choice function* (SCF) $f : S_k^n \to [k]$ of all the voters' rankings, where S_k denotes the set of all possible total orderings of the k alternatives. We call a collection of rankings by the voters a *ranking* *profile*. We say that an SCF is *manipulable* if there exists a ranking profile where a voter can achieve a more desirable outcome of the election according to her true preferences by voting in a way that does not reflect her true preferences.

For example, consider Borda voting, where each candidate receives a score which is the sum of its ranks and the candidate with the lowest score wins. If the individual rankings are: $(abcd), (cadb)$, then a is the winner, but if the second voter were to vote $(cdba)$ instead then c will become the winner, so the second voter will want to vote untruthfully.

Theorem 1.4 (Gibbard-Satterthwaite)**.** *Any Social Choice Function which is not a dictatorship (i.e., not a function of a single voter), and which allows at least three alternatives to be elected, is manipulable.*

This theorem has contributed to the realization that it is unlikely to expect truthfulness in voting. There are many proofs of the Gibbard-Satterthwaite theorem, but all are more complex than the proof of Arrow's theorem given above. We will not provide a proof of the theorem in these notes.

2 Modern Perspectives

Work since the 1980s addressed novel aspects of aggregation of votes. Condorcet Jury Theorem assumes a probability distribution over the voters but is restricted to a specific aggregation function (majority) while Arrow theorem considers general aggregation functions but involves no probability model. There are many interesting questions that can be asked by combining the two perspectives. First, it is natural to ask about aggregation properties of Boolean functions. The study of aggregation properties of Boolean functions was fundamental to the development of the area of "Analysis of Boolean Functions" since the 1980s, starting with [8] and [35]. Second, we can ask questions regarding the probability of manipulation and paradoxes, questions that were analyzed since the early 2000s, starting with the works [36, 37, 26].

In terms of mathematical theory, the main statement of quantitative social theory have the same form as statements in property testing **add citations** and some of the main statements in additive combinatorics. In terms of the techniques, we will see that the main techniques have discrete isopetimetric flavor **add citations**. We will cover the following topics:

We will start by studying the question of noise stability of Boolean functions [9, 50]. This will lead us to discussing some of the main analytical tools in the area, including notably hyper-contraction [10, 57, 28, 6], the invariance principle [50] and Borell's Gaussian noise-stability result [12].

We will then devote a considerable effort to proving quantitative versions of Arrow Theorem. We will do so by proving a Gaussian version of Arrow theorem, as well as a quantitive version using reverse hyper-contraction [11]. Combining the two we will prove a general quantitive Arrow theorem [45].

Different tools were used in proving different quantitive versions of the manipulation theorem. The first proof, which applies only to 3 alternatives, uses a reduction to a quantitive Arrow Theorem [26, 24]. We will follow later proofs which apply in the case of a larger number of coordinates and use reverse hyper-contraction as a main tool [31, 54]. The classical proofs of manipulation theorems often use long paths of voting profiles. The most general proof in [54] will quantify such arguments using geometric tools from the theory of Markov chains.

In the last part of the notes we will review some of the more classical results of the aggregation power of Boolean functions and also discuss some future directions including studying the same questions for non-product measures.

We note that much of the interest in quantitative social choice theory comes from artificial intelligence and computer science, where virtual elections are now an established tool in preference aggregation (see the survey by Faliszewski and Procaccia [21]). Many of the results in the study of social choice are negative: it is impossible to design a voting system that satisfies a few desired properties all at once.

Recall that Gibbard-Satterthwaite theorem states that any SCF which is not a dictatorship (i.e., not a function of a single voter), and which allows at least three alternatives to be elected, is manipulable. This has contributed to the realization that it is unlikely to expect truthfulness in voting. Consequently, there have been many branches of research devoted to understanding the extent of manipulability of voting systems, and to finding ways of circumventing the negative results.

One approach, introduced by Bartholdi, Tovey and Trick [5], suggests computational complexity as a barrier against manipulation: if it is computationally hard for a voter to manipulate, then she would just tell the truth (we refer to the survey by Faliszewski and Procaccia [21] for a detailed history of the surrounding literature). This is a worst-case approach, and while worst-case hardness of manipulation is a desirable property for a SCF to have, this does not tell us anything about *typical* instances of the problem—is it easy or hard to manipulate *on average*?

The quantitative versions of Arrow and the GS theorem we will see below show that typically ranking is paradoxical and manipulation is easy on average. This follows a long line of research which proved such results for restricted classes of SCFs, including, e.g., Kelly [39], Conitzer and Sandholm [15], and Procaccia and Rosenschein [61] (see also the survey [21]).

3 Reading

The recommended reading for various topics is as follows:

- Noise Stability: We follow the approach of [17, 18]. The original approach for the proofs of noise stability was developed in [49, 50], see also [59] as a general reference to analysis of Boolean functions from the theoretical computer science perspective. Some of the strongest results on noise stability follow [46].

- Arrow-Kalai Theorem was proved in [36]. It is based on the FKN Theorem which was proven in [25]. More elegant and general proofs can be found in [33].
- The proof of the Gaussian and general Arrow Theorem follows [45]. Tighter results were later obtained in [38]. The results were later generalized further in [52].
- For the proof of the quantitive versions of the manipulation theorem we will follow [31] and [54], though the first version for the case of 3 alternatives was established in [26, 24].
- The work on aggregation of binary functions is based on seminal work in the 1980s and 1990s including [8] and [35].
- We will finally discuss more general voter distributions and other models in voting such as [29] and [56].

4 A Preview — Noise Stability and the Probability of Paradox for the Majority Function

As a preview of what's to come, we will compute the asymptotic probability of a non-transitive outcome in Condorcet setup with 3 alternatives and where voters vote uniformly at random. Let us denote the Majority function by $m : \{-1,1\}^n \to \{-1,1\}$ and assume that the number of voters n is odd. Paradoxes seem more likely when there is no bias towards a particular candidate so we will consider voters who vote independently and where voter i votes uniformly at random from the 6 possible rankings. Recall that we encode the 6 possible rankings by vectors $(x, y, z) \in \{-1,+1\}^3 \setminus \{\pm(1,1,1)\}$. Here x is $+1/-1$ if a voter ranks a above/below b, y is $+1/-1$ if voter ranks b above/below c, z is $+1/-1$ if voter ranks c above/below a.

How do we analyze the probability of a paradox? The following simple fact was used in [36]: Since the binary predicate $\psi : \{-1,1\}^3 \to \{0,1\}, \psi(a,b,c) = 1(a = b = c)$ can be expressed as

$$\psi(a,b,c) = \frac{1}{4}(1 + ab + ac + bc),$$

we can write

$$\mathbb{P}[m(x) = m(y) = m(z)] = \frac{1}{4}(1 + \mathbb{E}[m(x)m(y)] + \mathbb{E}[m(x)m(z)] + \mathbb{E}[m(y)m(z)]),$$

which due to symmetry can be written as

$$\mathbb{P}[m(x) = m(y) = m(z)] = \frac{1}{4}(1 + 3\mathbb{E}[m(x)m(y)]).$$

Moreover, the uniform distribution over $\{\pm 1\}^3 \setminus \{\pm(1,1,1)\}$ satisfies $\mathbb{E}[x_i y_i] = \mathbb{E}[y_i z_i] = \mathbb{E}[z_i x_i] = -1/3$ and the n coordinates are independent. As we will see shortly, we call the the quantity $\mathbb{E}[m(x)m(y)]$ is the noise stability of m with noise parameter $-1/3$. Its asymptotic value as $n \to \infty$ is easy to compute using a 2-dimensional CLT to obtain:

$$\lim_{n\to\infty} \mathbb{E}[m(x)m(y)] = \mathbb{E}[\operatorname{sgn}(X)\operatorname{sgn}(Y)],$$

where $X, Y \sim N(0, \begin{pmatrix} 1 & -\frac{1}{3} \\ -\frac{1}{3} & 1 \end{pmatrix})$ and we can see that

$$\mathbb{E}[\operatorname{sgn}(X)\operatorname{sgn}(Y)] = 2\mathbb{P}[\operatorname{sgn}(X) = \operatorname{sgn}(Y)] - 1 = 1 - \frac{2\arccos(-1/3)}{\pi}$$

and

$$\lim_{n\to\infty} \mathbb{P}[m(x) = m(y) = m(z)] = 1 - \frac{3\arccos(-1/3)}{2\pi} \approx 0.088\ .$$

5 Noise Stability

5.1 Boolean Noise Stability

Consider the following thought experiment: suppose the voters in binary voting obtain independent uniform signals: $x_i = +$ or $x_i = -$ with probability $1/2$. This is the same setting as in Condorcet Jury Theorem except the voters are completely uninformed.

Now consider the following process that produces a vector y as a *noisy version* of x. For each i independently: let $y_i = x_i$ with probability $(1+\theta)/2$ and $y_i = -x_i$ with probability $(1-\theta)/2$, where $\theta \in [-1,1]$. We chose the parametrization so that $\mathbb{E}[x_i y_i] = \theta$.

How should we interpret y? A simple interpretation is as a noise process of voting machines. Suppose that when each voter votes, there is a small probability, say 0.01 that the voting machine records the opposite vote (independently for all voters and independently of the intended vote). In this case $\theta = 0.98$. Given a voting aggregation function $f : \{-1,1\}^n \to \{-1,1\}$, ideally we would like the quantity

$$\mathbb{P}[f(x) = f(y)] = \frac{1}{2}(1 + \mathbb{E}[f(x)f(y)])$$

to be as large as possible if $\theta > 0$ and as small as possible if $\theta < 0$. The quantity $\mathbb{E}[f(x)f(y)]$ is called the *noise stability* of f. More generally, following [9] we define:

Definition 5.1. For two functions $f, g : \{-1,1\}^n \to \mathbb{R}$ the (ρ-)*noisy inner product* of f and g denoted by $\langle f, g\rangle_\rho$ is defined by $\mathbb{E}[f(x)g(y)]$, where $((x_i, y_i) : 1 \le i \le n)$ are i.i.d. mean 0 ($\mathbb{E}[x_i] = \mathbb{E}[y_i] = 0$) and ρ-correlated ($\mathbb{E}[x_i y_i] = \rho$). The *noise stability* of f is its noisy inner product with itself: $\langle f, f\rangle_\rho$.

We can also write the noisy inner product in terms of the *noise operator* T_ρ,

Definition 5.2. The Markov operator T_ρ maps functions $f : \{-1,1\}^n \to \mathbb{R}$ to functions $T_\rho f : \{-1,1\}^n \to \mathbb{R}$. It is defined by:

$$(T_\rho f)(x) = \mathbb{E}[f(y)|x].$$

The noise operator is also known as the Bonami-Beckner operator and plays a key role in the theory of hyper-contraction [10, 6]. Note that

$$\langle f, g\rangle_\rho = \mathbb{E}[f(x)g(y)] = \mathbb{E}[f T_\rho g] = \mathbb{E}[g T_\rho f] = \langle Tf, g\rangle = \langle f, Tg\rangle$$

and that

$$\langle f, g\rangle := \mathbb{E}[f(x)g(x)] = \langle f, g\rangle_1.$$

Basic properties of this operator can be revealed using its eigenfunctions, i.e., the Fourier basis. The following proposition is straightforward to prove:

Proposition 5.3. *For $S \subset [n]$, write $x_S = \prod_{i\in S} x_i$, so $x_\emptyset \equiv 1$. Then:*

- *$(x_S : S \subset [n])$ is an orthonormal basis for the space of all functions $f : \{-1,1\}^n \to \mathbb{R}$.*
- *x_S is an eigenfunction of T_ρ which corresponds to the eigenvalue $\rho^{|S|}$: $T_\rho x_S = \rho^{|S|} x_S$.*

The following easy result is folklore (see e.g. [51]).

Theorem 5.4. *For every $\rho > 0$, for every n and for every $f, g : \{-1,1\}^n \to \{-1,1\}$ with $\mathbb{E}[f] = \mathbb{E}[g] = 0$, it holds that*

$$\langle f, g\rangle_\rho \le \langle x_1, x_1\rangle_\rho = \rho.$$

$$\langle f, g\rangle_{-\rho} \ge \langle x_1, x_1\rangle_{-\rho} = -\rho.$$

Moreover, the only optimizers are dictator functions, i.e., functions of the form $f(x) = g(x) = x_i$ or $f(x) = g(x) = -x_i$.

Proof. Note that $\mathbb{E}[f] = \langle f, 1\rangle = \langle f, x_\emptyset\rangle$. So if $\mathbb{E}[f] = 0$, then: $f = \sum_{S\neq\emptyset} \hat{f}(S) x_S$, where $\hat{f}(S) = \langle f, x_S\rangle$. Moreover,

$$T_\rho f(x) = \mathbb{E}[f(y)|x] = \sum_S \rho^{|S|} \hat{f}(S) x_S,$$

and similarly for g. Therefore:

$$\langle f, g\rangle_\rho = \langle T_\rho f, g\rangle = \langle \sum_{S\neq\emptyset} \rho^{|S|} \hat{f}(S) x_S, \sum_{S\neq\emptyset} \hat{g}(S) x_S\rangle \tag{2}$$

$$= \sum_{S\neq\emptyset} \rho^{|S|} \hat{f}(S)\hat{g}(S) \le \rho \sqrt{\sum_{S\neq\emptyset} \hat{f}^2(S) \sum_{S\neq\emptyset} \hat{g}^2(S)} = \rho \tag{3}$$

where the last inequality uses Cauchy-Schwarz and Parseval's identity

$$\sum_S \hat{f}^2(S) = \langle f, f \rangle = 1.$$

In the case of equality we conclude that $f = g$ must be a linear function, so $f = \sum_i a_i x_i$. Since f is Boolean,

$$1 \equiv f^2 = \sum_i a_i^2 + \sum_{i \neq j} a_i a_j x_i x_j$$

and therefore it must be the case that f is a dictator. The case where $\rho < 1$ is similar and is left as an exercise. □

The Theorem above allow a quick proof of a version of Arrow's Theorem by Kalai [36]:

Corollary 5.5. *In the context of Arrow Theorem if $\mathbb{E}[f] = \mathbb{E}[g] = \mathbb{E}[h] = 0$ and $\mathbb{P}[f(x) = g(y) = h(z)] = 0$ then f, g and h are all the same dictator.*

Proof. Use the previous theorem and:

$$\mathbb{P}[f(x) = g(y) = h(z)] = \frac{1}{4}\left(1 + \langle f, g\rangle_{-1/3} + \langle g, h\rangle_{-1/3} + \langle h, f\rangle_{-1/3}\right)$$

□

From the voting perspective, there is something a little unnatural about the result above. It says that if we want to maximize robustness, then a dictator should decide. From a mathematical perspective, it is disappointing that there is something special about $\mathbb{E}[f] = 0$. In particular the following is open:

Problem 5.6. *For a generic $\rho > 0, 0 < \mu < 1$ what is the value of:*

$$\lim_{n \to \infty} \max\Big(\langle f, f\rangle_\rho : f : \{-1,1\}^n \to \{-1,1\}, \mathbb{E}[f] = \mu\Big)$$

Here are some additional examples:

- Similar argument to the theorem shows that if $\rho > 0$ then $\langle f, f\rangle_\rho \geq \rho^n$ for $f : \{-1,1\}^n \to \{-1,1\}$. The parity function $x_{[n]}$ achieves equality.
- The asymptotic noise stability of Majority is given by Sheppard formula [63], i.e., $\mathbb{E}[\operatorname{sgn}(N)\operatorname{sgn}(M)]$, where (N, M) are ρ-correlated random variables:

  $$\mathbb{E}[\operatorname{sgn}(M)\operatorname{sgn}(N)] = 2\mathbb{P}[\operatorname{sgn}(M) = \operatorname{sgn}(N)] - 1 = 1 - \frac{2\arccos(\rho)}{\pi} := \kappa(\rho)$$

  In particular if $\rho = 1 - \varepsilon$, then $\mathbb{P}[f(x) \neq f(y)]$ is of order $\sqrt{\varepsilon}$. Compare this to a dictator where it is of order ε.

- If we consider $n = r^2$ where r is odd and the function f implements electoral college, i.e.,
$$f(x_1, \dots, x_n) = m\Big(m(x_1, \dots, x_r), \dots, m(x_{n-r+1}, x_n)\Big),$$
then it is easy to see that asymptotically the noise stability is given by $\kappa(\kappa(\rho))$. In particular if $\rho = 1 - \varepsilon$ for small ε then $\mathbb{P}[f(x) \neq f(y)]$ is of order $\varepsilon^{1/4}$.

- Let m_r be majority function on r voters and define $m_r^{(1)} = m_r$ and by induction:
$$m_r^{(h)}(x_1, \dots, x_{r^h}) = m_r^{(h-1)}\left(m_r(x_1, \dots, x_r), \dots, m_r(x_{r^h-r+1}, \dots, x_{r^h})\right).$$
This function is called the recursive majority function.

Exercise 5.7 ([48]). *Show that for every $\varepsilon < 0.5$, if r is large enough and $n_h = r^h$ and if $\rho = 1 - n_h^{-\varepsilon}$ then*
$$\lim_{h\to\infty} \mathbb{E}[m_r^{(h)}(x) m_r^{(h)}(y)] = 0.$$

5.2 Gaussian Noise Stability

We will now take a detour of considering analogous quantities defined in Gaussian space. We will later see that this is quite useful in the Boolean setting.

Definition 5.8. In Gaussian space, the (ρ-)noisy inner product of $\phi : \mathbb{R}^n \to \mathbb{R}$ and $\psi : \mathbb{R}^n \to \mathbb{R}$ denoted by $\langle \phi, \psi \rangle_\rho$ is defined as
$$\mathbb{E}[\phi(M)\psi(N)],$$
where $((M_i, N_i))_{i=1}^n$ are i.i.d. two-dimensional Gaussian vectors, such that N_i, M_i are standard (mean 0, variance 1) Gaussian random variables and $\mathbb{E}[N_i M_i] = \rho$. The *noise stability* of ϕ is its inner product with itself: $\langle \phi, \phi \rangle_\rho$.

We will generally use f, g etc. to denote functions over the Boolean cube and ϕ, ψ etc. for functions in $L_2(\mathbb{R}^n, \gamma)$. In particular, for $\mu \in [0,1]$ we write χ_μ for the indicator of the interval $(-\infty, \Phi^{-1}(\mu))$ whose Gaussian measure is μ.

We can now state Borell's [12] noise stability result:

Theorem 5.9. *For all $n \geq 1$, $\rho > 0$ and $\phi, \psi : \mathbb{R}^n \to [0,1]$, it holds that:*
$$\langle 1 - \chi_{1-\mathbb{E}\phi}, \chi_{\mathbb{E}\psi} \rangle_\rho \leq \langle \phi, \psi \rangle_\rho \leq \langle \chi_{\mathbb{E}\phi}, \chi_{\mathbb{E}\psi} \rangle_\rho$$
$$\langle 1 - \chi_{1-\mathbb{E}\phi}, \chi_{\mathbb{E}\psi} \rangle_{-\rho} \geq \langle \phi, \psi \rangle_{-\rho} \geq \langle \chi_{\mathbb{E}\phi}, \chi_{\mathbb{E}\psi} \rangle_{-\rho}$$

Borell [12] was interested in more general functionals of the heat equations and he showed that these functionals increase with respect to nonincreasing spherical rearrangement. The fact that half spaces are the unique optimizers of ρ-noisy inner product was proven in [47], where a robust version of the theorem is also proven. Tighter robust versions were later proven by Eldan [20]. Other alternative proofs and generalization of Borell's result include [32, 40].

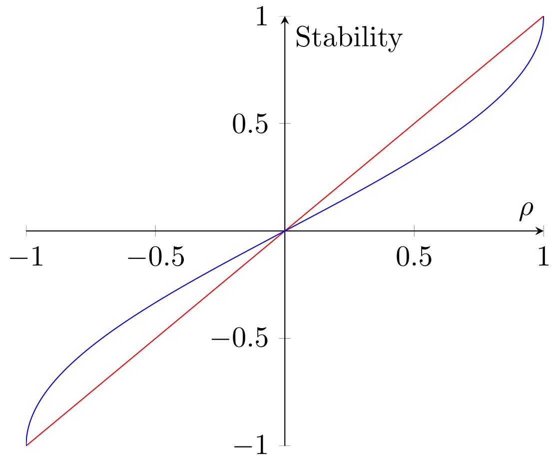

Figure 1: The noise stability of dictator and Gaussian half-space of measure 0.5. i.e., functions ρ and $1 - \arccos \rho / 2\pi$. Note that for every $0 < \rho < 1$, the dictator is more stable than the corresponding half-spaces and for every $-1 < \rho < 0$, it is less stable than the corresponding half-space

5.3 Gaussian and Boolean Noise Stability

By applying the CLT, it is easy to check that Gaussian noise stability provides a bound on the Boolean noise stability.

Proposition 5.10. *For every $\rho \in [-1, 1], \mu, \nu \in [0, 1]$, and for every*

$$s \in \Big[\langle 1 - \chi_{1-\mu}, \chi_\nu \rangle_\rho, \langle \chi_\mu, \chi_\nu \rangle_\rho\Big],$$

there exists sequence of Boolean functions $f_n, g_n : \{-1, 1\}^n \to \{0, 1\}$ such that $\mathbb{E}[f_n] \to \mu, \mathbb{E}[g_n] \to \nu$ and

$$\langle f_n, g_n \rangle_\rho \to s$$

Moreover, by Theorem 5.9 and Theorem 5.4 it follows that there is a strict inequality at $\mu = 1/2$. See Figure 1.

The proof of the proposition is standard using approximation of Gaussian random variables in terms of sums of independent Bernoullis:

Proof. To show that we can obtain the RHS of the interval, let

$$f_n = \chi_\mu(n^{-1/2} \sum_{i=1}^{n} x_i), \quad g_n = \chi_\nu(n^{-1/2} \sum_{i=1}^{n} x_i),$$

and apply the CLT. The proof of achievability of the LHS is identical. To obtain an intermediate point s take the inputs of f_n and g_n to be defined on overlapping blocks of bits, e.g.:

$$f_n = \chi_\mu(n^{-1/2} \sum_{i=\alpha n}^{(1+\alpha)n} x_i), \quad g_n = \chi_\nu(n^{-1/2} \sum_{i=1}^{n} x_i).$$

□

5.4 Smooth Boolean Functions

To better understand the connection between Boolean and Gaussian stability, we define two notions of smoothness, termed low influences and resilience for Boolean functions. We begin with the notion of influence, that was first defined in [8] and [35]. We will do so for general product probability spaces. To simplify notation we will often omit the sigma algebra and probability measure defined over a probability space Ω.

Definition 5.11. Consider a probability space Ω. For a function $f : \Omega^n \to \mathbb{R}$, we define the i'th influence of f as

$$I_i(f) = \mathbb{E}\big[\operatorname{Var}[f|x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n]\big],$$

where the expected value is with respect to the product measure on Ω^n. In the Boolean case with the uniform measure $f : \{-1,1\}^n \to \mathbb{R}$, the influence is equivalently defined as

$$I_i(f) = \mathbb{E}\big[\operatorname{Var}[f|x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n]\big] = \sum_{S:i\in S} \hat{f}^2(S).$$

Or as

$$I_i(f) = \mathbb{E}[|\partial_i f|^2],$$

where $(\partial_i f)(x_1, \ldots, x_n) = 0.5(f(x_1, \ldots, x_{i-1}, +, x_{i+1}, \ldots, x_n) - f(x_1, \ldots, x_{i-1}, -, x_{i+1}, \ldots, x_n))$ is the discrete i'th directional derivative.

An easy corollary of the definition is that for $|\rho| < 1$, it holds that $T_\rho f$ is small in the sense that the sum of its influences is bounded as a function of ρ only.

Lemma 5.12. *Let* $f : \{-1,1\}^n \to [-1,1]$ *and* $|\rho| < 1$ *then*

$$\sum_{i=1}^{n} I_i(T_\rho f) \le 1/(1 - |\rho|).$$

Proof.

$$\begin{aligned}\sum_{i=1}^{n} I_i(T_\rho f) &= \sum_{i=1}^{n}\sum_{S:i\in S}\widehat{T_\rho f}^2(S)\\ &= \sum_{S}|S|\widehat{T_\rho f}^2(S) = \sum_S |S|\rho^{|S|}\hat{f}^2(S)\\ &\le \max_k |\rho|^k k\sum_S \hat{f}^2(S) \le \max_k |\rho|^k k \le 1/(1-|\rho|)\end{aligned}$$

The proof follows. □

For a voting function $f : \{-1,1\}^n \to \{-1,1\}$, $I_i(f)$ is the probability that voter i is the deciding voter, given all other votes. A stronger notion of power of a voter or a small set of voters is that their vote affects the expected outcome on average. A function whose expectation is not affected by any small set of voters is called resilient. More formally,

Definition 5.13. We say that a function $f : \Omega^n \to \mathbb{R}$ is (r,α)-*resilient* if

$$\Big|\mathbb{E}\left[f|X_S = z\right] - \mathbb{E}[f]\Big| \le \alpha. \tag{4}$$

for all j, all sets S with $|S| \le r$ and all $z \in \Omega^S$.

Proposition 5.14. *If $f : \{-1,1\}^n \to \mathbb{R}$ satisfies*

$$\max(|\hat{f}(S)| : 0 < |S| \le r) \le 2^{-r}\alpha, \tag{5}$$

then f is (r,α)-resilient. In particular if f has all influences are bounded $4^{-r}\alpha^2$ then f is (r,α)-resilient.

Proof. The second statement follows from the first one immediately as for every non-empty S, we may choose $i \in S$ and then

$$\hat{f}^2(S) \le I_i(f) \le 4^{-r}\alpha^2,$$

as needed. For the first statement, assume (5). Then:

$$\Big|\mathbb{E}\left[f|X_S = z\right] - \mathbb{E}[f]\Big| = \Big|\mathbb{E}[\sum_{T\neq\emptyset}\hat{f}(T) z_{S\cap T} x_{T\setminus S}]\Big| = \Big|\sum_{\emptyset\neq T\subset S}\hat{f}(T) z_T\Big| \le 2^{|S|}2^{-r}\alpha \le \alpha.$$

□

Resilient functions have long been studied in the context of pseudo-randomness, see e.g. [13].

Thus, the statement that a function has a high influence variable means that there exists a voter i that can have a noticeable effect on the outcome *if voter i has access to all other votes cast.* The statement that a function is not resilient implies that there is a bounded set of voters who have noticeable effect on the outcome *on average*, i.e., with no access to other votes cast. Consider the following examples:

- Dictator has maximal influence of 1 (and all other 0). It is also not resilient for $r \geq 1, \alpha < 1$.
- Majority has all influences of order $n^{-1/2}$ and is also $(r, O(r/\sqrt{n}))$ resilient.
- An example of a resilient function with a high influence variable is the function

$$f(x) = x_1 \operatorname{sgn}(\sum_{i=2}^{n} x_i).$$

  Here coordinate 1 has influence 1 but the function is resilient. In terms of voting, voter 1 has a lot of power if she has access to all other votes cast (or the majority of the votes) but without access to this information, she is powerless. Moreover, every small set of k voters can change the expected value of f (by conditioning on their vote) by $O(kn^{-1/2})$. Another simple example is the parity function $\prod_{i=1}^{n} x_i$, which is $(r, 0)$ resilient for every $r < n$, but where all influences are 1.

The Majority is Stablest Theorem states that the extremal noise stability of low influence / resilient functions on the discrete cube is captured by Gaussian noise stability. Here are three increasingly stronger statements along this line:

Theorem 5.15 ([49, 50])**.** *For every $\varepsilon > 0, 0 \leq \rho < 1$, there exists a $\tau > 0$ for which the following holds. Let $f, g : \{-1,1\}^n \to [0,1]$ satisfy $\max(I_i(f), I_i(g)) < \tau$ for all i. Then*

$$\langle f, g \rangle_\rho \leq \langle \chi_{\mathbb{E}f}, \chi_{\mathbb{E}g} \rangle_\rho + \varepsilon.$$

This theorem is called Majority Is Stablest since $\langle \chi_{\mathbb{E}f}, \chi_{\mathbb{E}g} \rangle_\rho = \lim_{n\to\infty} \langle f_n, g_n \rangle_\rho$, where $f_n(x) = \chi_{\mathbb{E}f}(n^{-1/2} \sum_{i=1}^n x_i)$ and $g_n(x) = \chi_{\mathbb{E}g}(n^{-1/2} \sum_{i=1}^n x_i)$.

It turns out that for two functions, it is in fact enough that one of them is low influence to obtain the same results, i.e.:

Theorem 5.16 ([44], Prop 1.15)**.** *For every $\varepsilon > 0$ and $0 \leq \rho < 1$, there exists a $\tau(\rho, \varepsilon) > 0$ for which the following holds. Let $f, g : \{-1,1\}^n \to [0,1]$ be such that $\min(I_i(f), I_i(g)) < \tau$ for all i. Then*

$$\langle f, g \rangle_\rho \leq \langle \chi_{\mathbb{E}f}, \chi_{\mathbb{E}g} \rangle_\rho + \varepsilon, \tag{6}$$

where one can take

$$\tau = \varepsilon^{O\left(\frac{\log(1/\varepsilon)\log(1/(1-\rho))}{(1-\rho)\varepsilon}\right)}. \tag{7}$$

In particular the statement above holds when $\max_i I_i(f) < \tau$ and g is any *Boolean function bounded between 0 and 1.*

Moreover, one can replace the low influence condition by the condition that the function is resilient:

Theorem 5.17 ([46]). *For every $\varepsilon > 0, 0 \le \rho < 1$, there exist $r, \alpha > 0$ for which the following holds. Let $f : \{-1,1\}^n \to [0,1]$ be (r,α)-resilient and let $g : \{-1,1\}^n \to [0,1]$ be an arbitrary function. Then*

$$\langle f, g \rangle_\rho \le \langle \chi_{\mathbb{E}f}, \chi_{\mathbb{E}g} \rangle_\rho + \varepsilon.$$

One can take

$$r = O\left(\frac{1}{\varepsilon^2(1-\rho)\tau}\right), \alpha = O\left(\varepsilon 2^{-r}\right), \tag{8}$$

where τ is given by (7).

Note in particular that for our current bounds for τ and for fixed ρ, r is exponential in a polynomial in $1/\varepsilon$ and α is doubly exponential in a polynomial in $1/\varepsilon$. Similar statements for one function were proven before by [60] and appeared in [34].

6 Proof of Majority is Stablest

The original proof of the Majority is Stablest Theorem used a non-linear invariance principle to derive the theorem from Borell's result [49, 50]. We will follow a different proof that gives an independent proof of Borell's result. This proof from [17, 18] is based on induction on dimension in the discrete cube.

First, it will be useful, for $\rho \in [-1,1]$, to define $J_\rho : (0,1)^2 \to [0,1]$ by

$$J_\rho(x, y) := \langle \chi_x, \chi_y \rangle_\rho = \mathbb{P}[N \le \Phi^{-1}(x), M \le \Phi^{-1}(y)],$$

where N, M are jointly normally distributed random variables with the covariance matrix

$$\mathrm{Cov}(N, M) = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

Instead of just proving Borell's result, we will prove a functional form of the result, first stated and proved in [47]. In [47] it is proved that in the Gaussian setup where $\phi, \psi : \mathbb{R}^n \to [0,1]$, and N, M are jointly normal random variables with covariance $\begin{pmatrix} I_n & \rho I_n \\ \rho I_n & I_n \end{pmatrix}$, we have

$$\mathbb{E} J_\rho(\phi(N), \psi(M)) \le J_\rho(\mathbb{E}\phi, \mathbb{E}\psi). \tag{9}$$

Note that $J_\rho(0, x) = 0$ for all x and $J_\rho(1,1) = 1$. Therefore for any two sets A, B:

$$\mathbb{E} J_\rho(1_A(N), 1_B(M)) = \mathbb{P}[N \in A, M \in B].$$

Thus (9) applied to indicator functions, implies Borell's inequality [12] (Theorem 5.9).

Exercise 6.1. *Prove that (9) is in fact, equivalent to Theorem 5.9.*

To prove Majority is Stablest, we would like to show an analogous inequality on the cube, with $f, g : \{-1,1\}^n \to [0,1]$ and X, Y correlated points on $\{-1,1\}^n$. Our main observation is that while inequality (9) is not true in this setup, it is true with some extra error terms on the right hand side. These error terms will diminish in the low influence and Gaussian case. First, let us define the error term:

Definition 6.2. Let Ω be a probability space and $f : \Omega^n \to \mathbb{R}$. Consider the martingale defined as

$$f_0 = f, f_1 = \mathbb{E}[f|X_2, \ldots, X_n], \ldots, f_i = \mathbb{E}[f|X_{i+1}, \ldots, X_n], \ldots, f_n = \mathbb{E}[f],$$

and let

$$\Delta_m(f) = \sum_{i=1}^{m} \mathbb{E}[|f_i - f_{i-1}|^3],$$

for $m \leq n$.

Note that by orthogonality of martingale increments we know that

$$\sum_{i=1}^{n} \mathbb{E}[|f_i - f_{i-1}|^2] = \mathrm{Var}[f],$$

and by Jensen's inequality, it follows that

Exercise 6.3.

$$\mathbb{E}[|f_i - f_{i-1}|^2] \leq I_i(f).$$

We therefore expect $\Delta_n(f)$ to be small when the differences $|f_i - f_{i-1}|$ are typically small, i.e., f is smooth to changing one coordinate at a time in some average sense.

Definition 6.4. Let Ω_1 and Ω_2 be two sets and μ be a probability measure on $\Omega_1 \times \Omega_2$. We say that μ has Rényi correlation at most ρ if for every measurable (w.r.t μ) $f : \Omega_1 \to \mathbb{R}$ and $g : \Omega_2 \to \mathbb{R}$ with $\mathbb{E}_\mu f = \mathbb{E}_\mu g = 0$,

$$|\mathbb{E}_\mu[fg]| \leq \rho\sqrt{\mathbb{E}_\mu[f^2]\mathbb{E}_\mu[g^2]}.$$

For example, suppose that $\Omega_1 = \Omega_2$ and suppose (X, Y) are generated by the following procedure: first choose X according to some distribution ν. Then, with probability ρ, we set $Y = X$, and with probability $1 - \rho$, Y is chosen to be an independent sample from ν. If μ is the distribution of (X, Y), then it is easy to check that μ has Rényi correlation ρ. In particular, in the definition of noisy inner product, we let $(x, y) \in \{-1,1\}^2$ have $\mathbb{E}[y] = \mathbb{E}[y] = 0$ and $\mathbb{E}[xy] = \rho$ and (x, y) have Rényi correlation ρ.

We prove the following general theorem, which we will later use to derive both Borell's inequality and the "Majority is Stablest" theorem.

Theorem 6.5. *For any $\epsilon > 0$ and $0 < \rho < 1$, there is $C(\rho) > 0$ such that the following holds. Let μ be a measure on $\Omega_1 \times \Omega_2$ with Rényi correlation at most ρ and let $(X_i, Y_i)_{i=1}^n$ be i.i.d. variables with distribution μ. Then for any measurable functions $f : \Omega_1^n \to [\epsilon, 1-\epsilon]$ and $g : \Omega_2^n \to [\epsilon, 1-\epsilon]$,*

$$\mathbb{E}J_\rho(f(X), g(Y)) \leq J_\rho(\mathbb{E}f, \mathbb{E}g) + C(\rho)\epsilon^{-C(\rho)}(\Delta_n(f) + \Delta_n(g)).$$

To see that there isn't much difference between the case of $[0,1]$ valued functions and $[\varepsilon, 1-\varepsilon]$ valued functions, we note that

Proposition 6.6. *Consider the setup of Theorem . Let $0 < \varepsilon < 0.5$ and let*

$$\bar{f}(x) = \begin{cases} x, & \varepsilon \leq x \leq 1-\varepsilon, \\ \varepsilon, & x \leq \varepsilon, \\ 1-\varepsilon, & x \geq 1-\varepsilon. \end{cases}$$

and similarly $\bar{g}(x)$. Then $|\mathbb{E}\bar{f} - \mathbb{E}[f]| \leq \varepsilon$ and similarly for g. Moreover,

$$\left|\mathbb{E}J_\rho(f(X), g(Y)) - \mathbb{E}J_\rho(\bar{f}(X), \bar{g}(Y))\right| \leq 2\varepsilon$$

and $I_i(\bar{f}) \leq I_i(f)$ for all i and similarly for g.

Proof. The proof follows from the fact that J_ρ is $1 - Lip$ in each of its coordinates. □

6.1 The Base Case

We prove Theorem 6.6 by induction on n. In this section, we will prove the base case $n = 1$:

Claim 6.7. *For any $\epsilon > 0$ and $0 < \rho < 1$, there is a $C(\rho)$ such that for any two random variables $X, Y \in [\epsilon, 1-\epsilon]$ with correlation in $[-\rho, \rho]$,*

$$\mathbb{E}J_\rho(X, Y) \leq J_\rho(\mathbb{E}X, \mathbb{E}Y) + C(\rho)\epsilon^{-C(\rho)}(\mathbb{E}|X - \mathbb{E}X|^3 + \mathbb{E}|Y - \mathbb{E}Y|^3).$$

The proof of Claim 6.7 essentially follows from Taylor's theorem applied to the function J_ρ; the crucial point is that J_ρ satisfies a certain differential equation. Define the matrix $M_{\rho\sigma}(x, y)$ by

$$M_{\rho\sigma}(x, y) = \begin{pmatrix} \frac{\partial^2 J_\rho(x,y)}{\partial x^2} & \sigma\frac{\partial^2 J_\rho(x,y)}{\partial x \partial y} \\ \sigma\frac{\partial^2 J_\rho(x,y)}{\partial x \partial y} & \frac{\partial^2 J_\rho(x,y)}{\partial y^2} \end{pmatrix}.$$

Claim 6.8. *For any $(x, y) \in (0,1)^2$ and $0 \leq |\sigma| \leq \rho$, $M_{\rho\sigma}(x, y)$ is a negative semidefinite matrix. Likewise, if $|\sigma| \geq \rho$, then $M_{\rho\sigma}(x, y)$ is a positive semidefinite matrix.*

We will also use the fact that the third derivatives of J_ρ are bounded (at least, away from the boundary of $[0,1]^2$).

Claim 6.9. *For any $-1 < \rho < 1$, there exists $C(\rho) > 0$ such that for any $i, j \ge 0$, $i+j = 3$,*

$$\left| \frac{\partial^3 J_\rho(x,y)}{\partial x^i \partial y^j} \right| \le C(\rho)(xy(1-x)(1-y))^{-C(\rho)}.$$

Further, the function $C(\rho)$ can be chosen so that it is continuous for $\rho \in (-1,1)$.

Claims 6.8 and 6.9 follow from elementary calculus, and we defer their proofs (we note that Claim 6.8 is implicit in [47]). Now we will use them with Taylor's theorem to prove Claim 6.7.

Proof of Claim 6.7. Fix $\epsilon > 0$ and $\rho \in (0,1)$. Now let $C(\rho)$ be large enough so that all third derivatives of J_ρ are uniformly bounded by $C(\rho)\epsilon^{-C(\rho)}$ on the square $[\epsilon, 1-\epsilon]^2$ (such a $C(\rho)$ exists by Claim 6.9). Taylor's theorem then implies that for any $a, b, a+x, b+y \in [\epsilon, 1-\epsilon]$,

$$\begin{aligned} J_\rho(a+x, b+y) \le J_\rho(a,b) + x\frac{\partial J_\rho}{\partial x}(a,b) + y\frac{\partial J_\rho}{\partial y}(a,b) \\ + \frac{1}{2}(x\ y)\begin{pmatrix} \frac{\partial^2 J_\rho}{\partial x^2}(a,b) & \frac{\partial^2 J_\rho}{\partial x \partial y}(a,b) \\ \frac{\partial^2 J_\rho}{\partial x \partial y}(a,b) & \frac{\partial^2 J_\rho}{\partial y^2}(a,b) \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + C(\rho)\epsilon^{-C(\rho)}(|x|^3 + |y|^3). \quad (10) \end{aligned}$$

Now suppose that X and Y are random variables taking values in $[\epsilon, 1-\epsilon]$. If we apply (10) with $a = \mathbb{E}X$, $b = \mathbb{E}Y$, $x = X - \mathbb{E}X$, and $y = Y - \mathbb{E}Y$, and then take expectations of both sides, we obtain

$$\begin{aligned} \mathbb{E}J_\rho(X,Y) \le J_\rho(\mathbb{E}X, \mathbb{E}Y) + \frac{1}{2}\mathbb{E}\left[(\tilde{X}\ \tilde{Y})\begin{pmatrix} \frac{\partial^2 J_\rho}{\partial x^2}(a,b) & \frac{\partial^2 J_\rho}{\partial x \partial y}(a,b) \\ \frac{\partial^2 J_\rho}{\partial x \partial y}(a,b) & \frac{\partial^2 J_\rho}{\partial y^2}(a,b) \end{pmatrix}\begin{pmatrix} \tilde{X} \\ \tilde{Y} \end{pmatrix}\right] \\ + C(\rho)\epsilon^{-C(\rho)}(\mathbb{E}|\tilde{X}|^3 + \mathbb{E}|\tilde{Y}|^3), \quad (11) \end{aligned}$$

where $\tilde{X} = X - \mathbb{E}X$ and $\tilde{Y} = Y - \mathbb{E}Y$. Now, if X and Y have correlation $\sigma \in [0, \rho]$ then $\mathbb{E}\tilde{X}\tilde{Y} = \sigma\sqrt{\mathbb{E}\tilde{X}^2 \mathbb{E}\tilde{Y}^2}$, and so

$$\mathbb{E}\left[(\tilde{X}\ \tilde{Y})\begin{pmatrix} \frac{\partial^2 J_\rho}{\partial x^2}(a,b) & \frac{\partial^2 J_\rho}{\partial x \partial y}(a,b) \\ \frac{\partial^2 J_\rho}{\partial x \partial y}(a,b) & \frac{\partial^2 J_\rho}{\partial y^2}(a,b) \end{pmatrix}\begin{pmatrix} \tilde{X} \\ \tilde{Y} \end{pmatrix}\right] = (\sigma_X\ \sigma_Y)\begin{pmatrix} \frac{\partial^2 J_\rho}{\partial x^2}(a,b) & \sigma\frac{\partial^2 J_\rho}{\partial x \partial y}(a,b) \\ \sigma\frac{\partial^2 J_\rho}{\partial x \partial y}(a,b) & \frac{\partial^2 J_\rho}{\partial y^2}(a,b) \end{pmatrix}\begin{pmatrix} \sigma_X \\ \sigma_Y \end{pmatrix},$$

where $\sigma_X = \sqrt{\mathbb{E}\tilde{X}^2}$ and $\sigma_Y = \sqrt{\mathbb{E}\tilde{Y}^2}$. By Claim 6.8,

$$(\sigma_X\ \sigma_Y)\begin{pmatrix} \frac{\partial^2 J_\rho}{\partial x^2}(a,b) & \sigma\frac{\partial^2 J_\rho}{\partial x \partial y}(a,b) \\ \sigma\frac{\partial^2 J_\rho}{\partial x \partial y}(a,b) & \frac{\partial^2 J_\rho}{\partial y^2}(a,b) \end{pmatrix}\begin{pmatrix} \sigma_X \\ \sigma_Y \end{pmatrix} = (\sigma_X\ \sigma_Y)M_{\rho\sigma}(a,b)\begin{pmatrix} \sigma_X \\ \sigma_Y \end{pmatrix} \le 0.$$

Applying this to (11), we obtain

$$\mathbb{E}J_\rho(X,Y) \le J_\rho(\mathbb{E}X, \mathbb{E}Y) + C(\rho)\epsilon^{-C(\rho)}(\mathbb{E}|\tilde{X}|^3 + \mathbb{E}|\tilde{Y}|^3).$$ □

6.2 The inductive step:

Next, we prove Theorem 6.6 by induction.

Proof of Theorem 6.6. Note that the base case follows from 6.7 since $f(X_1), g(Y_1)$ have correlation between $-\rho$ and ρ by the assumption on the Renyi correlation between the X's and Y's.

We now prove the inductive claim: Assume that the theorem holds with n replaced by $n-1$. Consider $f : \Omega_1^n \to [\epsilon, 1-\epsilon]$ and $g : \Omega_2^n \to [\epsilon, 1-\epsilon]$.

Conditioning on (X_n, Y_n) we have

$$\mathbb{E}J_\rho(f(X), g(Y)) = \mathbb{E}[\mathbb{E}[J_\rho(f(X), g(Y))|X_n, Y_n]].$$

Applying the inductive hypothesis for $n-1$ conditionally on X_n and Y_n,

$$\begin{aligned}\mathbb{E}[J_\rho(f(X), g(Y))|X_n, Y_n] \le J_\rho(&\mathbb{E}[f|X_n], \mathbb{E}[g|Y_n]) \\ &+ C(\rho)\epsilon^{-C(\rho)}(\Delta_{n-1}(f) + \Delta_{n-1}(g)).\end{aligned} \quad (12)$$

On the other hand, the base case for $n = 1$ implies that

$$\mathbb{E}[J_\rho(\mathbb{E}[f|X_n], \mathbb{E}[g|Y_n])] \ \le \ J_\rho(\mathbb{E}f, \mathbb{E}g) \ + \ C(\rho)\epsilon^{-C(\rho)}(\Delta_1(\mathbb{E}[f|X_n]) \ + \ \Delta_1(\mathbb{E}[g|Y_n])). \quad (13)$$

Taking the expectation of (12) over X_n and Y_n and combining it with (13), we obtain

$$\begin{aligned}\mathbb{E}J_\rho(f(X), g(Y)) \ &\le \ J_\rho(\mathbb{E}f, \mathbb{E}g) + C(\rho)\epsilon^{-C(\rho)}\big(\mathbb{E}[\Delta_{n-1}(f) + \Delta_{n-1}(g)]\big) \\ &+ \ C(\rho)\epsilon^{-C(\rho)}\big(\Delta_1(\mathbb{E}[f|X_n]) + \Delta_1(\mathbb{E}[g|Y_n])\big).\end{aligned}$$

Finally, note that the definition of Δ_n implies that the right-hand side above is just

$$J_\rho(\mathbb{E}f, \mathbb{E}g) + C(\rho)\epsilon^{-C(\rho)}\big(\Delta_n(f) + \Delta_n(g)\big).$$ □

6.3 Proof of Borell's Theorem

We will now prove the following functional version of Borell's result. It is easy to see it implies Theorem 5.9.

Theorem 6.10. *Let $\rho \ge 0$ and N and M are Gaussian vectors with joint distribution*

$$(N, M) \sim \mathcal{N}\left(0, \begin{pmatrix} I_d & \rho I_d \\ \rho I_d & I_d \end{pmatrix}\right).$$

For any measurable $f, g : \mathbb{R}^d \to [0, 1]$,

$$\mathbb{E}J_\rho(f(N), g(M)) \le J_\rho(\mathbb{E}f, \mathbb{E}g).$$

Proof of Theorem 6.10. Let $n = md$ and define

$$G_{1,n} = \frac{1}{\sqrt{m}} \left(\sum_{i=1}^{m} x_i, \sum_{i=m+1}^{2m} x_i, \dots, \sum_{i=(d-1)m+1}^{md} x_i \right).$$

Define $G_{2,n}$ similarly by using y instead of x. In other words, $G_{1,n}$ and $G_{2,n}$ are vectors obtained by averaging the vectors x and y over consecutive blocks of size m. Define Z as

$$Z = (x_1, x_{m+1}, \dots, x_{(d-1)m+1}, y_1, y_{m+1}, \dots, y_{(d-1)m+1}).$$

Observe that $(G_{1,n}, G_{2,n})$ is distributed as sum of m independent copies of Z scaled by $1/\sqrt{m}$. Applying the Lindeberg-Feller central limit theorem [22], we obtain $(G_{1,n}, G_{2,n}) \to_D (N, M)$ as $m \to \infty$.

Suppose first that f and g are L-Lipschitz functions taking values in $[\epsilon, 1-\epsilon]$. Define $F, G : \{-1,1\}^n \to \mathbb{R}$ as

$$F(z) = f\left(\frac{1}{\sqrt{m}} \left(\sum_{i=1}^{m} z_i, \sum_{i=m+1}^{2m} z_i, \dots, \sum_{i=(d-1)m+1}^{md} z_i \right) \right),$$

and similarly for G using g. By Theorem 6.6,

$$\mathbb{E} J_\rho(F(x), G(y)) \le J_\rho(\mathbb{E}F, \mathbb{E}G) + C(\rho)\epsilon^{-C(\rho)}(\Delta_n(F) + \Delta_n(G)). \tag{14}$$

Since f is L-Lipschitz,

$$|F(x) - F(x \oplus e_j)| \le \frac{2L}{\sqrt{m}},$$

for every j, and similarly for g. Therefore:

$$\max(\Delta_n(F), \Delta_n(G)) \le \frac{8L^3 n}{m^{3/2}} = \frac{8L^3 d}{\sqrt{m}}.$$

Applying this to (14),

$$\mathbb{E} J_\rho(F(x), G(y)) \le J_\rho(\mathbb{E}F, \mathbb{E}G) + C(\rho)\epsilon^{-C(\rho)} \frac{16L^3 d}{\sqrt{m}},$$

and so the definition of F, G implies

$$\mathbb{E} J_\rho(f(G_{1,n}), g(G_{2,n})) \le J_\rho(\mathbb{E}f(G_{1,n}), \mathbb{E}g(G_{2,n})) + C(\rho)\epsilon^{-C(\rho)} \frac{16L^3 d}{\sqrt{m}}.$$

Taking $m \to \infty$, the multivariate central limit theorem and using the fact that J_ρ is Lipschitz implies that

$$\mathbb{E} J_\rho(f(N), g(M)) \le J_\rho(\mathbb{E}f(N), \mathbb{E}g(M)). \tag{15}$$

This establishes the theorem for functions f and g which are Lipschitz and take values in $[\epsilon, 1-\epsilon]$. But any measurable $f_1, f_2 : \mathbb{R}^d \to [0,1]$ can be approximated (say in $L^p(\mathbb{R}^d, \gamma_d)$) by Lipschitz functions with values in $[\epsilon, 1-\epsilon]$. Since neither the Lipschitz constant nor ϵ appears in (15), the general statement of the theorem follows from the dominated convergence theorem. □

6.4 Smoothness and hyper-contraction

It turns out that there is an effective bound on $\Delta_n(T_\eta f)$ in the case where f has low influences, where T_η is the noise operator and η is close to 1. In this section we show how the fact that the noise operator T_η is "hyper-contractive" allows to bound $\Delta_n(T_\eta f)$ when f has low influences. In the next section, we will show that the smoothness of J implies that only a small error results from replacing f by $T_\eta f$ for η close to 1.

Hypercontractivity is a key feature of many of the proofs in the analysis of Boolean functions starting with the KKL paper [35]. The proof will use the famous hyper-contractive theorem of Bonami and Beckner:

Theorem 6.11. *[10, 6] Let $f : \{-1,1\}^n \to \mathbb{R}$ and $1 \le q \le p$. Then, if $\rho^2 \le \frac{q-1}{p-1}$, then:*

$$\|T_\rho f\|_p \le \|f\|_q.$$

We will use this fact to analyze the differences $\mathbb{E}[|f_i - f_{i-1}|^3]$. In particular we will prove the following:

Lemma 6.12. *Let $f : \{-1,1\}^n \to [0,1]$ and let $0 < \rho < 1$. Then:*

- $(T_\rho f)_i = T_\rho(f_i)$,
- $$\mathbb{E}[|T_\rho(f_i - f_{i-1})|^3] \le I_i(f)^{p/2},$$
  where $p = \min(3, 1 + 1/\rho^2)$,
- $$\Delta_n(T_{\rho^2} f) \le \frac{1}{1-\rho} \max I_i(f)^{p/2-1}.$$

Proof. The fact that $(T_\rho f)_i = T_\rho(f_i)$ follows immediately from the definitions. Let $p = \min(3, 1 + 1/\rho^2)$. Then we apply Theorem 6.11 to obtain:

$$\begin{aligned}\mathbb{E}[|T_\rho(f_i - f_{i-1})|^3] &= \mathbb{E}[|T_\rho(f_i - f_{i-1})|^p |T_\rho(f_i - f_{i-1})|^{3-p}] \\ &\le \mathbb{E}[|T_\rho(f_i - f_{i-1})|^p] = \|T_\rho(f_i - f_{i-1})\|_p^p \le \|f_i - f_{i-1}\|_2^p \le I_i^{p/2},\end{aligned}$$

and the proof follows of the second statement follows. To prove the third statement observe that:

$$\begin{aligned}\Delta_n(T_{\rho^2} f) &= \sum_{i=1}^{n} \mathbb{E}[|T_\rho^2(f_i - f_{i-1})|^3] \\ &\leq \sum_{i=1}^{n} I_i(T_\rho f)^{p/2} \leq \max I_i(T_\rho f)^{p/2-1} \sum_{i=1}^{n} I_i(T_\rho f) \\ &\leq \frac{1}{1-\rho} \max I_i(f)^{p/2-1}.\end{aligned}$$

□

6.5 Proof of Majority is Stablest (Theorem 5.15)

Proof. Fix $\varepsilon > 0$. Our goal is to prove that:

$$\langle f, g\rangle_\rho \leq \langle \chi_{\mathbb{E}f}, \chi_{\mathbb{E}g}\rangle_\rho + \varepsilon = J_\rho(\mathbb{E}f, \mathbb{E}g) + \varepsilon, \tag{16}$$

when $\max(I_i(f), I_i(g)) \leq \tau$ for a small enough $\tau(\varepsilon)$.

Let $\varepsilon' = \varepsilon/10$. Let $\bar{f}$ be the rounding of f to the interval $[\varepsilon', 1-\varepsilon']$ and similarly define $\bar{g}$. We first note that it suffices to prove the desired claim 16 for $\bar{f}$ and $\bar{g}$ with error ε'. This is because

$$|\langle f, g\rangle_\rho - \langle \bar{f}, \bar{g}\rangle_\rho| \leq 2\varepsilon', \quad |\mathbb{E}f - \mathbb{E}f'| \leq \varepsilon', \quad |\mathbb{E}g - \mathbb{E}g'| \leq \varepsilon$$

and because $J_\rho(x, y)$ is 1-Lip in each of the coordinates x and y. Moreover, truncation reduces variance and therefore $I_i(\bar{f}) \leq I_i(f)$ for all i and similarly for g. For simplicity of notation we redefine $\varepsilon = \varepsilon'$ and our goal is now to prove 16 for functions bounded in the interval $[\varepsilon, 1-\varepsilon]$.

Again, define $\varepsilon' = \varepsilon/10$ and let $\rho = \eta^2\rho'$ for $\eta < 1$ be chosen so that

$$J_{\rho'}(\mathbb{E}f, \mathbb{E}g) \leq J_\rho(\mathbb{E}f, \mathbb{E}g) + \varepsilon'.$$

Then it suffices to prove that

$$\langle f, g\rangle_\rho = \langle T_\eta f, T_\eta g\rangle_{\rho'} \leq J_{\rho'}(\mathbb{E}f, \mathbb{E}g) + \varepsilon'.$$

and since $J_{\rho'}(x, y) \geq xy$ it suffices to prove that

$$\mathbb{E}[J_{\rho'}(T_\eta f, T_\eta g)] \leq J_{\rho'}(\mathbb{E}f, \mathbb{E}g) + \varepsilon'. \tag{17}$$

Again we have that $I_i(T_\eta f) \leq I_i(f)$ for all i and similarly for g. Renaming again $\rho = \rho'$ and $\varepsilon' = \varepsilon$ and applying Theorem 6.6 what remain to show is that

$$C(\rho)\varepsilon^{-C(\rho)}(\Delta_n(T_\eta f) + \Delta_n(T_\eta g)) \leq \varepsilon.$$

By Lemma 6.12 this can be bounded by:

$$C(\rho)\varepsilon^{-C(\rho)}C(\eta)\max I_i(f)^\eta$$

Thus for τ small enough, this is less than ε as needed. □

6.6 Proof for resilient functions

We now prove Theorem 5.17. First we note that in the inductive proof in subsection 6.2, if we stop k steps before the end we obtain:

Proposition 6.13. *For every $1 \le k \le n$, it holds that:*

$$\begin{aligned}\mathbb{E}J_\rho(f(X),g(Y)) &\le \mathbb{E}\Big[J_\rho(f_{n-k}(X),g_{n-k}(Y))\Big] \\ &+ C(\rho)\epsilon^{-C(\rho)}\mathbb{E}\Big[\Delta_{n-k}(f)+\Delta_{n-k}(g)\Big],\end{aligned}$$

where $f_{n-k} = \mathbb{E}[f|X_{n-k},\dots,X_k]$ and $g_{n-k} = \mathbb{E}[g|Y_{n-k},\dots,Y_n])\Big]$.

Proof. The proof will imitate the previous proof up to (17). Using the fact that

$$\sum_{i=1}^n I_i(T_\eta f) \le (1-\eta)^{-1},$$

It follows that there are most $k = O(\tau^{-1}(1-\eta)^{-1})$ coordinates of f and g with influence more than τ. Without loss of generality, assume that these are the the last k coordinates. Then apply Proposition 6.13 to conclude that:

$$\mathbb{E}[J_{\rho'}(T_\eta f, T_\eta g)] \le \mathbb{E}[J_\rho(T_\eta f_{n-k}(X), T_\eta g_{n-k}(Y))] + \kappa,$$

where the error term κ is bounded as before by

$$C(\rho)\varepsilon^{-C(\rho)}C(\eta)\max I_{1\le i\le n-k}(f)^\eta \le \varepsilon/2.$$

Now if the function f is $(k,\varepsilon/100)$ resilient then so is f_{n-k} and $T_\eta f_{n-k}$. Therefore

$$|f_{n-k} - \mathbb{E}[f]| \le \varepsilon/100,$$

and therefore since J is Lipchitz it holds that

$$\mathbb{E}[J_\rho(T_\eta f_{n-k}(X), T_\eta g_{n-k}(Y))] \le \mathbb{E}[J_\rho(\mathbb{E}f, T_\eta g_{n-k}(Y))] + \varepsilon/100 \le J_\rho(\mathbb{E}f,\mathbb{E}g) + \varepsilon/100,$$

where the ultimate inequality follows from the fact that J is concave in each of the coordinates. □

Note that the only place where we used the resilience of f in showing that

$$\mathbb{E}[J_\rho(T_\eta f_{n-k}(X), T_\eta g_{n-k}(Y))] \le J_\rho(\mathbb{E}f, \mathbb{E}g) + \varepsilon/100$$

More generally:

Exercise 6.14.

- *Prove that if the W_1 distance between the joint distribution of $(T_\eta f_{n-k}(X), T_\eta g_{n-k}(Y))$ and the product of the marginals is bounded by ε then*

$$\mathbb{E}[J_\rho(T_\eta f_{n-k}(X), T_\eta g_{n-k}(Y))] \le J_\rho(\mathbb{E}f, \mathbb{E}g) + \varepsilon.$$

- *Show that if the W_1 distance between the joint distribution of $(T_\eta f_{n-k}(X), T_\eta g_{n-k}(Y))$ and the product of the marginals is smaller than the corresponding distance between the joint distribution of $(f_{n-k}(X), g_{n-k}(Y))$ and the product of their marginals.*

- *Show that if $f, g : \{-1,1\}^k \to \{0,1\}$ and the distance between the joint distribution of $(f(X), g(Y))$ and the product of their marginals is more than ε then:*

$$\sum_{S \neq \emptyset} |\hat{f}(S)\hat{g}(S)| \ge |\sum_{S \neq \emptyset} \rho^{|S|} \hat{f}(S)\hat{g}(S)| \ge \varepsilon.$$

This exercise implies in particular the following Theorem:

Theorem 6.15. *For every $\varepsilon > 0, 0 < \rho < 1$, there exists k and α such that if $f, g : \{-1,1\}^n \to \{0,1\}$ satisfy*

$$\langle f, g \rangle_\rho > J_\rho(\mathbb{E}f, \mathbb{E}g) + \varepsilon,$$

then there exists a set S with $|S| \le k$ such that $\hat{f}(S)\hat{g}(S) \ge \alpha$.

6.7 Cross Influences

We now sketch an alterantive proof of Theorem 5.16 given Theorem 5.15.

Proof.

- We will use the fact that if $I_i(g) \le \tau'$, then there exists a function g' that does not depend on i such that $\mathbb{E}[|g - g'|^2] \le \tau'$, i.e., $g' = \mathbb{E}[g|x_{-i}]$ and that if we let $f' = \mathbb{E}[f|x_{-i}]$ then for all ρ: $\langle f', g' \rangle_\rho = \langle f, g' \rangle_\rho$.

- As before, it suffices to prove the theorem for $T_\gamma f$ and $T_\gamma g$ for some γ close to, but less than 1 incurring an error of 0.1ε.

- We know that if all the influences of both f and g are less than some τ then we obtain the right statement with error 0.1ε.

- By Lemma 5.12,

$$\sum_i I_i(T_\gamma f) \le \max_k k\gamma^k \le C(\gamma),$$

  and thus there are at most $C(\gamma)/\tau$ coordinates of f with influence greater than τ. Let us denote this set of coordinates by A_f. Let A_g be the corresponding set for g.

- Let us choose τ' so that $\sqrt{\tau'}C(\gamma)/\tau \le 0.1\varepsilon$. Suppose that $\min(I_i(f), I_i(g)) \le \tau'$ for all i. Note that this implies in particular that A_f and A_g are disjoint.

  Let $f' = \mathbb{E}[T_\gamma f | x_{-A_g}]$ and $g' = \mathbb{E}[T_\gamma f | x_{-A_f}]$. Then:

$$|\langle f', g' \rangle - \langle T_\gamma f, T_\gamma g \rangle| \le 0.2\varepsilon.$$

  Moreover, $\max(I_i(f'), I_i(g')) \le \tau$ so we can apply Theorem 5.15 to conclude.

□

6.8 Majority is the most stable resilient function

We now prove Theorem 5.17 using different ideas from additive combinatorics. The proof will use Decision Trees as well as the following standard estimates, see e.g. [50, Appendix B] and subsection 6.10.

Lemma 6.16. *Assume $\rho < 1$ and $\rho_1 < \rho_2 < 1$ then*

$$|\langle \chi_{\mu_1}, \chi_{\mu_2} \rangle_{\rho_1} - \langle \chi_{\mu_1}, \chi_{\mu_2} \rangle_{\rho_2}| \le \frac{10(\rho_2 - \rho_1)}{1 - \rho_2}$$

$$|\langle \chi_{\mu_1'}, \chi_{\mu_2'} \rangle_{\rho} - \langle \chi_{\mu_1}, \chi_{\mu_2} \rangle_{\rho}| \le |\mu_1 - \mu_1'| + |\mu_2 - \mu_2'|$$

Decision trees allow to express Boolean function by querying variables on at a time, where the order of queries may depend on the values of previous variables. In the context of low influence functions, decision trees are useful as they allow to express functions of the form $T_\gamma h$ as a bounded size decision tree, where almost all the leaves are low influence functions. We will use the following regularity lemma, see e.g. [55, 19].

Lemma 6.17. *For a function $f : \{-1,1\}^n \to \mathbb{R}$, let $I(f) = \sum I_i(f)$. Then for any $\tau > 0, \varepsilon > 0$ and any function f there exists a decision tree for f of depth at most*

$$d \le 2 + \frac{I}{\tau\varepsilon}$$

such that the probability of reaching a leaf with influence sum τ or more is bounded by ε.

Proof. The construction of the decision tree is standard. If a function f_{x_I} at a certain node x_I has all influences less than τ or if the node is at level d do nothing. Otherwise, condition on the variable j with the maximum influence in f_{x_I} and create two children f_{y_J} and f_{z_K} where $J = K = I \cup \{j\}$, $y_i = z_i = x_i$ for all $i \in I$ and $y_j = 0$ and $z_j = 1$. Since

$$I(f_{x_I}) = I_j(f_{x_I}) + \frac{1}{2}(I(f_{y_J}) + I(f_{z_K})),$$

it easily follows that if L is the set of leaves of the tree and if $D(\ell)$ denotes the depth of leaf ℓ then

$$I(f) \geq \tau \sum_{\ell \in L} 2^{-D(\ell)} D(\ell).$$

Therefore if p is the fraction of paths that reach level d then

$$I(f) \geq (d-1)\tau p \implies p \leq I/(d-1)\tau,$$

and taking $d-1$ to be the smallest integer that is greater or equal to $\frac{I}{\tau\varepsilon}$ we obtain the desired result. $\square$

The proof will be carried out via a sequence of reductions which will give the function f more and more structure. Fix $\varepsilon > 0$ and $0 \leq \rho' < 1$. Recall that we want to show that if f is $(d(\varepsilon, \rho'), \alpha(\varepsilon, \rho'))$ resilient then for all g bounded between 0 and 1,

$$\langle f, g \rangle_{\rho'} \leq \langle \chi_{\mu_f}, \chi_{\mu_g} \rangle_{\rho'} + \varepsilon, \tag{18}$$

where $\mu_f = \mathbb{E} f$ and similarly for g.

Lemma 6.18. *In order to prove (18) it suffices to prove that*

$$\langle T_\eta f, g \rangle_\rho \leq \langle \chi_{\mu_f}, \chi_{\mu_g} \rangle_\rho + \varepsilon/2. \tag{19}$$

for

$$\rho = (1 - 0.01\varepsilon)\rho' + 0.01\varepsilon, \quad \eta = \rho'/\rho = 1 - \Omega(\varepsilon(1-\rho')). \tag{20}$$

Proof. Write $\rho' = \rho\eta$, where $1 - \rho \geq (1-\rho')/2$ and $\eta < 1$. Note that $\langle T_\eta f, g \rangle_\rho = \langle f, g \rangle_{\rho'}$. Moreover, f and $T_\eta f$ have the same expected value. If we could establish (19) and

$$|\langle \chi_{\mu_f}, \chi_{\mu_g} \rangle_\rho - \langle \chi_{\mu_f}, \chi_{\mu_g} \rangle_{\rho'}| < \varepsilon/2, \tag{21}$$

then (18) would follow. Note that (21) follows from Lemma 6.16 when

$$\frac{10(\rho - \rho')}{1 - \rho} < \varepsilon/2.$$

We may thus choose ρ and η as in (20). $\square$

Lemma 6.19. *Let τ be chosen so that (6) holds with error 0.01ε for ρ. Then it suffices to prove (19) for a function $h = T_\eta f$ that has a decision tree of depth d and such that for at most 0.01ε fraction of the inputs a random path of the decision tree terminates at a node with some influence greater than τ. Moreover*

$$d = O\left(\frac{1}{\varepsilon^2(1-\rho)\tau}\right)$$

Proof. We note that the function $h = T_\eta f$ satisfies:

$$I(h) := \sum I_i(h) = \sum_S |S|\hat{f}^2(S)\eta^{2|S|} \leq \max_s s\eta^{2s} = O\left(\frac{1}{\varepsilon(1-\rho)}\right).$$

Apply Lemma 6.17 to obtain a decision tree for h where for at most 0.01ε fraction of the inputs, a random path of the decision tree terminates at a node with some influence greater than τ. Note that the depth of the tree satisfies

$$d \leq C(1 + \frac{I}{\tau\varepsilon}) \leq C(1 + \frac{1}{\varepsilon^2(1-\rho)\tau})$$

as needed. □

We now conclude the proof of Theorem 5.17.

Proof. Let h be a function such as in Lemma 6.19. Assume furthermore that f is $(d, 0.01\varepsilon 2^{-d})$ resilient. Note that this implies that $h = T_\eta f$ is also $(d, 0.01\varepsilon 2^{-d})$ resilient. This is because for every S,

$$\mathbb{E}[T_\eta f | X_S = z] = \mathbb{E}[\mathbb{E}[f | X_S = z']],$$

where $z' \sim \otimes_{i \in S}(\eta\delta_{z_i} + \frac{1-\eta}{2}(\delta_1 + \delta_{-1}))$, and therefore:

$$\sup_{|S| \leq d, z} \Big| \mathbb{E}\left[T_\eta f | X_S = z\right] - \mathbb{E}[f] \Big| \leq \sup_{|S| \leq d, z} \Big| \mathbb{E}\left[f | X_S = z\right] - \mathbb{E}[f] \Big|.$$

Let x, y be two ρ-correlated inputs. Then

$$\mathbb{E}[h(x)g(y)] = \sum_{x_I, y_I} \mathbb{P}[x_I]\mathbb{P}[y_I | x_I]\mathbb{E}[h(x)g(y) | x_I, y_I],$$

where x_I denotes a random leaf of the decision tree and y_I is chosen after x_I to be a ρ-correlated version of x_I. Let A denote the set of x_I for which f_{x_I} has all influences less than τ. Then:

$$\begin{aligned} \mathbb{E}[h(x)g(y)] &= \sum_{x_I, y_I} \mathbb{P}[x_I]\mathbb{P}[y_I | x_I]\mathbb{E}[h(x)g(y) | x_I, y_I] \\ &\leq 0.01\varepsilon + \sum_{x_I \in A} \mathbb{P}[x_I] \sum_{y_I} \mathbb{P}[y_I | x_I]\mathbb{E}[h(x)g(y) | x_I, y_I]. \end{aligned}$$

Write $\mu' = \mu_f + 0.01\varepsilon$ and $\mu(y_I) = \mathbb{E}[g(y)|y_I]$. Note that since h is $(d, 0.01\varepsilon 2^{-d})$-resilient it follows that for all leaves x_I it holds that $\mathbb{E}[h|x_I] \le \mu'$. Thus for $x_I \in A$ we can apply (6) to obtain that

$$\mathbb{E}[h(x)g(y)|x_I, y_I] \le \langle \chi_{\mu'}, \chi_{\mu(y_I)} \rangle_\rho + 0.01\varepsilon.$$

Plugging this back in we obtain the bound

$$\begin{aligned}
\mathbb{E}[h(x)g(y)] &\le 0.02\varepsilon + \sum_{x_I \in A} \mathbb{P}[x_I] \sum_{y_I} \mathbb{P}[y_I|x_I] \langle \chi_{\mu'}, \chi_{\mu(y_I)} \rangle_\rho \\
&\le 0.02\varepsilon + \sum_{x_I, y_I} \mathbb{P}[x_I] \mathbb{P}[y_I|x_I] \langle \chi_{\mu'}, \chi_{\mu(y_I)} \rangle_\rho \\
&= 0.02\varepsilon + \langle \chi_{\mu'}, (\sum_{x_I, y_I} \mathbb{P}[x_I] \mathbb{P}[y_I|x_I] \chi_{\mu(y_I)}) \rangle_\rho
\end{aligned}$$

Note that $\psi = \sum_{x_I, y_I} \mathbb{P}[x_I] \mathbb{P}[y_I|x_I] \chi_{\mu(y_I)}$ is a $[0,1]$-valued function with $\mathbb{E}[\psi] = \mathbb{E}[g]$. Thus by Theorem 5.9 it follows that

$$0.02\varepsilon + \langle \chi_{\mu'}, (\sum_{x_I, y_I} \mathbb{P}[x_I] \mathbb{P}[y_I|x_I] \chi_{\mu(y_I)}) \rangle_\rho \le 0.02\varepsilon + \langle \chi_{\mu'}, \chi_{\mu_g} \rangle_\rho \le 0.04\varepsilon + \langle \chi_{\mu_f}, \chi_{\mu_g} \rangle_\rho,$$

where the last inequality follows from Lemma 6.16. □

6.9 Majority is Most Predictable

Suppose n voters are to make a binary decision. Assume that the outcome of the vote is determined by a *social choice* function $f : \{-1,1\}^n \to \{-1,1\}$, so that the outcome of the vote is $f(x_1, \ldots, x_n)$ where $x_i \in \{-1,1\}$ is the vote of voter i. We assume that the votes are independent, each ± 1 with probability $\frac{1}{2}$. It is natural to assume that the function f satisfies $f(-x) = -f(x)$, i.e., it does not discriminate between the two candidates. Note that this implies that $\mathbb{E}[f] = 0$ under the uniform distribution. A natural way to try and predict the outcome of the vote is to sample a subset of the voters, by sampling each voter independently with probability ρ. Conditioned on a vector X of votes the distribution of Y, the sampled votes, is i.i.d. where $Y_i = X_i$ with probability ρ and $Y_i = *$ (for unknown) otherwise.

Conditioned on $Y = y$, the vector of sampled votes, the optimal prediction of the outcome of the vote is given by $\text{sgn}((Tf)(y))$ where

$$(Tf)(y) = \mathbb{E}[f(X)|Y = y]. \tag{22}$$

This implies that the probability of correct prediction (also called predictability) is given by

$$\mathbb{P}[f = \text{sgn}(Tf)] = \frac{1}{2}(1 + \mathbb{E}[f \ \text{sgn}(Tf)]).$$

For example, when $f(x) = x_1$ is the dictator function, we have $\mathbb{E}[f \text{ sgn}(Tf)] = \rho$ corresponding to the trivial fact that the outcome of the election is known when voter 1 is sampled and are ± 1 with probability 1/2 otherwise. The notion of predictability is natural in statistical contexts. It was also studied in a more combinatorial context in [**?**].

Similarly to the Majority is Stablest Theorem, we can prove:

Theorem 6.20 ("Majority Is Most Predictable")**.** *Let $0 \le \rho \le 1$ and $\varepsilon > 0$ be given. Then there exists a $\tau > 0$ such that if $f : \{-1,1\}^n \to [-1,1]$ satisfies $\mathbb{E}[f] = 0$ and $\mathrm{Inf}_i(f) \le \tau$ for all i, then*

$$\mathbb{E}[f \text{ sgn}(Tf)] \le \tfrac{2}{\pi} \arcsin \sqrt{\rho} + \varepsilon, \tag{23}$$

where T is defined in (22).

Similarly, in the same setup, for every $\varepsilon > 0$, there exists (r, δ) such that if f is (r, δ)-resilient and satisfies $\mathbb{E}[f] = 0$ then (23) holds.

We note that from the central limit theorem it follows that if $\mathrm{Maj}_n(x_1, \ldots, x_n) = \mathrm{sgn}(\sum_{i=1}^n x_i)$, then

$$\lim_{n\to\infty} \mathbb{E}[\mathrm{Maj}_n \, \mathrm{sgn}(T \, \mathrm{Maj}_n)] = \tfrac{2}{\pi} \arcsin \sqrt{\rho}.$$

Remark 6.21. Note that Theorem 6.20 proves a weaker statement than showing that Majority is the most predictable function. The statement only asserts that if a function has low enough influences than its predictability cannot be more than ε larger than the asymptotic predictability value achieved by the majority function when the number of voters $n \to \infty$. This slightly inaccurate title of the theorem is inline with previous results such as the "Majority is Stablest Theorem" (see below). Similar language may be used later when informally discussing statements of various theorems.

Remark 6.22. One may wonder if for a finite n, among *all* functions $f : \{-1,1\}^n \to \{-1,1\}$ with $\mathbb{E}[f] = 0$, majority is the most predictable function. Note that the predictability of the dictator function $f(x) = x_1$ is given by ρ, and $\frac{2}{\pi} \arcsin \sqrt{\rho} > \rho$ for $\rho \to 0$. Therefore when ρ is small and n is large the majority function is more predictable than the dictator function. However, note that when $\rho \to 1$ we have $\rho > \frac{2}{\pi} \arcsin \sqrt{\rho}$ and therefore for values of ρ close to 1 and large n the dictator function is more predictable than the majority function.

We note that the bound obtained in Theorem 6.20 is a reminiscent of the Majority is Stablest theorem [49, 50] as both involve the arcsin function. However, the two theorems are quite different. The Majority is Stablest theorem asserts that under the same condition as in Theorem 6.20 it holds that

$$\mathbb{E}[f(X)f(Y)] \le \tfrac{2}{\pi} \arcsin \rho + \varepsilon.$$

where $(X_i, Y_i) \in \{-1,1\}^2$ are i.i.d. with $\mathbb{E}[X_i] = \mathbb{E}[Y_i] = 0$ and $\mathbb{E}[X_i Y_i] = \rho$. Thus "Majority is Stablest" considers two correlated voting vectors, while "Majority is Most

Predictable" considers a sample of one voting vector. In fact, both results follow from the more general invariance principle presented here. We note a further difference between stability and predictability: It is well known that in the context of "Majority is Stablest", for *all* $0 < \rho < 1$, among all boolean functions with $\mathbb{E}[f] = 0$ the maximum of $\mathbb{E}[f(x)f(y)]$ is obtained for dictator functions of the form $f(x) = x_i$. As discussed above, for ρ close to 0 and large n, the dictator is less predictable than the majority function.

The proof of Theorem 6.20 follows the same lines of the Majority is Stablest Theorem. The basic space $\Omega = \Omega_1 \times \Omega_2$ where $(x, y) \in \Omega$ is distributed as follows. First x is distributed uniformly in $\{\pm 1\}$. Conditioned on x, $y = x$ with probability ρ and is equal to $*$ with probability $1 - \rho$. It is easy to check that the that the Renyi correlation between x and y is $\sqrt{\rho}$.

6.10 Facts regarding J_ρ

Here we collect various facts about the function

$$J_\rho(x, y) := \langle \chi_x, \chi_y \rangle_\rho = \Pr[X \le \Phi^{-1}(x), Y \le \Phi^{-1}(y)],$$

where $(X, Y) \sim \mathcal{N}(0, (\begin{smallmatrix} 1 & \rho \\ \rho & 1 \end{smallmatrix}))$. As is standard, we will use ϕ to denote the density of the standard normal distribution. These calculations all follow from elementary calculus.

Claim 6.8. *For any $(x, y) \in (0, 1)^2$ and $0 \le |\sigma| \le \rho$, $M_{\rho\sigma}(x, y)$ is a negative semidefinite matrix. Likewise, if $|\sigma| \ge \rho$, then $M_{\rho\sigma}(x, y)$ is a positive semidefinite matrix.*

Proof. Towards proving this, note that we can define $Y = \rho \cdot X + \sqrt{1 - \rho^2} \cdot Z$ where $Z \sim \mathcal{N}(0, 1)$ is an independent normal. Also, let us define $\Phi^{-1}(x) = s$ and $\Phi^{-1}(y) = t$. For $s, t \in \mathbb{R}$, define $K_\rho(s, t)$ as

$$K_\rho(s, t) = \Pr_{X,Y}[X \le s, Y \le t] = \Pr_{X,Z}[X \le s, Z \le (t - \rho \cdot X)/\sqrt{1 - \rho^2}].$$

Note that for the aforementioned relations between x, y, s and t, $K_\rho(s, t) = J_\rho(x, y)$. Note that

$$K_\rho(s, t) = \int_{s'=-\infty}^{s} \phi(s') \int_{t'=-\infty}^{(t-\rho \cdot s')/\sqrt{1-\rho^2}} \phi(t') dt' ds'. \tag{24}$$

This implies that

$$\frac{\partial K_\rho(s, t)}{\partial s} = \phi(s) \int_{t'=-\infty}^{(t-\rho \cdot s)/\sqrt{1-\rho^2}} \phi(t') dt'.$$

By chain rule, we get that

$$\frac{\partial J_\rho(x, y)}{\partial x} = \frac{\partial K_\rho(s, t)}{\partial s} \cdot \frac{\partial s}{\partial x}.$$

By elementary calculus, it follows that

$$\frac{d\Phi^{-1}(x)}{dx} = \frac{1}{\phi(\Phi^{-1}(x))} \quad \Rightarrow \quad \frac{\partial s}{\partial x} = \frac{1}{\phi(\Phi^{-1}(x))} = \frac{1}{\phi(s)}.$$

Thus,

$$\frac{\partial J_\rho(x,y)}{\partial x} = \int_{t'=-\infty}^{(t-\rho\cdot s)/\sqrt{1-\rho^2}} \phi(t')dt'.$$

Thus, we next get that

$$\begin{aligned}\frac{\partial^2 J_\rho(x,y)}{\partial x^2} &= \frac{\partial^2 J_\rho(x,y)}{\partial x \partial s} \cdot \frac{\partial s}{\partial x} \\ &= \phi\left(\frac{t-\rho\cdot s}{\sqrt{1-\rho^2}}\right)\cdot\frac{-\rho}{\sqrt{1-\rho^2}}\cdot\frac{1}{\phi(s)} = \phi\left(\frac{\Phi^{-1}(y)-\rho\cdot\Phi^{-1}(x)}{\sqrt{1-\rho^2}}\right)\cdot\frac{-\rho}{\sqrt{1-\rho^2}}\cdot\frac{1}{\phi(s)}.\end{aligned}$$

$$\frac{\partial^2 J_\rho(x,y)}{\partial x \partial y} = \frac{\partial^2 J_\rho(x,y)}{\partial x \partial t}\cdot\frac{\partial t}{\partial y} = \phi\left(\frac{\Phi^{-1}(y)-\rho\cdot\Phi^{-1}(x)}{\sqrt{1-\rho^2}}\right)\cdot\frac{1}{\sqrt{1-\rho^2}}\cdot\frac{1}{\phi(t)}.$$

Because we know that $(X,Y)\sim(Y,X)$, by symmetry, we can conclude that

$$\frac{\partial^2 J_\rho(x,y)}{\partial y^2} = \phi\left(\frac{\Phi^{-1}(x)-\rho\cdot\Phi^{-1}(y)}{\sqrt{1-\rho^2}}\right)\cdot\frac{-\rho}{\sqrt{1-\rho^2}}\cdot\frac{1}{\phi(t)}.$$

and likewise,

$$\frac{\partial^2 J_\rho(x,y)}{\partial y \partial x} = \phi\left(\frac{\Phi^{-1}(x)-\rho\cdot\Phi^{-1}(y)}{\sqrt{1-\rho^2}}\right)\cdot\frac{1}{\sqrt{1-\rho^2}}\cdot\frac{1}{\phi(s)}.$$

It is obvious now that

$$\frac{\partial^2 J_\rho(x,y)}{\partial x^2}\cdot\frac{\partial^2 J_\rho(x,y)}{\partial y^2} - \rho^2\left(\frac{\partial^2 J_\rho(x,y)}{\partial x \partial y}\right)^2 = 0.$$

Now, suppose that $|\sigma| \le |\rho|$. Then

$$\det(M_{\rho\sigma}(x,y)) = \frac{\partial^2 J_\rho(x,y)}{\partial x^2}\cdot\frac{\partial^2 J_\rho(x,y)}{\partial y^2} - \sigma^2\left(\frac{\partial^2 J_\rho(x,y)}{\partial x \partial y}\right)^2 \ge 0.$$

If $\rho \ge 0$ then the diagonal of $M_{\rho\sigma}(x,y)$ is non-positive, and it follows that $M_{\rho\sigma}(x,y)$ is negative semidefinite. If $\rho \le 0$ then the diagonal is non-negative and so $M_{\rho\sigma}(x,y)$ is positive semidefinite. □

Claim 6.9. *For any $-1 < \rho < 1$, there exists $C(\rho) > 0$ such that for any $i, j \geq 0$, $i+j = 3$,*

$$\left| \frac{\partial^3 J_\rho(x,y)}{\partial x^i \partial y^j} \right| \leq C(\rho)(xy(1-x)(1-y))^{-C(\rho)}.$$

Further, the function $C(\rho)$ can be chosen so that it is continuous for $\rho \in (-1, 1)$.

Proof. As before, we set $\Phi^{-1}(x) = s$ and $\Phi^{-1}(y) = t$. From the proof of Claim 6.8, we see that

$$\frac{\partial^2 J_\rho(x,y)}{\partial x^2} = \phi\left(\frac{\Phi^{-1}(y) - \rho \cdot \Phi^{-1}(x)}{\sqrt{1-\rho^2}}\right) \cdot \frac{-\rho}{\sqrt{1-\rho^2}} \cdot \frac{1}{\phi(s)}.$$

To compute the third derivatives of J, recalling that $\frac{\partial s}{\partial x} = \frac{1}{\phi(s)}$ and $\frac{\partial t}{\partial y} = \frac{1}{\phi(t)}$, we have

$$\begin{aligned} \frac{\partial^3 J_\rho(x,y)}{\partial x^3} &= \frac{\rho}{(1-\rho^2)^{3/2}} \frac{\rho t + (2\rho^2 - 1)s}{\phi(s)} \exp\Big(-\frac{t^2 - 2\rho st + (2\rho^2-1)s^2}{2(1-\rho^2)}\Big) \\ &= \frac{\sqrt{2\pi}\rho}{(1-\rho^2)^{3/2}} (\rho t + (2\rho^2-1)s) \exp\Big(-\frac{t^2 - 2\rho st + (3\rho^2-2)s^2}{2(1-\rho^2)}\Big). \qquad (25) \end{aligned}$$

Now, $\Phi^{-1}(x) \sim \sqrt{2\log x}$ as $x \to 0$; hence there is a constant C such that $\Phi^{-1}(x) \leq C\sqrt{\log x}$ for all $x \leq \frac{1}{2}$. Hence, $\exp(s^2) \leq x^{-C}$ for all $x \leq \frac{1}{2}$; by symmetry, $\exp(s^2) \leq (x(1-x))^{-C}$ for all $x \in (0,1)$. Therefore

$$\begin{aligned} \exp\Big(-\frac{t^2 - 2\rho st + (3\rho^2-2)s^2}{2(1-\rho^2)}\Big) &= e^{-\frac{t^2}{2(1-\rho^2)}} e^{\frac{\rho st}{1-\rho^2}} e^{\frac{(2-3\rho^2)s^2}{2(1-\rho^2}} \\ &\leq e^{-\frac{t^2}{2(1-\rho^2)}} e^{\frac{\rho(s^2+t^2)}{2(1-\rho^2)}} e^{\frac{(2-3\rho^2)s^2}{2(1-\rho^2}} \\ &\leq \big(x(1-x)y(1-y)\big)^{-\frac{\rho}{2(1-\rho^2)}} \big(x(1-x)\big)^{-\frac{2-3\rho^2}{2(1-\rho^2)}}. \qquad (26) \end{aligned}$$

Further, using $\exp(s^2) \leq x^{-C}$ and $\exp(t^2) \leq y^{-C}$, $\rho t + (2\rho^2 - 1)s \leq 4(xy)^{-C}$. As a consequence, applying this to (25), we see that there is a constant $C(\rho) > 0$,

$$\left| \frac{\partial^3 J_\rho(x,y)}{\partial x^3} \right| \leq C(\rho)\big(x(1-x)y(1-y)\big)^{-C(\rho)}.$$

The other third derivatives are similar:

$$\frac{\partial^3 J_\rho(x,y)}{\partial x^2 \partial y} = \frac{\sqrt{2\pi}\rho}{(1-\rho^2)^{3/2}} (t - 2\rho s) \exp\Big(-\frac{(2\rho^2-1)t^2 - 2\rho st + (2\rho^2-1)s^2}{2(1-\rho^2)}\Big).$$

By the same steps that led to (26), we get

$$\left| \frac{\partial^3 J_\rho(x,y)}{\partial x^2 \partial y} \right| \leq C(\rho)\big(x(1-x)y(1-y)\big)^{-C(\rho)}$$

(for a slightly different $C(\rho)$). The bounds on $\partial^3 J/\partial y^2 \partial x$ and $\partial^3 J/\partial x^3$ then follow because J is symmetric in x and y.

The fact that $C(\rho)$ can be chosen so that it is continuous for $\rho \in (-1,1)$ is obvious from the discussion above. □

Claim 6.23. *For any $x, y \in (0,1)$,*

$$\left|\frac{\partial J_\rho(x,y)}{\partial \rho}\right| \leq (1-\rho^2)^{-3/2}.$$

Proof. We begin from (24), but this time we differentiate with respect to ρ:

$$\frac{\partial K_\rho(s,t)}{\partial \rho} = -\frac{1}{(1-\rho^2)^{3/2}} \int_{s'=-\infty}^{s} \phi(s')\phi\left(\frac{t-\rho s'}{\sqrt{1-\rho^2}}\right) ds'.$$

Since Range$(\phi) \subset (0,1]$ and $\int_{s'} \phi(s')ds' = 1$, it follows that

$$\left|\frac{\partial K_\rho(s,t)}{\partial \rho}\right| \leq (1-\rho^2)^{-3/2}.$$

Since $\frac{\partial J_\rho(s,t)}{\partial \rho} = \frac{\partial K_\rho(\Phi^{-1}(x),\Phi^{-1}(y))}{\partial \rho}$, the proof is complete. □

We also state the following useful claim without a proof. The proof is obvious from the calculations in the proofs of Claim 6.8 and Claim 6.9.

Claim 6.24. *For any $\rho \in (-1,1)$, $\varepsilon > 0$ there exists a continuous function $\gamma(\rho,\varepsilon)$ such that for any $(x,y) \in [\varepsilon, 1-\varepsilon]^2$ and $1 \leq i+j \leq 3$,*

$$\left|\frac{\partial^{i+j} J_\rho(x,y)}{\partial^i x \partial^j y}\right| \leq \gamma(\rho,\varepsilon).$$

7 Paradoxes, Noise Stability and Reverse Hyper-Contraction

7.1 Probability of Paradox

Our next goal is to prove a quantitative version of Arrow Theorem following [45]. We will only discuss the case of 3 alternatives but will allow different function to determine different pairwise selections. Recall that we consider voters who vote independently and where voter i votes uniformly at random from the 6 possible rankings. Recall that we encode the 6 possible rankings by vectors $(x, y, z) \in \{-1, +1\}^3 \setminus \{\pm(1, 1, 1)\}$. Here x is $+1/-1$ if a voter ranks a above/below b, y is $+1/-1$ if voter ranks b above/below c, z is $+1/-1$ if voter ranks c above/below a. We will assume further that $f, g, h : \{-1, 1\}^n \to \{0, 1\}$ are the aggregation functions for the a vs. b, b vs. c and c vs. a preferences. We will again use the following observation used in [36]: Since the binary predicate $\psi(a, b, c) = 1(a = b = c)$ can be expressed as

$$\psi(a, b, c) = \frac{1}{4}(1 + ab + ac + bc - a - b - c),$$

we can write

$$\begin{aligned}
\mathbb{P}[f(x) = g(y) = h(z)] &= \frac{1}{4}\Big(1 + \mathbb{E}[f(x)g(y)] + \mathbb{E}[g(y)h(z)] + \mathbb{E}[h(z)f(x)] - \mathbb{E}[f] - \mathbb{E}[g] - \mathbb{E}[h]\Big) \\
&= \frac{1}{4}\Big(1 + \langle f, g\rangle_{-1/3} + \langle g, h\rangle_{-1/3} + \langle h, f\rangle_{-1/3} - \mathbb{E}[f] - \mathbb{E}[g] - \mathbb{E}[h]\Big)
\end{aligned}$$

where the last equality follows from the fact that the uniform distribution over $\{\pm 1\}^3 \setminus \{\pm(1, 1, 1)\}$, satisfies $E[x_i y_i] = -1/3$ and similarly for other pairs of coordinates.

To state a quantitive version, we will say that a function f is ε-close to a function g if $\mathbb{P}[f \neq g] \leq \varepsilon$. The quantitative version we we wish to prove is the following:

Theorem 7.1. *For every $\varepsilon > 0$, there exists $\delta(\varepsilon) > 0$ such that the following holds for every n: If*

$$\mathbb{P}[f(x) = g(y) = h(z)] < \delta,$$

then either two of the functions f, g, h are ε-close to constant functions of the opposite sign, or there exists a variable i such that f, g and h are all ε-close to the same dictator on voter i.

The main significance of Theorem 7.1 is that it is dimension independent. We get the same bound no matter what the dimension is. This shows that one cannot avoid the curse of paradoxes in voting by assuming the probability of a paradox vanishes as the number of voters grow. Our proof of the quantitative version will use the Majority is Stablest Theorem along with reverse hyper-contractive inequality which we discuss next. Before proving Theorem 7.1 we give a direct implication of the Majority is Stablest Theorem in the case where the functions $f = g = h$ are all balanced so $\mathbb{E}[f] = \mathbb{E}[g] = \mathbb{E}[h] = 0$. Using Majority is Stablest Theorem we obtain:

Theorem 7.2 ([36, 50]). *For every $\varepsilon > 0$, there exists a $\tau > 0$ such that if $f, g, h : \{-1,1\}^n \to \{0,1\}$ satisfy $\mathbb{E}[f] = \mathbb{E}[g] = \mathbb{E}[h] = 1/2$ and have all influences bounded above by τ then:*

$$\mathbb{P}[f(x) = g(y) = h(z)] \geq 3\langle \chi_{\frac{1}{2}}, \chi_{\frac{1}{2}} \rangle_{-\frac{1}{3}} - \frac{1}{2} - \varepsilon. \tag{27}$$

Again, the right hand side of equation (27) is the asymptotic probability that $\mathbb{P}[f(x) = g(y) = h(z)]$ when $f = g = h = \chi_{\frac{1}{2}}(n^{-\frac{1}{2}} \sum_{i=1}^n x_i)$ are all given by the same Majority function. Theorem 7.2 provides a surprising counter argument to Condorcet' arguments. Condorcet argued that pairwise ranking by Majority is problematic as it results in a paradox and Theorem 7.2 shows that in fact Majority asymptotically minimizes the probability of a paradox among low influence functions.

We also have the following strengthening of Theorem 7.2.

Theorem 7.3. *For every $\varepsilon > 0$, there exist $m, \beta > 0$ such that if $f, g, h : \{-1,1\}^n \to \{0,1\}$ satisfy $\mathbb{E}[f] = \mathbb{E}[g] = \mathbb{E}[h] = 1/2$ and f, g and h are all (m, β)-resilient then*

$$\mathbb{P}[f(x) = g(y) = h(z)] \geq 3\langle \chi_{\frac{1}{2}}, \chi_{\frac{1}{2}} \rangle_{-\frac{1}{3}} - \frac{1}{2} - \varepsilon.$$

7.2 Reverse Hyper-Contraction

Recall that the hyper-contractive theorem 6.11 states that for $f : \{-1,1\}^n \to \mathbb{R}$ and $1 \leq q \leq p$ and for any $\rho^2 \leq \frac{q-1}{p-1}$, it holds that

$$\|T_\rho f\|_p \leq \|f\|_q.$$

Borell [11] proved a reverse inequality:

Theorem 7.4. *[11] Let $f, g : \{-1,1\}^n \to \mathbb{R}_+$ and $1 > p > q$. Then, for any $0 \leq \rho^2 \leq \frac{1-p}{1-q}$,*

$$\|T_\rho f\|_q \geq \|f\|_p.$$

and for any $0 \leq \rho^2 \leq (1-p)(1-q)$

$$\langle f, g \rangle_\rho \geq \|f\|_p \|g\|_q$$

While the inequalities may seem like a curiosity, as p and q norms for $p, q < 1$ are rarely used (nor are they norms), the second inequality is quite helpful in some social choice proofs. For a more general discussion of reverse-hyper contraction and its applications see [51, 52].

An immediate computational corollary of Theorem 7.4 is the following lemma (as stated in [51]):

Lemma 7.5. *Let $x, y \in \{-1,1\}^n$ be distributed uniformly and (x_i, y_i) are independent. Assume that $\mathbb{E}[x(i)] = \mathbb{E}[y(i)] = 0$ for all i and that $\mathbb{E}[x(i)y(i)] = \rho \geq 0$. Let $B_1, B_2 \subset \{-1,1\}^n$ be two sets and assume that*

$$\mathbb{P}[B_1] \geq e^{-\alpha^2}, \quad \mathbb{P}[B_2] \geq e^{-\beta^2}.$$

Then:

$$\mathbb{P}[x \in B_1, y \in B_2] \geq \exp(-\frac{\alpha^2 + \beta^2 + 2\rho\alpha\beta}{1-\rho^2}).$$

In particular, if $\mathbb{P}[B_1] \geq \varepsilon$ and $\mathbb{P}[B_2] \geq \varepsilon$, then

$$\mathbb{P}[x \in B_1, y \in B_2] \geq \varepsilon^{\frac{2}{1-\rho}}.$$

We will need to generalize the result above to negative ρ and further to different ρ values for different bits.

Lemma 7.6. *Let $x, y \in \{-1,1\}^n$ be distributed uniformly and (x_i, y_i) are independent. Assume that $\mathbb{E}[x(i)] = \mathbb{E}[y(i)] = 0$ for all i and that $|\mathbb{E}[x(i)y(i)]| \leq \rho$. Let $B_1, B_2 \subset \{-1,1\}^n$ be two sets and assume that*

$$\mathbb{P}[B_1] \geq e^{-\alpha^2}, \quad \mathbb{P}[B_2] \geq e^{-\beta^2}.$$

Then:

$$\mathbb{P}[x \in B_1, y \in B_2] \geq \exp(-\frac{\alpha^2 + \beta^2 + 2\rho\alpha\beta}{1-\rho^2}).$$

In particular if $\mathbb{P}[B_1] \geq \varepsilon$ and $\mathbb{P}[B_2] \geq \varepsilon$, then:

$$\mathbb{P}[x \in B_1, y \in B_2] \geq \varepsilon^{\frac{2}{1-\rho}}. \tag{28}$$

Proof. Take z so that (x_i, z_i) are independent and $\mathbb{E}[z_i] = 0$ and $\mathbb{E}[x_i z_i] = \rho$. It is easy to see that there exists w_i independent of x, z s.t. the joint distribution of (x, y) is the same as $(x, z \cdot w)$, where $z \cdot w = (z_1 w_1, \ldots, z_n w_n)$. Now for each fixed w we have that

$$\mathbb{P}[x \in B_1, z \cdot w \in B_2] = \mathbb{P}[x \in B_1, z \in w \cdot B_2] \geq \exp(-\frac{\alpha^2 + \beta^2 + 2\rho\alpha\beta}{1-\rho^2}),$$

where $w \cdot B_2 = \{w \cdot w' : w' \in B_2\}$. Therefore taking expectation over w we obtain:

$$\mathbb{P}[x \in B_1, y \in B_2] = \mathbb{E}[\mathbb{P}[x \in B_1, z \cdot w \in B_2 \mid w]] \geq \exp(-\frac{\alpha^2 + \beta^2 + 2\rho\alpha\beta}{1-\rho^2})$$

as needed. The conclusion (28) follows by simple substitution (note the difference with Corollary 3.5 in [51] for sets of equal size which is a typo). □

Applying the CLT and using [12] one obtains the same result for Gaussian random variables.

Lemma 7.7. *Let N, M be $N(0, I_n)$ with $(N(i), M(i))_{i=1}^n$ independent. Assume that $|\mathbb{E}[N(i)M(i)]| \le \rho$. Let $B_1, B_2 \subset \mathbb{R}^n$ be two sets and assume that*

$$\mathbb{P}[B_1] \ge e^{-\alpha^2}, \quad \mathbb{P}[B_2] \ge e^{-\beta^2},$$

Then:

$$\mathbb{P}[N \in B_1, M \in B_2] \ge \exp(-\frac{\alpha^2 + \beta^2 + 2\rho\alpha\beta}{1-\rho^2}).$$

In particular if $\mathbb{P}[B_1] \ge \varepsilon$ and $\mathbb{P}[B_2] \ge \varepsilon$, then:

$$\mathbb{P}[N \in B_1, M \in B_2] \ge \varepsilon^{\frac{2}{1-\rho}}. \tag{29}$$

Proof. Fix the values of α and β and assume without loss of generality that $\max_i |\mathbb{E}[N(i)M(i)]|$ is obtained for $i = 1$. Then by [12] (see also [44]), the minimum of the quantity $\mathbb{P}[N \in B_1, M \in B_2]$ under the constraints on the measures given by α and β is obtained in one dimension, where B_1 and B_2 are intervals I_1, I_2. Look at random variables $x(i), y(i)$, where $\mathbb{E}[x(i)] = \mathbb{E}[y(i)] = 0$ and $\mathbb{E}[x(i)y(i)] = \mathbb{E}[M_1 N_1]$. Let $X_n = n^{-1/2} \sum_{i=1} x(i)$ and $Y_n = n^{-1/2} \sum_{i=1} y^(i)$. Then the CLT implies that

$$\mathbb{P}[X_n \in I_1] \to \mathbb{P}[N_1 \in B_1], \quad \mathbb{P}[Y_n \in I_2] \to \mathbb{P}[M_1 \in B_2],$$

and

$$\mathbb{P}[X_n \in I_1, Y_n \in I_2] \to \mathbb{P}[N_1 \in B_1, M_1 \in B_2].$$

The proof now follows from the previous lemma. □

7.3 The Gaussian Arrow Theorem

The next step is to consider a Gaussian version of the problem. The Gaussian version corresponds to a situation where the functions f, g, h can only "see" averages of large subsets of the voters. We thus define a 3 dimensional normal vector N. The first coordinate of N is supposed to represent the deviation of the number of voters where a ranks above b from the mean. The second coordinate is for b ranking above c and the last coordinate for c ranking above a.

Since averaging maintains the expected value and covariances, we define:

$$\mathbb{E}[N_1^2] = \mathbb{E}[N_2^2] = \mathbb{E}[N_3^2] = 1, \quad \mathbb{E}[N_1 N_2] = \mathbb{E}[N_2 N_3] = \mathbb{E}[N_3 N_1] = -1/3.$$

We let $N(1), \ldots, N(n)$ be independent copies of N. We write $\mathcal{N} = (N(1), \ldots, N(n))$ and for $1 \le i \le 3$ we write $\mathcal{N}_i = (N(1)_i, \ldots, N(n)_i)$. The Gaussian version of Arrow theorem states:

Theorem 7.8. *For every $\varepsilon > 0$ there exists a $\delta = \delta(\varepsilon) > 0$ such that the following hold. Let $\phi_1, \phi_2, \phi_3 : \mathbb{R}^n \to \{-1, 1\}$. Assume that for all $1 \le i \ne j \le 3$ and all $u \in \{-1, 1\}$ it holds that*

$$\mathbb{P}[\phi_i(\mathcal{N}_i) = u] + \mathbb{P}[\phi_j(\mathcal{N}_j) = -u] \ge 2\varepsilon. \tag{30}$$

Then with the setup given in (30) it holds that:

$$\mathbb{P}[\phi_1(\mathcal{N}_1) = \phi_2(\mathcal{N}_2) = \phi_3(\mathcal{N}_3)] \ge \delta.$$

Moreover, one may take $\delta = (\varepsilon/2)^{18}$.

We note that if $\mathbb{P}[\phi_i(\mathcal{N}_i) = u] + \mathbb{P}[\phi_j(\mathcal{N}_i) = -u] \le 2\varepsilon$, then one of the alternatives will be ranked the top/bottom with probability at least $1 - 2\varepsilon$. Therefore the theorem states that unless this is the case, the probability of a paradox is at least δ. Since the Gaussian setup excludes dictator functions in terms of the original vote, this result is to be expected in this case.

Proof. We will consider two cases: either all the functions ϕ_i satisfy $|\mathbb{E}\phi_i| \le 1 - \varepsilon$, or there exists at least one function with $|\mathbb{E}\phi_i| > 1 - \varepsilon$. Assume first that there exists a function ϕ_i with $|\mathbb{E}\phi_i| > 1 - \varepsilon$. Without loss of generality assume that $\mathbb{P}[\phi_1 = 1] > 1 - \varepsilon/2$. Note that by (30) it follows that $\mathbb{P}[\phi_1 = -1] + \mathbb{P}[\phi_2 = 1] \ge 2\varepsilon$ and therefore $\mathbb{P}[\phi_2 = 1] > \varepsilon$ and similarly $\mathbb{P}[\phi_3 = 1] > \varepsilon$. In particular, $\mathbb{P}[\phi_1 = 1, \phi_2 = 1] > \varepsilon/2$ We now look at the function $\psi = 1(\phi_1 = 1, \phi_2 = 1)$. Let

$$\mathcal{M}_1 = \frac{\sqrt{3}}{2}(\mathcal{N}_1 + \mathcal{N}_2), \quad \mathcal{M}_2 = \frac{\sqrt{3}}{2\sqrt{2}}(\mathcal{N}_1 - \mathcal{N}_2).$$

Then it is easy to see that $\mathcal{M}_2(i)$ is uncorrelated with and therefore independent of $\mathcal{N}_3(i), \mathcal{M}_1(i)$ for all i. Moreover, for all i the covariance between $\mathcal{M}_1(i)$ and $\mathcal{N}_3(i)$ is $1 - 1/\sqrt{3}$ and $1 - 1/\sqrt{3} > 1/3$. We may now apply Lemma 7.7 with the vectors

$$(\mathcal{N}_3(1), \dots, \mathcal{N}_3(n), Z_1, \dots, Z_n), \quad (\mathcal{M}_1(1), \dots, \mathcal{M}_1(n), \mathcal{M}_2(1), \dots, \mathcal{M}_2(n)),$$

where $Z = (Z_1, \dots, Z_n)$ is a normal Gaussian vector independent of anything else. We obtain:

$$\mathbb{P}[\phi_1(\mathcal{N}_1) = 1, \phi_2(\mathcal{N}_2) = 1, \phi_3(\mathcal{N}_3) = 1] = \mathbb{P}[\phi_3(\mathcal{N}_3, Z) = 1, \psi(\mathcal{M}_1, \mathcal{M}_2) = 1] \ge ((\varepsilon/2))^{\frac{2}{1/3}} \ge (\varepsilon/2)^6.$$

We next consider the case where all functions satisfy $|\mathbb{E}\phi_i| \le 1 - \varepsilon$. Notice that at least two of the functions obtain the same value with probability at least 1/2. Let's assume that $\mathbb{P}[\phi_1 = 1] \ge 1/2$ and $\mathbb{P}[\phi_2 = 1] \ge 1/2$. Then by Lemma 7.7 we obtain that

$$\mathbb{P}[\phi_1 = 1, \phi_2 = 1] \ge 1/8.$$

Again we define $\psi = 1(\phi_1 = 1, \phi_2 = 1)$. Since $\mathbb{P}[\phi_3 = 1] > \varepsilon/2$, we may apply Lemma 7.7 and obtain that:

$$\mathbb{P}[\phi_1 = 1, \phi_2 = 1, \phi_3 = 1] = \mathbb{P}[\phi_1 = 1, \psi = 1] \ge (\varepsilon/8)^6.$$

This concludes the proof. □

7.4 Arrow Theorem for Low Influence Functions

Theorem 7.9. *For every $\varepsilon > 0$ there exist a $\delta(\varepsilon) > 0$ and a $\tau(\delta) > 0$ such that the following holds. Let $f_1, f_2, f_3 : \{-1,1\}^n \to \{-1,1\}$. Assume that for all $1 \le i \ne j \le 3$ and all $u \in \{-1,1\}$ it holds that*

$$\mathbb{P}[f_i = u] + \mathbb{P}[f_j = -u] \ge 4\varepsilon \tag{31}$$

and for all j it holds that

$$|\{1 \le i \le 3 : I_j(f_i) > \tau\}| \le 1. \tag{32}$$

Then it holds that

$$\mathbb{P}[f_1(x) = f_2(y) = f_3(z))] \ge \delta.$$

Moreover, assuming the uniform distribution, one may take:

$$\delta = \frac{1}{8}(\varepsilon/2)^{20}, \quad \tau = \tau(\delta),$$

where

$$\tau(\delta) := \delta^{C\frac{\log(1/\delta)}{\delta}},$$

for some absolute constant C.

Proof. Let $\phi_1, \phi_2, \phi_3 : \mathbb{R} \to \{-1,1\}$ be of the form $\phi_i = \mathrm{sgn}(x - t_i)$, where t_i is chosen so that $\mathbb{E}[\phi_i] = \mathbb{E}[f_i]$ (where the first expected value is according to the Gaussian measure). Let $N_1, N_2, N_3 \sim N(0,1)$ be jointly Gaussian with $\mathbb{E}[N_i N_{i+1}] = -1/3$. From Theorem 7.8 it follows that:

$$P([\phi_1(N_1) = \phi_2(N_2) = \phi_3(N_3)] > 8\delta, \tag{33}$$

and from the MIST 5.16, it follows that by choosing C in the definition of τ large enough, we have:

$$\langle f_1, f_2\rangle_{-1/3} \ge \langle \phi_1, \phi_2\rangle_{-1/3} - \delta, \quad \langle f_2, f_3\rangle_{-1/3} \ge \langle \phi_2, \phi_3\rangle_{-1/3} - \delta, \quad \langle f_3, f_1\rangle_{-1/3} \ge \langle \phi_3, \phi_1\rangle_{-1/3} - \delta.$$

It now follows that:

$$\begin{aligned}
\mathbb{P}[f_1(x) = f_2(y) = f_3(z)] &= \frac{1}{4}\left(1 + \langle f_1, f_2\rangle_{-1/3} + \langle f_2, f_3\rangle_{-1/3} + \langle f_3, f_1\rangle_{-1/3}\right) \qquad (34)\\
&\ge \frac{1}{4}\left(1 + \langle \phi_1, \phi_2\rangle_{-1/3} + \langle \phi_2, \phi_3\rangle_{-1/3} + \langle \phi_3, \phi_1\rangle_{-1/3}\right) - 3\delta/4\\
&= \mathbb{P}[\phi_1(N_1) = \phi(N_2) = \phi_3(N_3)] - 3\delta/4 > 7\delta,
\end{aligned}$$

as needed. □

7.5 Arrow Theorem with at most one influential voter

In addition to the case where all the influences are small, we need to consider the case where one voter is influential. This includes the case of the dictator function. The theorem below extends the case of low influences to the case where there is only one influential variable.

Theorem 7.10. *For every $\varepsilon > 0$ there exists a $\delta(\varepsilon) > 0$ and a $\tau(\delta) > 0$ such that the following holds for all n. Let $f_1, f_2, f_3 : \{-1,1\}^n \to \{-1,1\}$. Assume that for all $1 \le i \le 3$ and $j > 1$ it holds that*

$$I_j(f_i) < 0.5\tau. \tag{35}$$

Then either

$$\mathbb{P}[f_1(x) = f_2(y) = f_3(z)] \ge \delta/6, \tag{36}$$

or there exist (f_1, f_2, f_3) which are 9ε-close to either a dictator function $\pm(x_i, x_i, x_i)$ or a function (g_1, g_2, g_3) where two of the g_i's are constant and of different signs. Moreover, one may take:

$$\delta = (\varepsilon/2)^{20}, \quad \tau = \tau(\delta).$$

Proof. Consider the functions f_i^b for $1 \le i \le 3$ and $b \in \{-1,1\}$ defined by

$$f_i^b(x_2, \ldots, x_n) = f_i(b, x_2, \ldots, x_n).$$

Note that for all $b \in \{-1,1\}$, for all $1 \le i \le 3$ and for all $j > 1$ it holds that $I_j(f_i^{b_i}) < \tau$ and therefore we may apply Theorem 7.9. We obtain that for every $b = (b_1, b_2, b_3) \notin \{(1,1,1), (-1,-1,-1)\}$ either:

$$\mathbb{P}[f_1^{b_1}(x) = f_2^{b_2}(y) = f_3^{b_3}(z)] \ge \delta, \tag{37}$$

or there exist a $u(b) \in \{-1,1\}$ and an $i = i(b)$ such that

$$\min(\mathbb{P}[f_i^{b_i} = u(b)], \mathbb{P}[f_{i+1}^{b_{i+1}} = -u(b)]) \ge 1 - 3\varepsilon. \tag{38}$$

Note that if there exists a vector $b = (b_1, b_2, b_3) \notin \{(1,1,1), (-1,-1,-1)\}$ for which (37) holds then (36) follows immediately.

It thus remains to consider the case where (38) holds for all 6 vectors b. In this case we will define new functions g_i as follows. We let $g_i(b, x_2, \ldots, x_n) = u$ if $\mathbb{P}[f_i^b = u] \ge 1 - 3\varepsilon$ for $u \in \{-1,1\}$ and $g_i(b, x_2, \ldots, x_n) = f_i(b, x_2, \ldots, x_n)$ otherwise. We let G be the social choice function defined by g_1, g_2 and g_3. From (38) it follows that for every $b = (b_1, b_2, b_3) \notin \{(1,1,1), (-1,-1,-1)\}$ there exist two functions g_i, g_{i+1} and a value u s.t. $g_i(b_i, x_2, \ldots, x_n)$ is the constant function u and $g_{i+1}(b_{i+1}, x_2, \ldots, x_n)$ is the constant function $-u$. So

$$P(g_1, g_2, g_3) = \mathbb{P}[(g_1, g_2, g_3) \in \{(1,1,1), (-1,-1,-1)\}] = 0,$$

and therefore by Theorem 1.3 (g_1, g_2, g_3) is of the required form. It is further easy to see that $\mathbb{P}[f_i \ne g_i] \le 3\varepsilon$ for all i, as needed. The proof follows. $\square$

7.6 Arrow Theorem with two influential voters

Our first application of reverse hyper-contraction is the following:

Lemma 7.11. *Suppose that* $I_1(f) > \varepsilon$ *and* $I_2(g) > \varepsilon$*. Let*

$$B = \{((x,y,z))_{i=3}^{n} : 1 \textit{ is pivotal for } f \textit{ and } 2 \textit{ is pivotal for } g)\}.$$

Then

$$\mathbb{P}[B] \geq \varepsilon^3.$$

Proof. Let

$$B_1 = \{((x,y,z))_{i=3}^{n} : 1 \text{ is pivotal for } f(\cdot,\cdot,x_3,\dots,x_n)\},$$

$$B_2 = \{((x,y,z))_{i=3}^{n} : 2 \text{ is pivotal for } g(\cdot,\cdot,y_3,\dots,y_n)\}.$$

Then $\mathbb{P}[B_1] \geq I_1(f) > \varepsilon$ and $\mathbb{P}[B_2] \geq I_2(g) > \varepsilon$, and our goal is to obtain a bound on $\mathbb{P}[B_1 \cap B_2]$. Note that the event B_1 is determined by x and the event B_2 is determined by y. So the proof follows immediately from Lemma 7.6 with $\rho = -1/3$. □

We can now prove the main result of the section.

Theorem 7.12. *Suppose that there exist voters* i *and* j *such that*

$$I_i(f) > \varepsilon, \quad I_j(g) > \varepsilon.$$

then $\mathbb{P}[f(x) = g(y) = h(z)] > \frac{1}{36}\varepsilon^3$*.*

Proof. Without loss of generality assume that $i = 1$ and $j = 2$. Let $B = B_1 \cap B_2$, where B_1 and B_2 are the events from Lemma 7.11. By the lemma we have $\mathbb{P}[B] \geq \varepsilon^3$. Note that conditioned on any $((x,y,z)_{i=3}^{n}) \in B$, the functions f, g and h on coordinates 1 and 2 satisfy the condition of Arrow Theorem 1.3. Thus with probability at least 1/36, the outcome is not transitive. Therefore:

$$\mathbb{P}[f(x) = g(y) = h(z)] \geq \frac{1}{36}\mathbb{P}[B] \geq \frac{1}{36}\varepsilon^3.$$

□

7.7 Proof of the quantitative Arrow Theorem

We are finally ready to prove Theorem 7.1 which we restate below:

Theorem 7.13. *For every* $\varepsilon > 0$*, there exists* $\delta(\varepsilon) > 0$ *such that the following holds for every* n*: If*

$$\mathbb{P}[f(x) = g(y) = h(z)] < \delta,$$

then either two of the functions f, g, h are ε-close to constant functions of the opposite sign, or there exists a variable i such that f, g and h are all ε-close to the same dictator on voter i. Moreover, one can take

$$\delta = \exp\left(-\frac{C}{\varepsilon^{21}}\right). \tag{39}$$

Proof. Let η be a small constant to be determined later. We will consider three cases:

- There exist two voters $i \neq j \in [n]$ and two different functions say f and g such that

$$I_i(f) > \eta, \quad I_j(g) > \eta. \tag{40}$$

- For pair of functions $k_1 \neq k_2 \in \{f, g, h\}$ and every $i \in [n]$, it holds that

$$\min(I_i(k_1), I_i(k_2)) < \eta. \tag{41}$$

- There exists a voter j' such that for all $j \neq j'$

$$\max(I_j(f), I_j(g), I_j(h)) < \eta. \tag{42}$$

First note that each (f, g, h) satisfies at least one of the three conditions (40), (41) or (42). Thus it suffices to prove the theorem for each of the three cases.

In (40), we have by Theorem 7.12 have that

$$\mathbb{P}[f(x) = g(y) = h(z)] > \frac{1}{36}\eta^3.$$

We thus obtain that $\mathbb{P}[f(x) = g(y) = h(z)] > \delta$ where δ is given in (39) by taking larger values C' for C.

In case (41), by Theorem 7.9 it follows that either there exists a function (g_1, g_2, g_3) which always puts a candidate at top / bottom and (f_1, f_2, f_3) is ε-close to (g_1, g_2, g_3) (if (31) holds), or $\mathbb{P}[f(x) = g(y) = h(z)] > C\varepsilon^{20} >> \delta$.

Similarly in the remaining case (42), we have by Theorem 7.10 that either (f_1, f_2, f_3) is ε-close to (g_1, g_2, g_3) with $\mathbb{P}[g_1(x) = g_2(y) = g_3(z)] = 0$ or $\mathbb{P}[f(x) = g(y) = h(z)] > C\varepsilon^{20} >> \delta$. The proof follows.

□

8 More general statements

In this section we discuss a more general statement of Arrow Theorem closer to his original formulation and its quantitative counterpart. This requires to introduce a number of additional definition. The reduction from the more general statements of Arrow Theorem to the 3 candidate case discussed above will be carried out in this section.

8.1 General Setup

Consider $A = \{a, b, \ldots, \}$, a set of $k \geq 3$ alternatives. A *transitive preference* over A is a ranking of the alternatives from top to bottom where ties are not allowed. Such a ranking corresponds to a *permutation* σ of the elements $1, \ldots, k$ where σ_i is the rank of alternative i. The set of all rankings will be denoted by S_k.

A *constitution* is a function F that associates to every n-tuple $\sigma = (\sigma(1), \ldots, \sigma(n))$ of transitive preferences (also called a *profile*), and every pair of alternatives a, b a preference between a and b. Some key properties of constitutions include:

- *Transitivity.* The constitution F is *transitive* if $F(\sigma)$ is transitive for all σ. In other words, for all σ and for all three alternatives a, b and c, if $F(\sigma)$ prefers a to b, and prefers b to c, it also prefers a to c. Thus F is transitive if and only if its image is a subset of the permutations on k elements.

- *Independence of Irrelevant Alternatives (IIA).* The constitution F satisfies the IIA property if for every pair of alternatives a and b, the social ranking of a vs. b (higher or lower) depends only on their relative rankings by all voters. The IIA condition implies that the pairwise preference between any pair of outcomes depends only on the individual pairwise preferences. Thus, if F satisfies the IIA property then there exists functions $f^{a>b}$ for every pair of candidates a and b such that

$$F(\sigma) = ((f^{a>b}(x^{a>b}) : \{a, b\} \in \binom{k}{2})$$

- *Unanimity.* The constitution F satisfies *Unanimity* if the social outcome ranks a above b whenever all individuals rank a above b.

- The constitution F is a *dictator* on voter i, if $F(\sigma) = \sigma(i)$, for all σ, or $F(\sigma) = -\sigma$, for all σ, where $-\sigma(i)$ is the ranking $\sigma_k(i) > \sigma_{k-1}(i) \ldots \sigma_2(i) > \sigma_1(i)$ by reversing the ranking $\sigma(i)$.

Arrow's theorem states [2, 3] that:

Theorem 8.1. *Any constitution on three or more alternatives which satisfies Transitivity, IIA and Unanimity is a dictatorship.*

It is possible to give a characterization of all constitutions satisfying IIA and Transitivity. Results of Wilson [65] provide a partial characterization for the case where voters are allowed to rank some alternatives as equal. In order to obtain a quantitative version of Arrow theorem, we give an explicit and complete characterization of all constitutions satisfying IIA and Transitivity in the case where all voters vote using a strict preference order. Write $\mathcal{F}_k(n)$ for the set of all constitutions on k alternatives and n voters satisfying IIA and Transitivity. For the characterization it is useful write $A >_F B$ for the statement that for all σ it holds that $F(\sigma)$ ranks all alternatives in A above all alternatives in B. We will further write F_A for the constitution F restricted to the alternatives in A. The IIA condition implies that F_A depends only on the individual rankings of the alternatives in the set A. The characterization of $\mathcal{F}_k(n)$ we prove is the following.

Theorem 8.2. *The class $\mathcal{F}_k(n)$ consist exactly of all constitutions F satisfying the following: There exist a partition of the set of alternatives into disjoint sets $A_1, \ldots, A_r$ such that:*

- $$A_1 >_F A_2 >_F \ldots >_F A_r,$$
- *For all A_s s.t. $|A_s| \geq 3$, there exists a voter j such that F_{A_s} is a dictator on voter j.*
- *For all A_s such that $|A_s| = 2$, the constitution F_{A_s} is an arbitrary non-constant function of the preferences on the alternatives in A_s.*

We note that for all $k \geq 3$ all elements of $\mathcal{F}_k(n)$ are not desirable as constitutions. Indeed elements of $F_k(n)$ either have dictators whose vote is followed with respect to some of the alternatives, or they always rank some alternatives on top some other. For a related discussion see [65]. The statement above follows easily from Theorem 3 in [65]. The exact formulation is taken from [43].

The main goal of the current section is to provide a quantitative version of Theorem 8.2 assuming voters vote independently and uniformly at random. Note that Theorem 8.2 above implies that if $F \notin \mathcal{F}_k(n)$ then $P(F) \geq (k!)^{-n}$. However if n is large and the probability of a non-transitive outcome is indeed as small as $(k!)^{-n}$, one may argue that a non-transitive outcome is so unlikely that in practice Arrow's theorem does not hold.

The main goal of this section is to prove the following statement:

Theorem 8.3. *For every number of alternatives $k \geq 1$ and $0.01 > \varepsilon > 0$, there exists a $\delta = \delta(\varepsilon)$, such that for every $n \geq 1$, if F is a constitution on n voters and k alternatives satisfying:*

- *IIA and*
- *$P(F) < \delta$,*

then there exists $G \in \mathcal{F}_k(n)$ satisfying $D(F, G) < k^2\varepsilon$. Moreover, one may take:

$$\delta = \exp\left(-\frac{C}{\varepsilon^{21}}\right), \tag{43}$$

for some absolute constant $0 < C < \infty$.

We therefore obtain the following:

Corollary 8.4. *For any number of alternatives $k \geq 3$ and $\varepsilon > 0$, there exists a $\delta = \delta(\varepsilon)$, such that for every n, if F is a constitution on n voters and k alternatives satisfying:*

- *IIA and*
- *F is $k^2\varepsilon$ far from any dictator, so $D(F,G) > k^2\varepsilon$ for any dictator G,*
- *For every pair of alternatives a and b, the probability that F ranks a above b is at least $k^2\varepsilon$,*

then the probability of a non-transitive outcome, $P(F)$, is at least δ, where $\delta(\varepsilon)$ may be taken as in (43).

Proof. Assume by contradiction that $P(F) < \delta$. Then by Theorem 8.3 there exists a function $G \in \mathcal{F}_{n,k}$ satisfying $D(F,G) < k^2\varepsilon$. Note that for every pair of alternatives a and b it holds that:

$$\mathbb{P}[G \text{ ranks } a \text{ above } b] \geq \mathbb{P}[F \text{ ranks } a \text{ above } b] - D(F,G) > 0.$$

Therefore for every pair of alternatives there is a positive probability that G ranks a above b. Thus by Theorem 8.3 it follows that G is a dictator which is a contradiction. □

Remark 8.5. Note that if $G \in \mathcal{F}_k(n)$ and F is any constitution satisfying $D(F,G) < k^2\varepsilon$ then $P(F) < k^2\varepsilon$.

Remark 8.6. The bounds stated in Theorem 8.3 and Corollary 8.4 in terms of k and ε is clearly not an optimal one. We expect that the true dependency has δ which is some fixed power of ε. Moreover we expect that the bound $D(F,G) < k^2\varepsilon$ should be improved to $D(F,G) < \varepsilon$.

8.1.1 Nisan's argument

N. Nisan argued in his blog [**?**] that the natural way to study quantitative versions of Arrow's theorem is to look at functions from S_k^n to S_k and check to what extent do they satisfy the IIA property. He defines a function to be η-IIA if for every two alternatives a and b it holds that $\mathbb{P}[F(\sigma) \neq \mathcal{F}(\tau)] \leq \eta$, where σ is uniformly chosen and τ is uniformly chosen conditioned on the a, b ranking at τ being identical to that of σ for all voters. In his blog Nisan sketches how a quantitative Arrow's theorem proven for the definition used here implies a quantitative Arrow's theorem for his definition. We briefly repeat the argument with some minor modifications and corrections.

Fixing alternatives a, b and writing $p_{a,b} : \{0,1\}^n \to \{0,1\}$ for the probability that given a vector of n binary preferences between a and b, F ranks a above b. If F satisfies the IIA property then $p_{a,b} \in \{0,1\}$ a.s. If F is η-IIA then $\mathbb{E}[2p_{a,b}(1-p_{a,b})] \leq \eta$, and therefore $\mathbb{E}[\min(p_{a,b}, 1-p_{a,b})] \leq \eta$.

Assume a quantitative Arrow theorem such as the one proven here with parameters ε, δ and suppose by contradiction that $F : S_k^n \to S_k$ is η-IIA and ε far from $\mathcal{F}_k(n)$ for some small η to be determined later. Define a function G as follows. Let $G(\sigma)$ rank a above b if for the majority of τ which agree with σ in the a, b orderings it holds that $F(\tau)$ ranks a above b. We note that for every pair of alternatives a, b it holds that

$$\mathbb{P}[F(\sigma), G(\sigma) \text{ have different order on } a, b] = \mathbb{E}[\min(p_{a,b}, 1 - p_{a,b})] \leq \eta.$$

By taking a union bound on all pairs of alternatives, this implies that $D(F, G) \leq \binom{k}{2}\eta \leq k^2\eta/2$. Note further that G satisfies the IIA property by definition. Since F is transitive and from the quantitative Arrow theorem proven here we conclude that

$$D(F, G) \geq \mathbb{P}[P(G)] \geq \delta.$$

and a contradiction is implied unless $k^2\eta/2 \geq \delta$. Thus the Arrow theorem for the η-IIA definition holds with $\eta(\varepsilon) = 2\delta/k^2$.(We briefly note that moving from F to G does not preserve the property of the function being balanced so in the setting of Kalai's theorem an additional argument is needed)

8.2 Proof of Theorem 8.3

Proof. The proof follows by applying Theorem 7.13 to triplets of alternatives. Assume $P(F) < \delta(\varepsilon)$.

Note that if $g_1, g_2 : \{-1, 1\}^n \to \{-1, 1\}$ are two different function each of which is either a dictator or a constant function than $D(g_1, g_2) \geq 1/2$. Therefore for all a, b it holds that $D(f^{a>b}, g) < \varepsilon/10$ for at most one function g which is either a dictator or a constant function. In case there exists such function we let $g^{a>b} = g$, otherwise, we let $g^{a>b} = f^{a>b}$.

Let G be the social choice function defined by the functions $g^{a>b}$. Clearly:

$$D(F, G) < \binom{k}{2}\varepsilon < k^2\varepsilon.$$

The proof would follow if we could show $P(G) = 0$ and therefore $G \in \mathcal{F}_k(n)$.

To prove that $G \in F_k(n)$ is suffices to show that for every set A of three alternatives, it holds that $G_A \in \mathcal{F}_3(n)$. Since $P(F) < \delta$ implies $P(F_A) < \delta$, Theorem 7.13 implies that there exists a function $H_A \in F_3(n)$ s.t. $D(H_A, F_A) < \varepsilon$. There are two cases to consider:

- H_A is a dictator. This implies that $f^{a>b}$ is ε close to a dictator for each a, b and therefore $f^{a>b} = g^{a>b}$ for all pairs a, b, so $G_A = H_A \in \mathcal{F}_3(n)$.
- There exists an alternative (say a) that H_A always ranks at the top/bottom. In this case we have that $f^{a>b}$ and $f^{c>a}$ are at most ε far from the constant functions 1 and -1 (or -1 and 1). The functions $g^{a>b}$ and $g^{c>a}$ have to take the same constant values and therefore again we have that $G_A \in \mathcal{F}_3(n)$.

The proof follows.

□

Remark 8.7. Note that this proof is generic in the sense that it takes the quantitative Arrow's result for 3 alternatives as a black box and produces a quantitative Arrow result for any $k \geq 3$ alternatives.

9 Low Influence Optimality for $k \geq 3$ alternatives

When we are considering $k \geq 3$ alternatives, we want to define more formally the possible outcome in Arrow voting. Since for every two alternatives, a winner is decided, the aggregation results in a *tournament* G_k on the set $[k]$. Recall that G_k is a *tournament* on $[k]$ if it is a directed graph on the vertex set $[k]$ such that for all $a, b \in [k]$ either $(a > b) \in G_k$ or $(b > a) \in G_k$. Given individual rankings $(\sigma_i)_{i=1}^n$ the tournament G_k is defined as follows.

Let $x^{a>b}(i) = 1$, if $\sigma_i(a) > \sigma_i(b)$, and $x^{a>b}(i) = -1$ if $\sigma_i(a) < \sigma_i(b)$. Note that $x^{b>a} = -x^{a>b}$.

The binary decision between each pair of candidates is performed via a anti-symmetric function $f : \{-1,1\}^n \to \{-1,1\}$ so that $f(-x) = -f(x)$ for all $x \in \{-1,1\}$. The tournament $G_k = G_k(\sigma; f)$ is then defined by letting $(a > b) \in G_k$ if and only if $f(x^{a>b}) = 1$.

Note that there are $2^{\binom{k}{2}}$tournaments while there are only $k! = 2^{\Theta(k \log k)}$ linear rankings. For the purposes of social choice, some tournaments make more sense than others.

Definition 9.1. We say that a tournament G_k is *linear* if it is acyclic. We will write $\mathrm{Acyc}(G_k)$ for the logical statement that G_k is acyclic. Non-linear tournaments are often referred to as non-rational in economics as they represent an order where there are 3 candidates a, b and c such that a is preferred to b, b is preferred to c and c is preferred to a.

We say that the tournament G_k is a *unique max tournament* if there is a candidate $a \in [k]$ such that for all $b \neq a$ it holds that $(a > b) \in G_k$. We write $\mathrm{UniqueBest}(G_k)$ for the logical statement that G_k has a unique max. Note that the unique max property is weaker than linearity. It corresponds to the fact that there is a candidate that dominates all other candidates.

A generalization of Borell's result along with a general invariance principle [42, 44] allows to prove the following [32]:

Theorem 9.2. *For any $k \geq 1$ and $\epsilon > 0$ there exists a $\tau(\epsilon, k) > 0$ such that for any anti-symmetric $f : \{-1,1\}^n \to \{0,1\}$ satisfying $\max_i \mathrm{Inf}_i f \leq \tau$,*

$$\mathbb{P}[\mathrm{UniqueBest}_k(f)] \leq \lim_{n\to\infty} \mathbb{P}[\mathrm{UniqueBest}_k(\mathrm{Maj}_n)] + \epsilon \qquad (44)$$

Other than the case $k = 3$, where the notions of unique-max and linear tournaments coincide, very little is known about which function maximizes the probability of a linear order. Even computing this probability for Majority provides a surprising result: Interestingly, we are not able to derive similar results for Acyc. We do calculate the probability that Acyc holds for majority.

Proposition 9.3. *We have*

$$\lim_{n\to\infty} \mathbb{P}[\mathrm{Acyc}(G_k(\sigma;\mathrm{Maj}_n))] = \exp(-\Theta(k^{5/3})). \tag{45}$$

We find this asymptotic behavior quite surprising. Indeed, given the previous results that the probability that there is a unique max is $k^{-1+o(1)}$, one may expect that the probability that the order is linear would be

$$k^{-1+o(1)}(k-1)^{-1+o(1)}\ldots = (k!)^{-1+o(1)}.$$

However, it turns out that there is a strong negative correlation between the event that there is a unique maximum among the k candidates and that among the other candidates there is a unique max. We note that results in economics [7] have shown that for majority vote the probability that the outcome will contain a Hamiltonian cycle when the number of voters goes to infinity is $1 - o_k(1)$.

Proof. We use the multi-dimensional CLT. Let

$$X_{a>b} = \frac{1}{\sqrt{n}}\left(|\{\sigma : \sigma(a) > \sigma(b)\}| - |\{\sigma : \sigma(b) > \sigma(a)\}|\right)$$

By the CLT at the limit the collection of variables $(X_{a>b})_{a\neq b}$ converges to a joint Gaussian vector $(N_{a>b})_{a\neq b}$ satisfying for all distinct a, b, c, d:

$$N_{a>b} = -N_{b>a}, \quad \mathrm{Cov}[N_{a>b}, N_{a>c}] = \frac{1}{3}, \quad \mathrm{Cov}[N_{a>b}, N_{c>d}] = 0.$$

and $N_{a>b} \sim N(0,1)$ for all a and b.

We are interested in providing bounds on

$$P[\forall a > b : N_{a>b} > 0]$$

as the probability that the resulting tournament is an order is obtained by multiplying by a $k! = \exp(\Theta(k \log k))$ factor.

We claim that there exist independent $N(0,1)$ random variables X_a for $1 \le a \le k$ and $Z_{a>b}$ for $1 \le a \neq b \le k$ such that

$$N_{a>b} = \frac{1}{\sqrt{3}}(X_a - X_b + Z_{a>b})$$

(where $Z_{a>b} = -Z_{b>a}$). This follows from the fact that the joint distribution of Gaussian random variables is determined by the covariance matrix (this is noted in the literature in [58]).

We now prove the upper bound. Let α be a constant to be chosen later. Note that for all α and large enough k it holds that:

$$P[|X_a| > k^\alpha] \leq \exp(-\Omega(k^{2\alpha})).$$

Therefore the probability that for at least half of the a's in the interval $[k/2, k]$ it holds that $|X_a| > k^\alpha$ is at most

$$\exp(-\Theta(k^{1+2\alpha})).$$

Let's assume that at least half of the a's in the interval $[k/2, k]$ satisfy that $|X_a| < k^\alpha$. We claim that in this case the number $H_{k/4}[-k^\alpha, k^\alpha]$ of pairs $a > b$ such that $X_a, X_b \in -[k^\alpha, k^\alpha]$ and $X_a - X_b < 1$ is $\Omega(k^{2-\alpha})$.

For the last claim partition the interval $[-k^\alpha, k^\alpha]$ into sub-intervals of length 1 and note that at least $\Omega(k)$ of the points belong to sub-intervals which contain at least $\Omega(k^{1-\alpha})$ points. This implies that the number of pairs $a > b$ satisfying $|X_a - X_b| < 1$ is $\Omega(k^{2-\alpha})$.

Note that for such pair $a > b$ in order that $N_{a>b} > 0$ we need that $Z_{a>b} > -1$ which happens with constant probability.

We conclude that given that half of the X's fall in $[-k^\alpha, k^\alpha]$ the probability of a linear order is bounded by

$$\exp(-\Omega(k^{2-\alpha})).$$

Thus overall we have bounded the probability by

$$\exp(-\Omega(k^{1+2\alpha})) + \exp(-\Omega(k^{2-\alpha})).$$

The optimal exponent is $\alpha = 1/3$ giving the desired upper bound.

For the lower bound we condition on X_a taking value in $(a, a+1)k^{-2/3}$. Each probability is at least $\exp(-O(k^{2/3}))$ and therefore the probability that all X_a take such values is

$$\exp(-O(k^{5/3})).$$

Moreover, conditioned on X_a taking such values the probability that

$$Z_{a>b} > X_b - X_a,$$

for all $a > b$ is at least

$$\left(\prod_{i=0}^{k-1} \Phi(i)^{k^{2/3}}\right)^k \geq \left(\prod_{i=0}^{\infty} \Phi(i)\right)^{k^{5/3}} = \exp(-O(k^{5/3})).$$

This proves the required result. $\square$

10 Manipulation, Isoperimetry and Concentration of Measure

10.1 Quantitive Manipulation and Isoperimetry

The GS Theorem, Theorem 1.4, states that under natural conditions, there exist profiles of voters such that at least one voter can manipulate. The basic approach for the proof is to view the theorem as an isoperimetric theorem. In classical Isoperimetric theory, the goal is to find conditions that establish large boundary between sets. In the context of manipulation we can consider a voter who can manipulate as a special boundary point and our goal is to prove that there are many boundary points.

It is natural to consider the following graph where the vertex set is S_k^n — the set of all voting profiles and there are edges between voting profiles that differ at a single voter. The statement of the GS theorem can be interpreted in terms of this graph: for certain natural partitions of S_k^n into k parts there is an edge of the graph between two different parts that corresponds to a manipulation. We will see that the existence of many edges between different parts of the graph follows from classical Isoperimetric theory.

Thus one may consider quantitative statements of the GS theorem as isoperimetric statements: It is not only the case that there are many edges between different parts of the partition, but it is also the case that many of these edges correspond to manipulation by one of the voters.

In the classical setup, isoperimetry and concentration of measure are closely related. In particular, as we will see below, for any set of fractional size at least ε in S_k^n, the set of profiles at graph distance at most $C(\varepsilon)\sqrt{n}$ contains almost the whole graph. One would hope that this statement implies that typically a small coalition can manipulate, but this is not known.

add examples of Plurality, single voter and maniuplation, also reverse tribes

10.2 A quantitive GS Theorem

Our goal in this section is to prove a quantitative version of the manipulation Theorem. We will mostly follow [53, 54] who proved a pretty general average manipulation theorem for a single voter. Some special cases of the Theorem were known before, in particular in the case of 3 alternatives this was proved by Friedgut, Kalai, Keller and Nisan [26, 24].

In this section we will prove that: if $k \geq 3$ and the SCF f is ε-far from the family of nonmanipulable functions, then the probability of a ranking profile being manipulable is bounded from below by a polynomial in $1/n$, $1/k$, and ε. We continue by first presenting our results, then discussing their implications, and finally we conclude this section by commenting on the techniques used in the proof.

10.3 Definition and Statement

Recall that our basic setup consists of n voters electing a winner among k alternatives via An SCF $f : S_k^n \to [k]$. We now define manipulability in more detail:

Definition 10.1 (Manipulation points)**.** Let $\sigma \in S_k^n$ be a ranking profile. Write $a \overset{\sigma_i}{>} b$ to denote that alternative a is preferred over b by voter i. A SCF $f \colon S_k^n \to [k]$ is *manipulable* at the ranking profile $\sigma \in S_k^n$ if there exists a $\sigma' \in S_k^n$ and an $i \in [n]$ such that σ and σ' only differ in the i^{th} coordinate and

$$f(\sigma') \overset{\sigma_i}{>} f(\sigma).$$

In this case we also say that σ is a *manipulation point* of f, and that (σ, σ') is *a manipulation pair* for f. We say that f is *manipulable* if it is manipulable at some point σ. We also say that σ is an r-*manipulation point* of f if f has a manipulation pair (σ, σ') such that σ' is obtained from σ by permuting (at most) r adjacent alternatives in one of the coordinates of σ. (We allow $r > k$—any manipulation point is an r-manipulation point for $r > k$.)

Let $M(f)$ denote the set of manipulation points of the SCF f, and for a given r, let $M_r(f)$ denote the set of r-manipulation points of f. When the SCF is obvious from the context, we write simply M and M_r.

We first recall Gibbard and Satterthwaite theorem (stated as Theorem 1.4 in the introduction):

Theorem 10.2 (Gibbard-Satterthwaite [27, 62])**.** *Any SCF $f \colon S_k^n \to [k]$ which takes at least three values and is not a dictator (i.e., not a function of only one voter) is manipulable.*

This theorem is tight in the sense that *monotone* SCFs which are dictators or only have two possible outcomes are indeed nonmanipulable (a function is non-monotone, and clearly manipulable, if for some ranking profile a voter can change the outcome from, say, a to b by moving a ahead of b in her preference). It is useful to introduce a refined notion of a dictator before defining the set of nonmanipulable SCFs.

Definition 10.3 (Dictator on a subset)**.** For a subset of alternatives $H \subseteq [k]$, let top_H be the SCF on one voter whose output is always the top ranked alternative among those in H.

Definition 10.4 (Nonmanipulable SCFs)**.** We denote by NONMANIP $\equiv$ NONMANIP (n, k) the set of nonmanipulable SCFs, which is the following:

$$\begin{aligned}\text{NONMANIP}(n,k) = &\{f : S_k^n \to [k] \mid f(\sigma) = \text{top}_H(\sigma_i) \text{ for some } i \in [n], H \subseteq [k], H \neq \emptyset\} \\ &\cup \{f : S_k^n \to [k] \mid f \text{ is a monotone function taking on exactly two values}\}.\end{aligned}$$

When the parameters n and k are obvious from the context, we omit them.

Another useful class of functions, which is larger than NONMANIP, but which has a simpler description, is the following.

Definition 10.5. Define, for parameters n and k that remain implicit:

$\overline{\text{NONMANIP}} = \{f : S_k^n \to [k] \mid f \text{ only depends on one coordinate or takes at most two values}\}.$

The notation should be thought of as "closure" rather than "complement".

As discussed previously, our goal is to study manipulability from a quantitative viewpoint, and in order to do so we need to define the distance between SCFs.

Definition 10.6 (Distance between SCFs)**.** The distance $\mathbf{D}(f,g)$ between two SCFs $f,g\colon S_k^n \to [k]$ is defined as the fraction of inputs on which they differ: $\mathbf{D}(f,g) = \mathbb{P}(f(\sigma) \neq g(\sigma))$, where $\sigma \in S_k^n$ is uniformly selected. For a class G of SCFs, we write $\mathbf{D}(f,G) = \min_{g\in G} \mathbf{D}(f,g)$.

The concepts of anonymity and neutrality of SCFs will be important to us, so we define them here.

Definition 10.7 (Anonymity)**.** A SCF is *anonymous* if it is invariant under changes made to the names of the voters. More precisely, a SCF $f : S_k^n \to [k]$ is anonymous if for every $\sigma = (\sigma_1,\dots,\sigma_n) \in S_k^n$ and every $\pi \in S_n$,

$$f(\sigma_1,\dots,\sigma_n) = f\left(\sigma_{\pi(1)},\dots,\sigma_{\pi(n)}\right).$$

Definition 10.8 (Neutrality)**.** A SCF is *neutral* if it commutes with changes made to the names of the alternatives. More precisely, a SCF $f : S_k^n \to [k]$ is neutral if for every $\sigma = (\sigma_1,\dots,\sigma_n) \in S_k^n$ and every $\pi \in S_k$,

$$f(\pi\circ\sigma_1,\dots,\pi\circ\sigma_n) = \pi(f(\sigma)).$$

Our goal is to sketch the proof of the following Theorem:

Theorem 10.9. *Suppose we have $n \geq 1$ voters, $k \geq 3$ alternatives, and a SCF $f : S_k^n \to [k]$ satisfying $\mathbf{D}(f, \overline{\text{NONMANIP}}) \geq \varepsilon$. Then*

$$\mathbb{P}(\sigma \in M(f)) \geq \mathbb{P}(\sigma \in M_4(f)) \geq p\left(\varepsilon, \frac{1}{n}, \frac{1}{k}\right) \tag{46}$$

for some polynomial p, where $\sigma \in S_k^n$ is selected uniformly.

An immediate consequence is that

$$\mathbb{P}\left((\sigma,\sigma') \text{ is a manipulation pair for } f\right) \geq q\left(\varepsilon, \frac{1}{n}, \frac{1}{k}\right)$$

for some polynomial q, where $\sigma \in S_k^n$ is uniformly selected, and σ' is obtained from σ by uniformly selecting a coordinate $i \in \{1,\dots,n\}$, uniformly selecting $j \in \{1,\dots,n-3\}$, and then uniformly randomly permuting the following four adjacent alternatives in σ_i: $\sigma_i(j), \sigma_i(j+1), \sigma_i(j+2)$, and $\sigma_i(j+3)$.

10.4 Proof Ideas

We first present our techniques that achieve a lower bound for the probability of manipulation that involves factors of $\frac{1}{k!}$ and then describe how a refined approach leads to a lower bound which has inverse polynomial dependence on k.

Rankings graph and applying the original Gibbard-Satterthwaite theorem. Consider the graph $G = (V, E)$ having vertex set $V = S_k^n$, the set of all ranking profiles, and let $(\sigma, \sigma') \in E$ if and only if σ and σ' differ in exactly one coordinate. The SCF $f : S_k^n \to [k]$ naturally partitions V into k subsets. Since every manipulation point must be on the boundary between two such subsets, we are interested in the size of such boundaries.

For two alternatives a and b, and voter i, denote by $B_i^{a,b}$ the boundary between $f^{-1}(a)$ and $f^{-1}(b)$ in voter i. A simple lemma tells us that at least two of the boundaries are large; in the following assume that these are $B_1^{a,b}$ and $B_2^{a,c}$. Now if a ranking profile σ lies on *both* of these boundaries, then applying the original Gibbard-Satterthwaite theorem to the restricted SCF on two voters where we fix all coordinates of σ except the first two, we get that there must exist a manipulation point which agrees with σ in all but the first two coordinates. Consequently, if we can show that the *intersection* of the boundaries $B_1^{a,b}$ and $B_2^{a,c}$ is large, then we have many manipulation points.

Fibers and reverse hypercontractivity. In order to have more "control" over what is happening at the boundaries, we partition the graph further—this idea is due to Friedgut et al. [26, 24]. Given a ranking profile σ and two alternatives a and b, σ induces a *vector of preferences* $x^{a,b}(\sigma) \in \{-1,1\}^n$ between a and b. For a vector $z^{a,b} \in \{-1,1\}^n$ we define the *fiber with respect to preferences between* a *and* b, denoted by $F\left(z^{a,b}\right)$, to be the set of ranking profiles for which the vector of preferences between a and b is $z^{a,b}$. We can then partition the vertex set V into such fibers, and work inside each fiber separately. Working inside a specific fiber is advantageous, because it gives us the extra knowledge of the vector of preferences between a and b.

We distinguish two types of fibers: large and small. We say that a fiber w.r.t. preferences between a and b is *large* if almost all of the ranking profiles in this fiber lie on the boundary $B_1^{a,b}$, and *small* otherwise. Now since the boundary $B_1^{a,b}$ is large, either there is big mass on the large fibers w.r.t. preferences between a and b or big mass on the small fibers. This holds analogously for the boundary $B_2^{a,c}$ and fibers w.r.t. preferences between a and c.

Consider the case when there is big mass on the large fibers of both $B_1^{a,b}$ and $B_2^{a,c}$. Notice that for a ranking profile σ, being in a fiber w.r.t. preferences between a and b only depends on the vector of preferences between a and b, $x^{a,b}(\sigma)$, which is a uniform bit vector. Similarly, being in a fiber w.r.t. preferences between a and c only depends on $x^{a,c}(\sigma)$. Moreover, we know the exact correlation between the coordinates of $x^{a,b}(\sigma)$ and $x^{a,c}(\sigma)$, and it is in exactly this setting where *reverse hypercontractivity* applies and shows that the *intersection* of the large fibers of $B_1^{a,b}$ and $B_2^{a,c}$ is also large. Finally, by the definition of a large fiber it follows that the intersection of the *boundaries* $B_1^{a,b}$ and $B_2^{a,c}$ is large as well, and we can finish the argument using the Gibbard-Satterthwaite theorem as above.

To deal with the case when there is big mass on the *small* fibers of $B_1^{a,b}$ we use various isoperimetric techniques, including the canonical path method. In particular, we use the fact that for a small fiber for $B_1^{a,b}$, the size of the boundary of $B_1^{a,b}$ in the small fiber is comparable to the size of $B_1^{a,b}$ in the small fiber itself, up to polynomial factors.

A refined geometry. Using this approach with the rankings graph above, our bound includes $\frac{1}{k!}$ factors. In order to obtain inverse polynomial dependence on k we use a refined approach. Instead of the rankings graph outlined above, we use an underlying graph with a different edge structure: $(\sigma, \sigma') \in E$ if and only if σ and σ' differ in exactly one coordinate, and in this coordinate they differ by a single adjacent transposition. In order to prove the refined result, we need to show that the geometric and combinatorial quantities such as boundaries and manipulation points are roughly the same in the refined graph as in the original rankings graph. In particular, this is where we need to analyze carefully functions of one voter, and ultimately prove a quantitative Gibbard-Satterthwaite theorem for one voter.

10.5 Isoperimetric Results

10.5.1 Boundaries and influences

For a general graph $G = (V, E)$, and a subset of the vertices $A \subseteq G$, we define the *edge boundary* of A as

$$\partial_e(A) = \{(u, v) \in E : u \in A, v \notin A\}.$$

We also define the *boundary* (or vertex boundary) of a subset of the vertices $A \subseteq G$ to be the set of vertices in A which have a neighbor that is not in A:

$$\partial(A) = \{u \in A : \text{ there exists } v \notin A \text{ such that } (u, v) \in E\}.$$

If $u \in \partial(A)$, we also say that u is *on* the edge boundary of A.

Definition 10.10 (Boundaries)**.** For a given SCF f and a given alternative $a \in [k]$, we define

$$H^a(f) = \{\sigma \in S_k^n : f(\sigma) = a\},$$

the set of ranking profiles where the outcome of the vote is a. The edge boundary of this set is denoted by $B^a(f) : B^a(f) = \partial_e(H^a(f))$. This boundary can be partitioned: we say that the edge boundary of $H^a(f)$ in the direction of the i^{th} coordinate is

$$B_i^a(f) = \left\{(\sigma, \sigma') \in B^a(f) : \sigma_i \neq \sigma_i'\right\}.$$

The boundary $B^a(f)$ can be therefore written as $B^a(f) = \cup_{i=1}^n B_i^a(f)$. We can also define the boundary between two alternatives a and b in the direction of the i^{th} coordinate:

$$B_i^{a,b}(f) = \left\{(\sigma, \sigma') \in B_i^a(f) : f(\sigma') = b\right\}.$$

We also say that $\sigma \in B_i^a(f)$ is *on* the boundary $B_i^{a,b}(f)$ if there exists σ' such that $(\sigma, \sigma') \in B_i^{a,b}(f)$.

We will need to generalize the definition of influences as follows:

Definition 10.11 (Influences)**.** We define the *influence* of the i^{th} coordinate on f as

$$\operatorname{Inf}_i(f) = \mathbb{P}\left(f(\sigma) \neq f\left(\sigma^{(i)}\right)\right) = \mathbb{P}\left(\left(\sigma, \sigma^{(i)}\right) \in \cup_{a=1}^k B_i^a(f)\right),$$

where σ is uniform on S_k^n and $\sigma^{(i)}$ is obtained from σ by rerandomizing the i^{th} coordinate. Similarly, we define the influence of the i^{th} coordinate with respect to a single alternative $a \in [k]$ or a pair of alternatives $a, b \in [k]$ as

$$\operatorname{Inf}_i^a(f) = \mathbb{P}\left(f(\sigma) = a, f\left(\sigma^{(i)}\right) \neq a\right) = \mathbb{P}\left(\left(\sigma, \sigma^{(i)}\right) \in B_i^a(f)\right),$$

and

$$\operatorname{Inf}_i^{a,b}(f) = \mathbb{P}\left(f(\sigma) = a, f\left(\sigma^{(i)}\right) = b\right) = \mathbb{P}\left(\left(\sigma, \sigma^{(i)}\right) \in B_i^{a,b}(f)\right),$$

respectively.

Clearly

$$\operatorname{Inf}_i(f) = \sum_{a=1}^k \operatorname{Inf}_i^a(f) = \sum_{a,b \in [k]: a \neq b} \operatorname{Inf}_i^{a,b}(f).$$

Most of the time the specific SCF f will be clear from the context, in which case we omit the dependence on f, and write simply $B^a \equiv B^a(f)$, $B_i^a \equiv B_i^a(f)$, etc.

10.5.2 Large boundaries

The following standard proposition bounds the total influence with respect to a given candidate from below by the variance with respect to that candidate.

Proposition 10.12. *For any $f \colon S_k^n \to [k]$ and $a \in [k]$,*

$$\sum_{i=1}^n \operatorname{Inf}_i^a(f) \geq \operatorname{Var}[1_{\{f(X)=a\}}] \tag{47}$$

where $X \in S_k^n$ is uniformly selected.

Proof. Create a random walk $X = X^{(0)}, \ldots, X^{(n)} = Y$ from X by re-randomizing the i:th coordinate in the i:th step, i.e. for $i \in [n]$, $X^{(i)} \in S_k^n$ is obtained by re-randomizing the i:th coordinate of $X^{(i-1)}$. Letting $g(x) = 1_{\{f(x)=a\}}$ and using that X, Y are independent and that if $g(X) \neq g(Y)$ then the value of g has to change at some edge on the path we have

$$\begin{aligned} 2\operatorname{Var}[1_{\{f(X)=a\}}] &= 2\operatorname{Var} g(X) = \mathbb{P}(g(X) \neq g(Y)) \leq \\ &\leq \mathbb{P}(\cup_{i\in[n]}\{g(X^{(i-1)}) \neq g(X^{(i)})\}) \leq \sum_{i=1}^{n} 2\operatorname{Inf}_i^a(f) \end{aligned}$$

□

Further, if a function is far from all constants not all such variances can be small:

Lemma 10.13. *For any $f\colon S_k^n \to [k]$,*

$$\mathbf{D}(f, \mathrm{CONST}) \leq \frac{q}{2}\sum_{a=1}^{k} \operatorname{Var}[1_{\{f(X)=a\}}] \tag{48}$$

Proof. For $a \in [k]$, let $\mu_a = \mathbb{P}(f(X) = a)$ and assume w.l.o.g. that $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_q$. Then,

$$\begin{aligned} \mathbf{D}(f, \mathrm{CONST}) = (1-\mu_1) \leq k\mu_1(1-\mu_1) = \frac{k}{2}\left(1-\mu_1^2-(1-\mu_1)^2\right) \leq \\ \leq \frac{k}{2}\left(1-\sum_{a=1}^{q}\mu_a^2\right) = \frac{k}{2}\sum_{a=1}^{q}\mu_a - \mu_a^2 = \frac{k}{2}\sum_{a=1}^{q}\operatorname{Var}[1_{\{f(X)=a\}}] \end{aligned}$$

□

The following lemma shows that there are some boundaries which are large

Lemma 10.14. *Fix $k \geq 3$ and $f\colon S_k^n \to [k]$ satisfying $\mathbf{D}(f, \overline{\mathrm{NONMANIP}}) \geq \varepsilon$. Then there exist distinct $i, j \in [n]$ and $\{a,b\}, \{c,d\} \subseteq [k]$ such that $c \notin \{a,b\}$ and*

$$\operatorname{Inf}_i^{a,b}(f) \geq \frac{2\varepsilon}{nk^2(k-1)} \quad \text{and} \quad \operatorname{Inf}_j^{c,d}(f) \geq \frac{2\varepsilon}{nk^2(k-1)}. \tag{49}$$

Proof. For $a \neq b$ let $A^{a,b} = \left\{i \in [n] \mid \operatorname{Inf}_i^{a,b} \geq \frac{2\varepsilon}{nk^2(k-1)}\right\}$.

We first claim that for all $\{a,b\}$ there exists $\{c,d\}$ such that $\{c,d\} \neq \{a,b\}$ and $A^{c,d} \neq \emptyset$. Note that f being ε-far from taking two values implies that we can find a $c \notin \{a,b\}$ such that $1 - \frac{\varepsilon}{q} \geq \mathbb{P}(f(X) = c) \geq \frac{\varepsilon}{k-2} \geq \frac{\varepsilon}{k}$. But then by Proposition 10.12 it holds that

$$\sum_{d\neq c}\sum_{i=1}^{n}\operatorname{Inf}_i^{c,d}(f) = \sum_{i=1}^{n}\operatorname{Inf}_i^{c}(f) \geq \operatorname{Var}[1_{\{f(X)=c\}}] \geq \frac{\varepsilon(1-\varepsilon/k)}{k} \geq \frac{\varepsilon(k-1)}{k^2}$$

hence there must exist some $d \neq c$ and $i \in [n]$ such that $\mathrm{Inf}_i^{c,d} \geq \frac{\varepsilon}{nk^2} \geq \frac{2\varepsilon}{nk^2(k-1)}$, and thus $A^{c,d} \neq \emptyset$.

We next claim that

$$|\cup_{a,b} A^{a,b}| \geq 2 \tag{50}$$

To see this, assume the contrary, i.e. $\cup_{a,b} A^{a,b} \subseteq \{i\}$ for some $i \in [n]$. Then for all $j \neq i$ it holds that

$$\mathrm{Inf}_j(f) = \sum_{c,d} \mathrm{Inf}_j^{c,d}(f) < \frac{k(k-1)}{2} \frac{2\varepsilon}{nk^2(k-1)} = \frac{\varepsilon}{nk} \tag{51}$$

For $\sigma \in S_n$, let $f_\sigma(x) = f(x_1, \ldots, x_{i-1}, \sigma, x_{i+1}, \ldots, x_n)$ and note that for $j \neq i$,

$$\mathrm{Inf}_j(f) = \frac{1}{k!} \sum_{\sigma \in S_n} \mathrm{Inf}_j(f_\sigma) \tag{52}$$

while $\mathrm{Inf}_i(f_\sigma) = 0$. Hence, by (51), we have

$$\varepsilon > k \sum_{j \neq i} \mathrm{Inf}_j(f) = \frac{k}{k!} \sum_{j=1}^{n} \sum_{\sigma} \mathrm{Inf}_j(f_\sigma) \geq \frac{2}{k!} \sum_{\sigma} \mathbf{D}(f_\sigma, \mathrm{CONST}) = 2\,\mathbf{D}(f, \mathrm{DICT}_i)$$

where the second inequality follows from Lemma 10.13 and Proposition 10.12. But this means that f is $\varepsilon/2$-close to a dictator, contradicting the assumption that $\mathbf{D}(f, \mathrm{NONMANIP}) \geq \varepsilon$.

Hence (50) holds. Therefore we can either find $i \neq j$ and $\{a,b\} \neq \{c,d\}$ such that $i \in A^{a,b}$ and $j \in A^{c,d}$ which proves the theorem, or we must have $|A^{a,b}| \geq 2$ for some $\{a,b\}$ while $A^{c,d} = \emptyset$ for any $\{c,d\} \neq \{a,b\}$. However, this contradicts the first claim in the proof. The result follows. □

10.5.3 General isoperimetric results

Our rankings graph is the Cartesian product of n complete graphs on $k!$ vertices. The edge-isoperimetric problem on the product of complete graphs was originally solved by Lindsey [41], implying the following result.

Corollary 10.15. *If $A \subseteq K_\ell \times \cdots \times K_\ell$ (n copies of the complete graph K_ℓ) and $|A| \leq \left(1 - \frac{1}{\ell}\right) \ell^n$, then $|\partial_e(A)| \geq |A|$.*

10.5.4 Fibers

In our proof we need to partition the graph even further—this idea is due to Friedgut, Kalai, Keller, and Nisan [26, 24].

Definition 10.16. For a ranking profile $\sigma \in S_k^n$ define the vector

$$x^{a,b} \equiv x^{a,b}(\sigma) = \left(x_1^{a,b}(\sigma), \ldots, x_n^{a,b}(\sigma)\right)$$

of preferences between a and b, where $x_i^{a,b}(\sigma) = 1$ if $a \overset{\sigma_i}{>} b$ and $x_i^{a,b}(\sigma) = -1$ otherwise.

Definition 10.17 (Fibers)**.** For a pair of alternatives $a, b \in [k]$ and a vector $z^{a,b} \in \{-1,1\}^n$, write

$$F\left(z^{a,b}\right) := \left\{\sigma : x^{a,b}(\sigma) = z^{a,b}\right\}.$$

We call the $F\left(z^{a,b}\right)$ *fibers* with respect to preferences between a and b.

So for any pair of alternatives a, b, we can partition the ranking profiles according to its fibers:

$$S_k^n = \bigcup_{z^{a,b} \in \{-1,1\}^n} F\left(z^{a,b}\right).$$

Given a SCF f, for any pair of alternatives $a, b \in [k]$ and $i \in [n]$, we can also partition the boundary $B_i^{a,b}(f)$ according to its fibers. There are multiple, slightly different ways of doing this, but for our purposes the following definition is most useful. Define

$$B_i\left(z^{a,b}\right) := \left\{\sigma \in F\left(z^{a,b}\right) : f(\sigma) = a, \text{ and there exists } \sigma' \text{ s.t. } \left(\sigma, \sigma'\right) \in B_i^{a,b}\right\},$$

where we omit the dependence of $B_i\left(z^{a,b}\right)$ on f. So $B_i\left(z^{a,b}\right) \subseteq F\left(z^{a,b}\right)$ is the set of vertices on the given fiber for which the outcome is a and which lies on the boundary between a and b in direction i. We call the sets of the form $B_i\left(z^{a,b}\right)$ *fibers for the boundary* $B_i^{a,b}$ (again omitting the dependence on f of both sets).

We now distinguish between small and large fibers for the boundary $B_i^{a,b}$.

Definition 10.18 (Small and large fibers)**.** We say that the fiber $B_i\left(z^{a,b}\right)$ is *large* if

$$\mathbb{P}\left(\sigma \in B_i\left(z^{a,b}\right) \,\middle|\, \sigma \in F\left(z^{a,b}\right)\right) \geq 1 - \frac{\varepsilon^3}{4n^3k^9}, \tag{53}$$

and *small* otherwise.

We denote by $\mathrm{Lg}\left(B_i^{a,b}\right)$ the union of large fibers for the boundary $B_i^{a,b}$, i.e.,

$$\mathrm{Lg}\left(B_i^{a,b}\right) := \left\{\sigma : B_i\left(x^{a,b}(\sigma)\right) \text{ is a large fiber, and } \sigma \in B_i\left(x^{a,b}(\sigma)\right)\right\}$$

and similarly, we denote by $\mathrm{Sm}\left(B_i^{a,b}\right)$ the union of small fibers.

We remark that what is important is that the fraction appearing on the right hand side of (53) is a polynomial of $\frac{1}{n}$, $\frac{1}{k}$ and ε—the specific polynomial in this definition is the end result of the computation in the proof.

Finally, for a voter i and a pair of alternatives $a, b \in [k]$, we define

$$F_i^{a,b} := \left\{\sigma : B_i\left(x^{a,b}(\sigma)\right) \text{ is a large fiber}\right\}.$$

So this means that

$$\mathbb{P}\left(\sigma \in \cup_{z^{a,b}} B_i\left(z^{a,b}\right) \middle| \sigma \in F_i^{a,b}\right) \geq 1 - \frac{\varepsilon^3}{4n^3k^9}. \tag{54}$$

10.5.5 Boundaries of boundaries

Finally, we also look at boundaries of boundaries. In particular, for a given vector $z^{a,b}$ of preferences between a and b, we can think of the fiber $F\left(z^{a,b}\right)$ as a subgraph of the original rankings graph. When we write $\partial\left(B_i\left(z^{a,b}\right)\right)$, we mean the boundary of $B_i\left(z^{a,b}\right)$ in the subgraph of the rankings graph induced by the fiber $F\left(z^{a,b}\right)$. That is,

$$\begin{aligned}\partial\left(B_i\left(z^{a,b}\right)\right) = \{\sigma \in B_i\left(z^{a,b}\right) : \exists\, \pi \in F\left(z^{a,b}\right) \setminus B_i\left(z^{a,b}\right) \text{ s.t.}\\ \sigma \text{ and } \pi \text{ differ in exactly one coordinate}\}.\end{aligned}$$

10.5.6 Dictators and miscellaneous definitions

For a ranking profile $\sigma = (\sigma_1, \ldots, \sigma_n)$ we sometimes write σ_{-i} for the collection of all coordinates except the i^{th} coordinate, i.e., $\sigma = (\sigma_i, \sigma_{-i})$. Furthermore, we sometimes distinguish two coordinates, e.g., we write $\sigma = \left(\sigma_1, \sigma_i, \sigma_{-\{1,i\}}\right)$.

Definition 10.19 (Induced SCF on one coordinate)**.** Let $f_{\sigma_{-i}}$ denote the SCF on one voter induced by f by fixing all voter preferences except the i^{th} one according to σ_{-i}. I.e.,

$$f_{\sigma_{-i}}(\cdot) := f(\cdot, \sigma_{-i}).$$

Recall Definition 10.3 of a dictator on a subset.

Definition 10.20 (Ranking profiles giving dictators on a subset)**.** For a coordinate i and a subset of alternatives $H \subseteq [k]$, define

$$D_i^H := \left\{\sigma_{-i} : f_{\sigma_{-i}}(\cdot) \equiv \text{top}_H(\cdot)\right\}.$$

Also, for a pair of alternatives a and b, define

$$D_i(a,b) := \bigcup_{H:\{a,b\}\subseteq H, |H|\geq 3} D_i^H.$$

11 Inverse polynomial manipulability for a fixed number of alternatives

We prove here the following theorem (Theorem 11.1 below), which is weaker than our main theorem, Theorem 10.9, in two aspects: first, the condition $\mathbf{D}(f, \text{NONMANIP}) \geq \varepsilon$ is replaced with the stronger condition $\mathbf{D}\left(f, \overline{\text{NONMANIP}}\right) \geq \varepsilon$, and second, we allow factors of $\frac{1}{k!}$ in our lower bounds for $\mathbb{P}(\sigma \in M(f))$. The advantage is that the proof of this statement is relatively simpler. We move on to getting a lower bound with polynomial dependence on k in the following sections, and finally we replace the condition $\mathbf{D}\left(f, \overline{\text{NONMANIP}}\right) \geq \varepsilon$ with $\mathbf{D}(f, \text{NONMANIP}) \geq \varepsilon$.

Theorem 11.1. *Suppose we have $n \geq 2$ voters, $k \geq 3$ alternatives, and a SCF $f : S_k^n \to [k]$ satisfying $\mathbf{D}\left(f, \overline{\text{NONMANIP}}\right) \geq \varepsilon$. Then*

$$\mathbb{P}(\sigma \in M(f)) \geq p\left(\varepsilon, \frac{1}{n}, \frac{1}{k!}\right), \tag{55}$$

for some polynomial p, where $\sigma \in S_k^n$ is selected uniformly.

An immediate consequence is that

$$\mathbb{P}\left(\left(\sigma, \sigma'\right) \text{ is a manipulation pair for } f\right) \geq q\left(\varepsilon, \frac{1}{n}, \frac{1}{k!}\right),$$

for some polynomial q, where $\sigma \in S_k^n$ is selected uniformly, and σ' is obtained from σ by uniformly selecting a coordinate $i \in \{1, \ldots, n\}$ and resetting the i^{th} coordinate to a random preference. In particular, the specific lower bound for $\mathbb{P}(\sigma \in M(f))$ implies that we can take $q\left(\varepsilon, \frac{1}{n}, \frac{1}{k}\right) = \frac{\varepsilon^5}{4n^8 k^{12} (k!)^5}$.

First we provide an overview of the proof of Theorem 11.1 in Section 11.1. In this overview we use adjectives such as "big", and "not too small" to describe probabilities—here these are all synonymous with "has probability at least an inverse polynomial of n, $k!$, and ε^{-1}".

11.1 Overview of proof

The tactic in proving Theorem 11.1 is roughly the following:

- By Lemma 10.14, we know that there are at least two boundaries which are big. W.l.o.g. we can assume that these are either $B_1^{a,b}$ and $B_2^{a,c}$, or $B_1^{a,b}$ and $B_2^{c,d}$ with $\{a, b\} \cap \{c, d\} = \emptyset$. Our proof works in both cases, but we continue the outline of the proof assuming the former case which is more difficult.
- We partition $B_1^{a,b}$ according to its fibers based on the preferences between a and b of the n voters. Similarly for $B_2^{a,c}$ and preferences between a and c.

- We can distinguish small and large fibers for these two boundaries. Now since $B_1^{a,b}$ is big, either the mass of small fibers, or the mass of large fibers is big. Similarly for $B_2^{a,c}$.

- Suppose first that there is big mass on large fibers in both $B_1^{a,b}$ and $B_2^{a,c}$. In this case the probability of our random ranking σ being in $F_1^{a,b}$ is big, and similarly for $F_2^{a,c}$. Being in $F_1^{a,b}$ only depends on the vector $x^{a,b}(\sigma)$ of preferences between a and b, and similarly being in $F_2^{a,c}$ only depends on the vector $x^{a,c}(\sigma)$ of preferences between a and c. We know the correlation between $x^{a,b}(\sigma)$ and $x^{a,c}(\sigma)$ and hence we can apply reverse hypercontractivity (Lemma 7.5), which tells us that the probability that σ lies in both $F_1^{a,b}$ and $F_2^{a,c}$ is big as well. If $\sigma \in F_1^{a,b}$, then voter 1 is pivotal between alternatives a and b with big probability, and similarly if $\sigma \in F_2^{a,c}$, then voter 2 is pivotal between alternatives a and c with big probability. So now we have that the probability that both voter 1 is pivotal between a and b and voter 2 is pivotal between a and c is big, and in this case the Gibbard-Satterthwaite theorem tells us that there is a manipulation point which agrees with this ranking profile in all except for perhaps the first two coordinates. So there are many manipulation points.

- Now suppose that the mass of small fibers in $B_1^{a,b}$ is big. By isoperimetric theory, the size of the boundary of every small fiber is comparable (same order up to $\text{poly}^{-1}\left(\varepsilon^{-1}, n, k!\right)$ factors) to the size of the small fiber. Consequently, the total size of the boundaries of small fibers is comparable to the total size of small fibers, which in this case has to be big.

  We then distinguish two cases: either we are on the boundary of a small fiber in the first coordinate, or some other coordinate. If σ is on the boundary of a small fiber in some coordinate $j \neq 1$, then the Gibbard-Satterthwaite theorem tells us that there is a manipulation point which agrees with σ in all coordinates except perhaps in coordinates 1 and j. If our ranking profile σ is on the boundary of a small fiber in the first coordinate, then either there exists a manipulation point which agrees with σ in all coordinates except perhaps the first, or the SCF on one voter that we obtain from f by fixing the votes of voters 2 through n to be σ_{-1} must be a dictator on some subset of the alternatives. So either we get sufficiently many manipulation points this way, or for many votes of voters 2 through n, the restricted SCF obtained from f by fixing these votes is a dictator on coordinate 1 on some subset of the alternatives.

  Finally, to deal with dictators on the first coordinate, we look at the boundary of the dictators. Since $\mathbf{D}\left(f, \overline{\text{NONMANIP}}\right) \geq \varepsilon$, the boundary is big, and we can also show that there is a manipulation point near every boundary point.

- If the mass of small fibers in $B_2^{a,c}$ is big, then we can do the same thing for this boundary.

11.2 Division into cases

For the remainder of Section 11, let us fix the number of voters $n \geq 2$, the number of alternatives $k \geq 3$, and the SCF f, which satisfies $\mathbf{D}\left(f, \overline{\text{NONMANIP}}\right) \geq \varepsilon$. Accordingly, we typically omit the dependence of various sets (e.g., boundaries between two alternatives) on f.

Our starting point is Lemma 10.14. W.l.o.g. we may assume that the two boundaries that the lemma gives us have $i = 1$ and $j = 2$, so the lemma tells us that

$$\mathbb{P}\left(\left(\sigma, \sigma^{(1)}\right) \in B_1^{a,b}\right) \geq \frac{2\varepsilon}{nk^3},$$

where σ is uniform on the ranking profiles, and $\sigma^{(1)}$ is obtained by rerandomizing the first coordinate. This also means that

$$\mathbb{P}\left(\sigma \in \cup_{z^{a,b}} B_1\left(z^{a,b}\right)\right) \geq \frac{2\varepsilon}{nk^3},$$

and similar inequalities hold for the boundary $B_2^{c,d}$. The following lemma is an immediate corollary.

Lemma 11.2. *Either*

$$\mathbb{P}\left(\sigma \in \text{Sm}\left(B_1^{a,b}\right)\right) \geq \frac{\varepsilon}{nk^3} \tag{56}$$

or

$$\mathbb{P}\left(\sigma \in \text{Lg}\left(B_1^{a,b}\right)\right) \geq \frac{\varepsilon}{nk^3}, \tag{57}$$

and the same can be said for the boundary $B_2^{c,d}$.

We distinguish cases based upon this: either (56) holds, or (56) holds for the boundary $B_2^{c,d}$, or (57) holds for both boundaries. We only need one boundary for the small fiber case, and we need both boundaries only in the large fiber case. So in the large fiber case we must differentiate between two cases: whether $d \in \{a, b\}$ or $d \notin \{a, b\}$. We will see that if $d \notin \{a, b\}$ then the large fiber case cannot occur—so this method of proof works as well.

In the rest of the section we first deal with the large fiber case, and then with the small fiber case.

11.3 Big mass on large fibers

We now deal with the case when

$$\mathbb{P}\left(\sigma \in \text{Lg}\left(B_1^{a,b}\right)\right) \geq \frac{\varepsilon}{nk^3} \tag{58}$$

and also

$$\mathbb{P}\left(\sigma \in \text{Lg}\left(B_2^{c,d}\right)\right) \geq \frac{\varepsilon}{nk^3}. \tag{59}$$

As mentioned before, we must differentiate between two cases: whether $d \in \{a, b\}$ or $d \notin \{a, b\}$.

11.3.1 Case 1

Suppose $d \in \{a, b\}$, in which case we may assume w.l.o.g. that $d = a$.

Lemma 11.3. *If*

$$\mathbb{P}\left(\sigma \in \mathrm{Lg}\left(B_1^{a,b}\right)\right) \geq \frac{\varepsilon}{nk^3} \qquad \textit{and} \qquad \mathbb{P}\left(\sigma \in \mathrm{Lg}\left(B_2^{a,c}\right)\right) \geq \frac{\varepsilon}{nk^3}, \tag{60}$$

then

$$\mathbb{P}\left(\sigma \in M\right) \geq \frac{\varepsilon^3}{2n^3 k^9 \left(k!\right)^2}. \tag{61}$$

Proof. By (60) we have that

$$\mathbb{P}\left(\sigma \in F_1^{a,b}\right) \geq \frac{\varepsilon}{nk^3} \qquad \text{and} \qquad \mathbb{P}\left(\sigma \in F_2^{a,c}\right) \geq \frac{\varepsilon}{nk^3}.$$

We know that $\left|\mathbb{E}\left(x_i^{a,b}\left(\sigma\right) x_i^{a,c}\left(\sigma\right)\right)\right| = 1/3$, and so by reverse hypercontractivity (Lemma 7.5) we have that

$$\mathbb{P}\left(\sigma \in F_1^{a,b} \cap F_2^{a,c}\right) \geq \frac{\varepsilon^3}{n^3 k^9}. \tag{62}$$

Recall that we say that σ is *on* the boundary $B_1^{a,b}$ if there exists σ' such that $(\sigma, \sigma') \in B_1^{a,b}$. If $\sigma \in F_1^{a,b}$, then with big probability σ is on the boundary $B_1^{a,b}$, and if $\sigma \in F_2^{a,c}$, then with big probability σ is on the boundary $B_2^{a,c}$. Using this and (62) we can show that the probability of σ lying on both the boundary $B_1^{a,b}$ and the boundary $B_2^{a,c}$ is big. Then we are done, because if σ lies on both $B_1^{a,b}$ and $B_2^{a,c}$, then by the Gibbard-Satterthwaite theorem there is a $\hat{\sigma}$ which agrees with σ on the last $n-2$ coordinates, and which is a manipulation point, and there can be at most $(k!)^2$ ranking profiles that give the same manipulation point. Let us do the computation:

$$\begin{aligned}
&\mathbb{P}\left(\sigma \text{ on } B_1^{a,b}, \sigma \text{ on } B_2^{a,c}\right) \geq \mathbb{P}\left(\sigma \text{ on } B_1^{a,b}, \sigma \text{ on } B_2^{a,c}, \sigma \in F_1^{a,b} \cap F_2^{a,c}\right) \\
&\geq \mathbb{P}\left(\sigma \in F_1^{a,b} \cap F_2^{a,c}\right) - \mathbb{P}\left(\sigma \in F_1^{a,b} \cap F_2^{a,c}, \sigma \text{ not on } B_1^{a,b}\right) - \mathbb{P}\left(\sigma \in F_1^{a,b} \cap F_2^{a,c}, \sigma \text{ not on } B_2^{a,c}\right).
\end{aligned}$$

The first term is bounded below via (62), while the other two terms can be bounded using (54):

$$\mathbb{P}\left(\sigma \in F_1^{a,b} \cap F_2^{a,c}, \sigma \text{ not on } B_1^{a,b}\right) \leq \mathbb{P}\left(\sigma \in F_1^{a,b}, \sigma \text{ not on } B_1^{a,b}\right) \leq \mathbb{P}\left(\sigma \text{ not on } B_1^{a,b} \,\middle|\, \sigma \in F_1^{a,b}\right) \leq \frac{\varepsilon^3}{4n^3 k^9},$$

and similarly for the other term. Putting everything together gives us

$$\mathbb{P}\left(\sigma \text{ on } B_1^{a,b}, \sigma \text{ on } B_2^{a,c}\right) \geq \frac{\varepsilon^3}{2n^3 k^9},$$

which by the discussion above implies (61). □

11.3.2 Case 2

Lemma 11.4. *If $d \notin \{a, b\}$, then* (58) *and* (59) *cannot hold simultaneously.*

Proof. Suppose on the contrary that (58) and (59) do both hold. Then

$$\mathbb{P}\left(\sigma \in F_1^{a,b}\right) \geq \frac{\varepsilon}{nk^3} \qquad \text{and} \qquad \mathbb{P}\left(\sigma \in F_2^{c,d}\right) \geq \frac{\varepsilon}{nk^3}$$

as before. Since $\{a, b\} \cap \{c, d\} = \emptyset$, $\left\{\sigma \in F_1^{a,b}\right\}$ and $\left\{\sigma \in F_2^{c,d}\right\}$ are independent events, and so

$$\mathbb{P}\left(\sigma \in F_1^{a,b} \cap F_2^{c,d}\right) = \mathbb{P}\left(\sigma \in F_1^{a,b}\right) \mathbb{P}\left(\sigma \in F_2^{c,d}\right) \geq \frac{\varepsilon^2}{n^2 k^6}.$$

In the same way as before, by the definition of large fibers this implies that

$$\mathbb{P}\left(\sigma \text{ on } B_1^{a,b}, \sigma \text{ on } B_2^{c,d}\right) \geq \frac{\varepsilon^2}{2n^2 k^6} > 0,$$

but it is clear that

$$\mathbb{P}\left(\sigma \text{ on } B_1^{a,b}, \sigma \text{ on } B_2^{c,d}\right) = 0,$$

since σ on $B_1^{a,b}$ and on $B_2^{c,d}$ requires $f(\sigma) \in \{a, b\} \cap \{c, d\} = \emptyset$. So we have reached a contradiction. □

11.4 Big mass on small fibers

We now deal with the case when (56) holds, i.e., when we have a big mass on the small fibers for the boundary $B_1^{a,b}$. We formalize the ideas of the outline described in Section 11.1 in a series of statements.

First, we want to formalize that the boundaries of the boundaries are big, when we are on a small fiber.

Lemma 11.5. *Fix coordinate 1 and the pair of alternatives a, b. Let $z^{a,b}$ be such that $B_1\left(z^{a,b}\right)$ is a small fiber for $B_1^{a,b}$. Then, writing $B \equiv B_1\left(z^{a,b}\right)$, we have*

$$\left|\partial_e(B)\right| \geq \frac{\varepsilon^3}{4n^3 k^9} |B|$$

and

$$\mathbb{P}(\sigma \in \partial(B)) \geq \frac{\varepsilon^3}{2n^4 k^9 k!} \mathbb{P}(\sigma \in B), \tag{63}$$

where both the edge boundary $\partial_e(B)$ and the boundary $\partial(B)$ are with respect to the induced subgraph $F\left(z^{a,b}\right)$, which is isomorphic to $K_{k!/2}^n$, the Cartesian product of n complete graphs of size $k!/2$.

Proof. We use Corollary 10.15 with $\ell = k!/2$ and the set A being either B or $B^c := F\left(z^{a,b}\right) \setminus B$. Suppose first that $|B| \le \left(1 - \frac{2}{k!}\right)(k!/2)^n$. Then $|\partial_e(B)| \ge |B|$. Suppose now that $|B| > \left(1 - \frac{2}{k!}\right)(k!/2)^n$. Since we are in the case of a small fiber, we also know that $|B| \le \left(1 - \frac{\varepsilon^3}{4n^3k^9}\right)(k!/2)^n$. Consequently, we get

$$|\partial_e(B)| = |\partial_e(B^c)| \ge |B^c| \ge \frac{\varepsilon^3}{4n^3k^9}|B|,$$

which proves the first claim.

A ranking profile in $F\left(z^{a,b}\right)$ has $(k!/2 - 1)\, n \le nk!/2$ neighbors in $F\left(z^{a,b}\right)$, which then implies (63). □

Corollary 11.6. *If* (56) *holds, then*

$$\mathbb{P}\left(\sigma \in \bigcup_{z^{a,b}} \partial\left(B_1\left(z^{a,b}\right)\right)\right) \ge \frac{\varepsilon^4}{2n^5k^{12}k!}.$$

Proof. Using the previous lemma and (56) we have

$$\begin{aligned}\mathbb{P}\left(\sigma \in \bigcup_{z^{a,b}} \partial\left(B_1\left(z^{a,b}\right)\right)\right) &= \sum_{z^{a,b}} \mathbb{P}\left(\sigma \in \partial\left(B_1\left(z^{a,b}\right)\right)\right) \ge \sum_{z^{a,b}:B_1\left(z^{a,b}\right)\subseteq \mathrm{Sm}\left(B_1^{a,b}\right)} \mathbb{P}\left(\sigma \in \partial\left(B_1\left(z^{a,b}\right)\right)\right)\\ &\ge \sum_{z^{a,b}:B_1\left(z^{a,b}\right)\subseteq \mathrm{Sm}\left(B_1^{a,b}\right)} \frac{\varepsilon^3}{2n^4k^9k!}\mathbb{P}\left(\sigma \in B_1\left(z^{a,b}\right)\right) = \frac{\varepsilon^3}{2n^4k^9k!}\mathbb{P}\left(\sigma \in \mathrm{Sm}\left(B_1^{a,b}\right)\right)\\ &\ge \frac{\varepsilon^4}{2n^5k^{12}k!}.\end{aligned}$$

□

Next, we want to find manipulation points on the boundaries of boundaries.

Lemma 11.7. *Suppose the ranking profile σ is on the boundary of a fiber for $B_1^{a,b}$, i.e.,*

$$\sigma \in \bigcup_{z^{a,b}} \partial\left(B_1\left(z^{a,b}\right)\right).$$

Then either $\sigma_{-1} \in D_1(a,b)$, or there exists a manipulation point $\hat{\sigma}$ which differs from σ in at most two coordinates, one of them being the first coordinate.

Proof. First of all, by our assumption that σ is on the boundary of a fiber for $B_1^{a,b}$, we know that $\sigma \in B_1\left(z^{a,b}\right)$ for some $z^{a,b}$, which means that there exists a ranking profile $\sigma' = (\sigma_1', \sigma_{-1})$ such that $(\sigma, \sigma') \in B_1^{a,b}$. We may assume $a \overset{\sigma_1}{>} b$ and $b \overset{\sigma_1'}{>} a$, or else either σ or σ' is a manipulation point.

Now since $\sigma \in \partial\left(B_1\left(z^{a,b}\right)\right)$ we also know that there exists a ranking profile $\pi = (\pi_j, \sigma_{-j}) \in F\left(z^{a,b}\right) \setminus B_1\left(z^{a,b}\right)$ for some $j \in [k]$. We distinguish two cases: $j \neq 1$ and $j = 1$.

Case 1: $\mathbf{j \neq 1}$. What does it mean for $\pi = (\pi_j, \sigma_{-j})$ to be on the same fiber as σ, but for π to not be in $B_1\left(z^{a,b}\right)$? First of all, being on the same fiber means that σ_j and π_j both rank a and b in the same order. Now $\pi \notin B_1\left(z^{a,b}\right)$ means that

- either $f(\pi) \neq a$;
- or $f(\pi) = a$ and $f(\pi'_1, \pi_{-1}) \neq b$ for every $\pi'_1 \in S_k$.

If $f(\pi) = b$, then either σ or π is a manipulation point, since the order of a and b is the same in both σ_j and π_j (since σ and π are on the same fiber).

Suppose $f(\pi) = c \notin \{a, b\}$. Then we can define a SCF function on two coordinates by fixing all coordinates except coordinates 1 and j to agree with the respective coordinates of σ—letting coordinates 1 and j vary we get a SCF function on two coordinates which takes on at least three values (a, b, and c), and does not only depend on one coordinate. Now applying the Gibbard-Satterthwaite theorem we get that this SCF on two coordinates has a manipulation point, which means that our original SCF f has a manipulation point which agrees with σ in all coordinates except perhaps in coordinates 1 and j.

So the final case is that $f(\pi) = a$ and $f(\pi'_1, \pi_{-1}) \neq b$ for every $\pi'_1 \in S_k$. In particular for $\tilde{\pi} := (\sigma'_1, \pi_{-1}) = \left(\pi_j, \sigma'_{-j}\right)$ we have $f(\tilde{\pi}) \neq b$. Now if $f(\tilde{\pi}) = a$ then either σ' or $\tilde{\pi}$ is a manipulation point, since the order of a and b is the same in both $\sigma'_j = \sigma_j$ and π_j. Finally, if $f(\tilde{\pi}) = c \notin \{a, b\}$, then we can apply the Gibbard-Satterthwaite theorem just like in the previous paragraph.

Case 2: $\mathbf{j = 1}$. We can again ask: what does it mean for $\pi = (\pi_1, \sigma_{-1})$ to be on the same fiber as σ, but for π to not be in $B_1\left(z^{a,b}\right)$? First of all, being on the same fiber means that σ_1 and π_1 both rank a and b in the same order (namely, as discussed at the beginning, ranking a above b, or else we have a manipulation point). Now $\pi \notin B_1\left(z^{a,b}\right)$ means that

- either $f(\pi) \neq a$;
- or $f(\pi) = a$ and $f(\pi'_1, \pi_{-1}) \neq b$ for every $\pi'_1 \in S_k$.

However, we know that $f(\sigma') = b$ and that σ' is of the form $\sigma' = (\sigma'_1, \sigma_{-1}) = (\sigma'_1, \pi_{-1})$, and so the only way we can have $\pi \notin B_1\left(z^{a,b}\right)$ is if $f(\pi) \neq a$.

If $f(\pi) = b$, then π is a manipulation point, since $a \overset{\pi_1}{>} b$ and $f(\sigma) = a$.

So the remaining case is if $f(\pi) = c \notin \{a, b\}$. This means that $f_{\sigma_{-1}}$ (see Definition 10.19) takes on at least three values. Denote by $H \subseteq [k]$ the range of $f_{\sigma_{-1}}$. Now either $\sigma_{-1} \in D_1^H \subseteq D_1(a, b)$, or there exists a manipulation point $\hat{\sigma}$ which agrees with σ in every coordinate except perhaps the first. □

Finally, we need to deal with dictators on the first coordinate.

Lemma 11.8. *Assume that* $\mathbf{D}\left(f, \overline{\text{NONMANIP}}\right) \geq \varepsilon$. *We have that either*

$$\mathbb{P}\left(\sigma_{-1} \in D_1(a,b)\right) \leq \frac{\varepsilon^4}{4n^5 k^{12} k!},$$

or

$$\mathbb{P}\left(\sigma \in M\right) \geq \frac{\varepsilon^5}{4n^7 k^{12} (k!)^4}. \tag{64}$$

Proof. Suppose $\mathbb{P}\left(\sigma_{-1} \in D_1(a,b)\right) \geq \frac{\varepsilon^4}{4n^5 k^{12} k!}$, which is the same as

$$\sum_{H:\{a,b\}\subseteq H, |H|\geq 3} \mathbb{P}\left(\sigma_{-1} \in D_1^H\right) \geq \frac{\varepsilon^4}{4n^5 k^{12} k!}. \tag{65}$$

Note that for every $H \subseteq [k]$ we have

$$\varepsilon \leq \mathbf{D}\left(f, \overline{\text{NONMANIP}}\right) \leq \mathbb{P}\left(f(\sigma) \neq \text{top}_H(\sigma_1)\right) \leq 1 - \mathbb{P}\left(D_1^H\right),$$

and so

$$\mathbb{P}\left(D_1^H\right) \leq 1 - \varepsilon. \tag{66}$$

The main idea is that (66) implies that the size of the boundary of D_1^H is comparable to the size of D_1^H, and if we are on the boundary of D_1^H, then there is a manipulation point nearby.

So first let us establish that the size of the boundary of D_1^H is comparable to the size of D_1^H. This is done along the same lines as the proof of Lemma 11.5.

Notice that $D_1^H \subseteq S_k^{n-1}$, where S_k^{n-1} should be thought of as the Cartesian product of $n-1$ copies of the complete graph on S_k. We apply Corollary 10.15 with $\ell = k!$ and with $n-1$ copies, and we see that if $\varepsilon \geq \frac{1}{k!}$, then $\left|\partial_e\left(D_1^H\right)\right| \geq \left|D_1^H\right|$. If $\varepsilon < \frac{1}{k!}$ and $1 - \frac{1}{k!} \leq \mathbb{P}\left(D_1^H\right) \leq 1 - \varepsilon$ then

$$\left|\partial_e\left(D_1^H\right)\right| = \left|\partial_e\left(\left(D_1^H\right)^c\right)\right| \geq \left|\left(D_1^H\right)^c\right| \geq \varepsilon \left|D_1^H\right|.$$

So in any case we have $\left|\partial_e\left(D_1^H\right)\right| \geq \varepsilon \left|D_1^H\right|$. Since σ_{-1} has $(n-1)(k!-1) \leq nk!$ neighbors in S_k^{n-1}, we have that

$$\mathbb{P}\left(\sigma_{-1} \in \partial\left(D_1^H\right)\right) \geq \frac{\varepsilon}{nk!} \mathbb{P}\left(\sigma_{-1} \in D_1^H\right).$$

Consequently, by (65), we have

$$\begin{aligned} \mathbb{P}\left(\sigma_{-1} \in \bigcup_{H:\{a,b\}\subseteq H, |H|\geq 3} \partial\left(D_1^H\right)\right) &= \sum_{H:\{a,b\}\subseteq H, |H|\geq 3} \mathbb{P}\left(\sigma_{-1} \in \partial\left(D_1^H\right)\right) \\ &\geq \sum_{H:\{a,b\}\subseteq H, |H|\geq 3} \frac{\varepsilon}{nk!} \mathbb{P}\left(\sigma_{-1} \in D_1^H\right) \geq \frac{\varepsilon^5}{4n^6 k^{12} (k!)^2}. \end{aligned}$$

Next, suppose $\sigma_{-1} \in \partial\left(D_1^H\right)$ for some H such that $\{a, b\} \subseteq H, |H| \geq 3$. We want to show that then there is a manipulation point "close" to σ_{-1} in some sense. To be more precise: for the manipulation point $\hat{\sigma}$, $\hat{\sigma}_{-1}$ will agree with σ_{-1} in all except maybe one coordinate.

If $\sigma_{-1} \in \partial\left(D_1^H\right)$, then there exist $j \in \{2, \ldots, n\}$ and σ'_j such that $\sigma'_{-1} := \left(\sigma'_j, \sigma_{-\{1,j\}}\right) \notin D_1^H$. That is, $f_{\sigma'_{-1}}(\cdot) \not\equiv \text{top}_H(\cdot)$. There can be two ways that this can happen—the two cases are outlined below. Denote by $H' \subseteq [k]$ the range of $f_{\sigma'_{-1}}$.

Case 1: $\mathbf{H}' = \mathbf{H}$. In this case we automatically know that there exists a manipulation point $\hat{\sigma}$ such that $\hat{\sigma}_{-1} = \sigma'_{-1}$, and so $\hat{\sigma}_{-1}$ agrees with σ_{-1} in all coordinates except coordinate j.

Case 2: $\mathbf{H}' \neq \mathbf{H}$. W.l.o.g. suppose $H' \setminus H \neq \emptyset$, and let $c \in H' \setminus H$. (The other case when $H \setminus H' \neq \emptyset$ works in exactly the same way.) First of all, we may assume that $f_{\sigma'_{-1}}(\cdot) \equiv \text{top}_{H'}(\cdot)$, because otherwise we have a manipulation point just like in Case 1.

We can define a SCF on two coordinates by fixing all coordinates except coordinate 1 and j to agree with σ_{-1}, and varying coordinates 1 and j. We know that the outcome takes on at least three different values, since $\sigma_{-1} \in D_1^H$, and $|H| \geq 3$.

Now let us show that this SCF is not a function of the first coordinate. Let σ_1 be a ranking which puts c first, and then a. Then $f(\sigma_1, \sigma_{-1}) = a$, but $f\left(\sigma_1, \sigma'_{-1}\right) = c$, which shows that this SCF is not a function of the first coordinate (since a change in coordinate j can change the outcome).

Consequently, the Gibbard-Satterthwaite theorem tells us that this SCF on two coordinates has a manipulation point, and therefore there exists a manipulation point $\hat{\sigma}$ for f such that $\hat{\sigma}_{-1}$ agrees with σ_{-1} in all coordinates except coordinate j.

Putting everything together yields (64). □

11.5 Proof of Theorem 11.1 concluded

Proof of Theorem 11.1. If (58) and (59) hold, then we are done by Lemmas 11.3 and 11.4.

If not, then either (56) holds, or (56) holds for the boundary $B_2^{c,d}$; w.l.o.g. assume that (56) holds.

By Corollary 11.6, we have

$$\mathbb{P}\left(\sigma \in \bigcup_{z^{a,b}} \partial\left(B_1\left(z^{a,b}\right)\right)\right) \geq \frac{\varepsilon^4}{2n^5 k^{12} k!}.$$

We may assume that $\mathbb{P}(\sigma_{-1} \in D_1(a, b)) \leq \frac{\varepsilon^4}{4n^5 k^{12} k!}$, since otherwise we are done by Lemma 11.8. Consequently, we then have

$$\mathbb{P}\left(\sigma \in \bigcup_{z^{a,b}} \partial\left(B_1\left(z^{a,b}\right)\right), \sigma_{-1} \notin D_1(a, b)\right) \geq \frac{\varepsilon^4}{4n^5 k^{12} k!}.$$

We can then finish our argument using Lemma 11.7:

$$\mathbb{P}\left(\sigma \in M\right) \geq \frac{1}{n\left(k!\right)^{2}}\mathbb{P}\left(\sigma \in \bigcup_{z^{a,b}} \partial\left(B_1\left(z^{a,b}\right)\right), \sigma_{-1} \notin D_1\left(a,b\right)\right) \geq \frac{\varepsilon^4}{4n^6k^{12}\left(k!\right)^3}. \qquad \square$$

12 An overview of the refined proof

In order to improve on the result of Theorem 11.1—in particular to get rid of the factor of $\frac{1}{(k!)^4}$—we need to refine the methods used in the previous section.

The key to the refined method is to consider the so-called *refined rankings graph* instead of the general rankings graph studied in Section 11. The vertices of this graph are again ranking profiles (elements of S_k^n), and two vertices are connected by an edge if they differ in exactly one coordinate, and by an adjacent transposition in that coordinate. Again, the SCF f naturally partitions the vertices of this graph into k subsets, depending on the value of f at a given vertex. Clearly a 2-manipulation point can only be on the edge boundary of such a subset in the refined rankings graph, and so it is important to study these boundaries.

One of the important steps of the proof in Section 11 is creating a configuration where we fix all but two coordinates, and the SCF f takes on at least three values when we vary these two coordinates—then we can define another SCF on two voters and k alternatives which must have a manipulation point by the Gibbard-Satterthwaite theorem. The advantage of the refined rankings graph is that we can create a configuration where we fix all but two coordinates, and in these two coordinates we also fix all but constantly many adjacent alternatives, and the SCF takes on at least three values when we vary these constantly many adjacent alternatives in the two coordinates. Then we can define another SCF on two voters and r alternatives, where r is a small constant, which must have a manipulation point by the Gibbard-Satterthwaite theorem. Since r is a constant, we only lose a constant factor in our estimates, not factors of $\frac{1}{k!}$.

We state the refined result in Theorem 16.1, which we also prove in Section 16. The proof of Theorem 16.1 follows the outline of the proof of Theorem 11.1: we know that there are at least two refined boundaries which are big

The difficulty is dealing with the case when we are on the boundary of a small fiber in the first coordinate. Suppose $\sigma = (\sigma_1, \sigma_{-1})$ is on such a boundary. We know that there are $k!$ ranking profiles which agree with σ in coordinates 2 through n. The difficulty comes from the fact that—in order to obtain a polynomial bound in k—we are only allowed to look at a polynomial number (in k) of these ranking profiles when searching for a manipulation point. If there is an r-manipulation point among them for some small constant r, then we are done. If this is not the case then σ is what we call a *local dictator* on some subset of the alternatives in coordinate 1. We say that σ is a local dictator on some subset $H \subseteq [k]$ of the alternatives in coordinate 1 if the alternatives in H are adjacent in σ_1, and permuting the alternatives in H in every possible way in the first coordinate, the outcome of the SCF f is always the top-ranked alternative in H.

So instead of dealing with dictators on some subset in coordinate 1, as in Section 11, we have to deal with *local dictators* on some subset in coordinate 1. This analysis involves essentially only the first coordinate, in essence proving a quantitative Gibbard-Satterthwaite theorem for one voter. This has not been studied in the literature before, and, moreover, we were not able to utilize previous quantitative Gibbard-Satterthwaite theorems to solve this problem easily. Hence we separate this argument from the rest of the proof of Theorem 16.1 and formulate a quantitative Gibbard-Satterthwaite theorem for one voter, Theorem 12.1, which is proven in Section 15. This proof forms the backbone for the proof of Theorem 16.1, which is then proven in Section 16.

Theorem 12.1. *Suppose $f : S_k \to [k]$ is a SCF on $n = 1$ voter and $k \geq 3$ alternatives which satisfies $\mathbf{D}(f, \text{NONMANIP}) \geq \varepsilon$. Then*

$$\mathbb{P}(\sigma \in M(f)) \geq \mathbb{P}(\sigma \in M_3(f)) \geq p\left(\varepsilon, \frac{1}{k}\right), \tag{67}$$

for some polynomial p, where $\sigma \in S_k$ is selected uniformly. In particular, we show a lower bound of $\frac{\varepsilon^3}{10^5 k^{16}}$.

13 Refined rankings graph—introduction and preliminaries

13.1 Transpositions, boundaries, and influences

Definition 13.1 (Adjacent transpositions)**.** Given two elements $a, b \in [k]$, the adjacent transposition $[a : b]$ between them is defined as follows. If $\sigma \in S_k$ has a and b adjacent, then $[a : b]\,\sigma$ is obtained from σ by exchanging a and b. Otherwise $[a : b]\,\sigma = \sigma$.

We let T denote the set of all $k(k-1)/2$ adjacent transpositions.

For $\sigma \in S_k^n$, we let $[a : b]_i\,\sigma$ denote the ranking profile obtained by applying $[a : b]$ on the i^{th} coordinate of σ while leaving all other coordinates unchanged.

Definition 13.2 (Boundaries)**.** For a given SCF f and a given alternative $a \in [k]$, we define

$$H^a(f) = \{\sigma \in S_k^n : f(\sigma) = a\},$$

the set of ranking profiles where the outcome of the vote is a. The edge boundary of this set (with respect to the underlying refined rankings graph) is denoted by $B^{a;T}(f)$: $B^{a;T}(f) = \partial_e(H^a(f))$. This boundary can be partitioned: we say that the edge boundary of $H^a(f)$ in the direction of the i^{th} coordinate is

$$B_i^{a;T}(f) = \left\{(\sigma, \sigma') \in B^{a;T}(f) : \sigma_i \neq \sigma'_i\right\}.$$

The boundary $B^a(f)$ can be therefore written as $B^{a;T}(f) = \cup_{i=1}^n B_i^{a;T}(f)$. We can also define the boundary between two alternatives a and b in the direction of the i^{th} coordinate:

$$B_i^{a,b;T}(f) = \left\{ \left(\sigma, \sigma'\right) \in B_i^{a;T}(f) : f\left(\sigma'\right) = b \right\}.$$

Moreover, we can define the boundary between two alternatives a and b in the direction of the i^{th} coordinate with respect to the adjacent transposition $z \in T$:

$$B_i^{a,b;z}(f) = \left\{ \left(\sigma, \sigma'\right) \in B_i^{a;T}(f) : \sigma' = z_i\sigma, f\left(\sigma'\right) = b \right\}.$$

We also say that σ is *on* the boundary $B_i^{a,b;z}(f)$ if $(\sigma, z_i\sigma) \in B_i^{a,b;z}(f)$. Clearly we have

$$B_i^{a,b;T}(f) = \bigcup_{z\in T} B_i^{a,b;z}(f).$$

Definition 13.3 (Influences)**.** Given $z \in T$, we define

$$\begin{aligned} \operatorname{Inf}_i^{a,b;z}(f) &= \mathbb{P}\left(f(\sigma) = a, f\left(\sigma^{(i)}\right) = b\right) \\ \operatorname{Inf}_i^{a;z}(f) &= \mathbb{P}\left(f(\sigma) = a, f\left(\sigma^{(i)}\right) \neq a\right) \\ \operatorname{Inf}_i^{a,b;T}(f) &= \sum_{z\in T} \operatorname{Inf}_i^{a,b;z}(f), \end{aligned}$$

where σ is uniformly distributed in S_k^n and $\sigma^{(i)}$ is obtained from σ by rerandomizing the i^{th} coordinate σ_i in the following way: with probability $1/2$ we keep it as σ_i, and otherwise we replace it by $z\sigma_i$.

Note that for $a \neq b$,

$$\operatorname{Inf}_i^{a,b;z}(f) = \frac{1}{2}\mathbb{P}\left(f(\sigma) = a, f(z_i\sigma) = b\right) = \frac{1}{2}\frac{\left|B_i^{a,b;z}(f)\right|}{(k!)^n}.$$

Again, most of the time the specific SCF f will be clear from the context, in which case we omit the dependence on f.

13.2 Canonical Paths and Group Actions

In order to derive the more refined result, we will need to consider in more detail the properties of the permutation group L_q with respect to adjacent transpositions. Again we use canonical paths arguments. We state the arguments in a more general setup.

Definition 13.4. Let L be a graph.

- Let $P_L(\ell)$ denote the set of paths of length at most ℓ in L and $P_L = \cup_{\ell \in \mathbb{N}} P_L(l)$ the set of paths of finite length.
- Let $L_1, L_2 \subseteq L$. A *canonical path map* on L from L_1 to L_2 of *length* ℓ is a map $\Gamma\colon L_1 \times L_2 \to P_L(\ell)$ which satisfies that $\Gamma(x, y)$ begins at x and ends at y for all $(x, y) \in L_1 \times L_2$.
- Given a canonical path map $\Gamma\colon L_1 \times L_2 \to P_L(\ell)$ and $0 \le i \le \ell$ we define the inverse image mapping of the i'th vertex, $\Gamma_i^{-1} : L \to 2^{L_1 \times L_2}$ as

  $$\Gamma_i^{-1}(z) = \{(x, y) \mid \text{length}(\Gamma(x, y)) \ge i, \Gamma(x, y)_i = z\}.$$

  Further, we let

  $$\Gamma^{-1}(z) = \cup_{i=0}^{\ell} \Gamma_i^{-1}(z)$$

- Given a group H acting on L we say that a canonical path map $\Gamma\colon L_1 \times L_2 \to P_L(\ell)$ is *H-invariant* if $HL_1 = L_1$ and $HL_2 = L_2$ and

  $$\Gamma(hx, hy) = h\Gamma(x, y),$$

  for all $h \in H$ and all $(x, y) \in L_1 \times L_2$.

We will use the following proposition. Recall that a group H acting on L is called *fixed-point-free* if for all $x \in L$ and all $h \in H$ different than the identity it holds that $hx \neq x$.

Proposition 13.5. *Let H be a fixed-point-free group acting on L and let $\Gamma\colon L_1 \times L_2 \to P_L(\ell)$ be a canonical path map that is H-invariant. Then for all $z \in L$ and $0 \le i \le l$ it holds that*

$$|\Gamma_i^{-1}(z)| \le \frac{|L_1||L_2|}{|H|} \tag{68}$$

and

$$|\Gamma^{-1}(z)| \le \frac{(\ell+1)|L_1||L_2|}{|H|} \tag{69}$$

Proof. Note that for all i,

$$|L_1 \times L_2| \ge \sum_{w} |\Gamma_i^{-1}(w)| \ge \sum_{h \in H} |\Gamma_i^{-1}(hz)| = |H||\Gamma_i^{-1}(z)|,$$

where the first inequality follows since the value of the i'th vertex partitions the set of paths of length at least i, the second inequality since H is fixed-point-free, and the final equality from the path being H-invariant. We thus obtain:

$$|\Gamma^{-1}(z)| \le \sum_{i=0}^{\ell} |\Gamma_i^{-1}(z)| \le \frac{(\ell+1)|L_1||L_2|}{|H|},$$

as needed. □

Two applications of the result above will be given for adjacent transpositions.

Definition 13.6. Given two elements $a, b \in [k]$ the *adjacent transposition* $[a : b]$ between them is defined as follows. If $\sigma \in S_k$ has a and b adjacent, then $[a : b]\sigma$ is obtained from σ be exchanging a and b. Otherwise, $[a : b]\sigma = \sigma$.

We let T denote the set of all $q(q-1)/2$ adjacent transpositions. Given $\mu \in T$, we define

$$\begin{aligned}
\mathrm{Inf}_i^{a,b;\mu}(f) &= \mathbb{P}(f(X) = a, f(X^{(i)}) = b) & (70)\\
\mathrm{Inf}_i^{a;\mu}(f) &= \mathbb{P}(f(X) = a, f(X^{(i)}) \neq a) & (71)\\
\mathrm{Inf}_i^{a,b;T}(f) &= \sum_{\mu \in T} \mathrm{Inf}_i^{a,b;\mu}(f) & (72)
\end{aligned}$$

where $X^{(i)}$ is obtained from X by re-randomizing the i:th coordinate X_i in the following way: with probability $1/2$ we keep it as X_i and otherwise we replace it by μX_i.

Finally for $\sigma \in S_k^n$ we will let $[\sigma : b]_i\, \sigma$ denote the element obtained by applying $[a : b]$ on the i:th coordinate of σ while leaving all other coordinates unchanged.

Proposition 13.7. *There exists a canonical path map $\Gamma \colon S_k \times S_k \to P_{S_k}(\ell)$ of length at most $\ell = k(k-1)/2 < k^2/2$, all of whose edges are adjacent transpositions such that for all μ it holds that:*

$$|\Gamma^{-1}(\mu)| \leq \frac{k^2 k!}{2} \tag{73}$$

Proof. Given $\sigma, \pi \in S_k$ consider the following canonical path starting at σ and ending at π. Take the element $\pi(1)$ ranked at the top for π and bubble it to the top by performing adjacent transpositions. Then take the element $\pi(2)$ ranked second for π and bubble it to the second position etc. Clearly the length of the path is at most $k(k-1)/2$. Let $H = \{\sigma \mapsto \tau\sigma \mid \tau \in S_k\}$ be the group of compositions with all possible permutations of the candidates. Since H is a fixed-point-free group acting on S_k and the described canonical path map is H-invariant the result follows from Proposition 13.5. □

Corollary 13.8. *For any $f \colon S_k^n \to [k]$, $a \in [k]$ and $i \in [n]$ it holds that*

$$\sum_{\mu \in T} \mathrm{Inf}_i^{a;\mu}(f) \geq \frac{1}{k^2}\,\mathrm{Inf}_i^a(f), \tag{74}$$

where T is the set of all adjacent transpositions.

Proof. This is a standard canonical path argument. Since both sides of the desired inequality involve averaging over all coordinates but the i'th coordinate, it follows that it suffices to prove the claim in the case where $i = n = 1$. Let $B = \{(u,v) \in S_k \times S_k \mid f(u) = a \neq f(v), \exists \mu \in T : v = \mu u\}$ and note that

$$\sum_{\mu \in T} \mathrm{Inf}_1^{a;\mu}(f) = \frac{|B|}{2k!}, \tag{75}$$

Consider the canonical path map Γ constructed in Proposition 13.7. Note that each canonical path between an element in $A := \{\sigma \in L_q \mid f(\sigma) = a\}$ and an element in A^c must pass via one of the edges in B. Define $h : A \times A^C \to B$ by letting $h(\sigma, \pi)$ be the first edge in B which $\Gamma(\sigma, \pi)$ passes through. Then by (73), for any $(u, v) \in B$,

$$|h^{-1}((u,v))| \leq |\Gamma^{-1}(u)| \leq \frac{k^2 k!}{2} \tag{76}$$

Thus

$$|B| \geq \frac{|A||A^c|}{k^2 k!/2} \tag{77}$$

Combining (75) and (77) we obtain:

$$\sum_{\mu \in T} \mathrm{Inf}_1^{a;\mu}(f) \geq \frac{1}{2k!} \frac{|A||A^c|}{k^2 k!/2} = \frac{1}{k^2} \frac{|A|}{k!} \frac{|A^c|}{k!} = \frac{1}{k^2} \mathrm{Inf}_1^a(f)$$

□

A second application of Proposition 13.7 is the following.

Proposition 13.9. *Fix two elements $a, b \in [k]$ and let $B \subseteq S_k$ denote the set of all permutations where a is ranked above b. Then there exists a canonical path map $\Gamma : B \times B \to P_B(k^2)$ consisting of adjacent transpositions such that all permutations along the path satisfy that a is ranked above b. Moreover for all μ it holds that:*

$$|\Gamma^{-1}(\mu)| \leq k^4 k!$$

Proof. $\Gamma(\sigma, \pi)$ is defined as follows. We look at all elements different than a, b, starting with the top one of π, and bubble each of them upwards to its position in π ignoring a, b. After we have done so, we have all elements but a, b ordered as in π, followed by a, followed by b. We now bubble a to its location in $y\pi$ and then bubble b. Note that the length of the path so defined is at most

$$\frac{k(k-1)}{2} + 2(k-1) = \frac{(k+4)(k-1)}{2} < k^2$$

The proof now follows from Proposition 13.5 by considering the group H which acts by permuting arbitrary all elements but those labeled by a and b:

$$|\Gamma^{-1}(\mu)| \leq \frac{k^2 |B|^2}{|H|} = \frac{k^2 (k!/2)^2}{(k-2)!} \leq k^4 k!$$

□

14 Refined Boundaries

Similarly to the previous construction we now define the i:th a-b boundary with respect to an adjacent swap $z \in T$ as

$$B_i^{a,b;z}(f) = \{(x, y) \mid f(x) = a, f(y) = b, x_i = zy_i, \forall j \neq i : x_j = y_j\},$$

and the boundary with respect to arbitrary adjacent swaps on the i:th coordinate as

$$B_i^{a,b;T}(f) = \bigcup_{z \in T} B_i^{a,b;z}(f)$$

Note that for $a \neq b$,

$$\mathrm{Inf}_i^{a,b;z}(f) = \frac{1}{2}\mathbb{P}(f(X) = a, f(zX) = b) = \frac{1}{2}\frac{|B_i^{a,b;z}(f)|}{(k!)^n} \tag{78}$$

14.1 Manipulation points on refined boundaries

The following three lemmas from Isaksson, Kindler and Mossel [30, 31] identify manipulation points on (or close to) these refined boundaries.

Lemma 14.1. *Fix $f : S_k^n \to [k]$, distinct $a, b \in [k]$ and $(\sigma, \pi) \in B_i^{a,b;T}$. Then either $\sigma_i = [a : b]\pi_i$, or one of σ and π is a 2-manipulation point for f.*

Proof. Suppose $\sigma_i = [c : d]\pi_i$ where $\{c, d\} \neq \{a, b\}$. Then an adjacent transposition of c and d will not change the order of a and b. Hence $b \overset{\sigma_i}{>} a$ iff $b \overset{\pi_i}{>} a$. But then either

1. $f(\pi) = b \overset{x_i}{>} a = f(\sigma)$ and σ is a 2-manipulation point or
2. $f(\sigma) = a \overset{y_i}{>} b = f(\pi)$ and π is a 2-manipulation point.

□

Lemma 14.2. *Fix $f : S_k^n \to [k]$ and points $\sigma, \pi, \mu \in S_k^n$ such that $(\sigma, \pi) \in B_i^{a,b;T}$, $(\mu, \pi) \in B_j^{c,b;T}$ where a, b, c are distinct and $i \neq j$. Then there exists a 3-manipulation point $\nu \in S_k^n$ for f such that $\nu_\ell = \pi_\ell$ for $\ell \notin \{i, j\}$ and ν_i is equal to σ_i or π_i except that the position of c may be shifted arbitrarily and ν_j is equal to μ_j or π_j except that the position of a may be shifted arbitrarily.*

Proof. By Lemma 14.1 we must have $\sigma_i = [a : b]\pi_i$ and $\mu_j = [c : b]\pi_j$, or σ, π or μ is a 2-manipulation point in which case we are done.

Now create a new triple (σ', π', μ') by starting from (σ, π, μ) and simultaneously in the i:th coordinate of σ, π and μ, bubbling c towards the pair ab until it becomes adjacent to the pair. Since c is never swapped with a or b during this process Lemma 14.1 implies that for any intermediate triple $(\widetilde{\sigma}, \widetilde{\pi}, \widetilde{\mu})$ we have $f(\widetilde{\sigma}) = a$, $f(\widetilde{\pi}) = b$ and $f(\widetilde{\mu}) \notin \{a, b\}$, or one of $\widetilde{\sigma}$, $\widetilde{\pi}$ and $\widetilde{\mu}$ is a 2-manipulation point. But since we also have $\widetilde{\mu} = [c : b]_j \widetilde{\pi}$, we must actually have $f(\widetilde{\mu}) = c$, or either $\widetilde{\pi}$ or $\widetilde{\mu}$ is a 2-manipulation point.

Similarly bubbling a towards the pair bc in coordinate j starting from (σ', π', μ') gives us σ'', π'', μ'' all having a, b, c adjacent in coordinates i and j such that $(\sigma'', \pi'') \in B_i^{a,b;[a:b]}$ and $(\mu', \pi'') \in B_j^{c,b;[c:b]}$. Note that σ'', π'', μ'' are equal except for a reordering of the blocks containing a, b, c in coordinates i and j.

Now arbitrary adjacent swapping of a, b, c in these coordinates of σ'', π'' and μ'' will keep the value of f in $\{a, b, c\}$, or give rise to a 2-manipulation point by Lemma 14.1. Thus we can define a social choice function with 2 voters and 3 candidates $f' : S^2_{\{a,b,c\}} \to \{a, b, c\}$ by letting $f'(\nu) = f(g(\nu))$, where $g(\nu) \in S_q^n$ is obtained from σ'' by simply reordering the two blocks of elements a, b, c in coordinates i and j to match ν_1 and ν_2, respectively. Since f' takes three values and is not a dictator, Gibbard-Satterthwaite (Theorem 10.2) implies that f' has a manipulation point and hence f has a 3-manipulation point satisfying our requirements. □

14.2 Large refined boundaries

The following lemma, again from Isaksson, Kindler and Mossel [30, 31], shows that there are large refined boundaries (or else we have a lot of 2-manipulation points automatically).

Lemma 14.3. *Fix $k \geq 3$ and $f : S_k^n \to [k]$ satisfying $\mathbf{D}(f, \overline{\text{NONMANIP}}) \geq \varepsilon$. Let σ be uniformly selected from S_k^n. Then either*

$$\mathbb{P}\left(\sigma \in M_2\left(f\right)\right) \geq \frac{4\varepsilon}{nk^7}, \tag{79}$$

or there exist distinct $i, j \in [n]$ and $\{a, b\}, \{c, d\} \subseteq [k]$ such that $c \notin \{a, b\}$ and

$$\mathrm{Inf}_i^{a,b;[a:b]}(f) \geq \frac{2\varepsilon}{nk^7} \quad \text{and} \quad \mathrm{Inf}_j^{c,d;[c:d]}(f) \geq \frac{2\varepsilon}{nk^7}. \tag{80}$$

Proof. First, suppose that $\mathrm{Inf}_i^{a,b;\nu} \geq \frac{2\varepsilon}{nk^7}$ for some i, $a \neq b$ and $\nu \neq [a : b]$. Then by Lemma 14.1 for any point $(\sigma, \sigma') \in B_i^{a,b;\nu}(f)$ at least one of σ or $\sigma' = \nu\sigma$ is a 2-manipulation point. Let $\widetilde{M}$ be the set of all such 2-manipulation points. Then

$$|\widetilde{M}| \geq |B_i^{a,b;\nu}(f)| = 2(k!)^n \,\mathrm{Inf}_i^{a,b;\nu}(f) \geq \frac{4\varepsilon}{nk^7}(k!)^n \tag{81}$$

Dividing with $(k!)^n$ gives (79). Thus, for the remainder of the proof we may assume that

$$\mathrm{Inf}_i^{a,b;\nu} < \frac{2\varepsilon}{nk^7} \quad , \quad \forall i \in [n], \{a, b\} \subseteq [k], \nu \neq [a : b] \tag{82}$$

Now, for $a \neq b$ let $A^{a,b} = \left\{ i \in [n] \mid \mathrm{Inf}_i^{a,b;[a:b]} \geq \frac{2\varepsilon}{nk^7} \right\}$.

We first claim that for all $\{a,b\}$ there exists $\{c,d\}$ such that $\{c,d\} \neq \{a,b\}$ and $A^{c,d} \neq \emptyset$. Note that f being ε-far from taking two values asserts that we can find a $c \notin \{a,b\}$ such that $1 - \frac{\varepsilon}{k} \geq \mathbb{P}(f(X) = c) \geq \frac{\varepsilon}{q-2} \geq \frac{\varepsilon}{k}$. But then, by Corollary 13.8 and Proposition 10.12,

$$\sum_{\mu \in T} \sum_{d \neq c} \sum_{i=1}^{n} \mathrm{Inf}_i^{c,d;\mu}(f) = \sum_{\mu \in T} \sum_{i=1}^{n} \mathrm{Inf}_i^{c;\mu}(f) \geq \frac{1}{k^2} \mathrm{Var}[1_{\{f(X)=c\}}] \geq \frac{\varepsilon(k-1)}{k^4}$$

hence there must exist some $\mu \in T$, $d \neq c$ and $i \in [n]$ such that $\mathrm{Inf}_i^{c,d;\mu} \geq \frac{\varepsilon}{nk^6}$. But by (82) we must have $w = [c:d]$, hence $A^{c,d} \neq \emptyset$.

We next claim that

$$|\cup_{a,b} A^{a,b}| \geq 2 \tag{83}$$

To see this, assume the contrary, i.e. $\cup_{a,b} A^{a,b} \subseteq \{i\}$ for some $i \in [n]$. Then, by Corollary 13.8, for all $j \neq i$ it holds that

$$\mathrm{Inf}_j(f) \leq k^2 \sum_{\nu \in T} \sum_{a} \mathrm{Inf}_j^{a;\nu}(f) = k^2 \sum_{\nu \in T, a, b > a} \mathrm{Inf}_j^{a,b;\nu}(f) \leq \frac{k^6}{2} \frac{2\varepsilon}{nk^7} = \frac{\varepsilon}{nk} \tag{84}$$

For $\bar{\sigma} \in L_q$, let $f_{\sigma(\bar{x})} = f(\sigma_1, \ldots, \sigma_{i-1}, \bar{\sigma}, \sigma_{i+1}, \ldots, \sigma_n)$ and note that for $j \neq i$,

$$\mathrm{Inf}_j(f) = \frac{1}{k!} \sum_{\bar{\sigma} \in S_k} \mathrm{Inf}_j(f_{\bar{\sigma}}) \tag{85}$$

while $\mathrm{Inf}_i(f_{\bar{\sigma}}) = 0$. Hence, by (84), we have

$$\varepsilon \geq k \sum_{j \neq i} \mathrm{Inf}_j(f) = \frac{k}{k!} \sum_{j=1}^{n} \sum_{\bar{\sigma}} \mathrm{Inf}_j(f_{\bar{\sigma}}) \geq \frac{2}{k!} \sum_{\bar{\sigma}} \mathbf{D}(f_{\bar{\sigma}}, \mathrm{CONST}) = 2\,\mathbf{D}(f, \mathrm{DICT}_i)$$

where the second inequality follows from Lemma 10.13 and Proposition 10.12. But this means that f is $\varepsilon/2$-close to a dictator, contradicting the assumption that $\mathbf{D}(f, \mathrm{NONMANIP}) \geq \varepsilon$.

Hence (83) holds. Therefore we can either find $i \neq j$ and $\{a,b\} \neq \{c,d\}$ such that $i \in A^{a,b}$ and $j \in A^{c,d}$ which proves the theorem, or we must have $|A^{a,b}| \geq 2$ for some $\{a,b\}$ while $A^{c,d} = \emptyset$ for any $\{c,d\} \neq \{a,b\}$. However, this contradicts the first claim in the proof. The result follows. □

14.3 Fibers

We again use fibers $F\left(z^{a,b}\right)$ as defined in Definition 10.17. However, we need more than this. We note that the following definitions only apply in Section 16, i.e., when we have at least two voters; in Section 15, when we have only one voter, things are simpler.

Given the result of Lemma 14.3, our primary interest is in the boundary $B_i^{a,b;[a:b]}$. For ranking profiles on this boundary, we know that the alternatives a and b are adjacent in coordinate i—so we know more than just the preference between a and b in coordinate i. Consequently we would like to divide the set of ranking profiles with a and b adjacent in coordinate i according to the preferences between a and b in all coordinates except coordinate i. The following definitions make this precise.

As done in Section 10.5.6 for ranking profiles, we can write $x_{-i}^{a,b} \equiv x_{-i}^{a,b}(\sigma)$ for the vector of preferences between a and b for all coordinates except coordinate i, i.e., the whole vector of preferences between a and b is $x^{a,b}(\sigma) = \left(x_i^{a,b}(\sigma), x_{-i}^{a,b}(\sigma)\right)$.

We can define $F\left(z_{-i}^{a,b}\right)$ analogously to $F\left(z^{a,b}\right)$:

$$F\left(z_{-i}^{a,b}\right) := \left\{\sigma : x_{-i}^{a,b}(\sigma) = z_{-i}^{a,b}\right\}.$$

We also define the subset of $F\left(z_{-i}^{a,b}\right)$ where a and b are adjacent in coordinate i, with a above b:

$$\bar{F}\left(z_{-i}^{a,b}\right) := \left\{\sigma \in F\left(z_{-i}^{a,b}\right) : a \text{ and } b \text{ are adjacent in coordinate } i, \text{ with } a \text{ above } b\right\}.$$

Given a SCF f, for any pair of alternatives $a, b \in [k]$ and coordinate $i \in [n]$, we can also partition the boundary $B_i^{a,b}(f)$ according to its fibers. There are multiple, slightly different ways of doing this, but for our purposes the following definition is most useful.

Define

$$B_i\left(z_{-i}^{a,b}\right) := \left\{\sigma \in \bar{F}\left(z_{-i}^{a,b}\right) : f(\sigma) = a, f\left([a:b]_i \sigma\right) = b\right\},$$

where we omit the dependence of $B_i\left(z_{-i}^{a,b}\right)$ on f. We call sets of the form $B_i\left(z_{-i}^{a,b}\right) \subseteq \bar{F}\left(z_{-i}^{a,b}\right)$ *fibers for the boundary* $B_i^{a,b;[a:b]}$.

We now distinguish between small and large fibers for the boundary $B_i^{a,b;[a:b]}$.

Definition 14.4 (Small and large fibers)**.** We say that the fiber $B_i\left(z_{-i}^{a,b}\right) \subseteq \bar{F}\left(z_{-i}^{a,b}\right)$ is *large* if

$$\mathbb{P}\left(\sigma \in B_i\left(z_{-i}^{a,b}\right) \,\middle|\, \sigma \in \bar{F}\left(z_{-i}^{a,b}\right)\right) \geq 1 - \gamma,$$

where $\gamma = \frac{\varepsilon^3}{10^3 n^3 k^{24}}$, and *small* otherwise.

As before, we denote by $\mathrm{Lg}\left(B_i^{a,b;[a:b]}\right)$ the union of large fibers for the boundary $B_i^{a,b;[a:b]}$, i.e.,

$$\mathrm{Lg}\left(B_i^{a,b;[a:b]}\right) := \bigcup_{B_i\left(z_{-i}^{a,b}\right) \text{ is a large fiber}} B_i\left(z_{-i}^{a,b}\right),$$

and similarly, we denote by $\mathrm{Sm}\left(B_i^{a,b;[a:b]}\right)$ the union of small fibers.

As in Definition 10.18, we remark that what is important is that γ is a polynomial of $\frac{1}{n}$, $\frac{1}{k}$ and ε—the specific polynomial in this definition is the end result of the computation in the proof.

The following definition is used in Section 16.3 in dealing with the large fiber case in the refined setting.

Definition 14.5. For a coordinate i and a pair of alternatives a and b, define $F_i^{a,b}$ to be the set of ranking profiles σ such that $x^{a,b}(\sigma)$ satisfies

$$\mathbb{P}\left(f(\tilde{\sigma}) = \mathrm{top}_{\{a,b\}}(\tilde{\sigma}_i) \,\middle|\, \tilde{\sigma} \in F\left(x_{-i}^{a,b}(\sigma)\right)\right) \geq 1 - 2k\gamma.$$

Clearly $F_i^{a,b}$ is the union of fibers of the form $F\left(z^{a,b}\right)$, and also $F\left(\left(1, x_{-i}^{a,b}\right)\right) \subseteq F_i^{a,b}$ if and only if $F\left(\left(-1, x_{-i}^{a,b}\right)\right) \subseteq F_i^{a,b}$.

14.4 Boundaries of boundaries

In the refined graph setting, just like in the general rankings graph setting, we also look at boundaries of boundaries.

For a given vector $z_{-i}^{a,b}$ of preferences between a and b, we can think of $\bar{F}\left(z_{-i}^{a,b}\right)$ as a subgraph of the original refined rankings graph S_k^n, i.e., two ranking profiles in $\bar{F}\left(z_{-i}^{a,b}\right)$ are adjacent if they differ by one adjacent transposition in exactly one coordinate. Since both of the ranking profiles are in $\bar{F}\left(z_{-i}^{a,b}\right)$, this adjacent transposition keeps the order of a and b in all coordinates, and moreover it keeps a and b adjacent in coordinate i.

We choose to slightly modify this graph: the vertex set is still $\bar{F}\left(z_{-i}^{a,b}\right)$, but we modify the edge set by adding new edges. Suppose $\sigma \in \bar{F}\left(z_{-i}^{a,b}\right)$ and

$$\sigma_i = \begin{pmatrix} \vdots \\ c \\ a \\ b \\ d \\ \vdots \end{pmatrix}; \qquad \sigma_i' = \begin{pmatrix} \vdots \\ a \\ b \\ c \\ d \\ \vdots \end{pmatrix}; \qquad \sigma_i'' = \begin{pmatrix} \vdots \\ c \\ d \\ a \\ b \\ \vdots \end{pmatrix}.$$

Define in this way $\sigma' = (\sigma_i', \sigma_{-i})$ and $\sigma'' = (\sigma_i'', \sigma_{-i})$, and add (σ, σ') and (σ, σ'') to the edge set. So basically, we consider the block of a and b in coordinate i as a single element, and connect two ranking profiles in $\bar{F}\left(z_{-i}^{a,b}\right)$ if they differ in an adjacent transposition in a single coordinate, allowing this transposition to move the block of a and b in coordinate i. We call this graph $G\left(z_{-i}^{a,b}\right) = \left(\bar{F}\left(z_{-i}^{a,b}\right), E\left(z_{-i}^{a,b}\right)\right)$, where $E\left(z_{-i}^{a,b}\right)$ is the edge set.

When we write $\partial_e\left(B_i\left(z_{-i}^{a,b}\right)\right)$, we mean the edge boundary of $B_i\left(z_{-i}^{a,b}\right)$ in the graph $G\left(z_{-i}^{a,b}\right)$, and similarly when we write $\partial\left(B_i\left(z_{-i}^{a,b}\right)\right)$, we mean the vertex boundary of $B_i\left(z_{-i}^{a,b}\right)$ in the graph $G\left(z_{-i}^{a,b}\right)$.

14.5 Local dictators, conditioning and miscellaneous definitions

In the general rankings graph setting we defined a dictator on a subset of the alternatives, but in the refined rankings graph setting we need to define so-called *local dictators*.

Definition 14.6 (Local dictators)**.** For a coordinate i and a subset of alternatives $H \subseteq [k]$, define LD_i^H to be the set of ranking profiles σ such that the alternatives in H form an adjacent block in σ_i, and permuting them among themselves in any order, the outcome of the SCF f is always the top ranked alternative among those in H. If $\sigma \in \mathrm{LD}_i^H$, then we call σ a local dictator on H in coordinate i.

Also, for a pair of alternatives a and b, define

$$\mathrm{LD}_i\left(a,b\right) := \bigcup_{c\notin\{a,b\}} \mathrm{LD}_i^{\{a,b,c\}},$$

the set of local dictators on three alternatives, two of which are a and b, in coordinate i.

In dealing with local dictators, we will condition on the top of a particular coordinate being fixed. We therefore introduce the following notation.

Definition 14.7 (Conditioning)**.** For any coordinate $i \in [n]$ and any vector $\mathbf{v}$ of alternatives we define

$$\mathbb{P}_i^{\mathbf{v}}\left(\cdot\right) := \mathbb{P}\left(\cdot \,|\, \left(\sigma_i\left(1\right),\ldots,\sigma_i\left(|\mathbf{v}|\right)\right) = \mathbf{v}\right),$$

where $|\mathbf{v}|$ denotes the length of the vector $\mathbf{v}$. E.g., $\mathbb{P}_1^{(a)}\left(\cdot\right) = \mathbb{P}\left(\cdot \,|\, \sigma_1\left(1\right) = a\right)$ and

$$\mathbb{P}_1^{(a,b,c)} = \mathbb{P}\left(\cdot \,|\, \left(\sigma_1\left(1\right),\sigma_1\left(2\right),\sigma_1\left(3\right)\right) = \left(a,b,c\right)\right).$$

We use the following notation in the proof of Theorem 14.8:

Theorem 14.8. *Suppose f is a SCF on n voters and $k \geq 3$ alternatives for which $\mathbf{D}\left(f,\overline{\mathrm{NONMANIP}}\right) \leq \alpha$. Then either*

$$\mathbf{D}\left(f,\mathrm{NONMANIP}\right) < 100n^4k^8\alpha^{1/3} \tag{86}$$

or

$$\mathbb{P}\left(\sigma \in M\left(f\right)\right) \geq \mathbb{P}\left(\sigma \in M_3\left(f\right)\right) \geq \alpha. \tag{87}$$

Definition 14.9 (Majority function)**.** For a function f whose domain X is finite and whose range is the set $\{a,b\}$, define $\mathrm{Maj}\left(f\right)$ by

$$\mathrm{Maj}\left(f\right) = \begin{cases} a & \text{if} \quad \#\left\{x\in X : f\left(x\right)=a\right\} \geq \#\left\{x\in X : f\left(x\right)=b\right\}, \\ b & \text{if} \quad \#\left\{x\in X : f\left(x\right)=a\right\} < \#\left\{x\in X : f\left(x\right)=b\right\}. \end{cases}$$

15 Quantitative Gibbard-Satterthwaite theorem for one voter

In this section we prove our quantitative Gibbard-Satterthwaite theorem for one voter, Theorem 12.1. As mentioned before, we present this proof before proving Theorem 16.1, because the proof of Theorem 16.1 follows the lines of this proof, with slight modifications needed to deal with having $n > 1$ coordinates.

For the remainder of this section, let us fix the number of voters to be 1, the number of alternatives $k \geq 3$, and the SCF f, which satisfies $\mathbf{D}(f, \text{NONMANIP}) \geq \varepsilon$. Accordingly, we typically omit the dependence of various sets (e.g., boundaries between two alternatives) on f.

An additional notational remark: since our SCF is on one voter only, we omit the subscripts that denote the coordinate we are on. E.g., we write simply $\text{Inf}^{a,b}$ instead of $\text{Inf}_1^{a,b}$, etc.

We present the proof in several steps.

15.1 Large boundary between two alternatives

The first thing we have to establish is a large boundary between two alternatives. This can be done just like in Lemma 14.3, except there are two small differences. On the one hand, the assumption of the lemma, namely that $\mathbf{D}(f, \text{NONMANIP}) \geq \varepsilon$, is weaker than that of the original lemma. On the other hand, here we only need one big boundary, unlike in Lemma 14.3, where there are two big boundaries in two different coordinates. The following lemma formulates what we need.

Lemma 15.1. *Recall that f is a SCF on 1 voter and $k \geq 3$ alternatives which satisfies $\mathbf{D}(f, \text{NONMANIP}) \geq \varepsilon$. Let $\sigma \in S_k$ be selected uniformly. Then either*

$$\mathbb{P}(\sigma \in M_2) \geq \frac{4\varepsilon}{k^6} \tag{88}$$

or there exist alternatives $a, b \in [k]$, $a \neq b$ such that

$$\text{Inf}^{a,b;[a:b]} \geq \frac{2\varepsilon}{k^6}. \tag{89}$$

Proof. The proof is just like the proof of Lemma 14.3. First, suppose that $\text{Inf}^{a,b;z} \geq \frac{2\varepsilon}{k^6}$ for some pair of alternatives $a \neq b$, and transposition $z \neq [a : b]$. Then by Lemma 14.1, for any point $(\sigma, \sigma') \in B^{a,b;z}$, at least one of σ or $\sigma' = z\sigma$ is a 2-manipulation point. Then

$$|M_2| \geq \left|B^{a,b;z}\right| = 2 \cdot k! \cdot \text{Inf}^{a,b;z} \geq \frac{4\varepsilon}{k^6} k!,$$

and dividing with $k!$ gives (88). So for the remainder of the proof we may assume that $\text{Inf}^{a,b;z} < \frac{2\varepsilon}{k^6}$ for every $a \neq b$ and $z \neq [a : b]$.

For every $a \in [k]$, $\mathbf{D}\left(f, \mathrm{top}_{\{a\}}\right) \geq \varepsilon$, so $\mathbb{P}\left(f\left(\sigma\right)=a\right) \leq 1-\varepsilon$. On the other hand, there exists an alternative, say $a \in [k]$, such that $\mathbb{P}\left(f\left(\sigma\right)=a\right) \geq \frac{1}{k}$. So for this alternative we have

$$\mathrm{Var}\left(\mathbf{1}\left[f\left(\sigma\right)=a\right]\right) \geq \frac{\varepsilon}{k},$$

and consequently using Corollary 13.8 and Proposition 10.12 we have

$$\sum_{w \in T} \sum_{b \neq a} \mathrm{Inf}^{a,b;w} = \sum_{w \in T} \mathrm{Inf}^{a;w} \geq \frac{1}{k^2} \mathrm{Var}\left(\mathbf{1}\left[f\left(\sigma\right)=a\right]\right) \geq \frac{\varepsilon}{k^3}.$$

Hence there must exist some $w \in T$ and $b \neq a$ such that $\mathrm{Inf}^{a,b;w} \geq \frac{2\varepsilon}{k^6}$, but by our assumption we must have $w = [a:b]$. $\square$

If (88) holds, then we are done, so in the following we assume that (89) holds.

We know that σ is on $B^{a,b;[a:b]}$ if $f\left(\sigma\right)=a$ and $f\left(\left[a:b\right]\sigma\right)=b$. We know that if $b \overset{\sigma}{>} a$, then σ is a 2-manipulation point, so if this happens in more than half of the cases when σ is on $B^{a,b;[a:b]}$, then we have

$$\mathbb{P}\left(\sigma \in M_2\right) \geq \frac{2\varepsilon}{k^6},$$

in which case we are again done. So we may assume in the following that

$$\mathbb{P}\left(\sigma \in B\right) \geq \frac{2\varepsilon}{k^6}, \tag{90}$$

where

$$B := \left\{\sigma : f\left(\sigma\right)=a, f\left(\left[a:b\right]\sigma\right)=b, a \overset{\sigma}{>} b\right\}.$$

15.2 Division into cases

We again divide into two cases.

We introduce the set $\bar{F}$ of permutations where a is directly above b:

$$\bar{F} := \left\{\sigma \in S_k : a \overset{\sigma}{>} b \text{ and } b \overset{\sigma'}{>} a, \text{ where } \sigma' = \left[a:b\right]\sigma\right\}.$$

One of the following two cases must hold.

Case 1: Small fiber case. We have

$$\mathbb{P}\left(\sigma \in B \,\middle|\, \sigma \in \bar{F}\right) \leq 1 - \frac{\varepsilon}{4k}. \tag{91}$$

Case 2: Large fiber case. We have

$$\mathbb{P}\left(\sigma \in B \,\middle|\, \sigma \in \bar{F}\right) > 1 - \frac{\varepsilon}{4k}. \tag{92}$$

15.3 Small fiber case

In this section we assume that (91) holds.

We first formalize that the boundary $\partial(B)$ of B is big (recall the definition of $\partial(B)$ from Section 14.4). The proof uses the canonical path method again.

Lemma 15.2. *If* (91) *holds, then*

$$\mathbb{P}(\sigma \in \partial(B)) \geq \frac{\varepsilon}{2k^4}\mathbb{P}(\sigma \in B). \tag{93}$$

Proof. Let $B^c = \bar{F} \setminus B$. For every $(\sigma, \sigma') \in B \times B^c$, we define a canonical path from σ to σ', which has to pass through at least one edge in $\partial_e(B)$. Then if we show that every edge in $\partial_e(B)$ lies on at most r canonical paths, then it follows that $|\partial_e(B)| \geq |B|\,|B^c|/r$.

So let $(\sigma, \sigma') \in B \times B^c$. We apply the path construction of Proposition 13.7, but considering the block formed by a and b as a single element. Since this path goes from σ (which is in B) to σ' (which is in B^c), it must pass through at least one edge in $\partial_e(B)$.

For a given edge $(\pi, \pi') \in \partial_e(B)$, at most how many possible $(\sigma, \sigma') \in B \times B^c$ pairs are there such that the canonical path between σ and σ' defined above passes through (π, π')? We learn from Proposition 13.7 that there are at most $(k-1)^2(k-1)!/2 < k^2(k-1)!/2$ possibilities for the pair (σ, σ').

Recall that $\left|\bar{F}\right| = (k-1)!$. By our assumption we have $|B| \leq \left(1 - \frac{\varepsilon}{4k}\right)(k-1)!$, and so $|B^c| \geq \frac{\varepsilon}{4k}(k-1)!$. Therefore

$$|\partial_e(B)| \geq \frac{|B|\,|B^c|}{\frac{k^2}{2}(k-1)!} \geq \frac{\varepsilon}{2k^3}|B|.$$

Now in G every ranking profile has $k-2 < k$ neighbors, which implies (93). □

Corollary 15.3. *If* (91) *holds, then*

$$\mathbb{P}(\sigma \in \partial(B)) \geq \frac{\varepsilon^2}{k^{10}}. \tag{94}$$

Proof. Combine Lemma 15.2 and (90). □

Next we want to find manipulation points on the boundary $\partial(B)$. The next lemma tells us that if we are on the boundary $\partial(B)$, then either we can find manipulation points easily, or we are at a local dictator on three alternatives.

Lemma 15.4. *Suppose that $\sigma \in \partial(B)$. Then*

- *either $\sigma \in \mathrm{LD}(a, b)$,*
- *or there exists $\hat{\sigma} \in M_3$ such that $\hat{\sigma}$ is equal to σ or $[a:b]\sigma$ except that the position of a third alternative c might be shifted arbitrarily.*

Proof. Since $\sigma \in \partial(B) \subseteq B$, we know that $f(\sigma) = a$, and if $\sigma' = [a:b]\,\sigma$, then $f(\sigma') = b$. Let $\pi \in B^c$ denote the ranking profile such that $(\sigma, \pi) \in \partial_e(B)$, and let $\pi' = [a:b]\,\pi$. Since $\pi \notin B$, $(f(\pi), f(\pi')) \neq (a, b)$. Then, by Lemma 14.1, if $f(\pi) \neq f(\pi')$, then one of π and π' is a 2-manipulation point. So assume $f(\pi) = f(\pi')$.

There are two cases to consider: either σ and π differ by an adjacent transposition not involving the block of a and b, or they differ by an adjacent transposition that moves the block of a and b.

In the former case, it is not hard to see that one of σ, σ', π, π' is a 2-manipulation point, by Lemma 14.1.

If σ and π differ by an adjacent transposition that involves the block of a and b, then there are again two cases to consider: either this transposition moves the block of a and b up in the ranking, or it moves it down.

If the block of a and b is moved up to get from σ to π, then we must have $f(\pi) = a$, or else σ or π is a 3-manipulation point. Then we must have $f(\pi') = f(\pi) = a$, in which case π' is a 3-manipulation point, since $f(\sigma') = b$.

The final case is when the block of a and b is moved down to get from σ to π, and a third alternative, call it c, is moved up, directly above the block of a and b. Now if $f(\pi) = d \notin \{a, b, c\}$, then σ or π is a 3-manipulation point. If $f(\pi) = f(\pi') = a$, then π' is a 3-manipulation point, whereas if $f(\pi) = f(\pi') = b$, then π is a 3-manipulation point. The remaining case is when $f(\pi) = f(\pi') = c$. Now if $f([b:c]\,\sigma) \neq a$ or $f([a:c]\,\sigma') \neq b$, then we again have a 3-manipulation point close to σ. Otherwise $\sigma \in \mathrm{LD}(a, b)$. $\square$

The following corollary then tells us that either we have found many 3-manipulation points, or we have many local dictators on three alternatives.

Corollary 15.5. *If* (91) *holds, then either*

$$\sum_{c \notin \{a,b\}} \mathbb{P}\left(\sigma \in \mathrm{LD}^{\{a,b,c\}}\right) = \mathbb{P}\left(\sigma \in \mathrm{LD}(a, b)\right) \geq \frac{\varepsilon^2}{2k^{10}} \tag{95}$$

or

$$\mathbb{P}\left(\sigma \in M_3\right) \geq \frac{\varepsilon^2}{4k^{12}}.$$

15.3.1 Dealing with local dictators

So the remaining case we have to deal with in this small fiber case is when (95) holds, i.e., we have many local dictators on three alternatives.

Lemma 15.6. *Suppose $\sigma \in \mathrm{LD}^{\{a,b,c\}}$ for some alternative $c \notin \{a, b\}$. Let σ' be equal to σ except that the block of a, b and c is moved to the top of the coordinate. Then*

- *either $\sigma' \in \mathrm{LD}^{\{a,b,c\}}$,*

- *or there exists a 3-manipulation point $\hat{\sigma}$ which agrees with σ except that the positions of a, b and c might be shifted arbitrarily.*

Proof. W.l.o.g. we may assume that in σ alternative a is ranked above b, which is ranked above c. Now move a to the top using a sequence of adjacent transpositions, all involving a; we call this procedure "bubbling" a to the top. If at any point during this the outcome of f is not a, then we have found a 2-manipulation point. Now bubble up b to right below a, and then bubble up c to be right below b. Again, if at any point during this the outcome of f is not a, then there is a 2-manipulation point. Otherwise we now have a, b and c at the top (in this order), with the outcome of f being a. Now permuting alternatives a, b and c at the top, we either have a 3-manipulation point, or $\sigma' \in \mathrm{LD}^{\{a,b,c\}}$. □

Corollary 15.7. *If* (95) *holds, then either*

$$\sum_{c \notin \{a,b\}} \mathbb{P}\left(\sigma \in \mathrm{LD}^{\{a,b,c\}}, \{\sigma(1), \sigma(2), \sigma(3)\} = \{a,b,c\}\right) \geq \frac{\varepsilon^2}{4k^{11}} \tag{96}$$

or

$$\mathbb{P}(\sigma \in M_3) \geq \frac{\varepsilon^2}{4k^{13}}.$$

Proof. Lemma 15.6 tells us that when we move the block of a, b, and c up to the top, we either encounter a 3-manipulation point, or we get a local dictator on $\{a,b,c\}$ at the top.

If we get a 3-manipulation point, by the describtion of this manipulation point in the lemma, there can be at most k^3 ranking profiles that give the same manipulation point.

If we arrive at a local dictator at the top, then there could have been at most k different places where the block of a, b and c could have come from. □

Now (96) is equivalent to

$$\sum_{c \notin \{a,b\}} \mathbb{P}\left(\sigma \in \mathrm{LD}^{\{a,b,c\}}, (\sigma(1), \sigma(2), \sigma(3)) = (a,b,c)\right) \geq \frac{\varepsilon^2}{24k^{11}}. \tag{97}$$

We know that

$$\mathbb{P}((\sigma(1), \sigma(2), \sigma(3)) = (a,b,c)) = \frac{1}{k(k-1)(k-2)} \leq \frac{6}{k^3},$$

and so (97) implies (recall Definition 14.7)

$$\sum_{c \notin \{a,b\}} \mathbb{P}^{(a,b,c)}\left(\sigma \in \mathrm{LD}^{\{a,b,c\}}\right) \geq \frac{\varepsilon^2}{144k^8}. \tag{98}$$

Now fix an alternative $c \notin \{a,b\}$ and define the graph $G_{(a,b,c)} = \left(V_{(a,b,c)}, E_{(a,b,c)}\right)$ to have vertex set

$$V_{(a,b,c)} := \{\sigma \in S_k : (\sigma(1), \sigma(2), \sigma(3)) = (a,b,c)\}$$

and for $\sigma, \pi \in V_{(a,b,c)}$ let $(\sigma, \pi) \in E_{(a,b,c)}$ if and only if σ and π differ by an adjacent transposition. So $G_{(a,b,c)}$ is the subgraph of the refined rankings graph induced by the vertex set $V_{(a,b,c)}$. (If $k = 3$ or $k = 4$, then this graph consists of only one vertex, and no edges.)

Let

$$T(a,b,c) := V_{(a,b,c)} \cap \mathrm{LD}^{\{a,b,c\}},$$

and let $\partial_e(T(a,b,c))$ and $\partial(T(a,b,c))$ denote the edge- and vertex-boundary of $T(a,b,c)$ in $G_{(a,b,c)}$, respectively.

The next lemma shows that unless $T(a,b,c)$ is almost all of $V_{(a,b,c)}$, the size of the boundary $\partial(T(a,b,c))$ is comparable to the size of $T(a,b,c)$. The proof uses a canonical path argument, just like in Lemma 15.2.

Lemma 15.8. *Let $c \notin \{a,b\}$ be arbitrary. Write $T \equiv T(a,b,c)$ for simplicity. If $\mathbb{P}^{(a,b,c)}(\sigma \in T) \leq 1 - \delta$, then*

$$\mathbb{P}^{(a,b,c)}(\sigma \in \partial(T)) \geq \frac{\delta}{k^3} \mathbb{P}^{(a,b,c)}(\sigma \in T). \tag{99}$$

Proof. Let $T^c = V_{(a,b,c)} \setminus T(a,b,c)$. For every $(\sigma, \sigma') \in T \times T^c$, we define a canonical path from σ to σ', which has to pass through at least one edge in $\partial_e(T)$. Then if we show that every edge in $\partial_e(T)$ lies on at most r canonical paths, then it follows that $|\partial_e(T)| \geq |T| |T^c| / r$.

So let $(\sigma, \sigma') \in T \times T^c$. We apply the path construction of Proposition 13.7, but only to alternatives $[k] \setminus \{a,b,c\}$.

The analysis of this construction is done in exactly the same way as in Lemma 15.2; in the end we get that there are at most $k^2(k-3)!$ paths that pass through a given edge in $\partial_e(T)$.

Recall that $\left|V_{(a,b,c)}\right| = (k-3)!$ and that by our assumption $|T| \leq (1-\delta)(k-3)!$, so $|T^c| \geq \delta(k-3)!$. Therefore

$$|\partial_e(T)| \geq \frac{|T| |T^c|}{k^2(k-3)!} \geq \frac{\delta}{k^2} |T|.$$

Now every vertex in $V_{(a,b,c)}$ has $k - 4 < k$ neighbors, which implies (99). □

The next lemma tells us that if σ is on the boundary of a set of local dictators on $\{a,b,c\}$ for some alternative $c \notin \{a,b\}$, then there is a 2-manipulation point $\hat{\sigma}$ which is close to σ.

Lemma 15.9. *Suppose $\sigma \in \partial(T(a,b,c))$ for some $c \notin \{a,b\}$. Then there exists $\hat{\sigma} \in M_2$ which equals $z\sigma$ for some adjacent transposition z that does not involve a, b or c, except that the order of the block of a, b and c might be rearranged.*

Proof. Let π be the ranking profile such that $(\sigma, \pi) \in \partial_e (T(a,b,c))$, and let z be the adjacent transposition in which they differ, i.e., $\pi = z\sigma$. Since $\pi \notin T(a,b,c)$, there exists a reordering of the block of a, b, and c at the top of π such that the outcome of f is not the top ranked alternative. Call the resulting vector π'. W.l.o.g. let us assume that $\pi'(1) = a$. Let us also define $\sigma' := z\pi'$. Now π' is a 2-manipulation point, since $f(\sigma') = a$. $\square$

The next corollary puts together Corollary 15.7 and Lemmas 15.8 and 15.9.

Corollary 15.10. *Suppose* (96) *holds. Then if for every $c \notin \{a,b\}$ we have $\mathbb{P}^{(a,b,c)}(\sigma \in T(a,b,c)) \leq 1 - \frac{\varepsilon}{100k}$, then*

$$\mathbb{P}(\sigma \in M_2) \geq \frac{\varepsilon^3}{10^5 k^{16}}.$$

Proof. We know that (96) implies

$$\sum_{c \notin \{a,b\}} \mathbb{P}^{a,b,c}(\sigma \in T(a,b,c)) \geq \frac{\varepsilon^2}{144k^8}.$$

Now using the assumptions, Lemma 15.8 with $\delta = \frac{\varepsilon}{100k}$, and Lemma 15.9, we have

$$\begin{aligned}\mathbb{P}(\sigma \in M_2) &\geq \sum_{c \neq \{a,b\}} \frac{1}{k^3}\mathbb{P}^{(a,b,c)}(\sigma \in M_2) \geq \sum_{c \notin \{a,b\}} \frac{1}{6k^4}\mathbb{P}^{(a,b,c)}(\sigma \in \partial(T(a,b,c))) \\ &\geq \sum_{c \notin \{a,b\}} \frac{\varepsilon}{600k^8}\mathbb{P}^{(a,b,c)}(\sigma \in T(a,b,c)) \geq \frac{\varepsilon^3}{86400k^{16}} \geq \frac{\varepsilon^3}{10^5 k^{16}}.\end{aligned}$$

$\square$

So again we are left with one case to deal with: if there exists an alternative $c \notin \{a,b\}$ such that $\mathbb{P}^{(a,b,c)}(\sigma \in T(a,b,c)) > 1 - \frac{\varepsilon}{100k}$. Define a subset of alternatives $K \subseteq [k]$ in the following way:

$$K := \{a,b\} \cup \left\{c \in [k] \setminus \{a,b\} : \mathbb{P}^{(a,b,c)}(\sigma \in T(a,b,c)) > 1 - \frac{\varepsilon}{100k}\right\}.$$

In addition to a and b, K contains those alternatives that whenever they are at the top with a and b, they form a local dictator with high probability.

So our assumption now is that $|K| \geq 3$.

Our next step is to show that unless we have many manipulation points, for any alternative $c \in K$, conditioned on c being at the top, the outcome of f is c with probability close to 1.

Lemma 15.11. *Let $c \in K$. Then either*

$$\mathbb{P}^{(c)}(f(\sigma) = c) \geq 1 - \frac{\varepsilon}{50k}, \tag{100}$$

or

$$\mathbb{P}(\sigma \in M_2) \geq \frac{\varepsilon}{100k^4}. \tag{101}$$

Proof. First assume that $c \notin \{a, b\}$.

Let σ be uniform according to $\mathbb{P}^{(c)}$, i.e., uniform on S_k conditioned on $\sigma(1) = c$. Define σ', where σ' is constructed from σ by first bubbling up alternative a to just below c, using adjacent transpositions, and then bubbling up b to just below a. Clearly σ' is distributed according to $\mathbb{P}^{(c,a,b)}$, i.e., it is uniform on S_k conditioned on $(\sigma(1), \sigma(2), \sigma(3)) = (c, a, b)$.

Since $c \in K$, we know that $\mathbb{P}^{(c,a,b)}\left(\sigma \in \mathrm{LD}^{\{a,b,c\}}\right) > 1 - \frac{\varepsilon}{100k}$. This also means that

$$\mathbb{P}^{(c)}\left(\sigma' \in \mathrm{LD}^{\{a,b,c\}}\right) > 1 - \frac{\varepsilon}{100k}.$$

Now we can partition the ranking profiles into three parts, based on the outcome of the SCF f at σ and σ':

$$\begin{aligned} I_1 &= \left\{\sigma : f(\sigma) = c, f\left(\sigma'\right) = c\right\} \\ I_2 &= \left\{\sigma : f(\sigma) \neq c, f\left(\sigma'\right) = c\right\} \\ I_3 &= \left\{\sigma : f\left(\sigma'\right) \neq c\right\}. \end{aligned}$$

If $\mathbb{P}^{(c)}(I_1) \geq 1 - \frac{\varepsilon}{50k}$, then (100) holds. Otherwise we have $\mathbb{P}^{(c)}(I_2 \cup I_3) \geq \frac{\varepsilon}{50k}$, and since $\mathbb{P}^{(c)}(I_3) \leq \frac{\varepsilon}{100k}$, we have $\mathbb{P}^{(c)}(I_2) \geq \frac{\varepsilon}{100k}$.

Now if $\sigma \in I_2$, then we know that there is a 2-manipulation point along the way as we go from σ to σ'. I.e., to every $\sigma \in I_2$ there exists $\hat{\sigma} \in M_2$ such that $\hat{\sigma}$ is equal to σ except perhaps a and b are shifted arbitrarily. So there can be at most k^2 ranking profiles σ giving the same 2-manipulation point $\hat{\sigma}$, and so we have

$$\mathbb{P}(\sigma \in M_2) \geq \frac{1}{k}\mathbb{P}^{(c)}(\sigma \in M_2) \geq \frac{1}{k^3}\mathbb{P}^{(c)}(I_2) \geq \frac{\varepsilon}{100k^4},$$

showing (101).

Now suppose $c \in \{a, b\}$, w.l.o.g. assume $c = a$. We know that $|K| \geq 3$ and so there exists an alternative $d \in K \setminus \{a, b\}$. We can then do the same thing as above, but we now bubble up b and d. □

We now deal with alternatives that are not in K: either we have many manipulation points, or for any alternative $d \notin K$, the outcome of f is *not* d with probability close to 1.

Lemma 15.12. *Let $d \notin K$. If $\mathbb{P}(f(\sigma) = d) \geq \frac{\varepsilon}{4k}$, then*

$$\mathbb{P}(\sigma \in M_2) \geq \frac{\varepsilon^2}{10^6 k^9}.$$

Proof. Let σ be such that $f(\sigma) = d$. Bubble up d to the top, and call this ranking profile σ'. Now if $f(\sigma') \neq d$, then we know that there exists a 2-manipulation point $\hat{\sigma}$ along the way, i.e., a 2-manipulation $\hat{\sigma}$ which agrees with σ except perhaps d is shifted arbitrarily. Consequently, either

$$\mathbb{P}(\sigma \in M_2) \geq \frac{\varepsilon}{8k^2},$$

in which case we are done, or

$$\mathbb{P}\left(\sigma : f\left(\sigma\right) = f\left(\sigma'\right) = d\right) \geq \frac{\varepsilon}{8k}.$$

Next, let us bubble up a to just below d, and then bubble up b to just below d. Denote this ranking profile by $\sigma^{(d,b,a)}$, and analogously define $\sigma^{(d,a,b)}, \sigma^{(a,b,d)}, \sigma^{(a,d,b)}, \sigma^{(b,a,d)}$, and $\sigma^{(b,d,a)}$. Either we encounter a 2-manipulation point $\hat{\sigma}$ along the way of bubbling up to $\sigma^{(d,b,a)}$ ($\hat{\sigma}$ agrees with σ except d is at the top, and a and b might be arbitrarily shifted), or the outcome of the SCF f is d all along. So we have that either

$$\mathbb{P}\left(\sigma \in M_2\right) \geq \frac{\varepsilon}{16k^3},$$

in which case we are done, or

$$\mathbb{P}\left(\sigma : f\left(\sigma\right) = f\left(\sigma'\right) = f\left(\sigma^{(d,b,a)}\right) = f\left(\sigma^{(d,a,b)}\right) = d\right) \geq \frac{\varepsilon}{16k}.$$

Now start from $\sigma^{(d,a,b)}$. First swap a and d to get $\sigma^{(a,d,b)}$, then swap d and b to get $\sigma^{(a,b,d)}$, and finally bubble d and b down to their original positions in σ, except for the fact that a is now at the top of the coordinate. Call this profile $\bar{\sigma}$. Since σ is uniformly distributed, $\bar{\sigma}$ is distributed according to $\mathbb{P}_1^{(a)}$, i.e., uniformly conditional on $\bar{\sigma}\left(1\right) = a$. Now note that one of the following three events has to happen. (These events are not mutually exclusive.)

$$\begin{aligned} I_1 &= \left\{f\left(\sigma^{(a,d,b)}\right) = f\left(\sigma^{(a,b,d)}\right) = a\right\} \\ I_2 &= \left\{f\left(\bar{\sigma}\right) \neq a\right\} \\ I_3 &= \{\sigma : \exists\ \hat{\sigma} \in M_2 \text{ which is equal to } \sigma \text{ except } a \text{ is shifted} \\ &\qquad \text{to the top, and } b \text{ and } d \text{ may be shifted arbitrarily}\}. \end{aligned}$$

Since $a \in K$, we know by Lemma 15.11 that (unless we already have enough manipulation points by the lemma) we must have

$$\mathbb{P}\left(f\left(\bar{\sigma}\right) \neq a\right) = \mathbb{P}^{(a)}\left(f\left(\bar{\sigma}\right) \neq a\right) \leq \frac{\varepsilon}{50k}.$$

Consequently

$$\mathbb{P}\left(I_1 \cup I_3, f\left(\sigma\right) = f\left(\sigma'\right) = f\left(\sigma^{(d,b,a)}\right) = f\left(\sigma^{(d,a,b)}\right) = d\right) \geq \frac{\varepsilon}{16k} - \frac{\varepsilon}{50k} = \frac{17\varepsilon}{400k},$$

and so either

$$\mathbb{P}\left(\sigma \in M_2\right) \geq \frac{17\varepsilon}{800k^3},$$

in which case we are done, or

$$\mathbb{P}\left(\sigma : f\left(\sigma^{(d,b,a)}\right) = f\left(\sigma^{(d,a,b)}\right) = d, f\left(\sigma^{(a,b,d)}\right) = f\left(\sigma^{(a,d,b)}\right) = a\right) \geq \frac{17\varepsilon}{800k}.$$

Next, we can do the same thing with b on top, and we ultimately get that either

$$\mathbb{P}\left(\sigma \in M_2\right) \geq \frac{\varepsilon}{1600k^3},$$

in which case we are done, or

$$\mathbb{P}^{(a,b,d)}\left(\sigma^{(a,b,d)} \in \mathrm{LD}^{\{a,b,d\}}\right) = \mathbb{P}\left(\sigma : \sigma^{(a,b,d)} \in \mathrm{LD}^{\{a,b,d\}}\right) \geq \frac{\varepsilon}{1600k}. \tag{102}$$

Define $G_{(a,b,d)}$ and $T_{(a,b,d)}$ analogously to $G_{(a,b,c)}$ and $T_{(a,b,c)}$, respectively.

Suppose that (102) holds. We also know that $d \notin K$, so Lemma 15.8 applies, and then Lemma 15.9 shows us how to find manipulation points. We can put these arguments together, just like in the proof of Corollary 15.10, to show what we need:

$$\begin{aligned}\mathbb{P}\left(\sigma \in M_2\right) &\geq \frac{1}{k^3}\mathbb{P}^{(a,b,d)}\left(\sigma \in M_2\right) \geq \frac{1}{6k^4}\mathbb{P}^{(a,b,d)}\left(\sigma \in \partial\left(T\left(a,b,d\right)\right)\right)\\ &\geq \frac{\varepsilon}{600k^8}\mathbb{P}^{(a,b,d)}\left(\sigma \in T\left(a,b,d\right)\right) \geq \frac{\varepsilon^2}{10^6 k^9}.\end{aligned}$$

□

Putting together the results of the previous lemmas, there is only one case to be covered, which is covered by the following final lemma. Basically, this lemma says that unless there are enough manipulation points, our function is close to a dictator on the subset of alternatives K.

Lemma 15.13. *Recall that we assume that $\mathbf{D}\left(f, \mathrm{NONMANIP}\right) \geq \varepsilon$. Furthermore assume that $|K| \geq 3$, for every $c \in K$ we have*

$$\mathbb{P}^{(c)}\left(f\left(\sigma\right) = c\right) \geq 1 - \frac{\varepsilon}{50k}, \tag{103}$$

and for every $d \notin K$ we have

$$\mathbb{P}\left(f\left(\sigma\right) = d\right) \leq \frac{\varepsilon}{4k}.$$

Then

$$\mathbb{P}\left(\sigma \in M_2\right) \geq \frac{\varepsilon}{4k^2}. \tag{104}$$

Proof. First note that

$$\mathbb{P}\left(f\left(\sigma\right) \neq \mathrm{top}_K\left(\sigma\right)\right) = \mathbb{P}\left(f\left(\sigma\right) \notin K\right) + \mathbb{P}\left(f\left(\sigma\right) \neq \mathrm{top}_K\left(\sigma\right), f\left(\sigma\right) \in K\right).$$

We know that

$$\varepsilon \leq \mathbf{D}\left(f, \mathrm{NONMANIP}\right) \leq \mathbb{P}\left(f\left(\sigma\right) \neq \mathrm{top}_K\left(\sigma\right)\right)$$

and also that

$$\mathbb{P}\left(f\left(\sigma\right)\notin K\right)\leq\left(k-\left|K\right|\right)\frac{\varepsilon}{4k}\leq\frac{\varepsilon}{2},$$

which together imply that

$$\mathbb{P}\left(f\left(\sigma\right)\neq\mathrm{top}_{K}\left(\sigma\right),f\left(\sigma\right)\in K\right)\geq\frac{\varepsilon}{2}.$$

Let σ be such that $f\left(\sigma\right)\neq\mathrm{top}_{K}\left(\sigma\right)$ and $f\left(\sigma\right)\in K$. Now bubble $\mathrm{top}_{K}\left(\sigma\right)$ up to the top in σ, call this ranking profile $\bar{\sigma}$. Clearly then $\mathrm{top}_{K}\left(\bar{\sigma}\right)=\mathrm{top}_{K}\left(\sigma\right)$.

There are two cases to consider. If $f\left(\sigma\right)\neq f\left(\bar{\sigma}\right)$, then there is a 2-manipulation point along the way from σ to $\bar{\sigma}$, i.e., a 2-manipulation point $\hat{\sigma}$ such that $\hat{\sigma}$ agrees with σ except perhaps some alternative c is arbitrarily shifted. Otherwise $f\left(\sigma\right)=f\left(\bar{\sigma}\right)$, and so $f\left(\bar{\sigma}\right)\neq\mathrm{top}_{K}\left(\bar{\sigma}\right)$.

Consequently we have that either (104) holds, or that

$$\mathbb{P}\left(\sigma:f\left(\bar{\sigma}\right)\neq\mathrm{top}_{K}\left(\bar{\sigma}\right)\right)\geq\frac{\varepsilon}{4}. \qquad (105)$$

By the construction of $\bar{\sigma}$, we know that $\bar{\sigma}$ is uniformly distributed conditional on $\bar{\sigma}\left(1\right)\in K$. Consequently, by (103), we have that

$$\mathbb{P}\left(\sigma:f\left(\bar{\sigma}\right)\neq\mathrm{top}_{K}\left(\bar{\sigma}\right)\right)\leq\frac{\varepsilon}{50k},$$

which contradicts with (105) since $\frac{\varepsilon}{50k}<\frac{\varepsilon}{4}$. □

This concludes the proof of the small fiber case.

15.4 Large fiber case

In this section we assume that (92) holds. We show that we either have a lot 2-manipulation points or we have a lot of local dictators on three alternatives.

Our first step towards this is the following lemma.

Lemma 15.14. *Suppose* (92) *holds. Then*

$$\mathbb{P}^{(a,b)}\left(\sigma\in B\right)\geq1-\frac{\varepsilon}{4}. \qquad (106)$$

Proof. Let $B^{c}=\bar{F}\setminus B$. Our assumption (92) implies that $\mathbb{P}\left(\sigma\in B^{c}\,\middle|\,\sigma\in\bar{F}\right)\leq\frac{\varepsilon}{4k}$, which means that $\left|B^{c}\right|\leq\frac{\varepsilon(k-1)!}{4k}$, and so

$$\mathbb{P}^{(a,b)}\left(\sigma\notin B\right)\leq\frac{\varepsilon\left(k-1\right)!}{4k\left(k-2\right)!}<\frac{\varepsilon}{4},$$

which is equivalent to (106). □

The next lemma (together with Section 15.3.1) concludes the proof in the large fiber case.

Lemma 15.15. *Suppose* (92) *holds and recall that our SCF f satisfies* $\mathbf{D}(f, \text{NONMANIP}) \geq \varepsilon$. *Then either*

$$\mathbb{P}(\sigma \in M_2) \geq \frac{\varepsilon}{4k^2} \tag{107}$$

or

$$\mathbb{P}(\sigma \in \text{LD}(a, b)) \geq \frac{\varepsilon}{4k^2}. \tag{108}$$

Proof. By Lemma 15.14 we know that (106) holds.

Let $\sigma \in S_k$ be uniform. Define σ' by being the same as σ except alternatives a and b are moved to the top of the coordinate: $\sigma'(1) = a$ and $\sigma'(2) = b$. Clearly σ' is distributed according to $\mathbb{P}^{(a,b)}(\cdot)$. Also define $\sigma'' = [a : b]\,\sigma'$.

We partition the set of ranking profiles S_k into three parts:

$$I_1 := \left\{\sigma \in S_k : f(\sigma) = \text{top}_{\{a,b\}}(\sigma), \left(f\left(\sigma'\right), f\left(\sigma''\right)\right) = (a, b)\right\}$$
$$I_2 := \left\{\sigma \in S_k : f(\sigma) \neq \text{top}_{\{a,b\}}(\sigma), \left(f\left(\sigma'\right), f\left(\sigma''\right)\right) = (a, b)\right\}$$
$$I_3 := \left\{\sigma \in S_k : \left(f\left(\sigma'\right), f\left(\sigma''\right)\right) \neq (a, b)\right\}.$$

By (106) we know that $\mathbb{P}(\sigma \in I_3) \leq \frac{\varepsilon}{4}$. We also know that $\mathbb{P}(\sigma \in I_1) \leq 1 - \varepsilon$, since $\mathbf{D}(f, \text{NONMANIP}) \geq \varepsilon$. Therefore we must have

$$\mathbb{P}(\sigma \in I_2) \geq \frac{3\varepsilon}{4} > \frac{\varepsilon}{2}.$$

Let us partition I_2 further, and write it as $I_2 = I_2' \cup \left(\cup_{c \notin \{a,b\}} I_{2,c}\right)$, where

$$I_2' := \left\{\sigma \in I_2 : f(\sigma) \neq \text{top}_{\{a,b\}}(\sigma), f(\sigma) \in \{a, b\}\right\}$$

and for any $c \notin \{a, b\}$,

$$I_{2,c} := \{\sigma \in I_2 : f(\sigma) = c\}.$$

Suppose $\sigma \in I_2'$. W.l.o.g. let us assume that a is ranked higher than b by σ, and therefore $f(\sigma) = b$, since $\sigma \in I_2'$. Then we can get from σ to σ' by first bubbling up a to the top, and then bubbling up b to just below a. Since $f(\sigma) = b$ and $f(\sigma') = a$, there must be a 2-manipulation point $\hat{\sigma}$ along the way, which is equal to σ except perhaps the positions of a and b are arbitrarily shifted.

Now suppose that $\sigma \in I_{2,c}$ for some $c \notin \{a, b\}$. We distinguish two cases: either c is ranked above both a and b in σ, or it is not.

If not, then say a is ranked above c in σ. Bubble a all the way to the top, and then bubble b as well, all the way to the top, just below a. Since $f(\sigma) = c$ and $f(\sigma') = a$, there must be a 2-manipulation point $\hat{\sigma}$ along the way, which is equal to σ except perhaps the positions of a and b are arbitrarily shifted.

If c is ranked above both a and b in σ, then the argument is similar. First bubble up a and b to just below c, and denote this ranking profile by $\tilde{\sigma}$, then permute these three alternatives arbitrarily, and then bubble a and b to the top. It is not hard to think through that either there is a 2-manipulation $\hat{\sigma}$ along the way, which is then equal to σ except perhaps the positions of a and b are arbitrarily shifted, or else $\tilde{\sigma} \in \mathrm{LD}^{\{a,b,c\}}$.

Combining these cases we see that either (107) or (108) must hold. □

So if (107) holds then we are done, and if (108) holds, then we refer back to Section 15.3.1, where we deal with the case of local dictators on three alternatives.

15.5 Proof of Theorem 12.1 concluded

Proof of Theorem 12.1. Our starting point is Lemma 15.1, which implies that (90) holds (unless we already have many 2-manipulation points). We then consider two cases, as indicated in Section 15.2.

We deal with the small fiber case—when (91) holds—in Section 15.3. First, Lemma 15.2, Corollary 15.3, Lemma 15.4 and Corollary 15.5 show that either there are many 3-manipulation points, or there are many local dictators on three alternatives. We then deal with the case of many local dictators in Section 15.3.1. Lemma 15.6, Corollary 15.7, Lemmas 15.8 and 15.9, Corollary 15.10, and Lemmas 15.11, 15.12 and 15.13 together show that there are many 3-manipulation points if there are many local dictators on three alternatives, and the SCF is ε-far from the family of nonmanipulable functions.

We deal with the large fiber case—when (92) holds—in Section 15.4. Here Lemma 15.15 shows that either we have many 2-manipulation points, or we have many local dictators on three alternatives. In this latter case we refer back to Section 15.3.1 to conclude the proof. □

16 Inverse polynomial manipulability for any number of alternatives

In this section we prove the theorem below, which is the same as our main theorem, Theorem 10.9, except that the condition of $\mathbf{D}(f, \mathrm{NONMANIP}) \geq \varepsilon$ from Theorem 10.9 is replaced with the stronger condition $\mathbf{D}\left(f, \overline{\mathrm{NONMANIP}}\right) \geq \varepsilon$.

Theorem 16.1. *Suppose we have $n \geq 2$ voters, $k \geq 3$ alternatives, and a SCF $f : S_k^n \to [k]$ satisfying $\mathbf{D}\left(f, \overline{\mathrm{NONMANIP}}\right) \geq \varepsilon$. Then*

$$\mathbb{P}(\sigma \in M(f)) \geq \mathbb{P}(\sigma \in M_4(f)) \geq p\left(\varepsilon, \frac{1}{n}, \frac{1}{k}\right), \tag{109}$$

for some polynomial p, where $\sigma \in S_k^n$ is selected uniformly. In particular, we show a lower bound of $\frac{\varepsilon^5}{10^9 n^7 k^{46}}$.

An immediate consequence is that

$$\mathbb{P}\left(\left(\sigma, \sigma'\right) \text{ is a manipulation pair for } f\right) \geq q\left(\varepsilon, \frac{1}{n}, \frac{1}{k}\right),$$

for some polynomial q, where $\sigma \in S_k^n$ is uniformly selected, and σ' is obtained from σ by uniformly selecting a coordinate $i \in \{1, \ldots, n\}$, uniformly selecting $j \in \{1, \ldots, n-3\}$, and then uniformly randomly permuting the following four adjacent alternatives in σ_i: $\sigma_i(j), \sigma_i(j+1), \sigma_i(j+2)$, and $\sigma_i(j+3)$. In particular, the specific lower bound for $\mathbb{P}(\sigma \in M_4(f))$ implies that we can take $q\left(\varepsilon, \frac{1}{n}, \frac{1}{k}\right) = \frac{\varepsilon^5}{10^{11} n^8 k^{47}}$.

For the remainder of the section, let us fix the number of voters $n \geq 2$, the number of alternatives $k \geq 3$, and the SCF f, which satisfies $\mathbf{D}\left(f, \overline{\text{NONMANIP}}\right) \geq \varepsilon$. Accordingly, we typically omit the dependence of various sets (e.g., boundaries between two alternatives) on f.

16.1 Division into cases

Our starting point in proving Theorem 16.1 is Lemma 14.3. Clearly if (79) holds then we are done, so in the rest of Section 16 we assume that this is not the case. Then Lemma 14.3 tells us that (80) holds, and w.l.o.g. we may assume that the two boundaries that the lemma gives us have $i = 1$ and $j = 2$. I.e., we have

$$\mathbb{P}\left(\sigma \text{ on } B_1^{a,b;[a:b]}\right) \geq \frac{4\varepsilon}{nk^7} \quad \text{and} \quad \mathbb{P}\left(\sigma \text{ on } B_2^{c,d;[c:d]}\right) \geq \frac{4\varepsilon}{nk^7},$$

where recall that σ is on $B_1^{a,b;[a:b]}$ if $f(\sigma) = a$ and $f([a:b]_1 \sigma) = b$. If σ is on $B_1^{a,b;[a:b]}$ and $b \overset{\sigma_1}{>} a$, then σ is a 2-manipulation point, so if this happens in more than half of the cases when σ is on $B_1^{a,b;[a:b]}$, then we have

$$\mathbb{P}(\sigma \in M_2) \geq \frac{2\varepsilon}{nk^7},$$

and we are done. Similarly in the case of the boundary between c and d in coordinate 2. So we may assume from now on that

$$\mathbb{P}\left(\sigma \in \cup_{z_{-1}^{a,b}} B_1\left(z_{-1}^{a,b}\right)\right) \geq \frac{2\varepsilon}{nk^7} \quad \text{and} \quad \mathbb{P}\left(\sigma \in \cup_{z_{-2}^{c,d}} B_2\left(z_{-2}^{c,d}\right)\right) \geq \frac{2\varepsilon}{nk^7}.$$

The following lemma is an immediate corollary.

Lemma 16.2. *Either*

$$\mathbb{P}\left(\sigma \in \mathrm{Sm}\left(B_1^{a,b;[a:b]}\right)\right) \geq \frac{\varepsilon}{nk^7} \tag{110}$$

or

$$\mathbb{P}\left(\sigma \in \mathrm{Lg}\left(B_1^{a,b;[a:b]}\right)\right) \geq \frac{\varepsilon}{nk^7}, \tag{111}$$

and the same can be said for the boundary $B_2^{c,d;[c:d]}$*.*

We distinguish cases based upon this: either (110) holds, or (110) holds for the boundary $B_2^{c,d;[c:d]}$, or (111) holds for both boundaries. We only need one boundary for the small fiber case, and we need both boundaries only in the large fiber case. So in the large fiber case we must differentiate between two cases: whether $d \in \{a, b\}$ or $d \notin \{a, b\}$. First of all, in the $d \notin \{a, b\}$ case the problem of finding a manipulation point with not too small (i.e., inverse polynomial in n, k and ε^{-1}) probability has already been solved by Isaksson, Kindler and Mossel [**?**], so we are primarily interested in the $d \in \{a, b\}$ case. But moreover, we will see that our method of proof works in both cases.

In the rest of the section we first deal with the small fiber case, and then with the large fiber case.

16.2 Small fiber case

We now deal with the case when (110) holds. We formalize the ideas of the outline in a series of statements.

First, we want to formalize that the boundaries of the boundaries are big in this refined graph setting as well, when we are on a small fiber. The proof uses the canonical path method, as successfully adapted to this setting by Isaksson, Kindler and Mossel [**?**], and is very similar to the proof of Lemma 15.2, with some necessary modifications due to the fact that we now have n coordinates.

Lemma 16.3. *Fix a coordinate and a pair of alternatives—for simplicity we choose coordinate 1 and alternatives a and b, but we note that this lemma holds in general, we do not assume anything special about these choices. Let* $z_{-1}^{a,b}$ *be such that* $B_1\left(z_{-1}^{a,b}\right)$ *is a small fiber for* $B_1^{a,b;[a:b]}$*. Then, writing* $B \equiv B_1\left(z_{-1}^{a,b}\right)$ *for simplicity, we have*

$$\mathbb{P}\left(\sigma \in \partial\left(B\right)\right) \geq \frac{\gamma}{2nk^5}\mathbb{P}\left(\sigma \in B\right). \tag{112}$$

Proof. Let $B^c = \bar{F}\left(z_{-1}^{a,b}\right) \setminus B$. For every $(\sigma, \sigma') \in B \times B^c$, we define a canonical path from σ to σ', which has to pass through at least one edge in $\partial_e\left(B\right)$. Then if we show that every edge in $\partial_e\left(B\right)$ lies on at most r canonical paths, then it follows that $\left|\partial_e\left(B\right)\right| \geq \left|B\right|\left|B^c\right|/r$.

So let $(\sigma, \sigma') \in B \times B^c$. We define a path from σ to σ' by applying a path construction in each coordinate one by one, and then concatenating these paths: first in the first coordinate we get from σ_1 to σ_1', while leaving all other coordinates unchanged, then in the second coordinate we get from σ_2 to σ_2', while leaving all other coordinates unchanged, and so on, finally in the last coordinate we get from σ_n to σ_n'. In the first coordinate we apply the path construction of Proposition 13.7, but considering the block formed by a and b as a single element; in all other coordinates we apply the path construction of Proposition 13.9. Since this path goes from σ (which is in B) to σ' (which is in B^c), it must pass through at least one edge in $\partial_e(B)$.

For a given edge $(\pi, \pi') \in \partial_e(B)$, at most how many possible $(\sigma, \sigma') \in B \times B^c$ pairs are there such that the canonical path between σ and σ' defined above passes through (π, π')? Let us differentiate two cases.

Suppose π and π' differ in the first coordinate. Then coordinates 2 through n of σ must agree with the respective coordinates of π, while coordinates 2 through n of σ' can be anything (up to the restriction given by $\sigma' \in B^c \subseteq \bar{F}\left(z_{-1}^{a,b}\right)$), giving $\left(\frac{k!}{2}\right)^{n-1}$ possibilities. Now fixing all coordinates except the first,Proposition 13.7 tells us that there are at most $(k-1)^2(k-1)!/2 < k^2(k-1)!$ possibilities for the pair (σ_1, σ_1'). So altogether there are at most $k^2(k-1)!\left(\frac{k!}{2}\right)^{n-1}$ paths that pass through a given edge in $\partial_e(B)$ in this case.

Suppose now that π and π' differ in the i^{th} coordinate, $i \neq 1$. Then the first $i-1$ coordinates of σ' must agree with the first $i-1$ coordinates of π, while coordinates $i+1, \dots, n$ of σ must agree with the respective coordinates of π. The first $i-1$ coordinates of σ, and coordinates $i+1, \dots, n$ of σ' can be anything (up to the restriction given by $\sigma, \sigma' \in \bar{F}\left(z_{-1}^{a,b}\right)$), giving $(k-1)!\left(\frac{k!}{2}\right)^{n-2}$ possibilities. Now fixing all coordinates except the i^{th} coordinate, [**?**, Proposition 6.6.] tells us that there are at most $k^4 k!$ possibilities for the pair (σ_i, σ_i'). So altogether there are at most $2k^4(k-1)!\left(\frac{k!}{2}\right)^{n-1}$ paths that pass through a given edge in $\partial_e(B)$ in this case.

So in any case, there are at most $2k^4(k-1)!\left(\frac{k!}{2}\right)^{n-1}$ paths that pass through a given edge in $\partial_e(B)$.

Recall that $\left|\bar{F}\left(z_{-1}^{a,b}\right)\right| = (k-1)!\left(\frac{k!}{2}\right)^{n-1}$, and also $|B^c| \geq \gamma(k-1)!\left(\frac{k!}{2}\right)^{n-1}$ since B is a small fiber. Therefore

$$|\partial_e(B)| \geq \frac{|B|\,|B^c|}{2k^4(k-1)!\left(\frac{k!}{2}\right)^{n-1}} \geq \frac{\gamma}{2k^4}|B|.$$

Now in $G\left(z_{-1}^{a,b}\right)$ every ranking profile has no more than nk neighbors, which implies (112). □

Corollary 16.4. *If* (110) *holds, then*

$$\mathbb{P}\left(\sigma \in \bigcup_{z_{-1}^{a,b}} \partial\left(B_1\left(z_{-1}^{a,b}\right)\right)\right) \geq \frac{\gamma\varepsilon}{2n^2k^{12}}.$$

Proof. Using the previous lemma and (110) we have

$$\begin{aligned}
\mathbb{P}\left(\sigma \in \bigcup_{z_{-1}^{a,b}} \partial\left(B_1\left(z_{-1}^{a,b}\right)\right)\right) &= \sum_{z_{-1}^{a,b}} \mathbb{P}\left(\sigma \in \partial\left(B_1\left(z_{-1}^{a,b}\right)\right)\right) \\
&\geq \sum_{z_{-1}^{a,b}: B_1\left(z_{-1}^{a,b}\right) \subseteq \mathrm{Sm}\left(B_1^{a,b;[a:b]}\right)} \mathbb{P}\left(\sigma \in \partial\left(B_1\left(z_{-1}^{a,b}\right)\right)\right) \\
&\geq \sum_{z_{-1}^{a,b}: B_1\left(z_{-1}^{a,b}\right) \subseteq \mathrm{Sm}\left(B_1^{a,b;[a:b]}\right)} \frac{\gamma}{2nk^5} \mathbb{P}\left(\sigma \in B_1\left(z^{a,b}\right)\right) \\
&= \frac{\gamma}{2nk^5} \mathbb{P}\left(\sigma \in \mathrm{Sm}\left(B_1^{a,b}\right)\right) \geq \frac{\gamma\varepsilon}{2n^2k^{12}}.
\end{aligned}$$

□

Next, we want to find manipulation points on the boundaries of boundaries.

Before we do this, let us divide the boundaries of the boundaries according to which direction they are in. If $\sigma \in \partial\left(B_1\left(z_{-1}^{a,b}\right)\right)$ for some $z_{-1}^{a,b}$, then we know that there exists a ranking profile π such that $(\sigma, \pi) \in \partial_e\left(B_1\left(z_{-1}^{a,b}\right)\right)$. We know that σ and π differ in exactly one coordinate, say coordinate j; in this case we say that σ is on the boundary of $B_1\left(z_{-1}^{a,b}\right)$ in direction j, and we write $\sigma \in \partial_j\left(B_1\left(z_{-1}^{a,b}\right)\right)$. (This notation should not be confused with that of the edge boundary.)

We can write the boundary of $B_1\left(z_{-1}^{a,b}\right)$ as a union of boundaries in the different directions:

$$\partial\left(B_1\left(z_{-1}^{a,b}\right)\right) = \cup_{j=1}^{n} \partial_j\left(B_1\left(z_{-1}^{a,b}\right)\right),$$

but note that this is not (necessarily) a disjoint union, as a ranking profile σ for which $\sigma \in \partial\left(B_1\left(z_{-1}^{a,b}\right)\right)$ might lie on the boundary in multiple directions.

In particular, we differentiate between the boundary in direction 1 and the boundary in all other directions. To this end we introduce the notation

$$\partial_{-1}\left(B_1\left(x_{-1}^{a,b}\right)\right) := \cup_{j=2}^{n} \partial_j\left(B_1\left(x_{-1}^{a,b}\right)\right).$$

With this notation we have the following corollary of Corollary 16.4.

Corollary 16.5. *If* (110) *holds, then either*

$$\mathbb{P}\left(\sigma \in \cup_{z_{-1}^{a,b}} \partial_{-1}\left(B_1\left(z_{-1}^{a,b}\right)\right)\right) \geq \frac{\gamma\varepsilon}{4n^2k^{12}} \tag{113}$$

or

$$\mathbb{P}\left(\sigma \in \cup_{z_{-1}^{a,b}} \partial_{1}\left(B_1\left(z_{-1}^{a,b}\right)\right)\right) \geq \frac{\gamma\varepsilon}{4n^2k^{12}}. \tag{114}$$

Lemma 16.6. *Suppose the ranking profile σ is on the boundary of a fiber for $B_1^{a,b;[a:b]}$ in direction $j \neq 1$, i.e.,*

$$\sigma \in \cup_{z_{-1}^{a,b}} \partial_{-1} \left(B_1 \left(z_{-1}^{a,b} \right) \right).$$

Then there exists a 3-manipulation point $\hat{\sigma}$ which agrees with σ in all coordinates except perhaps coordinate 1 and some coordinate $j \neq 1$; furthermore $\hat{\sigma}_1$ is equal to σ_1 or $[a:b]\,\sigma_1$, except that the position of a third alternative c might be shifted arbitrarily, and $\hat{\sigma}_j$ is equal to σ_j or $z\sigma_j$ for some adjacent transposition $z \in T$, except the position of b might be shifted arbitrarily.

Proof. Suppose $x_{-1}^{a,b}(\sigma) = z_{-1}^{a,b}$. Since $\sigma \in \partial \left(B_1 \left(z_{-1}^{a,b} \right) \right) \subseteq B_1 \left(z_{-1}^{a,b} \right)$, we know that $f(\sigma) = a$, and if $\sigma' = [a:b]_1\,\sigma$, then $f(\sigma') = b$.

Let $\pi = (\pi_j, \sigma_{-j})$ denote the ranking profile such that $(\sigma, \pi) \in \partial_e \left(B_1 \left(z_{-1}^{a,b} \right) \right)$. Let $\pi' := [a:b]_1\,\pi$. Since $\pi \notin B_1 \left(z_{-1}^{a,b} \right)$, $(f(\pi), f(\pi')) \neq (a,b)$. Then, by Lemma 14.1, if $f(\pi) \neq f(\pi')$, then one of π and π' is a 2-manipulation point.

So let us suppose that $f(\pi) = f(\pi')$. If $f(\pi') = a$, then one of σ' and π' is a 2-manipulation point by Lemma 14.1, since $\pi' = z_j \sigma'$ for some adjacent transposition $z \neq [a:b]$. If $f(\pi) = b$, then similarly one of σ and π is a 2-manipulation point.

Finally let us suppose that $f(\pi) = c$ for some $c \notin \{a, b\}$. In this case Lemma 14.2 tells us that there exists an appropriate 3-manipulation point $\hat{\sigma}$. □

Corollary 16.7. *If* (113) *holds, then*

$$\mathbb{P}(\sigma \in M_3) \geq \frac{\gamma \varepsilon}{8 n^3 k^{16}}. \tag{115}$$

Proof. Lemma 16.6 tells us that for every ranking profile σ which is on the boundary of a fiber for $B_1^{a,b;[a:b]}$ in some direction $j \neq 1$, there is a 3-manipulation point $\hat{\sigma}$ "nearby"; the lemma specifies what "nearby" means.

How many ranking profiles σ may give the same $\hat{\sigma}$? At most $2nk^4$, which comes from the following: σ and $\hat{\sigma}$ agree in all coordinates except maybe two, one of which is the first coordinate; there are $n-1 < n$ possibilities for the other coordinate; in the first coordinate, $\hat{\sigma}_1$ is either σ_1 or $[a:b]\,\sigma_1$ (giving 2 possibilities), while some alternative c ($k-2 < k$ possibilities) might be shifted arbitrarily (at most k possibilities); in the other coordinate $j \neq 1$, $\hat{\sigma}_j$ is equal to σ_j or $z\sigma_j$ for some adjacent transposition $z \in T$ (at most k possibilities), except b might be shifted arbitrarily (k possibilities).

So putting this result from Lemma 16.6 together with (113) yields (115). □

The remaining case we have to deal with is when (114) holds.

Lemma 16.8. *Suppose the ranking profile σ is on the boundary of a fiber for $B_1^{a,b;[a:b]}$ in direction 1, i.e.,*

$$\sigma \in \cup_{z_{-1}^{a,b}} \partial_1 \left(B_1 \left(z_{-1}^{a,b} \right) \right).$$

Then either $\sigma \in \mathrm{LD}_1(a,b)$, or there exists a 3-manipulation point $\hat{\sigma}$ which agrees with σ in all coordinates except perhaps in coordinate 1; furthermore $\hat{\sigma}_1$ is equal to σ_1, or $[a:b]\,\sigma_1$ except that the position of a third alternative c might be shifted arbitrarily.

Proof. Just like the proof of Lemma 15.4. □

The following corollary then tells us that either we have found many 3-manipulation points, or we have many local dictators on three alternatives in coordinate 1.

Corollary 16.9. *Suppose* (114) *holds. Then either*

$$\sum_{c \notin \{a,b\}} \mathbb{P}\left(\sigma \in \mathrm{LD}_1^{\{a,b,c\}}\right) = \mathbb{P}\left(\sigma \in \mathrm{LD}_1(a,b)\right) \geq \frac{\gamma\varepsilon}{8n^2k^{12}} \tag{116}$$

or

$$\mathbb{P}\left(\sigma \in M_3\right) \geq \frac{\gamma\varepsilon}{16n^2k^{14}}.$$

16.2.1 Dealing with local dictators

So the remaining case we have to deal with in this small fiber case is when (116) holds, i.e., we have many local dictators in coordinate 1.

Lemma 16.10. *Suppose $\sigma \in \mathrm{LD}_1^{\{a,b,c\}}$ for some alternative $c \notin \{a,b\}$. Define $\sigma' := (\sigma_1', \sigma_{-1})$ by letting σ_1' be equal to σ_1 except that the block of a, b and c is moved to the top of the coordinate. Then*

- *either $\sigma' \in \mathrm{LD}_1^{\{a,b,c\}}$,*
- *or there exists a 3-manipulation point $\hat{\sigma}$ which agrees with σ in all coordinates except perhaps in coordinate 1; furthermore $\hat{\sigma}_1$ is equal to σ_1 except that the position of a, b and c might be shifted arbitrarily.*

Proof. Just like the proof of Lemma 15.6. □

Corollary 16.11. *If* (116) *holds, then either*

$$\sum_{c \notin \{a,b\}} \mathbb{P}\left(\sigma \in LD_1^{\{a,b,c\}}, \{\sigma_1(1), \sigma_1(2), \sigma_1(3)\} = \{a,b,c\}\right) \geq \frac{\gamma\varepsilon}{16n^2k^{13}} \tag{117}$$

or

$$\mathbb{P}\left(\sigma \in M_3\right) \geq \frac{\gamma\varepsilon}{16n^2k^{15}}.$$

Proof. Just like the proof of Corollary 15.7. □

Now (117) is equivalent to

$$\sum_{c\notin\{a,b\}} \mathbb{P}\left(\sigma \in LD_1^{\{a,b,c\}}, (\sigma_1(1), \sigma_1(2), \sigma_1(3)) = (a,b,c)\right) \geq \frac{\gamma\varepsilon}{96n^2k^{13}}. \tag{118}$$

We know that

$$\mathbb{P}((\sigma_1(1), \sigma_1(2), \sigma_1(3)) = (a,b,c)) = \frac{1}{k(k-1)(k-2)} \leq \frac{6}{k^3},$$

and so (118) implies (recall Definition 14.7)

$$\sum_{c\notin\{a,b\}} \mathbb{P}_1^{(a,b,c)}\left(\sigma \in LD_1^{\{a,b,c\}}\right) \geq \frac{\gamma\varepsilon}{576n^2k^{10}}. \tag{119}$$

Now fix an alternative $c \notin \{a,b\}$ and define the graph $G_{(a,b,c)} = \left(V_{(a,b,c)}, E_{(a,b,c)}\right)$ to have vertex set

$$V_{(a,b,c)} := \{\sigma \in S_k^n : (\sigma_1(1), \sigma_1(2), \sigma_1(3)) = (a,b,c)\}$$

and for $\sigma, \pi \in V_{(a,b,c)}$ let $(\sigma,\pi) \in E_{(a,b,c)}$ if and only if σ and π differ in exactly one coordinate, and by an adjacent transposition in this coordinate. So $G_{(a,b,c)}$ is the subgraph of the refined rankings graph induced by the vertex set $V_{(a,b,c)}$.

Let

$$T_1(a,b,c) := V_{(a,b,c)} \cap LD_1^{\{a,b,c\}},$$

and let $\partial_e(T_1(a,b,c))$ and $\partial(T_1(a,b,c))$ denote the edge and vertex boundary of $T_1(a,b,c)$ in $G_{(a,b,c)}$, respectively.

The next lemma shows that unless $T_1(a,b,c)$ is almost all of $V_{(a,b,c)}$, the size of the boundary $\partial(T_1(a,b,c))$ is comparable to the size of $T_1(a,b,c)$.

Lemma 16.12. *Let $c \notin \{a,b\}$ be arbitrary. Write $T \equiv T_1(a,b,c)$ for simplicity. If $\mathbb{P}_1^{(a,b,c)}(\sigma \in T) \leq 1-\delta$, then*

$$\mathbb{P}_1^{(a,b,c)}(\sigma \in \partial(T)) \geq \frac{\delta}{nk^3}\mathbb{P}_1^{(a,b,c)}(\sigma \in T). \tag{120}$$

Proof. The proof is essentially the same as the proof of Lemma 15.8, with a slight modification to deal with n coordinates. Let $T^c = V_{(a,b,c)} \setminus T(a,b,c)$. For every $(\sigma,\sigma') \in T \times T^c$ we define a canonical path from σ to σ' by applying a path construction in each coordinate one by one, and then concatenating these paths. In all coordinates we apply the path construction of [**?**, Proposition 6.4.], but in the first coordinate we only apply it to alternatives $[k] \setminus \{a,b,c\}$.

The analysis of this construction is done in exactly the same way as in Lemma 16.3; in the end we get that $|\partial_e(T)| \geq \frac{\delta}{k^2}|T|$. Now every vertex in $V_{(a,b,c)}$ has no more than nk neighbors, which implies (120). □

The next lemma tells us that if σ is on the boundary of a set of local dictators on $\{a, b, c\}$ for some alternative $c \notin \{a, b\}$ in coordinate 1, then there is a 4-manipulation point $\hat{\sigma}$ which is close to σ. The proof is similar to that of Lemma 15.9, but we have to take care of all n coordinates.

Lemma 16.13. *Suppose $\sigma \in \partial (T_1 (a, b, c))$ for some $c \notin \{a, b\}$. We distinguish two cases, based on the number of alternatives.*

If $k = 3$, then there exists a (3-)manipulation point $\hat{\sigma}$ which differs from σ in at most two coordinates, one of them being the first coordinate.

If $k \geq 4$, then there exists a 4-manipulation point $\hat{\sigma}$ which differs from σ in at most two coordinates, one of them being the first coordinate; furthermore, $\hat{\sigma}_1$ is equal to σ_1 except that the order of the block of a, b and c might be rearranged and an additional alternative d might be shifted arbitrarily; and in the other coordinate, call it j, $\hat{\sigma}_j$ is equal to σ_j except perhaps a, b and c are shifted arbitrarily.

Proof. Let π be the ranking profile such that $(\sigma, \pi) \in \partial_e (T_1 (a, b, c))$, let j be the coordinate in which they differ, and let z be the adjacent transposition in which they differ, i.e., $\pi = z_j \sigma$. Since $\pi \notin T_1 (a, b, c)$, there exists a reordering of the block of a, b, and c at the top of π_1 such that the outcome of f is not the top ranked alternative in coordinate 1. Call the resulting vector π_1', and let $\pi' := (\pi_1', \pi_{-1})$. W.l.o.g. let us assume that $\pi_1' (1) = a$. Let us also define $\sigma' := z_j \pi'$. We distinguish two cases: $j = 1$ and $j \neq 1$.

If $j = 1$ (in which case we must have $k \geq 5$), π' is a 2-manipulation point, since $f (\sigma') = a$.

If $j \neq 1$, then there are various cases to consider. If the adjacent transposition z does not move a, then either π' or σ' is a 2-manipulation point. So let us suppose that $z = [a : d]$ for some $d \neq a$.

Clearly we must have $f (\pi') = d$, or else π' or σ' is a 2-manipulation point. Suppose first that $d \in \{b, c\}$. W.l.o.g. suppose that $d = b$.

Then take alternative c in coordinate j of both σ' and π', and bubble it to the block of a and b simultaneously in the two ranking profiles. If along the way the value of the outcome of the SCF f changes from a or b, respectively, then we have a 2-manipulation point by Lemma 14.1. Otherwise, we now have a, b, and c adjacent in both coordinates 1 and j. Now rearranging the order of the blocks of a, b, and c in these two coordinates (which can be done using adjacent transpositions), we either get a 2-manipulation point by Lemma 14.1, or we can define a new SCF on two voters and three alternatives, a, b, and c. This SCF takes on three values and it is also not hard to see that the outcome is not only a function of the first coordinate, so by the Gibbard-Satterthwaite theorem we know that this SCF has a manipulation point, which is a 3-manipulation point of the original SCF f.

Now let us look at the case when $d \notin \{b, c\}$. In this case we do something similar to what we just did in the previous paragraph. In both σ' and π', first bubble up alternative d in coordinate 1 up to the block of a, b, and c, and then bubble b and c in coordinate j to the block of a and d. All of this using adjacent transpositions. If the value of the outcome of the SCF f changes from a or d, respectively, at any time along the way, then we have a 2-manipulation point by Lemma 14.1. Otherwise, we now have a, b, c and d adjacent in both coordinates 1 and j, and we can apply the same trick to find a 4-manipulation point, using the Gibbard-Satterthwaite theorem. □

The next corollary puts together Corollary 16.11 and Lemmas 16.12 and 16.13.

Corollary 16.14. *Suppose* (117) *holds. Then if for every $c \notin \{a,b\}$ we have $\mathbb{P}_1^{(a,b,c)}(\sigma \in T_1(a,b,c)) \leq 1 - \frac{\varepsilon}{100k}$, then*

$$\mathbb{P}(\sigma \in M_4) \geq \frac{\gamma\varepsilon^2}{345600 n^4 k^{22}}.$$

Proof. We know that (117) implies

$$\sum_{c \notin \{a,b\}} \mathbb{P}_1^{a,b,c}(\sigma \in T_1(a,b,c)) \geq \frac{\gamma\varepsilon}{576 n^2 k^{10}}.$$

Now then using the assumptions, Lemma 16.12 with $\delta = \frac{\varepsilon}{100k}$ and Lemma 16.13, we have

$$\begin{aligned}
\mathbb{P}(\sigma \in M_4) &\geq \sum_{c \neq \{a,b\}} \frac{1}{k^3} \mathbb{P}_1^{(a,b,c)}(\sigma \in M_4) \geq \sum_{c \notin \{a,b\}} \frac{1}{6nk^8} \mathbb{P}_1^{(a,b,c)}(\sigma \in \partial(T_1(a,b,c))) \\
&\geq \sum_{c \notin \{a,b\}} \frac{\varepsilon}{600 n^2 k^{12}} \mathbb{P}_1^{(a,b,c)}(\sigma \in T_1(a,b,c)) \geq \frac{\gamma\varepsilon^2}{345600 n^4 k^{22}}.
\end{aligned}$$

□

So again we are left with one case to deal with: if there exists an alternative $c \notin \{a,b\}$ such that $\mathbb{P}_1^{(a,b,c)}(\sigma \in T_1(a,b,c)) > 1 - \frac{\varepsilon}{100k}$. Define a subset of alternatives $K \subseteq [k]$ in the following way:

$$K := \{a,b\} \cup \left\{ c \in [k] \setminus \{a,b\} : \mathbb{P}_1^{(a,b,c)}(\sigma \in T_1(a,b,c)) > 1 - \frac{\varepsilon}{100k} \right\}.$$

In addition to a and b, K contains those alternatives that whenever they are at the top of coordinate 1 with a and b, they form a local dictator with high probability.

So our assumption now is that $|K| \geq 3$.

Our next step is to show that unless we have many manipulation points, for any alternative $c \in K$, conditioned on c being at the top of the first coordinate, the outcome of f is c with probability close to 1.

Lemma 16.15. *Let $c \in K$. Then either*

$$\mathbb{P}_1^{(c)}(f(\sigma) = c) \geq 1 - \frac{\varepsilon}{50k}, \tag{121}$$

or

$$\mathbb{P}(\sigma \in M_2) \geq \frac{\varepsilon}{100k^4}. \tag{122}$$

Proof. Just like the proof of Lemma 15.11. □

We now deal with alternatives that are not in K: either we have many manipulation points, or for any alternative $d \notin K$, the outcome of f is *not* d with probability close to 1.

Lemma 16.16. *Let $d \notin K$. If $\mathbb{P}(f(\sigma) = d) \geq \frac{\varepsilon}{4k}$, then*

$$\mathbb{P}(\sigma \in M_4) \geq \frac{\varepsilon^2}{10^6 n^2 k^{13}}.$$

Proof. The proof is very similar to that of Lemma 15.12: we do the same steps in the first coordinate as done in the proof of Lemma 15.12, and the fact that we have n coordinates only matters at the very end.

Let σ be such that $f(\sigma) = d$. We will keep coordinates 2 through n to be fixed as σ_{-1} throughout the proof. By bubbling alternatives d, a, and b in the first coordinate, we can define σ', $\sigma^{(d,b,a)}$, $\sigma^{(d,a,b)}$, $\sigma^{(a,b,d)}$, $\sigma^{(a,d,b)}$, $\sigma^{(b,a,d)}$, and $\sigma^{(b,d,a)}$ just as in the proof of Lemma 15.12. Again, we can show that either

$$\mathbb{P}(\sigma \in M_2) \geq \frac{\varepsilon}{1600k^3},$$

in which case we are done, or

$$\mathbb{P}_1^{(a,b,d)}\left(\sigma^{(a,b,d)} \in LD_1^{\{a,b,d\}}\right) = \mathbb{P}\left(\sigma : \sigma^{(a,b,d)} \in LD_1^{\{a,b,d\}}\right) \geq \frac{\varepsilon}{1600k}. \tag{123}$$

Define $G_{(a,b,d)}$ and $T_{(a,b,d)}$ analogously to $G_{(a,b,c)}$ and $T_{(a,b,c)}$, respectively.

Suppose that (123) holds. We also know that $d \notin K$, so Lemma 16.12 applies, and then Lemma 16.13 shows us how to find manipulation points. We can put these arguments together, just like in the proof of Corollary 16.14, to show what we need:

$$\begin{aligned}\mathbb{P}(\sigma \in M_4) &\geq \frac{1}{k^3}\mathbb{P}_1^{(a,b,d)}(\sigma \in M_4) \geq \frac{1}{6nk^8}\mathbb{P}_1^{(a,b,d)}(\sigma \in \partial(T_1(a,b,d))) \\ &\geq \frac{\varepsilon}{600n^2k^{12}}\mathbb{P}_1^{(a,b,d)}(\sigma \in T_1(a,b,d)) \geq \frac{\varepsilon^2}{10^6 n^2 k^{13}}.\end{aligned}$$

□

Putting together the results of the previous lemmas, there is only one case to be covered, which is covered by the following final lemma. Basically, this lemma says that unless there are enough manipulation points, our function is close to a dictator in the first coordinate, on the subset of alternatives K.

Lemma 16.17. *Recall that we assume that $\mathbf{D}\left(f, \overline{\text{NONMANIP}}\right) \geq \varepsilon$. Furthermore assume that $|K| \geq 3$, for every $c \in K$ we have*

$$\mathbb{P}_1^{(c)}(f(\sigma) = c) \geq 1 - \frac{\varepsilon}{50k}, \tag{124}$$

and for every $d \notin K$ *we have*

$$\mathbb{P}\left(f\left(\sigma\right)=d\right)\leq\frac{\varepsilon}{4k}.$$

Then

$$\mathbb{P}\left(\sigma\in M_2\right)\geq\frac{\varepsilon}{4k^2}. \tag{125}$$

Proof. Just like the proof of Lemma 16.17. □

To conclude the proof in the small fiber case, inspect all the lower bounds for $\mathbb{P}\left(\sigma\in M_4\right)$ obtained in Section 16.2, and recall that $\gamma=\frac{\varepsilon^3}{10^3 n^3 k^{24}}$.

16.3 Large fiber case

We now deal with the large fiber case, when (111) holds for both boundaries, i.e., when

$$\mathbb{P}\left(\sigma\in\mathrm{Lg}\left(B_1^{a,b;[a:b]}\right)\right)\geq\frac{\varepsilon}{nk^7}$$

and

$$\mathbb{P}\left(\sigma\in\mathrm{Lg}\left(B_2^{c,d;[c:d]}\right)\right)\geq\frac{\varepsilon}{nk^7}.$$

We differentiate between two cases: whether $d\in\{a,b\}$ or $d\notin\{a,b\}$.

16.3.1 Case 1

Suppose $d\in\{a,b\}$, in which case w.l.o.g. we may assume that $d=a$. That is, in the rest of this case we may assume that

$$\mathbb{P}\left(\sigma\in\mathrm{Lg}\left(B_1^{a,b;[a:b]}\right)\right)\geq\frac{\varepsilon}{nk^7} \tag{126}$$

and

$$\mathbb{P}\left(\sigma\in\mathrm{Lg}\left(B_2^{a,c;[a:c]}\right)\right)\geq\frac{\varepsilon}{nk^7}. \tag{127}$$

First, let us look at only the boundary between a and b in direction 1. Let us fix a vector $z_{-1}^{a,b}$ which gives a large fiber $B_1\left(z_{-1}^{a,b}\right)$ for the boundary $B_1^{a,b;[a:b]}$, i.e., we know that

$$\mathbb{P}\left(\sigma\in B_1\left(z_{-1}^{a,b}\right)\,\middle|\,\sigma\in\bar{F}\left(z_{-1}^{a,b}\right)\right)\geq 1-\gamma. \tag{128}$$

Our basic goal in the following will be to show that conditional on the ranking profile σ being in the fiber $F\left(z_{-1}^{a,b}\right)$ (but not necessarily in $\bar{F}\left(z_{-1}^{a,b}\right)$), with high probability the outcome of the vote is $\mathrm{top}_{\{a,b\}}\left(\sigma_1\right)$, or else we have a lot of 2-manipulation points or local dictators on three alternatives in coordinate 1.

Our first step towards this is the following.

Lemma 16.18. *Suppose $z_{-1}^{a,b}$ gives a large fiber $B_1\left(z_{-1}^{a,b}\right)$ for the boundary $B_1^{a,b;[a:b]}$. Then*

$$\mathbb{P}_1^{(a,b)}\left(\sigma \in B_1\left(z_{-1}^{a,b}\right) \middle| \sigma \in F\left(z_{-1}^{a,b}\right)\right) \geq 1 - k\gamma. \tag{129}$$

Proof. We know that

$$\mathbb{P}\left(\left(\sigma_1\left(1\right), \sigma_1\left(2\right)\right) = (a,b) \middle| \sigma \in \bar{F}\left(z_{-1}^{a,b}\right)\right) = \frac{1}{k-1},$$

and so

$$\begin{aligned}\mathbb{P}_1^{(a,b)}&\left(\sigma \notin B_1\left(z_{-1}^{a,b}\right) \middle| \sigma \in F\left(z_{-1}^{a,b}\right)\right) = \mathbb{P}_1^{(a,b)}\left(\sigma \notin B_1\left(z_{-1}^{a,b}\right) \middle| \sigma \in \bar{F}\left(z_{-1}^{a,b}\right)\right) \\ &= (k-1)\,\mathbb{P}\left(\sigma \notin B_1\left(z_{-1}^{a,b}\right), \left(\sigma_1\left(1\right), \sigma_1\left(2\right)\right) = (a,b) \middle| \sigma \in \bar{F}\left(z_{-1}^{a,b}\right)\right) \leq (k-1)\,\gamma < k\gamma. \square\end{aligned}$$

The next lemma formalizes our goal mentioned above.

Lemma 16.19. *Suppose $z_{-1}^{a,b}$ gives a large fiber $B_1\left(z_{-1}^{a,b}\right)$ for the boundary $B_1^{a,b;[a:b]}$. Then either*

$$\mathbb{P}\left(f\left(\sigma\right) = \mathrm{top}_{\{a,b\}}\left(\sigma_1\right) \middle| \sigma \in F\left(z_{-1}^{a,b}\right)\right) \geq 1 - 2k\gamma \tag{130}$$

or

$$\mathbb{P}\left(\sigma \in M_2 \middle| \sigma \in F\left(z_{-1}^{a,b}\right)\right) \geq \frac{\gamma}{2k} \tag{131}$$

or

$$\mathbb{P}\left(\sigma \in LD_1\left(a,b\right) \middle| \sigma \in F\left(z_{-1}^{a,b}\right)\right) \geq \frac{\gamma}{2k}. \tag{132}$$

Proof. The proof of this lemma is essentially the same as that of Lemma 15.15, there are only two slight differences. First, we use Lemma 16.18 to know that (129) holds. Second, we take $\sigma \in F\left(z_{-1}^{a,b}\right)$ to be uniform, and we stay on the fiber $F\left(z_{-1}^{a,b}\right)$ throughout the proof: we modify only the first coordinate throughout the proof, in the same way as we did for Lemma 15.15. We omit the details. $\square$

Now this lemma holds for all vectors $z_{-1}^{a,b}$ which give a large fiber $B_1\left(z_{-1}^{a,b}\right)$ for the boundary $B_1^{a,b;[a:b]}$. By (126) we know that

$$\mathbb{P}\left(\sigma : B_1\left(x_{-1}^{a,b}\left(\sigma\right)\right) \text{ is a large fiber}\right) \geq \frac{\varepsilon}{nk^7}.$$

Now if (131) holds for at least a third of the vectors $z_{-1}^{a,b}$ that give a large fiber $B_1\left(z_{-1}^{a,b}\right)$, then it follows that

$$\mathbb{P}\left(\sigma \in M_2\right) \geq \frac{\gamma\varepsilon}{6nk^8}$$

and we are done. If (132) holds for at least a third of the vectors $z_{-1}^{a,b}$ that give a large fiber $B_1\left(z_{-1}^{a,b}\right)$, then similarly we have

$$\mathbb{P}\left(\sigma \in LD_1\left(a,b\right)\right) \geq \frac{\gamma\varepsilon}{6nk^8},$$

which means that (116) also holds, and so we are done by the argument in Section 16.2.1.

So the remaining case to consider is when (130) holds for at least a third of the vectors $z_{-1}^{a,b}$ that give a large fiber $B_1\left(z_{-1}^{a,b}\right)$.

We can go through this same argument for the boundary between a and c in direction 2 as well, and either we are done because

$$\mathbb{P}\left(\sigma \in M_2\right) \geq \frac{\gamma\varepsilon}{6nk^8}$$

or

$$\mathbb{P}\left(\sigma \in LD_2\left(a,c\right)\right) \geq \frac{\gamma\varepsilon}{6nk^8},$$

or for at least a third of the vectors $z_{-2}^{a,c}$ that give a large fiber $B_2\left(z_{-2}^{a,c}\right)$ we have

$$\mathbb{P}\left(f\left(\sigma\right) = \mathrm{top}_{\{a,c\}}\left(\sigma_2\right) \,\middle|\, \sigma \in F\left(z_{-2}^{a,c}\right)\right) \geq 1 - 2k\gamma.$$

So basically our final case is if

$$\mathbb{P}\left(\sigma \in F_1^{a,b}\right) \geq \frac{\varepsilon}{3nk^7} \tag{133}$$

and also

$$\mathbb{P}\left(\sigma \in F_2^{a,c}\right) \geq \frac{\varepsilon}{3nk^7}. \tag{134}$$

Notice that being in the set $F_1^{a,b}$ only depends on the vector $x^{a,b}\left(\sigma\right)$ of preferences between a and b, and similarly being in the set $F_2^{a,c}$ only depends on the vector $x^{a,c}\left(\sigma\right)$ of preferences between a and c. We know that $\left\{\left(x_i^{a,b}\left(\sigma\right), x_i^{a,c}\left(\sigma\right)\right)\right\}_{i=1}^{n}$ are independent, and for any given i we know that $\left|\mathbb{E}\left(x_i^{a,b}\left(\sigma\right) x_i^{a,c}\left(\sigma\right)\right)\right| = \frac{1}{3}$. Hence we can apply reverse hypercontractivity (Lemma **??**), to get the following result.

Lemma 16.20. *If* (133) *and* (134) *hold, then also*

$$\mathbb{P}\left(\sigma \in F_1^{a,b} \cap F_2^{a,c}\right) \geq \frac{\varepsilon^3}{27n^3k^{21}}. \tag{135}$$

Proof. See above. □

The next and final lemma then concludes that we have lots of manipulation points.

Lemma 16.21. *Suppose* (135) *holds. Then*

$$\mathbb{P}\left(\sigma \in M_3\right) \geq \frac{\varepsilon^3}{54n^3k^{27}} - \frac{9\gamma}{k^3}. \tag{136}$$

Proof. First let us define two events:

$$I_1 := \left\{\sigma : f\left(\sigma\right) = \mathrm{top}_{\{a,b\}}\left(\sigma_1\right)\right\}$$
$$I_2 := \left\{\sigma : f\left(\sigma\right) = \mathrm{top}_{\{a,c\}}\left(\sigma_2\right)\right\}.$$

Using similar estimates as previously in Lemma 11.3, we have

$$\mathbb{P}\left(\sigma \in I_1 \cap I_2 \cap F_1^{a,b} \cap F_2^{a,c}\right) \geq \mathbb{P}\left(\sigma \in F_1^{a,b} \cap F_2^{a,c}\right)$$
$$- \mathbb{P}\left(\sigma \notin I_1, \sigma \in F_1^{a,b} \cap F_2^{a,c}\right) - \mathbb{P}\left(\sigma \notin I_2, \sigma \in F_1^{a,b} \cap F_2^{a,c}\right).$$

The first term is bounded below via (135), while the other two terms can be bounded using the definition of $F_1^{a,b}$ and $F_2^{a,c}$, respectively:

$$\mathbb{P}\left(\sigma \notin I_1, \sigma \in F_1^{a,b} \cap F_2^{a,c}\right) \leq \mathbb{P}\left(\sigma \notin I_1, \sigma \in F_1^{a,b}\right) \leq \mathbb{P}\left(\sigma \notin I_1 \,\middle|\, \sigma \in F_1^{a,b}\right) \leq 2k\gamma,$$

and similarly for the other term. Putting everything together gives us

$$\mathbb{P}\left(\sigma \in I_1 \cap I_2 \cap F_1^{a,b} \cap F_2^{a,c}\right) \geq \frac{\varepsilon^3}{27n^3k^{21}} - 4k\gamma.$$

If $\sigma \in I_1 \cap I_2 \cap F_1^{a,b} \cap F_2^{a,c}$, then clearly we must have $f\left(\sigma\right) = a$, and therefore $x_1^{a,b}\left(\sigma\right) = 1$ and $x_2^{a,c}\left(\sigma\right) = 1$. Now define σ' from σ by bubbling up b in coordinate 1 to just below a, and bubbling up c in coordinate 2 to just below a. Either we encounter a 2-manipulation point along the way, or the outcome is still a: $f\left(\sigma'\right) = a$. If we encounter a 2-manipulation point along the way for at least half of such ranking profiles, then we are done:

$$\mathbb{P}\left(\sigma \in M_2\right) \geq \frac{1}{k^2}\left(\frac{\varepsilon^3}{54n^3k^{21}} - 2k\gamma\right) = \frac{\varepsilon^3}{54n^3k^{23}} - \frac{2\gamma}{k}.$$

Otherwise, we may assume that

$$\mathbb{P}\left(\sigma \in I_1 \cap I_2 \cap F_1^{a,b} \cap F_2^{a,c}, f\left(\sigma'\right) = a\right) \geq \frac{\varepsilon^3}{54n^3k^{21}} - 2k\gamma.$$

In this case define $\tilde{\sigma}' := [a:b]_1\, \sigma'$ and $\tilde{\sigma}'' := [a:c]_2\, \sigma'$. If $f\left(\tilde{\sigma}'\right) \notin \{a,b\}$ or $f\left(\tilde{\sigma}''\right) \notin \{a,c\}$, then we automatically have that one of $\sigma', \tilde{\sigma}', \tilde{\sigma}''$ is a 2-manipulation point. If $f\left(\tilde{\sigma}'\right) = b$ and $f\left(\tilde{\sigma}''\right) = c$, then by Lemma 14.2 we know that there exists a 3-manipulation point $\hat{\sigma}$

which agrees with σ except perhaps a, b, and c could be arbitrarily shifted in the first two coordinates. The final case is when $a \in \{f(\tilde{\sigma}'), f(\tilde{\sigma}'')\}$. But we now show that this has small probability, and therefore (136) follows.

First let us look at the case of $f(\tilde{\sigma}') = a$. We have

$$\begin{aligned}
&\mathbb{P}\left(\sigma \in I_1 \cap I_2 \cap F_1^{a,b} \cap F_2^{a,c}, f(\sigma') = a, f(\tilde{\sigma}') = a\right) \\
&= \sum_{z_{-1}^{a,b} : F\left(z_{-1}^{a,b}\right) \subseteq F_1^{a,b}} \mathbb{P}\left(\sigma \in I_1 \cap I_2 \cap F\left(\left(1, z_{-1}^{a,b}\right)\right) \cap F_2^{a,c}, f(\sigma') = a, f(\tilde{\sigma}') = a\right) \\
&= \sum_{z_{-1}^{a,b} : F\left(z_{-1}^{a,b}\right) \subseteq F_1^{a,b}} \mathbb{P}\left(\sigma \in I_1 \cap I_2 \cap F_2^{a,c}, f(\sigma') = a, f(\tilde{\sigma}') = a \,\middle|\, \sigma \in F\left(\left(1, z_{-1}^{a,b}\right)\right)\right) \mathbb{P}\left(\sigma \in F\left(\left(1, z_{-1}^{a,b}\right)\right)\right) \\
&\leq \sum_{z_{-1}^{a,b} : F\left(z_{-1}^{a,b}\right) \subseteq F_1^{a,b}} \mathbb{P}\left(\sigma : f(\tilde{\sigma}') = a \,\middle|\, \sigma \in F\left(\left(1, z_{-1}^{a,b}\right)\right)\right) \mathbb{P}\left(\sigma \in F\left(\left(1, z_{-1}^{a,b}\right)\right)\right).
\end{aligned}$$

Now we know that $\tilde{\sigma}' \in F\left(\left(-1, z_{-1}^{a,b}\right)\right) \subseteq F_1^{a,b}$, and we also know that

$$\mathbb{P}\left(f(\sigma) \neq b \,\middle|\, \sigma \in F\left(\left(-1, z_{-1}^{a,b}\right)\right)\right) \leq 4k\gamma.$$

The number of σ's that give the same $\tilde{\sigma}'$ is at most k^2, and so we can conclude that

$$\mathbb{P}\left(\sigma \in I_1 \cap I_2 \cap F_1^{a,b} \cap F_2^{a,c}, f(\sigma') = a, f(\tilde{\sigma}') = a\right) \leq 4k^3\gamma,$$

and similarly

$$\mathbb{P}\left(\sigma \in I_1 \cap I_2 \cap F_1^{a,b} \cap F_2^{a,c}, f(\sigma') = a, f(\tilde{\sigma}'') = a\right) \leq 4k^3\gamma,$$

which shows that

$$\mathbb{P}(\sigma \in M_3) \geq \frac{1}{k^6}\left(\frac{\varepsilon^3}{54n^3k^{21}} - 2k\gamma - 8k^3\gamma\right) \geq \frac{\varepsilon^3}{54n^3k^{27}} - \frac{9\gamma}{k^3}. \qquad \square$$

To conclude the proof in this case, recall that we have chosen $\gamma = \frac{\varepsilon^3}{10^3 n^3 k^{24}}$.

16.3.2 Case 2

First, as in the previous case, we can look at simply the boundary between a and b in direction 1, and conclude that either there are many manipulation points, or there are many local dictators, or (133) holds. This holds similarly for the boundary between c and d in direction 2. Finally, just as in Section 11.3.2, we can show that (133) and (134) cannot hold at the same time. We omit the details.

16.4 Proof of Theorem 16.1 concluded

Proof of Theorem 16.1. Our starting point is Lemma 14.3, which directly implies Lemma 16.2 (unless there are many 2-manipulation points, in which case we are done). We then consider two cases, as indicated in Section 16.1.

We deal with the small fiber case in Section 16.2. First, Lemmas 16.3, 16.6, and 16.8, and Corollaries 16.4, 16.5, 16.7, and 16.9 imply that either there are many 3-manipulation points, or there are many local dictators on three alternatives in coordinate 1. We then deal with the case of many local dictators in Section 16.2.1. Lemma 16.10, Corollary 16.11, Lemmas 16.12, 16.13, Corollary 16.14, and Lemmas 16.15, 16.16, and 16.17 together show that there are many 4-manipulation points if there are many local dictators on three alternatives, and the SCF is ε-far from the family of nonmanipulable functions.

We deal with the large fiber case in Section 16.3. Here Lemmas 16.18, 16.19, 16.20, and 16.21 show that if there are not many local dictators on three alternatives, then there are many 3-manipulation points. In the case when there are many local dictators, we refer back to Section 16.2.1 to conclude the proof. □

17 Reduction to distance from truly nonmanipulable SCFs

Proof of Theorem 14.8. Our assumption means that there exists a SCF $g \in \overline{\text{NONMANIP}}$ such that $\mathbf{D}(f, g) \leq \alpha$. We distinguish two cases: either g is a function of one coordinate, or g takes on at most two values.

Case 1. g is a function of one coordinate. In this case we can assume w.l.o.g. that g is a function of the first coordinate, i.e., there exists a SCF $h : S_k \to [k]$ on one coordinate such that for every ranking profile σ, we have $g(\sigma) = h(\sigma_1)$.

We know from the quantitative Gibbard-Satterthwaite theorem for one voter that for any β either $\mathbf{D}(h, \text{NONMANIP}(1,k)) \leq \beta$, or $\mathbb{P}(\sigma \in M_3(h)) \geq \frac{\beta^3}{10^5 k^{16}}$.

In the former case, we have that

$$\mathbf{D}(g, \text{NONMANIP}(n,k)) \leq \mathbf{D}(h, \text{NONMANIP}(1,k)) \leq \beta,$$

and so consequently

$$\mathbf{D}(f, \text{NONMANIP}(n,k)) \leq \alpha + \beta.$$

In the latter case, we have that

$$\mathbb{P}(\sigma \in M_3(g)) = \mathbb{P}(\sigma \in M_3(h)) \geq \frac{\beta^3}{10^5 k^{16}},$$

and so consequently

$$\mathbb{P}(\sigma \in M_3(f)) \geq \frac{\beta^3}{10^5 k^{16}} - 6nk\alpha,$$

since changing the outcome of a SCF at one ranking profile can change the number of 3-manipulation points by at most $6nk$. Now choosing $\beta = 100nk^6\alpha^{1/3}$ shows that either (86) or (87) holds.

Case 2. g is a function which takes on at most two values. W.l.o.g. we may assume that the range of g is $\{a,b\} \subset [k]$, i.e., for every ranking profile $\sigma \in S_k^n$ we have $g(\sigma) \in \{a,b\}$.

There is one thing we have to be careful about: even though g takes on at most two values, it is not necessarily a Boolean function, since the value of $g(\sigma)$ does not necessarily depend only on the Boolean vector $x^{a,b}(\sigma)$.

We now define a function $h : S_k^n \to \{a,b\}$ that is close in some sense to g and which can be viewed as a Boolean function $h : \{a,b\}^n \to \{a,b\}$ because $h(\sigma)$ depends on σ only through $x^{a,b}(\sigma)$. (The vector $x^{a,b}(\sigma) \in \{-1,1\}^n$ encodes which of a and b is preferred in each coordinate, and a vector in $\{a,b\}^n$ can encode the same information.) For a given ranking profile σ, let us consider the fiber on which it is on, $F\left(x^{a,b}(\sigma)\right)$, and let us define $g|_{F\left(x^{a,b}(\sigma)\right)}$ to be the restriction of g to ranking profiles in the fiber $F\left(x^{a,b}(\sigma)\right)$. Then define (see Definition 14.9)

$$h(\sigma) := \text{Maj}\left(g|_{F\left(x^{a,b}(\sigma)\right)}\right).$$

By definition, $h(\sigma)$ depends on σ only through $x^{a,b}(\sigma)$, so we may also view h as a Boolean function $h : \{a,b\}^n \to \{a,b\}$.

For any given $0 < \delta < 1$, we either have $\mathbf{D}(g,h) \leq \delta$, in which case $\mathbf{D}(f,h) \leq \alpha + \delta$, or if $\mathbf{D}(g,h) > \delta$, then we show presently that

$$\mathbb{P}(\sigma \in M_2(f)) \geq \frac{\delta}{4nk^5} - nk\alpha. \tag{137}$$

Choosing $\delta = 8n^2k^6\alpha$ then shows that either (87) holds, or $\mathbf{D}(f,h) \leq 9n^2k^6\alpha$.

Let us now show (137). We use a canonical path argument again, but first we divide the ranking profiles according to the fibers with respect to preference between a and b.

Let us consider an arbitrary fiber $F\left(z^{a,b}\right)$, and divide it into two disjoint sets: into those ranking profiles for which the outcome of g and h agree, and those for which these outcomes are different. I.e.,

$$F\left(z^{a,b}\right) = F^{\text{maj}}\left(z^{a,b}\right) \cup F^{\text{min}}\left(z^{a,b}\right),$$

where

$$\begin{aligned} F^{\text{maj}}\left(z^{a,b}\right) &= \left\{\sigma \in F\left(z^{a,b}\right) : g(\sigma) = h(\sigma)\right\}, \\ F^{\text{min}}\left(z^{a,b}\right) &= \left\{\sigma \in F\left(z^{a,b}\right) : g(\sigma) \neq h(\sigma)\right\}. \end{aligned}$$

By construction, we know that

$$\left|F^{\min}\left(z^{a,b}\right)\right| \leq \frac{1}{2}\left|F\left(z^{a,b}\right)\right| = \frac{1}{2}\left(\frac{k!}{2}\right)^n.$$

Now for every pair of profiles $(\sigma,\sigma') \in F^{\min}\left(z^{a,b}\right) \times F^{\mathrm{maj}}\left(z^{a,b}\right)$ define a canonical path from σ to σ' by applying a path construction in each coordinate one by one, and then concatenating these paths. In each coordinate we apply the path construction of [**?**, Proposition 6.6.]: we bubble up everything except a and b, and then finally bubble up the last two alternatives as well.

For a given edge $(\pi,\pi') \in F^{\min}\left(z^{a,b}\right) \times F^{\mathrm{maj}}\left(z^{a,b}\right)$ there are at most $2k^4\left(\frac{k!}{2}\right)^n$ possible pairs $(\sigma,\sigma') \in F^{\min}\left(z^{a,b}\right) \times F^{\mathrm{maj}}\left(z^{a,b}\right)$ such that the canonical path between σ and σ' defined above passes through (π,π'). (This can be shown just like in the previous lemmas, e.g., Lemma 16.3.) Consequently we have

$$\left|\partial_e\left(F^{\min}\left(z^{a,b}\right)\right)\right| \geq \frac{\left|F^{\min}\left(z^{a,b}\right)\right|\left|F^{\mathrm{maj}}\left(z^{a,b}\right)\right|}{2k^4\left(\frac{k!}{2}\right)^n} \geq \frac{\left|F^{\min}\left(z^{a,b}\right)\right|}{4k^4},$$

where the edge boundary $\partial_e\left(F^{\min}\left(z^{a,b}\right)\right)$ is defined via the refined rankings graph restricted to the fiber $F\left(z^{a,b}\right)$. Summing this over all fibers we have that

$$\sum_{z^{a,b}}\left|\partial_e\left(F^{\min}\left(z^{a,b}\right)\right)\right| \geq \sum_{z^{a,b}}\frac{\left|F^{\min}\left(z^{a,b}\right)\right|}{4k^4} \geq \frac{\delta}{4k^4}(k!)^n, \tag{138}$$

using the fact that $\mathbf{D}(g,h) > \delta$.

Now it is easy to see that if $(\sigma,\sigma') \in \partial_e\left(F^{\min}\left(z^{a,b}\right)\right)$ for some $z^{a,b}$, then either σ or σ' is a 2-manipulation point for g. In the refined rankings graph every vertex (ranking profile) has $n(k-1) < nk$ neighbors, so each 2-manipulation point can be counted at most nk times in the sum on the left hand side of (138), showing that

$$\mathbb{P}\left(\sigma \in M_2(g)\right) \geq \frac{\delta}{4nk^5},$$

from which (137) follows immediately, since changing the outcome of a SCF at one ranking profile can change the number of 2-manipulation points by at most nk.

So either we are done because (87) holds, or $\mathbf{D}(f,h) \leq 9n^2k^6\alpha$; suppose the latter case. Our final step is to look at h as a Boolean function, and use a result on testing monotonicity [**?**].

Denote by $\tilde{\mathbf{D}}$ the distance of h when viewed as a Boolean function from the set of monotone Boolean functions. Let $0 < \varepsilon < 1$ be arbitrary. Then either $\tilde{\mathbf{D}} \leq \varepsilon$, in which case $\mathbf{D}(h, \mathrm{NONMANIP}) \leq \tilde{\mathbf{D}} \leq \varepsilon$ and therefore $\mathbf{D}(f, \mathrm{NONMANIP}) \leq 9n^2k^6\alpha + \varepsilon$, or $\tilde{\mathbf{D}} > \varepsilon$. In the latter case we show that then

$$\mathbb{P}\left(\sigma \in M_2(f)\right) \geq \frac{2\varepsilon}{nk} - 9n^3k^7\alpha. \tag{139}$$

Choosing $\varepsilon = 5n^4k^8\alpha$ then shows that either (86) or (87) holds.

Let us now show (139). Let us view h as a Boolean function, and denote by $p(h)$ the fraction of pairs of strings, differing on one coordinate, that violate the monotonicity condition. Goldreich, Goldwasser, Lehman, Ron, and Samorodnitsky showed in [**?**, Theorem 2] that $p(h) \geq \frac{\tilde{\mathbf{D}}}{n}$.

Now going back to viewing h as a SCF on k alternatives, this tells us that there are at least $\frac{\varepsilon}{2}2^n$ pairs of fibers, which differ on one coordinate, that violate monotonicity. For each such pair of fibers, whenever a and b are adjacent in the coordinate where the two fibers differ, we get a 2-manipulation point. Such a 2-manipulation point can be counted at most n times in this way (since there are n coordinates where a and b can be adjacent). Consequently, we have

$$|M_2(h)| \geq \frac{\varepsilon}{2} \cdot 2^n \cdot 2(k-1)! \left(\frac{k!}{2}\right)^{n-1} \cdot \frac{1}{n} = \frac{2\varepsilon}{nk}(k!)^n,$$

i.e.,

$$\mathbb{P}(\sigma \in M_2(h)) \geq \frac{2\varepsilon}{nk},$$

from which (139) follows immediately, since changing the outcome of a SCF at one ranking profile can change the number of 2-manipulation points by at most nk. □

Proof of Theorem 10.9. First we argue without specific bounds. Suppose on the contrary that our SCF f does not have many 4-manipulation points. Then f is close to $\overline{\text{NONMANIP}}$ by Theorem 16.1. Consequently, by Theorem 14.8, f is close to NONMANIP, which is a contradiction.

Now we argue with specific bounds. Assume on the contrary that

$$\mathbb{P}(\sigma \in M_4(f)) < \frac{\varepsilon^{15}}{10^{39}n^{67}k^{166}}.$$

Then by Theorem 16.1 we have that $\mathbf{D}\left(f, \overline{\text{NONMANIP}}\right) < \frac{\varepsilon^3}{10^6 n^{12} k^{24}}$, and consequently by Theorem 14.8 we have $\mathbf{D}(f, \text{NONMANIP}) < \varepsilon$, which is a contradiction. □

18 Aggregation power of Boolean Functions

We now take a modern view of Condorcet Jury Theorem. First, recall the setting. There are two alternatives denoted $+$ and $-$, which are apiori equally likely to be the preferable alternative, and that each voter independently receives the correct information with probability $p > 1/2$ and incorrect information with probability $1-p$. We now denote the n signals by $x_1, \ldots, x_n$ and aggregate them via a boolean function $f : \{-1,1\}^n \to \{-1,1\}$. The following fact is a generalization of the second part of the Jury Theorem.

Theorem 18.1. *Let $0.5 < p \leq 1$. Then $f : \{-1,1\}^n \to \{-1,1\}$ maximizes $P[f(x) = s]$ if and only if $f(x) = +$ for all x such that $\sum_{i=1}^n x_i > 0$ and $f(x) = -$ for all x such that $\sum x_i < 0$.*

In particular when n is odd Majority is the only function that maximizes $P[f(x) = s]$.

This is a special case of the Neyman-Pearson lemma for hypothesis testing in statistics. The proof is easy. In our special case it immediately follows from the fact that

$$\frac{P[s = +|x_1, \ldots, x_n]}{P[s = -|x_1, \ldots, x_n]} = \frac{P[x_1, \ldots, x_n|s = +]}{P[x_1, \ldots, x_n|s = -]} = \left(\frac{p}{1-p}\right)^{n_+(x) - n_-(x)}$$

where $n_+(x)$ is the number of $+$s in x and similarly n_-.

Assume for a moment that $s = +$. Therefore, our interest is in the quantity $\mathbb{P}[f = +]$. If f is a monotone function, we have the following Russo's formula, where $\mathbb{P}_p$ denotes the product measure with marginals $\mathbb{P}_p[x_i = +] = p$:

Proposition 18.2. *Let $f : \{-1,1\}^n \to \{-1,1\}$ be monotone then*

$$\left.\frac{d}{dp}\right|_{p=1/2} \mathbb{P}_p[f = 1] = \sum_{i=1}^n I_i(f).$$

One short proof of Proposition 18.2 is to consider an n variable function $P_{(p_1,\ldots,p_n)}[f = 1]$ and applying the chain rule.

Lemma 18.3. *In the setting of Condorcet voting with uniform prior on s, among all monotone functions majority maximizes*

$$\left.\frac{d}{dp}\right|_{p=1/2} \mathbb{P}_p[f = s] = \sum_{i=1}^n I_i(f)$$

Proof. Consider jurors who receive the correct signal with probability p. Then for any monotone function f we have that

$$\mathbb{P}[f = s] = 0.5(\mathbb{P}_p[f = +] + \mathbb{P}_{1-p}[f = -]) = 0.5 + 0.5(\mathbb{P}_p[f = +] - \mathbb{P}_{1-p}[f = +]).$$

We know that for every p, among all monotone functions, the expression above is maximized for the Majority function. Taking the derivative with respect to p concludes the proof. □

By examining the proof of Theorem 18.1 more carefully, we see that the Majorities are the only maximizers of the influence sum.

Interestingly, dictators are the minimizers in their aggregation power. In particular we have:

Proposition 18.4. *For any Boolean function $f : \{-1,1\}^n \to \{-1,1\}$ with $\mathbb{E}[f] = 0$ it holds that $\sum I_i(f) \geq 1$. Equality holds if and only if f is a dictator.*

Proof. This follows from the fact that

$$\sum_i I_i(f) = \sum_S |S|\hat{f}^2(S).$$

□

In fact more is true:

Proposition 18.5. *Among all $f : \{-1,1\}^n \to \{-1,1\}$ which are monotone and satisfy $\mathbb{E}_{1/2}[f] = 0$, dictators are the only minimizers of $\mathbb{E}_p[f]$ for all $0.5 < p < 1$.*

Proof. Left as an exercise to the reader. □

19 The minimal Influence Problem and Coalitions

Poincare's inequality:

$$\sum_{i=1}^{n} I_i(f) = \sum_S |S|\hat{f}^2(S) \geq \mathrm{Var}[f]$$

implies that any function $f : \{-1,1\}^n \to \mathbb{R}$ has a variable i with $I_i(f) \geq \mathrm{Var}[f]/n$. This in turn implies the following:

Lemma 19.1. *If $f : \{-1,1\}^n \to \{-1,1\}$ is a monotone function with $\mathbb{E}[f] = 0$, then for any $\varepsilon > 0$, there exists a set $|S| \leq \varepsilon n$ such that $\mathbb{E}[f|\forall i \in S, x_i = 1] \geq \varepsilon/4$.*

That is, even a small linear fraction of the voters can change the expected value of the function noticeably.

Proof. Construct S iteratively. Let $f_1 = f$ and let $S_0 = \emptyset$. At stage i of the construction, let j be the variable with most influence of the function f_i and let $S_i = S_{i-1} \cup \{j\}$. Let f_{i+1} be defined as f_i conditioned on $x_j = 1$. If $\mathbb{E}[f_{i+1}] \geq 0.5$ we stop. Otherwise, we continue. Note that at each step we have $\mathbb{E}[f_{i+1} - f_i] \geq I_j(f_i) \geq 0.25/n$. The proof follows with $\delta(\varepsilon) = \varepsilon/4$. □

It is natural to ask if the claim and proof are tight. In particular, it is natural to ask if one can improve the bound $I_i(f) \geq \mathrm{Var}[f]/n$. For general function the answer is no, as the function $f = n^{-1/2}\sum x_i$ shows. For Boolean functions there is an example that is almost tight i.e., the tribes function.

Example 19.2. Let $m = 2^r$ and let $n = m\log_2 m = mr$ and define the tribes function $f : \{0,1\}^n \to \{0,1\}$ by

$$f(x) = (x_1 \wedge \ldots \wedge x_r) \vee (x_{r+1} \wedge \ldots \wedge x_{2r}) \vee \cdots \vee (x_{n-r+1} \wedge \ldots \wedge x_n)$$

This example from [8] can be described as follows. The function is 1 if there is a decision to go to war and 0 otherwise. In order to go to war there has to be at least one tribe where all members decide to go to war, i.e, there has to exist a k so that $x_{kr+1} = \ldots = x_{(k+1)r} = 1$. It is easy to see that the number of tribes that go to war is $\mathrm{Bin}(m, 1/m)$ which is approximately Poisson. In particular,

$$\mathbb{E}[f] = \mathbb{P}[\mathrm{Bin}(m, 1/m) \geq 1] \in [0.1, 0.9],$$

for $r \geq 1$. For $\partial_1 f = 1$, i.e., that the first coordinate is influential, we need that none of the tribes but the first tribe wants to go to war and that in tribe 1, everyone but the first voter wants to go to war. Thus

$$I_i(f) = \mathbb{P}[\partial_1 f = 1] = 2^{-r+1}\mathbb{P}[\mathrm{Bin}(m-1, 1/m) = 0] \leq 2^{-r+1} = \frac{2}{m} \leq \frac{2\log_2(n)}{n}.$$

It is also easy to see from the equation above that in fact $I_i(f)$ is of order $\frac{\log_2(n)}{n}$.

Thus, we already know that the minimal influence is of order at least $\mathrm{Var}[f]/n$ and we have an example where it is of order $\mathrm{Var}[f]\log n/n$. The KKL Theorem proven in the next section shows that the $\log n$ factor is in fact necessary.

Theorem 19.3. *There exists a constant c such that for all $f : \{-1,1\}^n \to \{-1,1\}$ it holds that*

$$\min_i I_i(f) \geq c\frac{\log n}{n}\mathrm{Var}[f]$$

Repeating the proof of Lemma 19.1 we obtain the following:

Lemma 19.4. *Let $f : \{-1,1\}^n \to \{-1,1\}$ be a monotone function with $\mathbb{E}[f] = 0$, then for any $\varepsilon > 0$, there exists a set $|S| \leq c(\varepsilon)n/\log n$ such that $\mathbb{E}[f|\forall i \in S, x_i = 1] \geq 1 - \varepsilon$.*

Interestingly, the tribes example is far from tight for Lemma 19.4 as it suffices to take S of size $\log n$ (e.g., a tribe) to fix the value of the function to 1. There is a famous example by Ajtai and Linial where a coalition of size of order $n/\log^2 n$ is needed to shift the expected value of f [1] by a constant.

20 Semi-Groups, Talagrand and the KKL Theorem

In this section we will write $P_t = T_{e^{-t}}$ so that P_t is a *semi-group*, so $P_{s+t} = P_s P_t$.

Lemma 20.1.

$$P_t f = \sum_S \hat{f}(S)e^{-t|S|}x_S, \quad (\nabla_i f)(x) = \frac{1}{2}(f(x_{-i}, 1) - f(x_{-i}, -1)) = \sum_{S: i \in S}\hat{f}(S)x_{S\setminus\{i\}}$$

Note that

$$\nabla_i(P_t f) = \sum_{S:i\in S} \hat{f}(S) e^{-t|S|} x_{S\setminus\{i\}} = e^{-t} P_t \nabla f_i.$$

and that

$$\frac{d}{dt} P_t f = -\sum_S |S| \hat{f}(S) e^{-t|S|} x_S = -\sum_{i=1}^n x_i \nabla_i(P_t f) = -e^{-t} \sum_{i=1}^n x_i P_t \nabla_i f$$

so

$$\frac{d}{dt} \mathbb{E}[f P_t g] = -e^{-t} \sum_{i=1}^n \mathbb{E}[f x_i P_t \nabla_i g] = -e^{-t} \sum_{i=1}^n \mathbb{E}[\nabla_i f P_t \nabla_i g]$$

Claim 20.2. *If f and g are monotone functions then $\mathbb{E}[f P_t g]$ is monotonically decreasing in t.*

Proof. Assume f and g are increasing then the proof follows since $\nabla_i f$ and $\nabla_i g$ are always non-negative. □

Claim 20.3. *Let $f_1, \dots, f_k$ be positive and monotone increasing then $\mathbb{E}[P_t f_1 P_t f_2 \dots P_t f_k]$ is monotone decreasing.*

Proof. Write the derivative as a sum of partial derivatives. □

Remark 20.4. The proofs are so easy that the claims can be generalized much further. We need MC that respects and such the generator only moves to comparable elements.

We now turn to applications a la Cordero-Ersquin and Ledoux [16] of the derivative formula which can be thought of as an integration by parts formula.

Claim 20.5.

$$Cov(f,g) = \int_0^\infty e^{-t} \sum_{i=1}^n \mathbb{E}[\nabla_i f P_t \nabla_i g] dt$$

Proof.

$$Cov(f,g) = \mathbb{E}[f P_0 g] - \mathbb{E}[f P_\infty g] = -\int_0^\infty \frac{d}{dt} \mathbb{E}[f P_t g] dt = \int_0^\infty e^{-t} \sum_{i=1}^n \mathbb{E}[\nabla_i f P_t \nabla_i g] dt$$

as needed. □

We can now prove some useful corollaries.

Claim 20.6.

$$|Cov(f,g)| \leq \sum_{i=1}^n \sqrt{I_i(f) I_i(g)}$$

Proof. Note that

$$|\mathbb{E}[\nabla_i f P_t \nabla_i g]| \leq \|\nabla_i f\|_2 \|P_t \nabla_i g\|_2 \leq \|\nabla_i f\|_2 \|\nabla_i g\|_2 = \sqrt{I_i(f) I_i(g)}.$$

The proof follows by integration. □

Remark 20.7. The proof again generalizes assuming only the Cordero Eresquin setup and spectral gap.

If instead of the contractive estimate above we use the hyper-contractive inequality:

$$\|E[\nabla_i f P_t \nabla_i g]| \leq \|\nabla_i f\|_{1+e^{-t}} \|\nabla_i g\|_{1+e^{-t}}$$

We get the following bound:

$$|Cov(f,g)| \leq \int_0^\infty e^{-t} \sum_{i=1}^n \|\nabla_i f\|_{1+e^{-t}} \|\nabla_i g\|_{1+e^{-t}} dt = \int_1^2 \sum_{i=1}^n \|\nabla_i f\|_v \|\nabla_i g\|_v dv$$

Now By Holder inequality

$$\|f\|_v \leq \|f\|_2^{1-\theta(v)} \|f\|_1^{\theta(v)} = \|f\|_2 \left(\frac{\|f\|_1}{\|f\|_2} \right)^{\theta(v)},$$

where $v\theta + 0.5(1-\theta)v = 1$ so $0.5\theta v = 1 - 0.5v$ so $\theta = 2/v - 1$. Writing

$$r_i = \frac{\|\nabla_i f\|_1 \|\nabla_i g\|_1}{\|\nabla_i f\|_2 \|\nabla_i g\|_2},$$

we see that

$$|Cov(f,g)| \leq \sum_{i=1}^n \|\nabla_i f\|_2 \|\nabla_i g\|_2 \int_1^2 r_i^{\theta(v)} dv$$

Note that $\theta(v) = 2/v - 1 \geq 1 - v/2$ so

$$\int_1^2 r_i^{\theta(v)} dv \leq \int_1^2 r_i^{1-v/2} dv = 2 \int_0^{1/2} r_i^v dv = \frac{1}{\ln r_i}(r_i - 1) \leq \frac{5}{1 + \ln(1/r_i)}$$

We thus obtain

Claim 20.8.

$$|Cov(f,g)| \leq 5 \sum_{i=1}^n \frac{\sqrt{I_i(f) I_i(g)}}{1 + \ln\left(\|\nabla_i f\|_2 \|\nabla_i g\|_2 / (\|\nabla_i f\|_1 \nabla\|_i g\|_2)\right)}$$

Remark 20.9. This is of course just a slight generalization of the proof by [16] of Talagrands's result [**?**].

$$Var[f] \leq 5 \sum_{i=1}^{n} \frac{I_i(f)}{1 + \ln\left(\|\nabla_i f\|_2 / (\|\nabla_i f\|_1)\right)}$$

We claim that furthermore the last inequality implies the KKL Theorem that for Boolean functions $f : \{-1,1\}^n \to \{0,1\}$ there exists an i such that $I_i(f) \geq cVar[f]\log n/n$. This is true since for Boolean functions $\nabla_i f \in \{1, 0, -1\}$ and therefore

$$\|\nabla_i f\|_2 = \sqrt{I_i(f)}, \quad \|\nabla_i f\|_1 = I_i(f).$$

So for such function we get that

$$Var[f] \leq 10 \sum_{i=1}^{n} \frac{I_i(f)}{1 - \ln I_i(f)}$$

So the there exist an i such that

$$\frac{I_i(f)}{1 - \ln I_i(f)} \geq cVar[f]/n$$

and therefore $I_i(f) \geq cVar[f]\log n/n$.

In the case where f and g are monotone in the same direction and moreover f and g are Boolean we can apply reverse-hyper-contraction instead to

$$E[\nabla_i f P_t \nabla_i g] \geq \|f_i\|_{1-e^{-t}} \|g_i\|_{1-e^{-t}} = (I_i(f)I_i(g))^{1/(1-e^{-t})}$$

A careful analysis of the integral lets us conclude that

$$Cov(f,g) \geq c \sum_{i=1}^{n} \frac{I_i(f)I_i(g)}{\sqrt{1 + 1/\log I_i(f)}\sqrt{1 + 1/\log I_i(g)}}$$

This is following the proof in Keller-M-Sen in the Gaussian case.

Remark 20.10. This again works in the general setup of C.E + the right kind of monotonicity.

References

[1] M. Ajtai and N. Linial. The influence of large coalitions. *Combinatorica*, 13(2):129–145, 1993.

[2] K. Arrow. A difficulty in the theory of social welfare. *J. of Political Economy*, 58:328–346, 1950.

[3] K. Arrow. *Social choice and individual values.* John Wiley and Sons, 1963.

[4] S. Barbera. Pivotal voters: A new proof of arrow's theorem. *Economics Letter*, 6:13–16, 1980.

[5] J. J. Bartholdi, C. A. Tovey, and M. A. Trick. The computational difficulty of manipulating an election. *Social Choice and Welfare*, 6(3):227–241, 1989.

[6] W. Beckner. Inequalities in Fourier analysis. *Ann. of Math. (2)*, 102(1):159–182, 1975.

[7] C. E. Bell. A random voting graph almost surely has a hamiltonian cycle when the number of alternatives is large. *Econometrica*, 49(6):1597–1603, 1981.

[8] M. Ben-Or and N. Linial. Collective coin flipping. In S. Micali, editor, *Randomness and Computation.* Academic Press, New York, 1990.

[9] I. Benjamini, G. Kalai, and O. Schramm. Noise sensitivity of boolean functions and applications to percolation. *Inst. Hautes Études Sci. Publ. Math.*, 90:5–43, 1999.

[10] A. Bonami. Étude des coefficients de Fourier des fonctions de $L^p(G)$. *Ann. Inst. Fourier (Grenoble)*, 20(fasc. 2):335–402 (1971), 1970.

[11] C. Borell. Positivity improving operators and hypercontractivity. *Math. Zeitschrift*, 180(2):225–234, 1982.

[12] C. Borell. Geometric bounds on the Ornstein-Uhlenbeck velocity process. *Z. Wahrsch. Verw. Gebiete*, 70(1):1–13, 1985.

[13] B. Chor, O. Goldreich, J. Hasted, J. Freidmann, S. Rudich, and R. Smolensky. The bit extraction problem or t-resilient functions. In *— 26th Annual Symposium on Foundations of Computer Science*, pages 396–407. IEEE, 1985.

[14] J.-A.-N. Condorcet. *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix.* De l'Imprimerie Royale, 1785.

[15] V. Conitzer and T. Sandholm. Nonexistence of Voting Rules That Are Usually Hard to Manipulate. In *Proceedings of the 21st National Conference on Artificial Intelligence*, volume 21, pages 627–634, 2006.

[16] D. Cordero-Erausquin and M. Ledoux. Hypercontractive measures, talagrand's inequality, and influences. In *Geometric Aspects of Functional Analysis*, volume 2050 of *Lecture Notes in Mathematics*. 2012.

[17] A. De, E. Mossel, and J. Neeman. Majority is stablest : Discrete and sos. In *STOC (Symposium on Theory of Computing)*, pages 477–486, 2013.

[18] A. De, E. Mossel, and J. Neeman. Majority is stablest: Discrete and sos. *Theory of Computing*, 12(4):1–50, 2016.

[19] I. Diakonikolas, D. M. Kane, and J. Nelson. Bounded independence fools degree-2 threshold functions. In *Foundations of Computer Science (FOCS), 2010 51st Annual IEEE Symposium on*, pages 11–20. IEEE, 2010.

[20] R. Eldan. A two-sided estimate for the gaussian noise stability deficit. *Inventiones mathematicae*, 201(2):561–624, 2015.

[21] P. Faliszewski and A. Procaccia. AI's War on Manipulation: Are We Winning? *AI Magazine*, 31(4):53–64, 2010.

[22] W. Feller. *An introduction to probability theory and its applications. Vol. I.* Third edition. John Wiley & Sons Inc., New York, 1968.

[23] E. Friedgut. Boolean functions with low average sensitivity depend on few coordinates. *Combinatorica*, 18(1):474–483, 1998.

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

[25] E. Friedgut, G. Kalai, and A. Naor. Boolean functions whose fourier transform is concentrated on the first two levels. *Advances in Applied Mathematics*, 29(3):427–437, 2002.

[26] E. Friedgut, G. Kalai, and N. Nisan. Elections can be manipulated often. In *Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 243–249, 2009.

[27] A. Gibbard. Manipulation of voting schemes: a general result. *Econometrica*, 41(4):587–601, 1973.

[28] L. Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.*, 97(4):1061–1083, 1975.

[29] O. Häggström, G. Kalai, and E. Mossel. A law of large numbers for weighted majority. *Advances in Applied Mathematics*, 37(1):112–123, 2006.

[30] M. Isaksson, G. Kindler, and E. Mossel. The geometry of manipulation - a quantitative proof of the gibbard satterthwaite theorem. In *Foundations of Computer Science (FOCS)*, pages 319–328, 2010.

[31] 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.

[32] M. Isaksson and E. Mossel. New maximally stable Gaussian partitions with discrete applications. *Israel Journal of Mathematics*, 189:347–396, 2012.

[33] J. Jendrej, K. Oleszkiewicz, and J. O. Wojtaszczyk. On some extensions of the fkn theorem. *Theory of Computing*, 11(1):445–469, 2015.

[34] C. Jones. A noisy-influence regularity lemma for boolean functions, 2016. arXiv preprint arXiv:1610.06950.

[35] J. Kahn, G. Kalai, and N. Linial. The influence of variables on boolean functions. In *Proceedings of the 29th Annual Symposium on Foundations of Computer Science*, pages 68–80, 1988.

[36] G. Kalai. A Fourier-theoretic perspective on the Condorcet paradox and Arrow's theorem. *Adv. in Appl. Math.*, 29(3):412–426, 2002.

[37] G. Kalai. Social Indeterminacy. *Econometrica*, 72:1565–1581, 2004.

[38] N. Keller. A tight quantitative version of arrow?s impossibility theorem. *Journal of the European Mathematical Society*, 14(5):1331–1355, 2012.

[39] J. Kelly. Almost all social choice rules are highly manipulable, but a few aren't. *Social Choice and Welfare*, 10(2):161–175, 1993.

[40] G. Kindler, N. Kirshner, and R. O'Donnell. Gaussian noise sensitivity and fourier tails. *Israel Journal of Mathematics*, 225(1):71–109, 2018.

[41] J. Lindsey. Assignment of Numbers to Vertices. *American Mathematical Monthly*, pages 508–516, 1964.

[42] E. Mossel. Gaussian bounds for noise correlation of functions and tight analysis of long codes. In *Foundations of Computer Science, 2008 (FOCS 08)*, pages 156–165. IEEE, 2008.

[43] E. Mossel. Arrow's impossibility theorem without unanimity. Posted on Arxiv 0901.4727, 2009.

[44] E. Mossel. Gaussian bounds for noise correlation of functions. *GAFA*, 19:1713–1756, 2010.

[45] E. Mossel. A quantitative arrow theorem. *Probability Theory and Related Fields*, 154(1):49–88, 2012.

[46] E. Mossel. Gaussian Bounds for Noise Correlation of Resilient Functions. ArXiv e-prints 1704.04745, 2019.

[47] E. Mossel and J. Neeman. Robust optimality of Gaussian noise stability. *J. Eur. Math. Soc. (JEMS)*, 17(2):433–482, 2015.

[48] E. Mossel and R. O'Donnell. On the noise sensitivity of monotone functions. *Random Structures Algorithms*, 23(3):333–350, 2003.

[49] E. Mossel, R. O'Donnell, and K. Oleszkiewicz. Noise stability of functions with low influences: invariance and optimality (extended abstract). In *46th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2005), 23-25 October 2005, Pittsburgh, PA, USA, Proceedings*, pages 21–30. IEEE Computer Society, 2005.

[50] E. Mossel, R. O'Donnell, and K. Oleszkiewicz. Noise stability of functions with low influences: invariance and optimality. *Annals of Mathematics*, 171(1):295–341, 2010.

[51] E. Mossel, R. O'Donnell, O. Regev, J. E. Steif, and B. Sudakov. Non-interactive correlation distillation, inhomogeneous Markov chains, and the reverse Bonami-Beckner inequality. *Israel J. Math.*, 154:299–336, 2006.

[52] E. Mossel, K. Oleszkiewicz, and A. Sen. On reverse hypercontractivity. *Geometric and Functional Analysis*, 23(3):1062–1097, 2013.

[53] E. Mossel and M. Z. Racz. A quantitative gibbard-satterthwaite theorem without neutrality. In H. J. Karloff and T. Pitassi, editors, *STOC. Proceedings of the 44th Symposium on Theory of Computing Conference, STOC 2012, New York, NY, USA, May 19 - 22, 2012*, pages 1041–1060. ACM, 2012.

[54] E. Mossel and M. Z. Racz. A quantitative gibbard-satterthwaite theorem without neutrality. *Combinatorica*, 35(3):317–387, 2015.

[55] E. Mossel and O. Schramm. Representations of general functions using smooth functions, 2008. Unpublished manuscript.

[56] I. Nehama. Approximately classic judgement aggregation. *Annals of Mathematics and Artificial Intelligence*, 68(1-3):91–134, 2013.

[57] E. Nelson. The free Markoff field. *J. Functional Analysis*, 12:211–227, 1973.

[58] R. G. Niemi and H. F. Weisberg. A mathematical solution for the probability of paradox of voting. *Behavioral Science*, 13:317–323, 1968.

[59] R. O'Donnell. *Analysis of boolean functions*. Cambridge University Press, 2014.

[60] R. O'Donnell, R. Servedio, L.-Y. Tan, and A. Wan. A regularity lemma for low noisy-influences, 2010. Unpublished manuscript.

[61] A. Procaccia and J. Rosenschein. Junta Distributions and the Average-case Complexity of Manipulating Elections. *Journal of Artificial Intelligence Research*, 28:157–181, 2007.

[62] M. A. Satterthwaite. Strategy-proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions. *J. of Economic Theory*, 10:187–217, 1975.

[63] W. Sheppard. On the application of the theory of error to cases of normal distribution and normal correlation. *Phil. Trans. Royal Soc. London*, 192:101–168, 1899.

[64] Wikipedia contributors. Marquis de condorcet— Wikipedia, the free encyclopedia, 2019.

[65] R. Wilson. Social choice theory without the pareto principle. *Journal of Economic Theory*, 5(3):478–486, 1972.