An algebraic proof of a robust social choice impossibility theorem

Dvir Falik
Institute of Computer Science
Hebrew University
Jerusalem, Israel
dvir.falik@mail.huji.ac.il

Ehud Friedgut
Institute of Mathematics
Hebrew University
Jerusalem, Israel
ehudf@math.huji.ac.il

***Abstract*—An important element of social choice theory are impossibility theorems, such as Arrow's theorem [1] and Gibbard-Satterthwaite's theorem [2], [3], which state that under certain natural constraints, social choice mechanisms are impossible to construct. In recent years, beginning in Kalai [4], much work has been done in finding *robust* versions of these theorems, showing that impossibility remains even when the constraints are *almost* always satisfied. In this work we present an Algebraic scheme for producing such results. We demonstrate it for a variant of Arrow's theorem, found in Dokow and Holzman [5].**

***Keywords*-Social Choice; Arrow's theorem; Robust impossibility theorems; Discrete Fourier analysis; Representation theory;**

I. Introduction

Social choice deals with aggregation of opinions of individuals in a society into a single opinion. There are several important *impossibilty* theorems in the field, stating that aggregation mechanisms satisfying some natural conditions, are dictatorial (dependent on the opinion of only one voter).

First amongst these theorems was Arrow's theorem. Let there be a set of n individuals, who wish to decide on the ranking of m alternatives. Each individual has his own full ranking of the alternatives. Let L_m be the set of full transitive linear orders on $[m]$ and O_m be the set of all anti symmetric relations on $[m]$. A social welfare function (SWF) is a function $f : L_m^n \to O_m$, that maps the individual rankings of the n voters into an aggregated relation. A SWF that always returns a transitive order (is into L_m) is called consistent.

Definition 1.1: A SWF f is called Independent of Irrelvant Alternatives (IIA), iff for every 2 alternatives a, b, the aggregated prefernece between a and b is only dependent on the individual prefernces between a and b.

Theorem 1.2: (Arrow) For $m \geq 3$, every function that agrees with unanimous votes, is consistent and IIA is dictatorial.

Another theorem of similar flavour is Gibbard-Satterthwaite's theorem (GS), known to be strongly connected to Arrow's theorem. It deals with a setting in which, the voters only wish to choose one of the m alternatives. A social choice function (SCF) is a function $f : L_m^n \to [m]$, that maps the individual rankings of n voters into an aggregated choice. GS deals with the game-theoretic notion of strategy proofness, where no voter has an incentive to misreport her true opinion and obtain a better result from her perspective.

For the formal definition of strategy-proofness, we introduce some notations. For a profile $x \in L_m^n, x = (x_1, ..., x_n)$ and a voter $i \in [n]$, we denote $x = (x^{-i}, x_i)$, where x^{-i} indicates the votes of all voters except the i'th. For $y \in L_m$, we use $<_y$ to indicate the corresponding order. We similarly define $>_y, \geq_y, \leq_y$.

Definition 1.3: A SCF f is called strategy-proof, iff

$$\forall i \in [n], x^{-i} \in L_m^{n-1}, x_i, y \in L_m, f(x^{-i}, x_i) \geq_{x_i} f(x^{-i}, y)$$

Theorem 1.4: (Gibbard-Satterthwaite) for $m \geq 3$, a social aggregator $f : L_m^n \to [m]$ that is onto and strategy-proof is dictatorial.

The connection between the notions of strategy proofness and IIA was demonstrated in [6], [7], and the connection between the proofs of these theorems was demonstrated in, e.g., [8]. However a single scheme that deals with both the different settings (SWF vs. SCF) and the different constraints (IIA vs. strategy-proofness) has not been shown.

In recent years, there was work done in providing analytical proofs for these theorems, and finding *robust* versions of them - i.e. showing that aggregators that almost satisfy the constraints (consistency, IIA, strategy-proofness) are close to fitting the classification (dictatorial).

The relaxation of the consistency constraint in Arrow's theorem was initiated in Kalai[4] and culminated in Mossel[9]. See also [10], [11], [12]. The same was done for several examples in the judgement aggregation setting in [13]. All of these works assumed complete IIA and relaxed consistency, despite the fact that consistency is a more natural constraint than IIA. However, a tradeoff of transforming almost-IIA and consistent functions to IIA and almost consistent functions has been established in, e.g., [13].

Relaxing the strategy proofness constraint in GS has some important computational implications. A robust Gibbard-Satterthwaite result with appropriate parameters, will show that a random manipulation attempt on a far from dictatorial

SCF has a non-negligible probability of being beneficial. Therefore, we would be able to deduce that the problem of finding a beneficial manipulation is average-case easy, despite the fact that it is NP hard for many natural SCF's. In [14] such a result was achieved, but only for functions with $m = 3$, and in [15] only for neutral functions with $m > 3$. See also [16].

Our work continues this line of research. We present an algebraic approach to these types of questions that can tackle many of them. In this work we will apply this scheme on a variant of Arrow's theorem. Our work relaxes an independency constraint similar to IIA and is the first work to tackle an independency type of constraint directly.

The techniques are novel. The algebraization of the problem uses tensor algebra. The diagonalization involved uses representation theory of the Symmetric Group. These techniques can be generalized to a wide class of problems, not limited to Social Choice. Specifically, they are a generalization of techniques used, up to now, mainly for Boolean functions.

II. Results

In this paper we present a *robust* impossibility theorem in the flavour of Arrow's and Gibbard-Satterthwaite's theorems. We analyze a constraint that is a variant of IIA. We show the impossibility of constructing functions that are far from dictatorial and almost always conform to this constraint. We present a single proof that deals with functions with a spectrum of ranges, from functions returning a full ranking (SWFs) to functions returning one alternative (SCFs), and a plethora of ranges in between. The result can be interpreted as *a 2-query dictatorship tester with full completeness.*

In our settings, we deal with aggregation of rankings of m alternatives. A ranking is a permutation $x \in \mathbb{S}_m$. We will use the convention $x(rank) = name$.

For presentation sake, we shall begin with the definition of the constraint when used for functions returning a full ranking (SWFs) and state the corresponding impossibility theorem without robustness. The more complicated definitions and theorems will follow.

A. Main Theorem

Definition 2.1: A social aggregator $f : \mathbb{S}_m^n \to \mathbb{S}_m$ satisfies Independence of Rankings (IR) iff the aggregated ranking of the j'th alternative is dependent only on the individual rankings of the j'th alternative.

$$\forall x, y \in \mathbb{S}_m^n, j \in [m],$$

$$\left(\forall i \in [n], x_i^{-1}(j) = y_i^{-1}(j)\right) \Rightarrow f(x)^{-1}(j) = f(y)^{-1}(i)$$

This constraint requires independency of rankings instead of independency of pairwise preferences required in IIA. This constraint was discussed in [5], in the context of non-binary judgement aggregation.

As in IIA, this definition compares voting profiles which may differ in any number of votes. Throughout this paper, we shall use an alternative, equivalent definition, that compares inputs that differ only in at most one vote, similarly to strategy-proofness:

Definition 2.2: A social aggregator $f : \mathbb{S}_m^n \to \mathbb{S}_m$ satisfies Independence of Rankings (IR) iff

$$\forall i \in [n], j \in [m], x^{-i} \in \mathbb{S}_m^{n-1}, x_i, y_i \in \mathbb{S}_m,$$

$$x_i^{-1}(j) = y_i^{-1}(j) \Rightarrow f(x^{-i}, x_i)^{-1}(j) = f(x^{-i}, y_i)^{-1}(j)$$

It is easy to show, via a hybrid argument, that these two definitions are equivalent, i.e., a function conforms to definition 2.1 iff it conforms to 2.2. The corresponding impossibility theorem is

Theorem 2.3: For $m \geq 3$, a social aggregator $f : \mathbb{S}_m^n \to \mathbb{S}_m$ that is IR is either a constant function or dictatorial of the following form: there exists a voter i and a constant permutation y of the rankings such that $f(x) = y \circ x_i$.

In [5], a similar impossibility theorem was shown, using purely combinatorial arguments. There are several differences between theorem 2.3 and the result in [5]. That result required additional constraints on the function. Those constraints also limited the constant permutation y from being anything but the identity permutation. It is important to note, however, that [5] dealt with a much wider setting.

B. Robust Impossibility Theorem

A robust impossibility theorem means that when the constraint is *almost* satisfied, then the function is *almost* dictatorial. A theorem relaxing the independency constraint (IR or IIA) in the context of SWF has not been obtained before.

To ease notation, we define the predicate

$$Q(x, i, j, y_i, f) =$$

$$\neg\left(x_i^{-1}(j) = y_i^{-1}(j) \Rightarrow f(x^{-i}, x_i)^{-1}(j) = f(x^{-i}, y_i)^{-1}(j)\right)$$

Where $x \in \mathbb{S}_m^n, i \in [n], j \in [m], y_i \in \mathbb{S}_m,\ f : \mathbb{S}_m^n \to \mathbb{S}_m$. We also define the quantity

$$IR(f) = \sum_{i \in [n], j \in [m]} Pr_{x^{-i} \in \mathbb{S}_m^{n-1}, x_i, y_i \in \mathbb{S}_m}\left(Q(x, i, j, y_i, f)\right)$$

$IR(f)$ measure the normalized number of unsatisfied IR constraints.

Definition 2.4: A social aggregator $f : \mathbb{S}_m^n \to \mathbb{S}_m$ is called $\epsilon - IR$ iff

$$IR(f) \leq \epsilon$$

Theorem 2.5: For $m \geq 3$, a social aggregator $f : \mathbb{S}_m^n \to \mathbb{S}_m$ that is ϵ-IR is $O(poly(m)\epsilon))$ close to a function that is either a constant function or dictatorial of the following form: there exists a voter i and a constant permutation y of the rankings such that $f(x) = y \circ x_i$.

C. A Spectrum of Ranges

As stated earlier, we will also deal with a setting where the aggregated opinion is not a full ranking, but a partial ranking. Let $H \subseteq \mathbb{S}_m$ be a subgroup of $\mathbb{S}_m$. We call it a *fixing* subgroup if it fixes some subset of m. An H-social aggregator (H-SA) is a function $f : \mathbb{S}_m^n \to \mathbb{S}_m/H$. Many types of functions fall under this scheme. Examples are:

- For H as the trivial group, H-social aggregators are SWFs.
- For H as the group of permutations fixing the element 1, H-social aggregators are SCFs.
- For H as the group of permutations fixing the set $\{1, 2, 3\}$, H-social aggregators are functions returning triumvirates.
- For H as the group of permutations fixing the sets $\{1\}$ and $\{2, 3\}$, H-social aggregators are functions returning a president and two vice-presidents.

The definition of IR can be extended to H-social aggregators in the following manner:

Definition 2.6: Let $H \subseteq \mathbb{S}_m$ be a fixing subgroup of $\mathbb{S}_m$. For H_1 a coset of H in $\mathbb{S}_m$, define the j-profile of H_1 as the multiset $H_1^{-1}(j) = \left\{y^{-1}(j) | y \in H_1\right\}$.

The definition of the predicate Q is naturally extended H-SA's, by taking definition 2.6 into account:

$$Q(x, i, j, y_i, f) =$$

$$\neg \left(x_i^{-1}(j) = y_i^{-1}(j) \Rightarrow f(x^{-i}, x_i)^{-1}(j) = f(x^{-i}, y_i)^{-1}(j)\right)$$

Definition 2.7: Let $H \subseteq \mathbb{S}_m$ be a fixing subgroup of $\mathbb{S}_m$. An H-social aggregator f satisfies Independence of Rankings (IR) iff the aggregated j-profile is dependent only on the individual rankings of the j'th alternative.

$$\forall i \in [n], j \in [m], x^{-i} \in \mathbb{S}_m^{n-1}, x_i, y_i \in \mathbb{S}_m,$$

$$\neg Q(x, i, j, y_i, f)$$

We shall leave the exact definitions of IR(f) and of an ϵ-IR H-social aggregator to a later part of the paper. The impossibility theorems also extend to H-social aggregators.

Theorem 2.8: Let $H \subseteq \mathbb{S}_m$ be a fixing subgroup of $\mathbb{S}_m$. For $m \geq 3$, an H-social aggregator f that is IR is either a constant function or dictatorial of the following form: there exists a voter i and a constant permutation y of the rankings such that $f(x) = Hy \circ x_i$.

Theorem 2.9: Let $H \subseteq \mathbb{S}_m$ be a fixing subgroup of $\mathbb{S}_m$. For $m \geq 3$, an H-social aggregator f that is ϵ-IR is $O_H(poly(m)\epsilon)$ close to a function that is either a constant function or dictatorial of the following form: there exists a voter i and a constant permutation y of the rankings such that $f(x) = y \circ x_i$.

III. Structure of the Proof

We give here a short exposition of the proof. For simplicity, we shall refer here to the basic form of the Main theorem (theorem 2.3), where the function is a SWF. To simplify the notation, we shall also use in this section definition 2.1 for IR, instead of 2.2, which is the definition we shall use in the rest of the paper.

We shall treat this problem as a contraint satisfaction problem (CSP). We shall use definition

$$\texttt{Find all functions } f : \mathbb{S}_m^n \to \mathbb{S}_m \texttt{ s.t.}$$
$$\texttt{IR: } \forall j \in [m], x, y \in \mathbb{S}_m^n,$$
$$x^{-1}(j) = y^{-1}(j) \Rightarrow (f(x))^{-1}(j) = (f(y))^{-1}(j)$$

A CSP has a generic algebraic encoding. The function f can be encoded as a function returning a vector in $\mathbb{R}^{\mathbb{S}_m}$, which is the characteristic vector of the singleton $\{f(x)\}$. This encoding can be interpreted as a tensor $F \in R^{\mathbb{S}_m^n \times \mathbb{S}_m}$, with 2 indices $x \in \mathbb{S}_m^n, v \in \mathbb{S}_m$

$$F_{x,v} = \mathbf{1}_{v=f(x)}.$$

The constraints can be algebraically encoded using a matrix that represents their truth table, or, in our case, since we want to count the number of constraints unsatisfied, the truth table of their negation. We use a block matrix. For every alternative $j \in [m]$, we define a constraint matrix L^j. For every two inputs $x, y \in \mathbb{S}_m^n$, the (x, y)'th entry of the matrix will be a matrix in $\mathbb{R}^{\mathbb{S}_m \times \mathbb{S}_m}$. This matrix will be the truthtable of the negation of the constraints concerning x, y and j. Explicitly,

$$\left(\left(\left(L^j\right)_{xy}\right)_{v_x v_y}\right) = \mathbf{1}_{\neg\left(x^{-1}(j)=y^{-1}(j) \Rightarrow (v_x^{-1}(j)=v_y^{-1}(j))\right)}$$

where $j \in [m], x, y \in \mathbb{S}_m^n$ and $v_x, v_y \in \mathbb{S}_m$.

We can use these block matrices in a quadratic form to count the number of unsatisfied constraints. Define $L = \sum_j L^j$. Since the L^j's are the truth tables of the negation of the constraints, it follows that the quadratic form

$$FLF^t$$

counts the number of constraints unsatisfied.

The CSP under this encoding takes the form

$$\texttt{Find all } F \in \mathbb{R}^{\mathbb{S}_m^n \times \mathbb{S}_m} \texttt{ s.t.}$$
$$\texttt{Consistency: } \forall x \in \mathbb{S}_m^n,$$
$$F_{\mathbf{x}*} \texttt{ is a char. vec. of a singleton}$$
$$\texttt{IR: } FLF^t = 0$$

The proof unfolds as follows:

- We show that $L \succeq 0$ (PSD), i.e. $FLF^t \geq 0$ for every F. Therefore, the functions that satisfy IR are precisely the kernel of L.
- Explicitly find the kernel of L, by diagonalizing it.

- Show that all consistent functions in the kernel of L are dictatorships.

Naturally, functions that are ϵ-IR are L_2 close to the kernel of L, depending on the spectral gap of L. We shall generalize a theorem of Friedgut, Kalai and Naor [17] to prove that such functions, that are also consistent, are L_2 close to dictatorships, to obtain the robustness result.

For H-SA, we shall encode f to return characteristic vectors of cosets of H, normalized so that their L_1 norm equals 1. We shall call such vectors H coset vectors. As a tensor F, this encoding takes the form:

$$F_{x,v} = \frac{1}{|H|}\mathbf{1}_{v\in f(x)}$$

The same Laplacian L encodes the notions of IR and ϵ-IR for H-SA, but the consistency constraint changes. The algebraic CSP for H-SAs is

`Find all` $F \in \mathbb{R}^{\mathbb{S}_m^n \times \mathbb{S}_m}$ `s.t.`

`Consistency:` $\forall x \in \mathbb{S}_m^n, F_{\mathbf{x}*}$ `is an` H`-coset vector`

`IR:` $\sum_j FL^jF = 0$

The introduction of H-SAs does not insert any new elements to the proof.

In subsection V-A, we shall construct and diagonalize L for functions with 1 voter. We shall then use this construction and its disgonalization to construct a Laplacian for any number of voters n, in subsection V-B, and prove the result.

IV. Representation Theory

A representation is a Homomorphism ρ from a group G to $GL_d(\mathbb{C})$, the group of complex d-dimensional square matrices. d is called the dimension of the representation $d(\rho)$. A representation is called irreducible if it is not similar to a direct sum of 2 representations.

For a finite group, there is a one-to-one correspondence between the conjugacy classes of the group and irreducible representations (up to similarity). The conjugacy classes of $\mathbb{S}_m$ have a certain natural ordering which we will not discuss here. We shall denote the number of conjugacy classes of $\mathbb{S}_m$ as $[\mathbb{S}_m]$. For a conjugacy class $k \in [[\mathbb{S}_m]]$, its corresponding irreducible representation will be denoted as ρ^k.

It is known that in the symmetric group, we can choose a basis under which all representations have real, unitary matrices as values. We assume we use such a basis.

The defining representation of the symmetric group $\mathbb{S}_m$ is the permutation representation P of dimension m.

$$P(x)_{ij} = \mathbf{1}_{x(i)=j}$$

In this work we also use two irrreducible representation corresponding to the first 2 conjugacy classes, in their natural order. These are the trivial representation ρ^0, which is of dimension 1 - $\rho^0(x) = 1$, and ρ^1 which is of dimension $m-1$.

It is known that P is similar to $\rho^0 \oplus \rho^1$, i.e., there exists an orthonormal $m\times m$ matrix H such that

$$P(x) = H\left(\rho^0(x)\oplus\rho^1(x)\right)H^t$$

Schur's orthonormality states that the vectors of the form $\rho^{\mathbf{k}}_{ij}$ are orthogonal, (but not necessarily orthonormal)

$$\sum_x \rho^{\mathbf{k_1}}(x)_{i_1j_1}\rho^{\mathbf{k_2}}(x)_{i_2j_2} = \delta_{\mathbf{k_1k_2}}\delta_{\mathbf{i_1i_2}}\delta_{\mathbf{j_1j_2}}\frac{m!}{d(\rho^{\mathbf{k_1}})}$$

It is easy to show that the first column of H is the normalized all ones vector. This is due to the fact that the all ones vector is an eigenvector of all permutation matrices. Therefore, H is of the following form (where C is a $m\times(m-1)$ matrix, and the C_i's are its rows)

$$H = \begin{pmatrix} \frac{1}{\sqrt{m}} & \\ \vdots & C \\ \frac{1}{\sqrt{m}} & \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{m}} & C_1 \\ \vdots & \vdots \\ \frac{1}{\sqrt{m}} & C_m \end{pmatrix} \quad (1)$$

V. The Proof

In this section we present a more detailed version of the proof, divided into lemmas. The actual proofs of the lemmas will appear in section VI.

A note on notation: The constructions used in the proof use block matrices, which are essentialy tensors. There are dedicated notation systems for use with tensor operations. Such are Einstein's convention, and Penrose's graphical notation. Each method has its own benefits and drawbacks. Since these dedicated notations are seldom used in the CS community, we choset to present the proof using the classical matrix notation, despite it being less suitable for our needs. This choice made a few of the statements more cumbersome. Whenever the phrase "matrix parsed as a vector" or the trace operator appear in the text, they are merely artifacts of the notation.

A. One Voter Functions

We begin with the analysis of social welfare functions on one voter $f:\mathbb{S}_m\to\mathbb{S}_m$. We construct in this section a constraint quadratic form as mentioned in section III, and diagonalize it.

Let X^j be the matrix $X\in\mathbb{R}^{\mathbb{S}_m\times\mathbb{S}_m}$,

$$X^j_{xy} = \mathbf{1}_{x^{-1}(j)=y^{-1}(j)} = \mathbf{1}_{x^{-1}y(j)=j}.$$

Let $\bar{X}^j$ be its complement $\bar{X}^j{}_{xy} = 1 - X^j_{xy}$. We will use the vector encoding described in section III, $F_{x,v} = \mathbf{1}_{v=f(x)}$ (or the corresponding definition for H-social aggregators).

$IR(f)$ is equivalent to a quadratic form based on the truthtable of the constraints, given in the following lemma:

Lemma 5.1: Let $L'^j = X^j \otimes \bar{X}^j$, and $L' = \sum_j L'^j$, then

$$IR(f) = \frac{1}{|\mathbb{S}_m|^2}FL'F^t$$

X^j is the adjacency matrix of the Caley-like graph[1] $\Gamma(\mathbb{S}_m, \mathbb{S}^j_m)$, where $\mathbb{S}^j_m$ is the subgroup of S_m of permutations fixing j. A quadratic form based on the Laplacian of that same graph is more suitable for our purposes, because it is PSD. The Laplacian of that graph, Y, is given by

$$Y^j = (m-1)!I - X^j$$

The corresponding quadratic form is given in the following lemma. The quadratic forms given in lemmas 5.1 and 5.2 are equivalent when F represents a consistent function. This quadratic form also works for H-social aggregators.

Lemma 5.2: Let $L''^j = Y^j \otimes X^j$, , and $L'' = \sum_j L''^j$, then

$$IR(f) = \frac{1}{|\mathbb{S}_m|^2} F L'' F^t$$

For H-SA, $IR(f)$ is defined via the quadratic form defined in lemma 5.2:

$$IR(f) = \sum_{j \in [m]} \mathbb{E}_{x,y \in \mathbb{S}_m} \mathbf{1}_{x^{-1}y(j)=(j)} \mathbb{E}_{h_1,h_2 \in H}$$

$$\left(\mathbf{1}_{(f(h_1x))^{-1}(f(h_2x))(j)=j} - \mathbf{1}_{(f(h_1x))^{-1}(f(h_2y))(j)=j} \right).$$

We still need to show that the definition of IR H-SAs (definition 2.7) is equivalent to the condition $IR(f) = 0$ (according to lemma 5.2).

Claim 5.3:: Let $H \subseteq \mathbb{S}_m$ be a fixing subgroup of $\mathbb{S}_m$. An H-social aggregator f satisfies IR iff $IR(f) = 0$.

Since X is the adjacency matrix of a Caley-like graph, it can be decomposed via the representations of the symmetric group. It turns out that all the relevant information lies in its ρ^1 component. This leads to a simplified quadratic form, used with a different encoding for f. Let $g : \mathbb{S}^1_m \to \mathbb{R}^{(m-1)\times(m-1)}$ be a an encoding of f such that $g(x) = \rho^1(f(x))$. A vector form of g is a block vector G whose each entry is a $m-1 \times m-1$ matrix $G_x = g(x)$. (For H-SAs, $g(x) = \mathbb{E}_{y \in f(x)} \rho^1(y)$).

The corresponding quadratic form is

Lemma 5.4: Let $L^j = Y^j \otimes D^j$, where $D^j = C^t_j C_j$ (See 1 for the definition of C), and $L = \sum_j L^j$, then

$$IR(f) = \frac{1}{|\mathbb{S}_m|^2} tr(GLG^t)$$

We partially diagonalize L in the following lemma, decomposing it to eigenspaces.

Lemma 5.5:

$$L = \sum_{r \in [[\mathbb{S}_m]]} d(\rho^r) \tilde{tr} \left((I \otimes \rho^r)\, \widehat{L}(r) \left(I \otimes (\rho^r)^t\right) \right)$$

[1] We use the term Caley graph even though $\mathbb{S}^j_m$ is not a set of generators as this property is irrelevant for our purposes.

where for a block matrix M, $\tilde{tr}(M)$ is the matrix $\tilde{tr}(M)_{xy} = tr(M_{xy})$ and

$$\widehat{L}(0) = I \cdot 0 \;,\; \widehat{L}(r > 1) = I \otimes I \cdot \frac{1}{m}$$

$$\widehat{L}(1) = \frac{1}{m-1} \left(\frac{m-1}{m} I \otimes I - \sum_j D^j \otimes D^j \right)$$

$\widehat{L}(1)$ is the only non diagonal term in lemma 5.5. The diagonalization of $\widehat{L}(1)$ is given by:

Lemma 5.6: $\widehat{L}(1)$ has 3 orthogonal eigenspaces whose dimensions are $1, m-1, (m-1)^2 - m$. Denote their corresponding basis matrices as U_0, U_1, U_2. The corresponding eigenvalues are $0, \frac{1}{m(m-1)}, \frac{1}{m}$. The eigenvectors are vectors of size $(m-1)\cdot(m-1)$. When read as a $(m-1)\times(m-1)$ matrix, U_0 is the identity matrix.

This diagonalization proves that L is PSD and determines its kernel, which is the space of IR functions. This is summarized in this corollary:

Corollary 5.7:: If $g : \mathbb{S}_m \to \mathbb{R}^{m-1\times m-1}$ is the ρ^1 encoding of an IR function, then there exist $(m-1)(m-1)$ sized vectors a and b (that can be view as $(m-1)\times(m-1)$ matrices A and B) such that

$$g_x = b\rho^0(x) + \tilde{tr}\left((aU^t_0)(I \otimes \rho^1(x)\right) = B + A\rho^1(x)$$

B. Many Voter Functions

The quadratic form for $IR(f)$ for functions on n voters is constructed using the quadratic form for 1 voter, in the following lemma:

Lemma 5.8: For a function $f : \mathbb{S}^n_m \to \mathbb{S}_m$, let G be as before, the encoding of f:

$$G_x = \rho^1(f(x)) = \mathbb{E}_{y \in f(x)} \rho^1(y)$$

Let

$$L^{n,j,i} = I^{\otimes i-1} \otimes Y^j \otimes I^{\otimes n-i} \otimes D^j$$

$$L^{n,j} = \sum_i L^{n,j,i} \quad L^n = \sum_j L^{n,j}$$

Then the number of unsatisfied constraints is

$$IR(f) = \frac{1}{|\mathbb{S}_m|^{n+1}} G L^n G$$

We can diagonalize L^n based on our diagonalization of L:

Corollary 5.9:: The diagonalization of L^n is given by:

$$L^n = \sum_{\bar{r} \in [[\mathbb{S}_m]]^n} d(\rho^{\bar{r}}) \tilde{tr} \left(\left(\rho^{\bar{r}}\right) \widehat{L^n}(\bar{r}) \left(\rho^{\bar{r}}\right)^t \right)$$

The $\widehat{L^n}$ coefficients are deduced from the 1 voter $\widehat{L}$'s.

The $\widehat{L^n}$ coefficients are matrices which are not necessarily diagonal. The following lemma partly characterizes their diagonalization, in a manner that suffices for our needs.

Lemma 5.10:

1) If there exists any coordinate $i \in [n]$ for which $r_i > 1$ then $\widehat{L^n}(\bar{r}) \succeq \frac{1}{m} I \otimes I$
2) Otherwise, if there exist at least 2 coordinates $i \in [n]$ for which $r_i = 1$ then $\widehat{L^n}(\bar{r}) \succeq O\left(\frac{1}{m^2}\right) I \otimes I$
3) Otherwise, if there exists exactly 1 coordinate $i \in [n]$ for which $r_i = 1$ then $\widehat{L^n}(\bar{r}) = \widehat{L}(1)$ has a 0 eigenvalue corrsponding to the eigenvector U_0 (as shown in lemma 5.6).
4) Otherwise, $\widehat{L^n}(\bar{r}) = 0$.

The kernel of L^n fully characterizes the functions that are IR:

Corollary 5.11::

- The kernel of L is only functions of the form $g(x_1, ..., x_n) = B + \sum_{i=1}^{n} A^i \cdot \rho^1(x_i)$
- The spectral gap is $\frac{1}{O(m^2)}$.

To finish the proof of theorem 2.3, we need to show that the intersection of the kernel of L^n with the consistency constraint, includes only dictatorships. We don't need to use the consistency constraint to its full capacity. All we need to use is the quadratic constraint that $\forall x, g(x)g^t(x) = M$, where M is some constant matrix. This constraint is valid for any H. For instance, if f is a SWF (H is the trivial group), then since ρ^1 is unitary, $\forall x, g(x)g^t(x) = I$.

Corollary 5.12:: IR functions which are consistent are dictatorships.

C. Robustness

The spectral gap of L^n determines the correlation between $IR(F)$ and the L_2 distance between f and the kernel of L^n.

Corollary 5.13:: If $IR(f) \leq \epsilon$, then there exists a function g in the kernel of L^n such that $\|f - g\|_2^2 \leq O(m^2)\epsilon$.

ϵ-IR functions are L_2-close to the kernel of L^n. That kernel is a linear subspace of the functions whose Fourier coefficients are supported on the first 2 levels. Consistent functions obey some pointwise quadratic constraint.

In [17]. it was shown that functions on n Boolean variables, whose Fourier coefficients are concentrated on the first 2 levels, and whose output is Boolean (which is a pointwise quadratic constraint) are close to being dictatorships. We adapt this theorem to our setting. From it we deduce our main thoerem. Its proof is outlined in section VII.

VI. Proofs for section V

A. Proofs of the Lemmas for 1 voter functions

Proof of lemma 5.1: This is the straightforward definition of the anti-constraints. Recall that $p \not\Rightarrow q$ is true iff $p = true$ and $q = false$.

$$FL'F^t = \sum_{jxv_xyv_y} F_{xv_x} X^j_{xy} \bar{X^j}_{v_xv_y} F_{yv_y} =$$

$$\sum_{jxv_xyv_y} \mathbf{1}_{v_x=f(x)} \mathbf{1}_{x^{-1}(j)=y^{-1}(j)} \mathbf{1}_{v_x^{-1}(j)\neq v_y^{-1}(j)} \mathbf{1}_{v_y=f(y)} =$$

$$\sum_{jxy} \mathbf{1}_{x^{-1}(j)=y^{-1}(j)} \mathbf{1}_{f(x)^{-1}(j)\neq f(y)^{-1}(j)}$$

∎

Proof of Lemma 5.2: Essentialy, one needs to show that the expressions

$$\sum_{jxyv_xv_y} F_{xv_x}\left(X^j_{xy} \cdot 1\right) F_{yv_y}$$

and

$$\sum_{jxyv_xv_y} F_{xv_x}\left((m-1)!\delta_{xy} \otimes X^j_{v_xv_y}\right) F_{yv_y}$$

are equivalent when F is consistent. Since $\forall x, \sum_v F_{xv} = 1$, the first expression translates to

$$\sum_{jxy} X^j_{xy} = m \cdot m! \cdot (m-1)! = m!^2$$

Since the diagonal of X^j, for every j, is all ones, the second expression is

$$\sum_{jxyv_xv_y} F_{xv_x}\left((m-1)!\delta_{xy} \otimes X^j_{v_xv_y}\right) F_{yv_y} =$$

$$\sum_{jxv_xv_y} F_{xv_x}\left((m-1)! X^j_{v_xv_y}\right) F_{xv_y} =$$

$$\sum_{jx} ((m-1)!1) = m \cdot m! \cdot (m-1)! = m!^2$$

∎

Before we carry on, this is good place to recall H and C from equation 1. Since H is orthonormal, we can easily deduce the following:

Claim 6.1::

$$\begin{aligned} CC^t &= I - \frac{J}{m} \\ C^tC &= I \\ \mathbf{1}C &= 0 \end{aligned}$$

Proof of Lemma 5.4: We begin by writing an explicit expression for X. We use P, the defining representation of $\mathbb{S}_m$. A permutation x has a fixed point j iff $(P_x)_{jj}$ is 1. Therefore,

$$X^j_{xy} = (P_{x^{-1}y})_{jj}$$

Recall that $P_x = H(\rho^0 \oplus \rho^1)H^t$, and that H is orthonrmal. Therefore

$$X^j_{xy} = (P_{xy^{-1}})_{jj} = \left(H\left(\rho^0(xy^{-1}) \oplus \rho^1(xy^{-1})\right)H^t\right)_{jj}$$

We will use the vectors defined by the representations ρ^0 and ρ^1 as diagonalization for L. First though, we wish to simplify the expression a bit. Recall $H = \left(\frac{\mathbf{1}}{\sqrt{m}}\; C\right)$, Therefore

$$X^j_{xy} = \left(\frac{\mathbf{1}\rho^0(xy^{-1})\mathbf{1}^t}{m} + C\rho^1(xy^{-1})C^t\right)_{jj}$$

We will now show that we can eliminate the summand corrsponding to ρ^0 in the inner matrices in L, assuming our function F is consistent. Recall $\rho^0_x = 1_x$.

$$L^j_{xyv_xv_y} = \left(Y^j \otimes X^j\right)_{xyv_xv_y} =$$

$$Y^j_{xy} \cdot \left(\frac{\mathbf{1}\rho^0(v_xv_y^{-1})\mathbf{1}^t}{m} + C\rho^1(v_xv_y^{-1})C^t\right)_{jj}$$

In our quadratic form, the summand corresponding to ρ^0 is

$$\sum_{jxyv_xv_y} F_{xv_x}\left(Y^j_{xy}\cdot\left(\frac{\mathbf{1}\rho^0(v_xv_y^{-1})\mathbf{1}^t}{m}\right)_{jj}\right)F_{yv_y}$$

Since F is consistent, and $\rho^0(v_xv_y^{-1}) = \rho^0(v_x)\rho^0(v_y^{-1}) = \mathbf{1}_{v_x}\mathbf{1}_{v_y}$, this is equal to

$$\sum_{jxy}\mathbf{1}_x\left(Y^j_{xy}\frac{1_j1_j}{m}\right)\mathbf{1}_y = 0$$

Therefore, removing this expression from L'' yields

$$IR(f) = \frac{\sum_{jxyv_xv_y} F_{xv_x}Y^j_{xy}\left(C\rho^1(v_x)\rho^1(v_y^{-1})C^t\right)_{jj}F_{yv_y}}{|S_m|^2}$$

We can denote $G_x = \sum_v F_{xv}\rho^1(v)$, and the qudaratic form becomes

$$IR(f) = \frac{\sum_{jxy} tr\left(G_x\left(Y^j_{xy}(CC^t)_{jj}\right)G_y\right)}{|S_m|^2}$$

■

Proof of claim 5.3 In the proof, we shall use the expression for $IR(f)$ from lemma 5.4. It is easier to understand the action of L when decomposing it to matrices that act on pairs of inputs. Define the matrix $Z^{j,x,y}$ to be the Laplacian of the graph whose vertices are the elements of $\mathbb{S}_m$ and has at most one edge, connecting x with y iff $x^{-1}(j) = y^{-1}(j)$. Clearly, $Y^j = \sum_{(x,y)\in\binom{\mathbb{S}_m}{2}} Z^{j,x,y}$. It is also easy to see that the diagonalization of the Z's is given by $Z^{j,x,y} = \mathbf{1}_{xy^-1(j)=j}d^{x,y^t}d^{x,y}$ where $d^{x,y}$ is a vector that has 1 in x, -1 in y and 0 otherwise. We shall now use this composition to show how the expression $tr\left(GL^jG^t\right)$ compares pairs of inputs.

$$tr\left(GL^jG^t\right) = \sum_{(x,y)\in\binom{\mathbb{S}_m}{2}} tr\left(G\left(Z^{j,x,y,}\otimes D^j\right)G^t\right) =$$

$$\sum_{xy^{-1}(j)=j} tr\left(G\left(d^{x,y^t}\otimes C^{j^t}\right)\cdot\left(d^{x,y}\otimes C^j\right)G^t\right) =$$

$$\sum_{xy^{-1}(j)=j} tr\left(\left(g(x)-g(y)\right)C^{j^t}\cdot C^j\left(g(x)-g(y)\right)^t\right) =$$

$$\sum_{xy^{-1}(j)=j}\left\langle C^j\left(g(x)-g(y)\right)^t, C^j\left(g(x)-g(y)\right)^t\right\rangle$$

To finish the proof, we shall show that there is a one-to-one mapping betwen the j-profile of the coset $f(x)$ and the term $C^jg^t(x)$. Clearly, the normalized characteristic vector of the j-profile of $g(x)$ is $e_j\left(\mathbb{E}_{y\in g(x)}P^t_y\right)$, where e_j is the j'th unit vector. Since the matrix H is regular (orthonormal), multiplying this vector by H is a one-to-one mapping, which gives:

$$\left(e_j\mathbb{E}_{y\in g(x)}P^t_y\right)H = \left(e_jH\left(1\oplus g^t(x)\right)H^t\right)H =$$

$$e_jH\left(1\oplus g^t(x)\right) = \begin{pmatrix}\frac{1}{\sqrt{m}} & C_j\end{pmatrix}\left(1\oplus g^t(x)\right) =$$

$$\begin{pmatrix}\frac{1}{\sqrt{m}} & C^jg^t(x)\end{pmatrix}$$

■

Proof of lemma 5.5: We transform the space $R^{\mathbb{S}_m}$ using the vectors defined by the irreps of $\mathbb{S}_m$.

Recall that

$$L = \sum_j\left((m-1)!I - X^j\right)\otimes\left(C^t_jC_j\right)$$

The outer matrix can be decomposed using the irreps of $\mathbb{S}_m$:

$$(m-1)!I_{xy} = \frac{(m-1)!}{m!}\left(\sum_{r\in[[\mathbb{S}_m]]} d(\rho^r)\rho^r(x)\rho^r(y^{-1})\right)$$

$$X^j_{xy} = \frac{\rho^0(x)\rho^0(y^{-1})}{m} + C_j\rho^1(x)\rho^1(y^{-1})C^t_j$$

Summing these decompositions, and using the fact that $\sum_j C^t_jC_j = C^tC = I$ yields the result.

■

Proof of lemma 5.6: Denote by E, a $m\times(m-1)^2$ matrix whose j's row is $C_j\otimes C_j$. We need to diagonalize

$$\sum_j D^j\otimes D^j = \sum_j\left(C^t_jC_j\right)\otimes\left(C^t_jC_j\right) =$$

$$\sum_j\left(C^t_j\otimes C^t_j\right)\left(C_j\otimes C_j\right) = E^tE$$

The nonzero eigenvalues of E^tE are the nonzero eigenvalues of EE^t (this can be deduced from the SVD decomposition of E). Recall $CC^t = I - \frac{J}{m}$. Therefore,

$$\left(EE^t\right)_{ij} = \left(C_i\otimes C_i\right)\left(C^t_j\otimes C^t_j\right) = \left(C_iC^t_j\right)^2 = \left(CC^t\right)^2_{ij} =$$

$$\left(\delta_{ij}-\frac{1}{m}\right)^2 = \left(1-\frac{2}{m}\right)\delta_{ij} + \left(\frac{1}{m}\right)^2$$

Therefore, $EE^t = \frac{m-2}{m}I + \frac{J}{m^2}$, and its eigenvalues are $\frac{m-1}{m}$ at multiplicity 1 and $\frac{m-2}{m}$ at multiplicity $m-1$.

We can verify that the eigenvector of E^tE corresponding to the $\frac{m-1}{m}$ eigenvalue is U_0, which is the identity matrix parsed as a vector. We need to use two simple facts:

- For 3 matrices A, B and C, the term $A \cdot B \cdot C^t$ when parsed as a vector, is equal to the term $(A \otimes C) \cdot B$, when B is parsed as a vector.
- For a matrix C whose rows are $\{C_j\}_j$, the term $\sum_j C_j \otimes C_j$ equals to the term $C^t C$ parsed as a row vector. This is because $C^t C = \sum_j C_j^t C_j$.

Therefore, we get that

$$(E \cdot U_0)_j = (C_j \otimes C_j) U_0 = C_j I C_j^t = C_j C_j^t = \left(CC^t\right)_{jj} = \left(I - \frac{J}{m}\right)_{jj} = \frac{m-1}{m}$$

Which means that $EU_0 = \mathbf{1}\frac{m-1}{m}$ and

$$\left(E^t E\right) U_0 = E^t \mathbf{1} \cdot \frac{m-1}{m} = \frac{m-1}{m} \left(\sum_j C_j \otimes C_j\right)^t$$

Since $\sum_j C_j \otimes C_j$ is $C^t C$ parsed as a vector and $C^t C = I$ and U_0 is I parsed as a vector, we get that $\left(E^t E\right) U_0 = \frac{m-1}{m} U_0$.
∎

Proof of corollary 5.7: We have shown that eigenspaces of L with eigenvalue 0 are ρ^0 and $\rho^1 U^0$. All the other eigenvalues are positive, so L is PSD.

There is the small matter of G being a block vector instead of a vector and the usage of the trace operator. Observe that when the blocks of G are parsed as vectors, the quadratic form $tr\left(GLG^t\right)$ translates to $G\left(I \otimes L\right) G^t$. Therefore, if G is an IR function, it is of the form of a member of the kernel of $I \otimes L$, parsed as a block vector. The diagonalization of L implies that G's Fourier coefficients $\widehat{G}(k) = \mathbb{E}_x G_x \rho^k(x)$ have the following form:
$\widehat{G}(0)$ can be anything, $\widehat{G}(1)$ must be of the form aU^{0^t} (a a column vector), and for $k > 1$, $\widehat{G}(k) = 0$. Reverse Fourier transform yields:

$$G_{xkl} = B_{kl}\rho^0(x) + \sum_{ts} \rho^1_{ts}(x)\delta_{sl}A_{kt} = 1_x B_{kl} + \sum_t A_{kt}\rho^1_{tl}(x)$$

This means that the function g is $g(x) = B + A \cdot \rho^1(x)$ (where A and B are $(m-1) \times (m-1)$ matrices), so g is a linear function in $\rho^1(x)$. It can be shown, using Schur's lemma, that if g is a consistent function, then either B is 0, or A is 0. We will not show that here. We will show it using a different technique in the n voter section. ∎

B. Proofs of Lemmas for many voter functions

Proof of lemma 5.8: We only check inputs where there exists $i \in [n]$ such that the i'th vote has changed and the other votes remain the same. The $Y^j \otimes D^j$ term accounts for the ith vote that changed and the I terms account for the other votes not changing. ∎

Proof of lemma 5.10:

1) Let i be such that $r_i > 1$, then, by lemma 5.10
$$\widehat{L^{n,i}}(\bar{r}) = \frac{1}{m}\bigotimes_{k=1}^{n} I_{d(r_k)} \otimes I_{m-1}$$
Since $\widehat{L^n}(\bar{r}) = \sum_j \widehat{L^{n,j}}(\bar{r})$ and for every $j, \bar{r}$, $\widehat{L^{n,j}}(\bar{r}) \succeq 0$, we have $\widehat{L^n}(\bar{r}) \succeq \widehat{L^{n,i}}(\bar{r})$.
2) We shall focus only on the case where there are exactly 2 distinct i and j such that $r_i = r_j = 1$, because this case produces the minimal eigenvalue. If there are more than 2 such indices, $\widehat{L^n}(\bar{r})$ is equivalent to the case of 2 indices, tensored by identity and summed with other PSD matrices.
Recall the diagonalization of $\widehat{L}(1)$ (from lemma 5.6)
$$\widehat{L}(1) \succeq \frac{1}{m(m-1)}\left(I \otimes I - \frac{1}{m-1}U^0 U^{0^t}\right)$$
From this, it is easy to deduce the following:
$$\widehat{L^n}(\bar{r}) = \widehat{L^{n,i}}(\bar{r}) + \widehat{L^{n,j}}(\bar{r}) \succeq \frac{1}{m(m-1)}\left(2I \otimes I \otimes I - \frac{1}{m-1}\left(A^i A^{i^t} + A^j A^{j^t}\right)\right)$$
Where A^i is a $(m-1)^3 \times (m-1)$ matrix of the form $A_{(pqr)s} = \delta_{pq}\delta_{rs}$ (p, q, r and s are indices going from 1 to $m-1$. (pqr) forms the row index of A^i and s is its column index). Likewise, A^j is a $(m-1)^3 \times (m-1)$ matrix of the form $A_{(pqr)s} = \delta_{pr}\delta_{qs}$. Denote the RHS as R.
Denote $B = A^i + A^j$ and $C = A^i - A^j$. It is easy to see that $A^i A^{i^t} + A^j A^{j^t} = \frac{1}{2}\left(BB^t + CC^t\right)$. It is also easy to verify that $B^t C = 0$ and therefore B and C are orthogonal. Therefore, B and C are bases to orthogonal eigenspaces of R. After normalization, we get that the minimal eigenvalue of R, corresponding to the columns of B, is $\frac{1}{m(m-1)}(2 - \frac{m}{m-1}) = \frac{m-2}{m(m-1)^2}$.
Important Note: Notice that if $m = 2$, R has eigenvalues equal to 0, and therefore we cannot deduce that the function is a dictatorship for $m = 2$.

Items **3** and **4** are trivial.

Proof of corollary 5.12: We need to show that only one of $B, A_1, ..., A_n$ is not zero.

- We show that the function g satisfies a quadratic constraint. Let M_H be $M_H = \mathbb{E}_{x \in H}\rho^1(x)$. Since g is consistent, it must be of the form $\forall x, \exists y, g(x) = M_H \rho^1(y)$. Therefore,
$$g(x)g(x)^t = M_H \rho^1(y)\rho^1(y)^t M_H^t = M_H M_H^t = \mathbb{E}_{x \in H}\rho^1(x)\mathbb{E}_{y \in H}\rho^1(y^{-1}) = \mathbb{E}_{x,y \in H}\rho^1(xy^{-1}) = \mathbb{E}_{z \in H}\rho^1(z) = M_H$$
.

- We show that w.l.o.g., we may assume that $\mathbb{E}_x g(x) = 0$ and therefore $B = 0$. Indeed, we introduce a dummy variable $y \in \mathbb{S}_m$ and define $g'(x_1, x_2, ..., x_n, y) = g(x_1 y^{-1}, x_2 y^{-1}, ..., x_n y^{-1})\rho^1(y)$. This makes the function g' neutral. It is easy to show that $\mathbb{E}_{x,y} g'(x'y) = 0$ and that if $IR(g) \le \epsilon$ then $IR(g') \le \epsilon$, using our expression for IR(f). After proving the claim assuming $B = 0$, we apply it to g' and get our result, without this assumption.
- Assuming $B = 0$,
$$g(x)g(x)^t = \sum_{ij} A^i \rho^1(x_i) {\rho^1}^t(x_j) {A^j}^t$$
The summand for i, j, translates to
$$\left(A^i \rho^1(x_i) {\rho^1}^t(x_j) {A^j}^t\right)_{tu} = \sum_{p,q,r,s} \rho^1_{pq}(x_i)\rho^1_{rs}(x_j)\left(A^i_{tp} A^j_{ru} \delta_{qs}\right)$$
From this expansion we may deduce the Fourier coefficient of gg^t at $\bar{r}$ where $\bar{r}$ is 1 at i and j and 0 otherwise:
$$\widehat{gg^t}_{tpruqs}(\bar{r}) = (m-1)^2\left(\left(A^i_{tp} A^j_{ru} \delta_{qs}\right) + \left(A^j_{tp} A^i_{ru} \delta_{qs}\right)\right) \quad (2)$$
Since gg^t is a constant function and the Fourier expansion is unique, we get that $\widehat{gg^t}(\bar{r})$ must be 0. Assume by contradiction that there exist indices tp and ru where $A^i_{tp} \neq 0, A^j_{ru} \neq 0$. 2 implies that $A^j_{tp} A^i_{ru} = -A^i_{tp} A^j_{ru} \neq 0$. Therefore, $A^j_{tp} \neq 0$ and $\widehat{gg^t}_{tptpqq}(\bar{r}) \simeq A^i_{tp} A^j_{tp} \neq 0$

∎

VII. Adapted version of FKN[17]

The adapted version of FKN's theorem is:

Theorem 7.1: Let $Lin(\mathbb{S}^n_m)$ be the space of functions of the form $\sum_i A^i \rho^1(x_i)$. Let $f : \mathbb{S}^n_m \to \mathbb{R}^{m-1 \times m-1}$ be a function such that

- $\mathbb{E} f = 0$
- There exists a matrix M such that $tr(M) = 1$ and $\forall x \in \mathbb{S}^n_m, f(x)f^t(x) = M$
- f is ϵ close to $Lin(\mathbb{S}^n_m)$.

then f is $O(m^4\epsilon)$ close to a function of the form $A^i \rho^1(x_i)$. (distances are L_2).

We use the theorem with g under the assumption the $\mathbb{E} g = 0$ (see the proof of 5.12) and that g is normalized so that $tr(M) = 1$. The normalization is dependent on H. Also, taking into account corollary 5.11 makes the final distance at most $O(m^6\epsilon)$.

We give here an outline of the proof. The complete proof will appear in the journal version.

Outline of the Proof:

The proof follows the second proof shown in [17], and does not vary from it much. Since we are dealing with matrices, there are a couple of times when we need to use union bounds on the entries of the matrices, giving a $O(m^4)$ penalty. It is possible this could be optimized and reduced.

The proof follows these lines: We denote $f = f_{\|} + f_{\perp}$, where $f_{\|}$ is f's projection onto $lin(\mathbb{S}^n_m)$. We define the function $r(x) = M - f_{\|}(x) f^t_{\|}(x)$, which measures the distance between $f_{\|}$ and consistency. Since f is consistent and close to $f_{\|}$, r is typically close to 0 (compare this to the proof of corollary 5.12). We use the information that r is quadratic to show that it does not have large deviations, and therefore *very* close to 0. To this end, we bound its second moment by its 4'th moment, using the Beckner-Bonami(BB) inequality. This cannot happen if more than one of the A^i coefficients is large. This is the original proof of [17]. In most stages of the proof, we simply use the entrywise scalar functions. However, the fact that g returns matrices does come into play.

A note regarding Beckner-Bonami's inequality: The version of BB's inequality we use takes the form:

Lemma 7.2: [Beckner (1975), Bonami (1970)] Let G be a finite group and $f : G^n \to \mathbb{R}$ be a function whose Fourier coefficients are supported on the lowest k levels (rank k). Let $p > 2$. Then
$$\|f\|_p \le c^k \|f\|_2.$$

In our case, we apply this inequlaity on the entries of r (r_{ij}). We use $p = 2, q = 4$ and the rank of r is 2. Unfortunately, we use $G = \mathbb{S}_m$ and because it is a large group, c in the inequality as stated is exponential in m. However, we have more information on the function r which we can use. r is not only of rank 2, but its Fourier coefficients are supported on combinations of ρ^1, we use this to get a polynomial c.

In [18], optimal constants are obtained for BB-type inequalities for simple random variables. In Theorem 3.1 in that paper, it is shown that the optimal constant depends only on the measure of the smallest atom of the random variable. It is shown via an optimization argument using Lagrange multipliers, that the optimum is obtained when the random variable has exactly 2 different values, and the result follows.

We reprove this theorem with the added constraint that the Fourier coefficients are supported on ρ^1 and $p = 2$. We show that the conclusion still holds, and the optimum is still obtained when the function has 2 distinct values. We then use a lemma from [19] that states that functions on the Symmetric group who obtain only 2 distinct values and whose Fourier coefficients are supported on ρ^1 are a linear combination of the characteristic vectors of cosets of the subgroups $\mathbb{S}^j_m$. This means that the smallest measure of an atom is $\frac{1}{m}$ and we are done. ∎

VIII. Future work

In upcoming future work, we intend to show the application of this scheme for proofs of a **robust version of Arrow's theorem** with relaxed IIA and a **robust version**

of Gibbard-Satterthwaite's theorem with relaxed startegy proofness for any SCF's on any number of alternatives.

Some interesting future reasearch directions may include an application of this technique to the generalized problem, known as judgement aggregation. In that setting, we have a permissible opinion space $X \subseteq [k]^m$ and for some such X, for functions $f : X^n \to X$, independency implies dictatorship. It is known for which X this holds, but there are no known robust versions for general opinion spaces X.

As is, our proof could be generalized for groups other than $\mathbb{S}_m$, and also other types of independency (such as independency of ranking of k-tuples).

As mentioned earlier, our result can be interpreted as a 2 query *dictatorship tester*, for functions $\mathbb{S}_m^n \to \mathbb{S}_m/H$. It may be interesting to see whether this has any computational implications.

References

[1] K. J. Arrow, *Social Choice and Individual Values*. New York: John Wiley and sons, 1963.

[2] A. Gibbard, "Manipulation of voting schemes: a general result," *Econometrica*, vol. 41, no. 4, pp. 587–601, 1973.

[3] M. A. Satterthwaite, "Strategy-proofness and arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions," *Journal of Economic Theory*, vol. 10, pp. 187–217, April 1975.

[4] G. Kalai, "A fourier-theoretic perspective for the condorcet paradox and arrow's theorem," *Advances in Applied Mathematics*, vol. Volume 29, Issue 3, pp. 412–426, 2001.

[5] E. Dokow and R. Holzman, "Aggregation of non-binary evaluations," *Advances in Applied Mathematics*, vol. 45, pp. 487–504, 2010.

[6] K. Nehring and C. Puppe, "The structure of strategy-proof social choice part i: General characterization and possibility results on median spaces," *Journal of Economic Theory*, vol. 135, pp. 269–305, 2007.

[7] F. Dietrich and C. List, "Strategy-proof judgement aggregation," *Economics and Philosophy*, vol. 23, pp. 269–300, 2007.

[8] P. J. Reny, "Arrows theorem and the gibbard-satterthwaite theorem: A unified approach," *Economics Letters*, pp. 99–105, 2001.

[9] E. Mossel, "A quantitative arrow theorem," 2011, arxiv 0903.2574.

[10] ——, "Arrow's impossibility theorem without unanimity," *CoRR*, vol. abs/0901.4727, 2009.

[11] N. Keller, "A tight quantitative version of arrow's impossibility theorem," *CoRR*, vol. abs/1003.3956, 2010.

[12] ——, "On the probability of a rational outcome for generalized social welfare functions on three alternatives," *Journal of Combinatorial Theory*, vol. Ser. A, 117(4), 2010.

[13] I. Nehama, "Approximate judgement aggregation," *CoRR*, vol. abs/1008.3829, 2010.

[14] E. Friedgut, G. Kalai, and N. Nisan, "Elections can be manipulated often," in *FOCS*. IEEE Computer Society, 2008, pp. 243–249.

[15] M. Isaksson, G. Kindler, and E. Mossel, "The geometry of manipulation: A quantitative proof of the gibbard-satterthwaite theorem," in *FOCS*. IEEE Computer Society, 2010, pp. 319–328.

[16] L. Xia, "A sufficient condition for voting rules to be frequently manipulable," in *In Proceedings of the Ninth ACM Conference on Electronic Commerce (EC*, 2008.

[17] E. Friedgut, G. Kalai, and A. Naor, "Boolean functions whose fourier transform is concentrated on the first two levels," *Adv. in Appl. Math*, vol. 29, 2002.

[18] P. Wolff, "Hypercontractivity of simple random variables," *Studia Math.*, vol. 180, iss. 3, pp. 219–236, 2007.

[19] D. Ellis, H. Pilpel, and E. Friedgut, "Intersecting families of permutations," *J. Amer. Math. Soc.*, vol. 24, pp. 649–682, 2011.