WHAT MAJORITY DECISIONS ARE POSSIBLE WITH POSSIBLE ABSTAINING

PAUL LARSON, NICK MATTEO, AND SAHARON SHELAH

Abstract. Suppose we are given a family of choice functions on pairs from a given finite set. The set is considered as a set of alternatives (say candidates for an office) and the functions as potential "voters." The question is, what choice functions agree, on every pair, with the majority of some finite subfamily of the voters? For the problem as stated, a complete characterization was given in Shelah [2009], but here we allow each voter to abstain. There are four cases.

Roslanowski and Shelah [1999]

1. Introduction

Condorcet's "paradox" demonstrates that given three candidates A, B, and C, majority rule may result in the society preferring A to B, B to C, and C to A [Condorcet, 1785]. McGarvey [1953] proved a far-reaching extension of Condorcet's paradox: For every asymmetric relation R on a set X of $\mathbf{n}$ candidates, there are m linear order relations on X: $R_1, R_2, \ldots, R_m$, with R as their strict simple majority relation. I.E. for every $a, b \in X$,

$$a \; R \; b \iff |\{i : a \; R_i \; b\}| > \frac{m}{2}.$$

In other words, given any set of choices from pairs from a set of $\mathbf{n}$ candidates, there is a population of m voters, all with simple linear-order preferences among the candidates, who will yield the given outcome for each pair in a majority-rule election between them.

McGarvey's proof gave $m = \mathbf{n}(\mathbf{n}-1)$. Stearns [1959] found a construction with $m = \mathbf{n}$ and noticed that a simple counting argument implies that m must be at least $\frac{\mathbf{n}}{\log \mathbf{n}}$. Erdős and Moser [1964] were able to give a construction with $m = O\left(\frac{\mathbf{n}}{\log \mathbf{n}}\right)$. Alon [2002] showed that there is a constant $c_1 > 0$ such that, given any asymmetric relation R, there is some m and linear orders $R_1, \ldots, R_m$ with

$$a \; R \; b \iff |\{i : a \; R_i \; b\}| > \frac{m}{2} + c_1\sqrt{m}$$

and that this is not true for any $c_2 > c_1$.

Gil Kalai asked to what extent the assertion of McGarvey's theorem holds if we replace the linear orders by an arbitrary isomorphism class of choice functions on pairs of elements. Namely, when can we guarantee that every asymmetric relation

Date: October 5, 2020.
2000 *Mathematics Subject Classification.* Primary 91B14; Secondary 05C20.
Key words and phrases. choice function, majority decision, Condorcet's paradox, tournament.
The research of the first author is supported by NSF grant DMS-0801009.
Partially supported by the United States-Israel Binational Science Foundation. This is publication number 965 in the third author's publication list.

R on X could result from a finite population of voters, each using a given kind of asymmetric relation on X?

Of course we must define what we mean by "kind" of asymmetric relation. Let $\binom{X}{k}$ denote the family of subsets of X with k elements:

$$\binom{X}{k} = \{Y \subseteq X : |Y| = k\}.$$

Either an individual voter's preferences among candidates, or the outcomes which would result from each two-candidate election, may be represented as

- An asymmetric relation R where $a\ R\ b$ iff a beats b; it is possible that $a \not R\ b$, $b \not R\ a$, and $a \neq b$.
- A choice function defined on some subfamily of $\binom{X}{2}$, choosing the winner in each pair. Such a choice function is called "full" if its domain is all of $\binom{X}{2}$, and "partial" otherwise.
- An oriented graph, i.e. a directed graph with nodes X and edges $a \to b$ when a beats b.

We shall treat these representations as largely interchangeable throughout this paper. Total asymmetric relations, full choice functions, and tournaments (complete oriented graphs) all correspond to the case of no abstaining. $\mathrm{Tor}(c)$ will denote the oriented graph associated with a choice function c. For any set X, $c \mapsto \mathrm{Tor}(c)$ is a bijection of full choice functions onto tournaments on X, and a bijection of all choice functions onto oriented graphs on X.

1.1. *Hypothesis.* Assume

(a) X is a finite set with $\mathbf{n} = |X|$ and $\mathbf{n} \geq 3$.
(b) $\mathfrak{C}$ is the set of choice functions on pairs in X;

$$\mathfrak{C} = \left\{c\colon Y \to X : Y \subseteq \binom{X}{2}, \forall\{x,y\} \in Y\ c\{x,y\} \in \{x,y\}\right\}.$$

When $c\{x,y\}$ is not defined it is interpreted as abstention or having no preference.

1.2. **Definition.** (a) $\mathrm{Per}(X)$ is the set of permutations of X.

(b) Choice functions c and d are *symmetric* iff there is $\sigma \in \mathrm{Per}(X)$ such that

$$d\{\sigma(x), \sigma(y)\} = \sigma(x) \iff c\{x,y\} = x;$$

we write $d = c^\sigma$.

(c) A set of choice functions $\mathscr{C} \subseteq \mathfrak{C}$ is *symmetric* iff it is closed under permutations of X. So for each $\sigma \in \mathrm{Per}(X)$, if $c \in \mathscr{C}$ then $c^\sigma \in \mathscr{C}$.

Choice functions c and d are symmetric iff $\mathrm{Tor}(c)$ and $\mathrm{Tor}(d)$ are isomorphic graphs. Note that symmetry of choice functions is an equivalence relation. These symmetric sets of choice functions are in fact what was meant by "kind of asymmetric relation."

The main result of Shelah [2009] pertained to full choice functions for the voters. It was shown that an arbitrary choice function d could result from a symmetric set $\mathscr{C}$, i.e. for each d there is a finite set $\{c_1, \ldots, c_m\} \subseteq \mathscr{C}$ such that

$$d\{x,y\} = x \iff |\{i : c_i\{x,y\} = x\}| > \frac{m}{2},$$

iff for some $c \in \mathscr{C}$ and $x \in X$,

$$|\{y : c\{x, y\} = y\}| \neq \frac{\mathbf{n} - 1}{2}.$$

We shall call this condition "imbalance."

1.3. **Definition.** For a choice function c,

(a) for any pair (x, y) in X^2, the *weight* of x over y for c is

$$W_y^x(c) = \begin{cases} 1 & \text{if } c\{x, y\} = x \\ 0 & \text{if } \{x, y\} \notin \operatorname{dom} c \\ -1 & \text{if } c\{x, y\} = y \end{cases}$$

(b) c is *balanced* iff

$$\forall x \in X, \sum_{y \in X} W_y^x(c) = 0.$$

That is, $|\{y : c\{x, y\} = x\}| = |\{y : c\{x, y\} = y\}|$, for every x in X.

(c) c is *imbalanced* iff c is not balanced.

(d) c is *pseudo-balanced* iff every edge of $\operatorname{Tor}(c)$ belongs to a directed cycle.

Gil Kalai further asked whether the number m can be given bounds in terms of $\mathbf{n}$, and what is the result of demanding a "non-trivial majority," e.g. 51%. We shall consider loose bounds while addressing the general case, when voters are permitted to abstain.

To determine what symmetric sets of partial choice functions could produce an arbitrary outcome, we shall characterize the set of all possible outcomes of a symmetric set of choice functions, the "majority closure" maj-cl($\mathscr{C}$). $\mathscr{C}$ is a satisfactory class for this extension to McGarvey's theorem iff maj-cl($\mathscr{C}$) $= \mathfrak{C}$.

1.4. **Definition.** For $\mathscr{C} \subseteq \mathfrak{C}$ let maj-cl($\mathscr{C}$) be the set of $d \in \mathfrak{C}$ such that, for some set of weights $\{r_c \in [0, 1]_{\mathbb{Q}} : c \in \mathscr{C}\}$ with $\sum_{c \in \mathscr{C}} r_c = 1$,

$$d\{x, y\} = x \iff \sum_{c \in \mathscr{C}} W_y^x(c) r_c > 0.$$

Here and throughout this paper, $[0, 1]_{\mathbb{Q}}$ refers to the rationals in the unit interval of the real line. Note that we cannot assume that maj-cl(maj-cl($\mathscr{C}$)) = maj-cl($\mathscr{C}$), and in fact we shall see this is not true. We must show that each function d in the majority closure, thus defined, can in fact be the outcome of a finite collection of voters using choice functions in $\mathscr{C}$; this is Claim 2.6.

Shelah [2009] gave a characterization of maj-cl($\mathscr{C}$) for symmetric sets of full choice functions; there are just two cases. maj-cl($\mathscr{C}$) $= \mathfrak{C}$ iff some $c \in \mathscr{C}$ is imbalanced; if every $c \in \mathscr{C}$ is balanced, then maj-cl($\mathscr{C}$) is the set of all pseudo-balanced functions. The situation with possible abstention is more complicated; there are four cases.

In addition to the representations as relations, choice functions, or oriented graphs discussed above, we shall also find it convenient to map each choice function c to a sequence in $[-1, 1]_{\mathbb{Q}}$ indexed by X^2, the "probability sequence"

$$\operatorname{pr}(c) = \langle W_y^x(c) : (x, y) \in X^2 \rangle.$$

Let $\operatorname{pr}(\mathscr{C}) = \{\operatorname{pr}(c) : c \in \mathscr{C}\}$. $\operatorname{pr}(X)$ will denote the set of all sequences $\bar{t}$ in $[-1, 1]_{\mathbb{Q}}^{X^2}$ such that $t_{x,y} = -t_{y,x}$; $\operatorname{pr}(X)$ contains $\operatorname{pr}(\mathfrak{C})$.

2. Basic Definitions and Facts

2.1. **Definition.** For a probability sequence $\bar{t} \in \mathrm{pr}(X)$,

(a) $\bar{t}$ is *balanced* iff for each $x \in X$

$$\sum_{y \in X} t_{x,y} = 0.$$

(b) $\mathrm{maj}(\bar{t})$ is the $c \in \mathfrak{C}$ such that $c\{x, y\} = x \iff t_{x,y} > 0$.

Note that maj and pr are mutually inverse functions from $\mathfrak{C}$ to $\mathrm{pr}(\mathfrak{C})$; $c = \mathrm{maj}(\mathrm{pr}(c))$ and $\bar{t} = \mathrm{pr}(\mathrm{maj}(\bar{t}))$ for all $c \in \mathfrak{C}$ and $\bar{t} \in \mathrm{pr}(\mathfrak{C})$. maj is also defined on the much larger set $\mathrm{pr}(X)$, which it maps onto $\mathfrak{C}$. The reason for reusing the name "balanced" is clear:

2.2. **Claim.** *If $c \in \mathfrak{C}$ is a balanced choice function, then* $\mathrm{pr}(c)$ *is a balanced probability sequence. If $\bar{t} \in \mathrm{pr}(\mathfrak{C})$ is a balanced probability sequence, then* $\mathrm{maj}(\bar{t})$ *is a balanced choice function.*

Proof. Suppose $c \in \mathfrak{C}$ is balanced. For every $x \in X$,

$$\sum_{y \in X} W^x_y(c) = 0.$$

$\mathrm{pr}(c)_{x,y} = W^x_y(c)$, so

$$\sum_{y \in X} \mathrm{pr}(c)_{x,y} = 0,$$

i.e. $\mathrm{pr}(c)$ is balanced.

Now suppose $\bar{t} \in \mathrm{pr}(\mathfrak{C})$ is balanced. $\bar{t} = \mathrm{pr}(c)$ for some $c \in \mathfrak{C}$, and $c = \mathrm{maj}(\bar{t})$.

$$\sum_{y \in X} t_{x,y} = 0,$$

but $t_{x,y} = W^x_y(c)$, so

$$\sum_{y \in X} W^x_y(c) = 0$$

and c is balanced. $\square$

However, it is not the case that $\mathrm{maj}(\bar{t})$ is balanced for every balanced $\bar{t} \in \mathrm{pr}(X)$. For instance, choose $x, z \in X$ and define a sequence $\bar{t}$ which has $t_{z,x} = 1$, $t_{x,z} = -1$; for all $y \notin \{x, z\}$

$$t_{x,y} = t_{y,z} = \frac{1}{\mathbf{n} - 2},$$
$$t_{y,x} = t_{z,y} = \frac{-1}{\mathbf{n} - 2};$$

and for all other pairs (u, v), $t_{u,v} = 0$. One may check that $\bar{t}$ is balanced, yet $\mathrm{maj}(\bar{t})$ has $\sum_{y \in X} W^x_y(\mathrm{maj}(\bar{t})) = \mathbf{n} - 3 > 0$ for any X with at least 4 elements.

2.3. **Definition.** For a choice function $c \in \mathfrak{C}$,

(a) c is *partisan* iff there is nonempty $W \subsetneq X$ such that

$$c\{x, y\} = x \iff x \in W \text{ and } y \notin W.$$

(b) c is *tiered* iff there is a partition $\{X_1, X_2, \ldots, X_k\}$ of X such that $c\{x, y\} = x$ iff $x \in X_i$, $y \in X_j$, and $i > j$. We say c is k-tiered where the partition has k sets. (So a partisan function is 2-tiered.)
(c) c is *chaotic* iff it is both imbalanced and not partisan.

2.4. **Definition.** For a subset $\mathscr{C} \subseteq \mathfrak{C}$,

(a) $\mathscr{C}$ is *trivial* iff $\mathscr{C} = \emptyset$ or $\mathscr{C} = \{c\}$ where $\operatorname{dom} c = \emptyset$, i.e. c makes no decisions.
(b) $\mathscr{C}$ is *balanced* iff every $c \in \mathscr{C}$ is balanced; $\mathscr{C}$ is *imbalanced* iff $\mathscr{C}$ is not balanced.
(c) $\mathscr{C}$ is *partisan* iff every $c \in \mathscr{C}$ is partisan; $\mathscr{C}$ is *nonpartisan* iff $\mathscr{C}$ is not partisan.
(d) $\mathscr{C}$ is *chaotic* iff there is some chaotic $c \in \mathscr{C}$.
(e) pr-cl($\mathscr{C}$) is the convex hull of pr($\mathscr{C}$), i.e.

$$\text{pr-cl}(\mathscr{C}) = \left\{ \sum_{i=1}^{k} r_i \bar{t}_i : k \in \mathbb{N}, r_i \in [0,1]_{\mathbb{Q}}, \sum_{i=1}^{k} r_i = 1, \bar{t}_i \in \text{pr}(\mathscr{C}) \right\}.$$

We can now establish some straightforward results which allow us to describe the possible outcomes due to voters chosen from a given symmetric set more explicitly.

2.5. **Claim.** *If $\mathscr{C}$ is a symmetric subset of $\mathfrak{C}$, then d is the strict simple majority outcome of some finite set $\{c_1, \ldots, c_m\}$ chosen from $\mathscr{C}$ iff*

$$d\{x, y\} = x \iff \sum_{i=1}^{m} W_y^x(c_i) > 0.$$

Proof. d is the strict simple majority outcome iff

$$\begin{aligned} d\{x, y\} = x &\iff |\{i : c_i\{x, y\} = x\}| > |\{i : c_i\{x, y\} = y\}| \\ &\iff \sum_{i=1}^{m} W_y^x(c_i) > 0. \end{aligned}$$

□

2.6. **Claim.** *Suppose $\mathscr{C}$ is a symmetric subset of $\mathfrak{C}$. Then $d \in$ maj-cl($\mathscr{C}$) iff d is a strict simple majority outcome of some $\{c_1, \ldots, c_m\} \subseteq \mathscr{C}$.*

Proof. If $d \in \mathfrak{C}$ and there is a set $\{c_1, \ldots, c_m\} \subseteq \mathscr{C}$ with d as their strict simple majority outcome, then for each $c \in \mathscr{C}$ let r_c be the number of times that c appears among the c_i, divided by m:

$$r_c = \frac{|\{i : c_i = c\}|}{m}.$$

Clearly $\forall c \in \mathscr{C}$ $0 \le r_c \le 1$ and $\sum_{c \in \mathscr{C}} r_c = 1$. Furthermore,

$$d\{x, y\} = x \iff \sum_{i=1}^{m} W_y^x(c_i) > 0 \iff \sum_{i=1}^{m} \frac{W_y^x(c_i)}{m} > 0 \iff \sum_{c \in \mathscr{C}} W_y^x(c) r_c > 0.$$

Conversely, suppose $d \in$ maj-cl($\mathscr{C}$). There are $r_c = \frac{a_c}{b_c}$ for each $c \in \mathscr{C}$ with $\sum r_c = 1$, and

$$d\{x, y\} = x \iff \sum_{c \in \mathscr{C}} W_y^x(c) r_c > 0.$$

Let $m = \operatorname{lcm}\{b_c : c \in \mathscr{C}\}$. Now construct a finite set $\{c_1, \ldots, c_m\}$ by taking $a_c \frac{m}{b_c}$ copies of each $c \in \mathscr{C}$. $\frac{m}{b_c}$ is an integer since $b_c \mid m$. Now for each $c \in \mathscr{C}$,

$$\sum_{c_i = c} W_y^x(c_i) = W_y^x(c) \cdot |\{i : c_i = c\}| = W_y^x(c) a_c \frac{m}{b_c} = W_y^x(c) r_c m.$$

So

$$\sum_{i=1}^{m} W_y^x(c_i) = \sum_{c \in \mathscr{C}} \sum_{c_i = c} W_y^x(c_i) = \sum_{c \in \mathscr{C}} W_y^x(c) r_c m.$$

$$\sum_{c \in \mathscr{C}} W_y^x(c) r_c m > 0 \iff \sum_{c \in \mathscr{C}} W_y^x(c) r_c > 0,$$

so d is the simple majority relation of the c_i. $\square$

2.7. **Claim.** *For any symmetric* $\mathscr{C} \subseteq \mathfrak{C}$, $\text{maj-cl}(\mathscr{C}) = \{\operatorname{maj}(\bar{t}) : \bar{t} \in \text{pr-cl}(\mathscr{C})\}$.

Proof. Suppose $d \in \text{maj-cl}(\mathscr{C})$. Then there are r_c per $c \in \mathscr{C}$ with $\sum_{c \in \mathscr{C}} r_c = 1$ and

$$d\{x, y\} = x \iff \sum_{c \in \mathscr{C}} r_c W_y^x(c) > 0.$$

Let $\bar{t} = \sum_{c \in \mathscr{C}} r_c \operatorname{pr}(c)$ be a probability sequence in $\text{pr-cl}(\mathscr{C})$. Then

$$t_{x,y} = \sum_{c \in \mathscr{C}} r_c \operatorname{pr}(c)_{x,y} = \sum_{c \in \mathscr{C}} r_c W_y^x(c).$$

So

$$\operatorname{maj}(\bar{t})\{x, y\} = x \iff \sum_{c \in \mathscr{C}} r_c W_y^x(c) > 0,$$

i.e. $d = \operatorname{maj}(\bar{t})$.

Suppose $d = \operatorname{maj}(\bar{t})$ for some $\bar{t} \in \text{pr-cl}(\mathscr{C})$. Then for some r_c with $\sum_{c \in \mathscr{C}} r_c = 1$, $d = \operatorname{maj}\left(\sum_{c \in \mathscr{C}} r_c \operatorname{pr}(c)\right)$.

$$d\{x, y\} = x \iff \sum_{c \in \mathscr{C}} r_c \operatorname{pr}(c)_{x,y} > 0 \iff \sum_{c \in \mathscr{C}} r_c W_y^x(c) > 0.$$

So $d \in \text{maj-cl}(\mathscr{C})$. $\square$

3. A Characterization

3.1. Main Claim. *Given a symmetric* $\mathscr{C} \subseteq \mathfrak{C}$*,*

(i) $\mathscr{C}$ *is trivial* $\iff$ maj-cl$(\mathscr{C})$ *is trivial.*
(ii) $\mathscr{C}$ *is balanced but nontrivial* $\iff$ maj-cl$(\mathscr{C})$ *is all pseudo-balanced choice functions on* X*.*
(iii) $\mathscr{C}$ *is partisan and nontrivial* $\iff$ maj-cl$(\mathscr{C})$ *is all tiered choice functions on* X*.*
(iv) $\mathscr{C}$ *is imbalanced and nonpartisan* $\iff$ maj-cl$(\mathscr{C}) = \mathfrak{C}$.

Proof. No nontrivial $\mathscr{C}$ is both balanced and partisan, so the sets

$$\{\mathscr{C} \subset \mathfrak{C} : \mathscr{C} \text{ is trivial}\},$$
$$\{\mathscr{C} \subset \mathfrak{C} : \mathscr{C} \text{ is symmetric, nontrivial, and balanced}\},$$
$$\{\mathscr{C} \subset \mathfrak{C} : \mathscr{C} \text{ is symmetric, nontrivial, and partisan}\}, \text{ and}$$
$$\{\mathscr{C} \subseteq \mathfrak{C} : \mathscr{C} \text{ is symmetric, imbalanced, and nonpartisan}\}$$

partition the family of all symmetric subsets of $\mathfrak{C}$. Thus proving each forward implication in the claim will give the reverse implications.

1. If no-one in $\mathscr{C}$ makes any choices, no combination can have a majority choice.
2. This is the content of Section 4, below.
3. Suppose $\mathscr{C}$ is partisan and $d \in$ maj-cl$(\mathscr{C})$. Then there is a finite set $\{c_1, \dots, c_m\} \subseteq \mathscr{C}$ with d as the strict simple majority outcome. For each $x \in X$, let k_x be the number of c_i such that x is in the "winning" partite set. For any $(x, y) \in X^2$, let k be the number of c_i where x and y are both in the winning subset; the number of c_i with $c_i\{x, y\} = x$ is $k_x - k$, and the number with $c_i\{x, y\} = y$ is $k_y - k$, so

$$d\{x, y\} = x \iff k_x - k > k_y - k \iff k_x > k_y.$$

Partition X into subsets for each value of k_x,

$$X = \bigcup\{\{x \in X : k_x = k\} : k \in \{k_x : x \in X\}\}.$$

These subsets form the tiers, so d is a tiered function.

For the reverse inclusion, suppose d is an arbitrary tiered function with tiers $\{X_1, \dots, X_k\}$. Let $c \in \mathscr{C}$. For each $x \in X$, let $\Gamma_x \subset \mathrm{Per}(X)$ be all permutations holding x fixed and c_x be a choice function in $\mathscr{C}$ with x in its winning subset. (If y is in the winning subset of c, $c^{(x,y)}$ will suffice for c_x.) Now construct $\{c_1, \dots, c_m\}$ by taking, for $1 \le i \le k$, each $x \in X_i$, and each $\sigma \in \Gamma_x$, i copies of c_x^σ. Suppose $x \in X_i$, $y \in X_j$, and $i > j$ (so $d\{x, y\} = x$). Then for any $z \in X_i \setminus \{x\}$, any $z \in X_j \setminus \{y\}$, and any z in other tiers, x and y are chosen by equal numbers of $\{c_z^\sigma : \sigma \in \Gamma_z\}$. So they are chosen by an equal number of $\{c_1, \dots, c_m\}$ except among the functions c_x^σ and c_y^σ; there are i such functions c_x^σ for each $\sigma \in \Gamma_x$, and j such functions c_y^σ for each $\sigma \in \Gamma_y$. The number of permutations fixing x and the number fixing y are the same, $|\Gamma_x| = |\Gamma_y|$. The number of permutations fixing x and putting y in the losing subset and the number fixing y and putting x in the losing subset is the same, l. There are il functions which choose x over y, and jl which choose y over x; $il > jl$ so the strict simple majority outcome of the c_i chooses x.

If $i = j$, then $il = jl$, so x and y tie in the majority outcome. Thus d is the strict simple majority outcome of $\{c_1, \dots, c_m\}$ and $d \in$ maj-cl$(\mathscr{C})$.

4. If $\mathscr{C}$ is a chaotic set, the conclusion follows from Section 5, below.

Otherwise, $\mathscr{C}$ is imbalanced and nonpartisan, but not chaotic. In this case, $\mathscr{C}$ contains imbalanced functions which are all partisan, and nonpartisan functions which are all balanced. By symmetry, $\mathscr{C}$ contains a nontrivial balanced symmetric subset $\mathscr{B}$ and a nontrivial partisan symmetric subset $\mathscr{P}$.

Let $d \in \mathfrak{C}$. Suppose $b \to a$ is any edge of d. For each $z \in X \setminus \{a, b\}$, by Claim 4.5 there is a voter population T_z from $\mathscr{B}$ such that a beats z, z beats b, b beats a, and no other pairs are decided; moreover, the same number of voters, say m, choose the winner in each pair, and no dissenting votes occur. In the combination $\bigcup_z T_z$, m voters choose a over each z, and each z over b. $m(\mathbf{n} - 2)$ voters choose b over a.

Let $c \in \mathscr{P}$ be a partisan function, and l be the size of the set of winning candidates under c. Let $A \subset \mathscr{P}$ be the set of functions symmetric to c so that a is on the losing side, $B \subset \mathscr{P}$ be the functions symmetric to c so that b is on the winning side, and $C = A \cap B$—the permutations with b winning and a losing.

$$|A| = \binom{\mathbf{n}-1}{l}, \quad |B| = \binom{\mathbf{n}-1}{l-1}, \quad |C| = \binom{\mathbf{n}-2}{l-1}.$$

Any candidates besides a or b are tied over all the voters of A, B, or C, since each is selected in the winning set an equal number of times. Furthermore, a does not beat b in any choice function in A, B, or C.

Now we seek to take appropriate numbers of copies of the sets A, B, and C so that, for some constant k, any $z \in X \setminus \{a, b\}$ defeats a by k votes, and loses to b by k votes. Say we have k_0 copies of A, k_1 copies of B, and k_2 copies of C in a population D. Now for each $z \in X \setminus \{a, b\}$,

$$\begin{aligned}
|\{c \in D : c\{z, a\} = z\}| &= k_0\binom{\mathbf{n}-2}{l-1} + k_1\binom{\mathbf{n}-3}{l-2} + k_2\binom{\mathbf{n}-3}{l-2}, \\
|\{c \in D : c\{z, a\} = a\}| &= k_1\binom{\mathbf{n}-3}{l-2}, \\
|\{c \in D : c\{b, z\} = b\}| &= k_0\binom{\mathbf{n}-3}{l-1} + k_1\binom{\mathbf{n}-2}{l-1} + k_2\binom{\mathbf{n}-3}{l-1}, \\
|\{c \in D : c\{b, z\} = z\}| &= k_0\binom{\mathbf{n}-3}{l-1}.
\end{aligned}$$

So we solve $k_0\binom{\mathbf{n}-2}{l-1} + k_2\binom{\mathbf{n}-3}{l-2} = k_1\binom{\mathbf{n}-2}{l-1} + k_2\binom{\mathbf{n}-3}{l-1}$, with coefficients not all zero. There are solutions

$$\begin{aligned}
k_0 &= \mathbf{n} - 2l, & k_1 &= 0, & k_2 &= \mathbf{n} - 2, & &\text{if } l \leq \frac{\mathbf{n}}{2}, \\
k_0 &= 0, & k_1 &= 2l - \mathbf{n}, & k_2 &= \mathbf{n} - 2, & &\text{if } l > \frac{\mathbf{n}}{2}.
\end{aligned}$$

Let $k = k_0\binom{\mathbf{n}-2}{l-1} + k_2\binom{\mathbf{n}-3}{l-2}$, the number of votes for b over z, or z over a, for any $z \in X \setminus \{a, b\}$. Now take the union of k copies of the population $\bigcup_z T_z$, mk_0 copies of A, mk_1 copies of B, and mk_2 copies of C for our voter population. b defeats any z mk times among the partisan functions, and z defeats b km times among the balanced functions, so they tie; similarly, a and z tie. Thus we are left with $b \to a$. The union of such populations for each edge in d yields d as the majority outcome.

Hence $\mathfrak{C} \subseteq \text{maj-cl}(\mathscr{C})$. □

4. Balanced Choices

4.1. **Definition.** For a choice function $c \in \mathfrak{C}$,

(a) c is *triangular* iff for some $\{x,y,z\} \in \binom{X}{3}$, $c\{x,y\} = x$, $c\{y,z\} = y$, $c\{z,x\} = z$, and for any $\{u,v\} \not\subset \{x,y,z\}$, $\{u,v\} \notin \operatorname{dom} c$. We write $c^{x,y,z}$ for such c.

(b) c is *cyclic* iff for some $\{x_1,\dots,x_k\} \in \binom{X}{k}$,
- $c\{x_i,x_j\} = x_i$ iff $j \equiv i+1 \mod k$.
- No other pair $\{x,y\}$ is in $\operatorname{dom} c$.

We write $c^{x_1,\dots,x_k}$ for such c. (Triangular functions are a specific case of cyclic functions.)

(c) Let $\mathrm{pr}^{\mathrm{bl}}$ be the set of all balanced $\bar{t} \in \mathrm{pr}(X)$.

Note that if $\mathscr{C}$ is symmetric and $c^{x,y,z} \in \mathscr{C}$, then $c^{u,v,w} \in \mathscr{C}$ for any $\{u,v,w\} \in \binom{X}{3}$; similarly for cyclic functions.

4.2. **Claim.** *If a choice function $c = \mathrm{maj}(\bar{t})$ for some $\bar{t} \in \mathrm{pr}^{\mathrm{bl}}$, then c is pseudo-balanced.*

Proof. Assume $c = \mathrm{maj}(\bar{t})$ for some $\bar{t} \in \mathrm{pr}^{\mathrm{bl}}$. Let $x \to y$ be any edge of $\mathrm{Tor}(c)$; then $t_{x,y} > 0$. Suppose $x \to y$ is in no directed cycle. Let Y be the set of $z \in X$ with a winning chain to x, i.e.

$$Y = \bigcup \left\{ \{z_1,\dots,z_k\} \in \binom{X \setminus \{y\}}{k} : k < \mathbf{n}, \forall i < k \ c\{z_i,z_{i+1}\} = z_i, z_k = x \right\}.$$

Let $\overline{Y} = X \setminus Y$. Note that $y \in \overline{Y}$ and $x \in Y$, so Y and $\overline{Y}$ partition X into nonempty sets. Suppose $z \in Y$ and $v \in \overline{Y}$. Say $z, z_1, \dots, z_k$ is a winning chain from z to x. If $t_{z,v} < 0$, then $v, z, z_1, \dots, z_k$ forms a winning chain from v to x. If v is y, the chain forms a directed cycle with the edge $x \to y$ and we are done. If $v \neq y$, this contradicts that $v \notin Y$. So we assume that $t_{z,v} \geq 0$.

Now

$$\sum_{z \in Y, v \in \overline{Y}} t_{z,v} > 0,$$

since every $t_{z,v} \geq 0$, and $t_{x,y} > 0$ is among them. For each $u \in \overline{Y}$ let

$$r_u = \sum_{z \in Y} t_{z,u},$$
$$\bar{r}_u = \sum_{v \in \overline{Y}} t_{v,u}.$$

Since $\bar{t}$ is balanced, $r_u + \bar{r}_u = 0$, so

$$\sum_{u \in \overline{Y}} (r_u + \bar{r}_u) = 0 = \sum_{u \in \overline{Y}} r_u + \sum_{u \in \overline{Y}} \bar{r}_u.$$

We have seen that the first summand is positive, so the second summand is negative. But it is zero because for each pair $(u,v) \in \overline{Y}^2$ we have $t_{u,v} + t_{v,u} = 0$. This is a contradiction; so $x \to y$ must be in some directed cycle. $\square$

4.3. **Claim.** $\mathrm{pr}^{\mathrm{bl}}$ *is a convex subset of* $\mathrm{pr}(X)$.

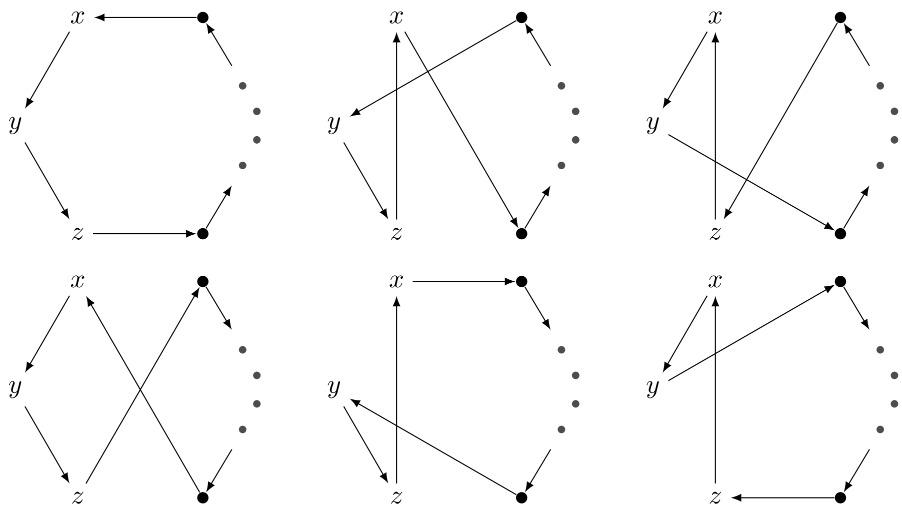

FIGURE 1. Constructing a triangle from permutations of a cycle, as in Claim 4.5. Note that every edge is matched by an opposing one, except for xy, yz, and zx.

Proof. Suppose $\bar{t}$ and $\bar{s}$ are balanced probability sequences, and $a \in [0,1]_{\mathbb{Q}}$. Then $\bar{u} = a\bar{t} + (1-a)\bar{s}$ has, for any $x \in X$,

$$\begin{aligned}\sum_{y\in X} u_{x,y} &= \sum_{y\in X} (at_{x,y} + (1-a)s_{x,y}) \\ &= a\sum_{y\in X} t_{x,y} + (1-a)\sum_{y\in X} s_{x,y} \\ &= a0 + (1-a)0 \;=\; 0.\end{aligned}$$

So $\bar{u}$ is a balanced probability sequence. □

4.4. **Claim.** *If $\mathscr{C} \subset \mathfrak{C}$ is symmetric, balanced, and nontrivial, then every $d \in$ maj-cl$(\mathscr{C})$ is pseudo-balanced.*

Proof. By Claim 2.2, every $\bar{t} \in \mathrm{pr}(\mathscr{C})$ is balanced. So by Claim 4.3, the convex hull pr-cl$(\mathscr{C})$ is contained in $\mathrm{pr}^{\mathrm{bl}}$. Any $d \in$ maj-cl$(\mathscr{C})$ is maj$(\bar{t})$ for some $\bar{t} \in$ pr-cl$(\mathscr{C})$, by Claim 2.7, hence is pseudo-balanced, by Claim 4.2. □

For the reverse inclusion, that every pseudo-balanced function is in the majority closure, we shall consider the graph interpretation. First we shall show that every triangular function is in the majority closure of a balanced symmetric set.

4.5. **Claim.** *If $\mathscr{C} \subset \mathfrak{C}$ is symmetric, balanced, and nontrivial, then for any $\{x, y, z\} \in \binom{X}{3}$, the triangular function $c^{x,y,z} \in$ maj-cl$(\mathscr{C})$. Moreover, there is a voter population generating $c^{x,y,z}$ such that the same number of votes choose x over y, y over z, or z over x, and there are no opposing votes in any of these cases.*

Proof. Let $c_0 \in \mathscr{C}$ and let m be the size of the smallest directed cycle in Tor(c_0). Index the distinct elements of this cycle, in order, as $x_1, x_2, \ldots, x_m$. Let $c = c_0^{\sigma}$ where $\sigma \in$ Per(X) takes x_1 to x, x_2 to y, and x_3 to z. Now consider the cycle in

$\mathrm{Tor}(c)$, so $x_1 = x$, $x_2 = y$, and $x_3 = z$. We define

$$\begin{aligned} c_1 &= c, \\ c_2 &= c^\rho \text{ where } \rho \text{ takes } x \mapsto y, y \mapsto z, z \mapsto x, \\ c_3 &= c^\rho \text{ where } \rho \text{ takes } x \mapsto z, y \mapsto x, z \mapsto y. \end{aligned}$$

Let $\Gamma_{x,y,z} \subset \mathrm{Per}(X)$ be the permutations fixing x, y, and z, and take $\{c_i^\sigma : 1 \leq i \leq 3, \sigma \in \Gamma_{x,y,z}\}$ as our finite set of voters. We claim the majority outcome of this set is $c^{x,y,z}$.

If $u, v \in X \setminus \{x, y, z\}$, then equal numbers of permutations $\sigma \in \Gamma_{x,y,z}$ have $c_i^\sigma\{u, v\} = u$ and $c_i^\sigma\{u, v\} = v$, so they are tied. If u is among $\{x, y, z\}$ and v is not, then there are some out-edges from u in $\mathrm{Tor}(c)$. Since c is balanced, there are an equal number of in-edges to u. For each $i \in \{1, 2, 3\}$, v will occupy the in-edges which are not on the cycle in as many permutations of c_i as permutations where it occupies out-edges which are not on the cycle. If $m = 3$, then v occupies no edges on the cycle. Otherwise, v occupies an out-edge from u along the cycle only in

- c_1^σ, for some set of $\sigma \in \Gamma_{x,y,z}$, if $u = z$; then v occupies an in-edge to z on the cycle in c_3 for an equal number of permutations.
- c_2^σ, for some set of $\sigma \in \Gamma_{x,y,z}$, if $u = x$; then v occupies an in-edge to x on the cycle in c_1 for an equal number of permutations.
- c_3^σ, for some set of $\sigma \in \Gamma_{x,y,z}$, if $u = y$; then v occupies an in-edge to y on the cycle in c_2 for an equal number of permutations.

Thus u and v are tied. Otherwise, $\{u, v\} \subset \{x, y, z\}$. For each $i \in \{1, 2, 3\}$, there are $|\Gamma_{x,y,z}|$ many c_i^σ. $c_i^\sigma\{x, y\} = x$ if $i = 1$ or $i = 3$; this is two-thirds of the voters, so x beats y. Similarly, $c_i^\sigma\{y, z\} = y$ if $i = 1$ or $i = 2$, and $c_i^\sigma\{z, x\} = z$ if $i = 2$ or $i = 3$. Thus the strict simple majority outcome is $c^{x,y,z}$.

The collection $\{c_i^\sigma : 1 \leq i \leq 3, \sigma \in \Gamma_{x,y,z}\}$ is the voter population in the claim. There are no opposing votes between x, y, or z because such a vote would imply an edge between two vertices of the cycle which is not itself on the cycle, contradicting our choice of the smallest cycle. $\square$

4.6. **Claim.** *If $\mathscr{C} \subseteq \mathfrak{C}$ and $c^{x,y,z} \in \text{maj-cl}(\mathscr{C})$ for any $\{x, y, z\} \in \binom{X}{3}$, then for any $k \geq 3$ and $\{x_1, \ldots, x_k\} \in \binom{X}{k}$, $c^{x_1,\ldots,x_k} \in \text{maj-cl}(\mathscr{C})$.*

Proof. If $k = 3$, we have $c^{x_1,x_2,x_3} \in \text{maj-cl}(\mathscr{C})$ by assumption.

Suppose that $3 < k \leq \mathbf{n}$ and for any $k-1$ distinct candidates $\{x_1, \ldots, x_{k-1}\} \subset X$ we have $c^{x_1,\ldots,x_{k-1}} \in \text{maj-cl}(\mathscr{C})$. Say $\{c_1, \ldots, c_{m_0}\}$ is a set of m_0 voters with $c^{x_1,\ldots,x_{k-1}}$ as their strict simple majority outcome. Let

$$l_0 = \sum_{i=1}^{m_0} W_{x_1}^{x_{k-1}}(c_i).$$

$l_0 > 0$ since $c^{x_1,\ldots,x_{k-1}}\{x_1, x_{k-1}\} = x_{k-1}$. By hypothesis $c^{x_1,x_{k-1},x_k} \in \text{maj-cl}(\mathscr{C})$. Say $\{c'_1, \ldots, c'_{m_1}\}$ is a set of m_1 voters with c^{x_1,x_{k-1},x_k} as their strict simple majority outcome. Let

$$l_1 = \sum_{i=1}^{m_1} W_{x_{k-1}}^{x_1}(c'_i).$$

Like l_0, l_1 is a positive integer. Let $L = \mathrm{lcm}(l_0, l_1)$, and take $\frac{L}{l_0}$ copies of each c_i, and $\frac{L}{l_1}$ copies of each c'_i, to make a finite set of voters $\{d_1, \ldots, d_m\}$ where $m =$

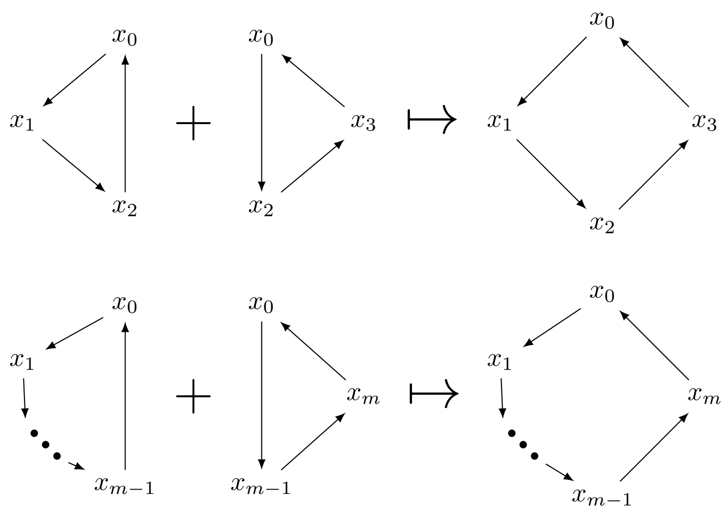

FIGURE 2. Constructing a cycle from triangles, as in Claim 4.6.

$\frac{L}{l_0}m_0 + \frac{L}{l_1}m_1$. Let d be their strict simple majority outcome. For any $\{u, v\} \in \binom{X}{2}$, if $\{u, v\} \not\subset \{x_1, \ldots, x_k\}$ then u ties v among the c_i and the c_i', so $\{u, v\} \notin \operatorname{dom} d$. Any two points on the cycle $(x_1, \ldots, x_{k-1})$ which are not adjacent are tied among the c_i and among the c_i'. Any point on the cycle besides x_1 or x_{k-1} is also tied with x_k among the c_i and among the c_i'. Any pair of consecutive points (x_i, x_j) on the cycle, besides (x_{k-1}, x_1), are tied among the c_i', but the majority of the c_i pick x_i. So $d\{x_i, x_j\} = x_i$. x_{k-1} and x_k tie among the c_i, but the majority of the c_i' pick x_{k-1}, so $d\{x_{k-1}, x_k\} = x_{k-1}$. Similarly $d\{x_k, x_1\} = x_k$. The only remaining pair to consider is $\{x_1, x_{k-1}\}$. $d\{x_1, x_{k-1}\}$ is defined iff

$$\sum_{i=1}^{m} W^{x_1}_{x_{k-1}}(d_i) \neq 0.$$

The left hand side is equal to

$$\begin{aligned} \frac{L}{l_0}\sum_{i=1}^{m_0} W^{x_1}_{x_{k-1}}(c_i) + \frac{L}{l_1}\sum_{i=1}^{m_1} W^{x_1}_{x_{k-1}}(c_i') \\ = \frac{L}{l_0}(-l_0) + \frac{L}{l_1}(l_1) = 0. \end{aligned}$$

Thus $\{x_1, x_{k-1}\} \notin \operatorname{dom} d$, and $d\{x_i, x_j\} = x_i$ iff $j \equiv i + 1 \bmod k$, so $d = c^{x_1,\ldots,x_k} \in$ maj-cl$(\mathscr{C})$.

By induction on k, any cyclic choice function on X is in maj-cl$(\mathscr{C})$. □

4.7. **Claim.** *If $\mathscr{C} \subset \mathfrak{C}$ is symmetric, balanced, and nontrivial, then every pseudo-balanced choice function $d \in \mathfrak{C}$ is in* maj-cl$(\mathscr{C})$.

Proof. Suppose d is an arbitrary pseudo-balanced choice function. Every edge of Tor(d) is on a directed cycle, so consider the decomposition of Tor(d) into cycles $C_{x,y}$ for each edge $x \to y$ in Tor(d). Let $d_{x,y}$ be the cyclic function corresponding to $C_{x,y}$ for each edge $x \to y$ in Tor(d). By Claim 4.5, $\mathscr{C}$ meets the requirements of Claim 4.6, so every cyclic function on X is in maj-cl$(\mathscr{C})$; in particular, each

$d_{x,y} \in \text{maj-cl}(\mathscr{C})$. If we combine all the finite sets of voters which yielded each $d_{x,y}$, we get another finite set of voters from $\mathscr{C}$, since there are finitely many edges in $\text{Tor}(d)$.

For any $\{u, v\} \in \binom{X}{2} \setminus \operatorname{dom} d$, the edge $u \to v$ is not in any cycle $C_{x,y}$. $\sum W_v^u(c_i) = 0$ over the c_i yielding any $d_{x,y}$, so $\sum W_v^u(c_i) = 0$ over all our voters and u and v tie in their strict simple majority outcome.

For any $\{u, v\} \in \binom{X}{2}$ with $d\{u, v\} = u$, the edge $u \to v$ is in $\text{Tor}(d)$. Since each cycle $C_{x,y}$ has edges only from $\text{Tor}(d)$, each population of c_i yielding some $d_{x,y}$ has $\sum W_v^u(c_i) \geq 0$, and in particular the population yielding $d_{u,v}$ has $\sum W_v^u(c_i) > 0$, so over all our voters $\sum W_v^u(c_i) > 0$.

Thus the strict simple majority outcome is d, and $d \in \text{maj-cl}(\mathscr{C})$. $\square$

Combining Claims 4.4 and 4.7, we have that $\text{maj-cl}(\mathscr{C})$ is the set of all pseudo-balanced choice functions whenever $\mathscr{C}$ is symmetric, balanced, and nontrivial.

5. Chaotic Choices

5.1. **Definition.** For a choice function $c \in \mathfrak{C}$,

(a) For $x \in X$, the *valence* of x with c is

$$\mathrm{val}_c(x) = \sum_{y \in X} W_y^x(c).$$

(b) For $\ell \in \{-1, 0, 1\}$, let

$$V_\ell(c) = \{(\mathrm{val}_c(x) - \ell, \mathrm{val}_c(y) + \ell)) : \{x, y\} \in \binom{X}{2},\ W_y^x(c) = \ell\}.$$

These are the valence pairs, with winners on the right in V_{-1}, ties in V_0, winners on the left in V_1, and with 1 subtracted from the valence of the winner. In fact,

$$V_\ell(c) = \left\{ \sum_{z \in X \setminus \{x,y\}} (W_z^x(c), W_z^y(c)) : \{x, y\} \in \binom{X}{2},\ W_y^x(c) = \ell \right\}.$$

(c) For a subset A of $\mathbb{Q} \times \mathbb{Q}$, let $\mathrm{conv}(A)$ be the convex hull of A in $\mathbb{Q} \times \mathbb{Q}$.

(d) Let

$$V^*(c) = \{a\bar{v}_1 + (1-a)\bar{v}_0 : \bar{v}_1 \in \mathrm{conv}\,(V_1(c))\,, \bar{v}_0 \in \mathrm{conv}\,(V_0(c))\,, a \in (0,1]_{\mathbb{Q}}\}.$$

These are the convex hulls of the "winning" valence pairs, together with ties, requiring some contribution from a non-tied pair.

5.2. **Claim.** *$(k_0, k_1) \in V_\ell(c) \iff (k_1, k_0) \in V_{-\ell}(c)$, for any $\ell \in \{-1, 0, 1\}$.*

Proof. A pair $(\mathrm{val}_c(u) + 1, \mathrm{val}_c(v) - 1)$ is in $V_{-1}(c)$ iff $W_v^u(c) = -1 \iff W_u^v(c) = 1$, so $(\mathrm{val}_c(v) - 1, \mathrm{val}_c(u) + 1) \in V_1(c)$.

A pair $(\mathrm{val}_c(u), \mathrm{val}_c(v))$ is in $V_0(c)$ iff $\{u, v\} = \{v, u\} \notin \mathrm{dom}\, c$, so $(\mathrm{val}_c(v), \mathrm{val}_c(u)) \in V_0(c)$. □

5.3. **Claim.** *An imbalanced $c \in \mathfrak{C}$ is partisan iff*

- *$V_1(c)$ lies on a line parallel to the line $y = x$, and*
- *$V_0(c)$ is contained in the line $y = x$.*

Proof. Suppose that $V_0(c)$ is all on $y = x$. Thus, candidates are only ever tied with others of the same valence. Suppose further that $V_1(c)$ is on a line $y = x - b$. Then whenever $c\{w, v\} = w$, $\mathrm{val}_c(w) - \mathrm{val}_c(v) = b + 2$ is constant; the valence of any winner is always $b + 2$ more than the defeated. Since c is imbalanced, there is such a pair $\{w, v\}$. If any candidate z has a valence different from that of w, it cannot be tied with w, so $\{w, z\} \in \mathrm{dom}\, c$. Therefore the valences differ by $b + 2$, but if $\mathrm{val}_c(z) \neq \mathrm{val}_c(v)$, then z can neither tie nor be comparable to v, a contradiction. So every candidate has the valence of w or the valence of v. If two candidates with the same valence do not tie, then $b = -2$ and $\mathrm{val}_c(v) = \mathrm{val}_c(w)$. In this case, every candidate has the same valence, which would mean that c is balanced; a contradiction. Furthermore, every candidate with the high valence must defeat everyone of the low valence, since they cannot be tied. This is the definition of a partisan function.

Conversely, suppose that c is partisan. Then two elements are tied only if they are both in the winning subset, or both in the losing subset; in either case, they have the same valence, so $V_0(c)$ is contained in the line $y = x$. $V_1(c)$ is the single point $(\mathrm{val}_c(w)-1, \mathrm{val}_c(v)+1)$ for any w in the winning subset and v in the losing subset. Naturally, this point is contained in a line parallel to $y = x$. $\square$

5.4. **Claim.** *If $c \in \mathfrak{C}$ is imbalanced, then a point of $V^*(c)$ lies above the line $y = -x$, and a point of $V^*(c)$ lies below it.*

Proof. Let v_1 and v_2 be two candidates in X having the highest valences with c. Since c is imbalanced, $\mathrm{val}_c(v_1) > 0$. The sum of the corresponding pair in any $V_\ell(c)$ is $\mathrm{val}_c(v_1) + \mathrm{val}_c(v_2)$. If this is less than or equal to 0, then $\mathrm{val}_c(v_2) \leq -1$. Thus all other valences are at most -1, so the average valence is strictly smaller than the average of $\mathrm{val}_c(v_1)$ and $\mathrm{val}_c(v_2)$, at most 0. This is a contradiction, since the average valence is always 0. Therefore, the sum of the valence pair for v_1 and v_2 in any $V_\ell(c)$ satisfies $y + x > 0$. If the pair is in $V_1(c)$, it is in $V^*(c)$. If it is in $V_{-1}(c)$, then its reflection in $V_1(c)$ (hence in $V^*(c)$) still satisfies $x + y > 0$. If it is in $V_0(c)$, then points arbitrarily close to the pair, along any line to a point of $V_1(c)$, are in $V^*(c)$. Some of these points are above $y = -x$.

Let u_1 and u_2 be two candidates in X having the smallest valences with c. By imbalance, $\mathrm{val}_c(u_1) < 0$. Suppose $\mathrm{val}_c(u_1) + \mathrm{val}_c(u_2) \geq 0$. Then the average of these two valences is at least 0, and the average over all candidates is larger; it is strictly larger, since $\mathrm{val}_c(v_1) > 0$ figures in the average. This is a contradiction, so the sum of the valence pair for u_1 and u_2 in any $V_\ell(c)$ satisfies $y + x < 0$. If this pair is in $V_{-1}(c)$, then its reflection in $V_1(c)$ still satisfies $x + y < 0$. Otherwise it is in $V^*(c)$, or arbitrarily close points are in $V^*(c)$. $\square$

To establish that every choice function is in the majority closure of a chaotic symmetric set, we shall first establish that some c therein satisfies a rather abstruse condition which we'll call "valence-imbalance."

5.5. **Definition.** A choice function $c \in \mathfrak{C}$ is *valence-imbalanced* iff $(0,0)$ can be represented as $r_{-1}\bar{v}_{-1} + r_0\bar{v}_0 + r_1\bar{v}_1$ where

(i) Each $\bar{v}_\ell$ is a pair in $\mathrm{conv}\,(V_\ell(c))$,
(ii) $r_{-1}, r_0, r_1 \in [0,1]_{\mathbb{Q}}$,
(iii) $r_{-1} + r_0 + r_1 = 1$,
(iv) $r_{-1} \neq r_1$.

Note that $(0,0) \in V^*(c)$ implies that c is valence-imbalanced.

5.6. **Claim.** *If $\bar{u}$ is strictly between two points of* $\mathrm{conv}\,(V_1(c))$ *on a line segment not parallel to $y = x$, then the nearest point on $y = x$ to $\bar{u}$ is of the form $r_{-1}\bar{v}_{-1} + r_1\bar{v}_1$, where $\bar{v}_\ell \in \mathrm{conv}\,(V_\ell(c))$, $r_\ell \in [0,1]_{\mathbb{Q}}$, $r_{-1} + r_1 = 1$, but $r_{-1} \neq r_1$.*

Proof. Say $\bar{u} = (a, a+b)$. If $b = 0$, the assertion holds. Otherwise, let $\bar{w} = \left(a + \frac{b}{2}, a + \frac{b}{2}\right)$; $\bar{w}$ is the nearest point to $\bar{u}$ on $y = x$.

There are $\bar{p}_0 = (a_0, b_0)$ and $\bar{p}_1 = (a_1, b_1)$ in $\mathrm{conv}\,(V_1(c))$ so that $\bar{u}$ is strictly between $\bar{p}_0$ and $\bar{p}_1$. If the line segment Λ between $\bar{p}_0$ and $\bar{p}_1$ is perpendicular to $y = x$, then either $\bar{w}$ is in Λ, and the assertion holds, or $\bar{w}$ is between $\bar{p}_1$ and the reflection of $\bar{p}_0$, $\bar{q}_0 = (b_0, a_0)$. So for some r_{-1} and r_1, $r_{-1}\bar{q}_0 + r_1\bar{p}_1 = \bar{w} = \frac{1}{2}\bar{q}_0 + \frac{1}{2}\bar{p}_0$. If $r_{-1} = r_1 = \frac{1}{2}$, then $\bar{p}_1 = \bar{p}_0$, contradicting that $\bar{u}$ is strictly between them. So $r_{-1} \neq r_1$.

If Λ is not perpendicular to the line $y = x$, then we may choose the points $\bar{p}_0$ and $\bar{p}_1$ so that $a_0 + b_0 < 2a + b < a_1 + b_1$, and on the same side of $y = x$. The reflection of Λ over $y = x$ lies on a line, say L. L contains $\bar{v} = (a+b, a)$, $\bar{q}_0 = (b_0, a_0)$,

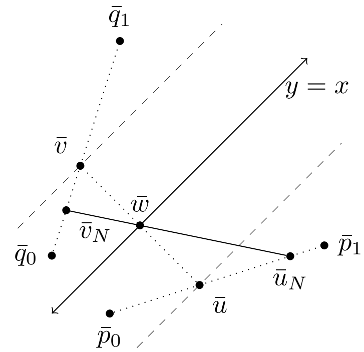

FIGURE 3. The situation of Claim 5.6. $\bar{u}_N$ and $\bar{v}_N$ cannot be equidistant from $\bar{w}$, since one lies within the dashed parallel lines, and one lies without.

and $\bar{q}_1 = (b_1, a_1)$, all in $\operatorname{conv}(V_{-1}(c))$; so $\bar{v}$ lies strictly between $\bar{q}_0$ and $\bar{q}_1$ within $\operatorname{conv}(V_{-1}(c))$.

Consider the sequence

$$\bar{u}_n = \frac{1}{n}\bar{p}_1 + \left(1 - \frac{1}{n}\right)\bar{u},$$

for $n \geq 1$, approaching $\bar{u}$ as $n \to \infty$. For each n, let $\bar{v}_n$ be the intersection of the line through $\bar{u}_n$ and $\bar{w}$ with the line L. (If necessary, define the $\bar{v}_n$ only for $n > j$, if the line from $\bar{u}_j$ through $\bar{w}$ is parallel to L; at most one such j can exist.) As $n \to \infty$, $\bar{v}_n$ approaches $\bar{v}$. So for some $N \in \mathbb{N}$ and all $n \geq N$, $\bar{v}_n$ is in the interval $(\bar{q}_0, \bar{q}_1)$, an open neighborhood of $\bar{v}$ in L. So

$$\bar{w} = r_{-1}\bar{u}_N + r_1\bar{v}_N,$$

for some r_{-1} and r_1 in $[0,1]_{\mathbb{Q}}$ with $r_{-1} + r_1 = 1$.

If $r_{-1} = r_1 = \frac{1}{2}$, then $\bar{v}_N = \bar{w} + (\bar{w} - \bar{u}_N)$ is just as far from the line $y = x$ as $\bar{u}_N = \bar{w} - (\bar{w} - \bar{u}_N)$. But notice, either

- both the interval $(\bar{u}, \bar{p}_1)$ on Λ and the interval $(\bar{v}, \bar{q}_1)$ on L are farther from $y = x$ than $\bar{u}$ is, while the intervals $(\bar{p}_0, \bar{u})$ and $(\bar{q}_0, \bar{v})$ are closer, or
- the intervals $(\bar{u}, \bar{p}_1)$ and $(\bar{v}, \bar{q}_1)$ are closer to $y = x$ than $\bar{u}$ is, while the intervals $(\bar{p}_0, \bar{u})$ and $(\bar{q}_0, \bar{v})$ are farther.

$\bar{u}_N$ is in the half-plane $y + x > 2a + b$, and $\bar{w}$ is the sole intersection of the line $\{\bar{u}_N + t(\bar{w} - \bar{u}_N) : t \in \mathbb{Q}\}$ with the line $y + x = 2a + b$. Thus $\bar{v}_N$ is in the half-plane $y + x < 2a + b$, hence in the interval $(\bar{q}_0, \bar{v})$ of L. But then the distance from $y = x$ to $\bar{u}$ is strictly between the distances from $y = x$ to $\bar{v}_N$ and to $\bar{u}_N$, contradicting that these distances are equal.

So $r_{-1} \neq r_1$. $\square$

5.7. **Claim.** *If c is chaotic, then c is valence-imbalanced.*

Proof. We consider five cases.

Case 1. $V_0(c) \setminus \{(0,0)\}$ is not contained in $y = x$, nor in one open half-plane of $y = -x$.

By Claim 5.4, there is a point $\bar{v}$ of $V^*(c)$ on the line $y = -x$. There is a point $(k_0, -k_0) \in \operatorname{conv}(V_0(c))$, $k_0 \neq 0$. Either $(0,0)$ is between $\bar{v}$ and $(k_0, -k_0)$ or it is between $\bar{v}$ and $(-k_0, k_0)$; either way, it is in $V^*(c)$, so c is valence-imbalanced.

Case 2. $V_0(c) \setminus \{(0,0)\}$ is not contained in $y = x$, but is contained in one open half-plane of $y = -x$.

There is $(k_0, k_1) \in V_0(c)$ such that $k_0 \neq k_1$ and $k_0 \neq -k_1$. By Claim 5.4, there is a point $\bar{v} \in V_1(c)$ strictly on the other side of $y = -x$. If $\bar{v}$ is on $y = x$, then $(0,0)$ is on the line segment between $\bar{v}$ and $\bar{w} = \left(\frac{k_0+k_1}{2}, \frac{k_0+k_1}{2}\right)$, hence in $V^*(c)$, so c is valence-imbalanced. If $\bar{v}$ is not on $y = x$, then it has a reflection $\bar{u} \in V_{-1}(c)$. $\operatorname{conv}\{\bar{v}, \bar{u}, \bar{w}\}$ contains an open disc about $(0,0)$. Since there are points in $\operatorname{conv}(V_0(c))$ arbitrarily close to $\bar{w}$ on the line between (k_0, k_1) and (k_1, k_0), we may choose (k_0', k_1') so $k_0' \neq k_1'$ and $(0,0) \in \operatorname{conv}\{\bar{v}, \bar{u}, (k_0', k_1')\}$.

Suppose c is valence-balanced. Say $\bar{v}$ is (v_0, v_1). Then

$$(0,0) = r\bar{v} + (1-2r)(k_0', k_1') + r\bar{u} = r(v_0 + v_1, v_1 + v_0) + (1-2r)(k_0', k_1'),$$

so $k_0' = \frac{-r(v_0+v_1)}{1-2r} = k_1'$, a contradiction. Therefore c is valence-imbalanced.

Case 3. $V_0(c)$ is contained in $y = x$, and $V_1(c)$ is contained in a line parallel to $y = x$.

Then c is partisan by Claim 5.3, contradicting that c is chaotic.

Case 4. $V_0(c)$ is contained in $y = x$, $\operatorname{conv}(V_1(c))$ contains a line segment not parallel to $y = x$, and $V_1(c)$ has points on either side of $y = -x$.

Then there is a point of $y = -x$ strictly between two points of $\operatorname{conv}(V_1(c))$ such that the line segment between them is not parallel to $y = x$. By Claim 5.6, $(0,0) = r_{-1}\bar{v}_{-1} + r_1\bar{v}_1$ where $\bar{v}_\ell \in \operatorname{conv}(V_\ell(c))$, $r_\ell \in [0,1]_{\mathbb{Q}}$, $r_{-1} + r_1 = 1$, but $r_{-1} \neq r_1$, i.e. c is valence-imbalenced.

Case 5. $V_0(c)$ is contained in $y = x$, $\operatorname{conv}(V_1(c))$ contains a line segment not parallel to $y = x$, and $V_1(c)$ is entirely on one side of $y = -x$.

Let $\bar{v}$ be strictly between two points of $V_1(c)$ on a line segment not parallel to $y = x$. By Claim 5.6, the nearest point $\bar{w}$ to $\bar{v}$ on the line $y = x$ is $r_{-1}\bar{v}_{-1} + r_1\bar{v}_1$ where $\bar{v}_\ell \in \operatorname{conv}(V_\ell(c))$, $r_\ell \in [0,1]_{\mathbb{Q}}$, $r_{-1} + r_1 = 1$, but $r_{-1} \neq r_1$. By Claim 5.4, $V_0(c)$ has a point $\bar{v}_0 = (k,k)$ strictly on the other side of $y = -x$ from $\bar{v}$, and hence from $\bar{w}$. So for some $r \in [0,1]_{\mathbb{Q}}$,

$$(0,0) = r\bar{v}_0 + (1-r)(r_{-1}\bar{v}_{-1} + r_1\bar{v}_1) = (1-r)r_{-1}\bar{v}_{-1} + r\bar{v}_0 + (1-r)r_1\bar{v}_1,$$

and $(1-r)r_{-1} \neq (1-r)r_1$, so c is valence-imbalanced. $\square$

5.8. **Claim.** *If $\mathscr{C} \subseteq \mathfrak{C}$ is symmetric and $c \in \mathscr{C}$ is valence-imbalanced, then for each $\{x,y\} \in \binom{X}{2}$ there is $d \in \text{maj-cl}(\mathscr{C})$ such that* $\operatorname{dom} d = \{\{x,y\}\}$.

Proof. Suppose $\{x,y\} \in \binom{X}{2}$. There are $\bar{v}_\ell \in \operatorname{conv}(V_\ell(c))$ and $r_\ell \in [0,1]_{\mathbb{Q}}$, $r_{-1} + r_0 + r_1 = 1$, $r_{-1} \neq r_1$, so $r_{-1}\bar{v}_{-1} + r_0\bar{v}_0 + r_1\bar{v}_1 = (0,0)$.

$$\bar{v}_{-1} = \sum_{\bar{v} \in V_{-1}(c)} s_{\bar{v}}^{-1}\bar{v}, \quad \bar{v}_0 = \sum_{\bar{v} \in V_0(c)} s_{\bar{v}}^{0}\bar{v}, \quad \bar{v}_1 = \sum_{\bar{v} \in V_1(c)} s_{\bar{v}}^{1}\bar{v},$$

where $\sum_{\bar{v}\in V_\ell(c)} s^\ell_{\bar{v}} = 1$ for each ℓ. Each $\bar{v} \in V_\ell(c)$ is $(\mathrm{val}_c(u^\ell_{\bar{v}}) - \ell, \mathrm{val}_c(w^\ell_{\bar{v}}) + \ell)$ for some $\{u^\ell_{\bar{v}}, w^\ell_{\bar{v}}\} \in \binom{X}{2}$. Let $\sigma^\ell_{\bar{v}}$ be a permutation taking $u^\ell_{\bar{v}}$ to x and $w^\ell_{\bar{v}}$ to y, i.e.

$$\sigma^\ell_{\bar{v}} = \begin{cases} (x,y) & \text{if } u^\ell_{\bar{v}} = y, w^\ell_{\bar{v}} = x, \\ (x, w^\ell_{\bar{v}}, y) & \text{if } u^\ell_{\bar{v}} = y, w^\ell_{\bar{v}} \neq x, \\ (u^\ell_{\bar{v}}, x)(w^\ell_{\bar{v}}, y) & \text{otherwise.} \end{cases}$$

Let $\Gamma_{x,y}$ be all the permutations of X which fix x and y, and let

$$\bar{t} = \frac{1}{|\Gamma_{x,y}|} \sum_{\ell\in\{-1,0,1\}} r_\ell \sum_{\bar{v}\in V_\ell(c)} s^\ell_{\bar{v}} \sum_{\tau\in\Gamma_{x,y}} \mathrm{pr}\left(c^{\tau\sigma^\ell_{\bar{v}}}\right).$$

Of course, the sum of the coefficients is

$$\sum_{\substack{\ell\in\{-1,0,1\}\\ \bar{v}\in V_\ell(c)\\ \tau\in\Gamma_{x,y}}} \frac{r_\ell s^\ell_{\bar{v}}}{|\Gamma_{x,y}|} = \frac{|\Gamma_{x,y}|}{|\Gamma_{x,y}|} \sum_{\ell\in\{-1,0,1\}} r_\ell = 1.$$

Since r_ℓ, $s^\ell_{\bar{v}}$, and $\frac{1}{|\Gamma_{x,y}|}$ are in $[0,1]_{\mathbb{Q}}$, $\frac{r_\ell s^\ell_{\bar{v}}}{|\Gamma_{x,y}|} \in [0,1]_{\mathbb{Q}}$. So $\bar{t} \in \text{pr-cl}(\mathscr{C})$.

Consider $t_{x,y}$. Each τ fixes x and y, so

$$\begin{aligned} c^{\tau\sigma^\ell_{\bar{v}}}\{x,y\} = x &\iff \\ c^{\sigma^\ell_{\bar{v}}}\{x,y\} = x &\iff \\ c\{u^\ell_{\bar{v}}, w^\ell_{\bar{v}}\} = u^\ell_{\bar{v}} &\iff \ell = 1. \end{aligned}$$

So $W^x_y(c^{\tau\sigma^\ell_{\bar{v}}}) = \ell$. Thus

$$\begin{aligned} t_{x,y} &= \frac{1}{|\Gamma_{x,y}|} \sum_{\ell\in\{-1,0,1\}} r_\ell \sum_{\bar{v}\in V_\ell(c)} s^\ell_{\bar{v}} \sum_{\tau\in\Gamma_{x,y}} \ell \\ &= \frac{|\Gamma_{x,y}|}{|\Gamma_{x,y}|} \sum_{\ell\in\{-1,0,1\}} \ell\, r_\ell = r_1 - r_{-1}. \end{aligned}$$

Since $r_{-1} \neq r_1$, $t_{x,y} \neq 0$.

Consider $t_{x,u}$ where $u \notin \{x,y\}$.

$$\begin{aligned} c^{\tau\sigma^\ell_{\bar{v}}}\{x,u\} = x &\iff \\ c^{\sigma^\ell_{\bar{v}}}\{x,\tau^{-1}(u)\} = x &\iff \\ c\{u^\ell_{\bar{v}}, u^*\} = u^\ell_{\bar{v}}, \end{aligned}$$

where

$$u^* = \left(\sigma^\ell_{\bar{v}}\right)^{-1}(\tau^{-1}(u)) = \begin{cases} \tau^{-1}(u) & \text{if } \tau^{-1}(u) \notin \{u^\ell_{\bar{v}}, w^\ell_{\bar{v}}\}, \\ x & \text{if } \tau^{-1}(u) = u^\ell_{\bar{v}}, \\ y & \text{if } \tau^{-1}(u) = w^\ell_{\bar{v}}. \end{cases}$$

u^* is never $u^\ell_{\bar v}$ or $w^\ell_{\bar v}$. For a given ℓ and $\bar v$, u^* varies equally over all elements of X besides $\{u^\ell_{\bar v}, w^\ell_{\bar v}\}$, so

$$\sum_{\tau\in\Gamma_{x,y}} W^x_u(c^{\tau\sigma^\ell_{\bar v}}) = \sum_{\tau\in\Gamma_{x,y}} W^{u^\ell_{\bar v}}_{u^*}(c) = \sum_{\tau\in\Gamma_{x,y,u}} \sum_{u^*\in X\setminus\{u^\ell_{\bar v},w^\ell_{\bar v}\}} W^{u^\ell_{\bar v}}_{u^*}(c)$$
$$= |\Gamma_{x,y,u}| \left(\mathrm{val}_c(u^\ell_{\bar v}) - W^{u^\ell_{\bar v}}_{u^\ell_{\bar v}}(c) - W^{u^\ell_{\bar v}}_{w^\ell_{\bar v}}(c)\right) = |\Gamma_{x,y,u}| \left(\mathrm{val}_c(u^\ell_{\bar v}) - \ell\right).$$

This is because, for each $u^* \in X\setminus\{u^\ell_{\bar v}, w^\ell_{\bar v}\}$, there are $|\Gamma_{x,y,u}|$ permutations τ which take u to u^*. $W^u_u(c) = 0$ for any u, and by choice of $u^\ell_{\bar v}$ and $w^\ell_{\bar v}$, $W^{u^\ell_{\bar v}}_{w^\ell_{\bar v}}(c) = \ell$.

Recall that the $u^\ell_{\bar v}$ were chosen so that $\mathrm{val}_c(u^\ell_{\bar v}) - \ell$ is the first coordinate of $\bar v$. Recall also that the r_ℓ, $s^\ell_{\bar v}$, and valence pairs $\bar v$ were chosen so

$$\sum_{\ell\in\{-1,0,1\}} r_\ell \sum_{\bar v\in V_\ell(c)} s^\ell_{\bar v}\bar v = (0,0).$$

$$\begin{aligned} t_{x,u} &= \frac{1}{|\Gamma_{x,y}|} \sum_{\ell\in\{-1,0,1\}} r_\ell \sum_{\bar v\in V_\ell(c)} s^\ell_{\bar v} \sum_{\tau\in\Gamma_{x,y}} W^x_u\left(c^{\tau\sigma^\ell_{\bar v}}\right) \\ &= \frac{1}{|\Gamma_{x,y}|} \sum_{\ell\in\{-1,0,1\}} r_\ell \sum_{\bar v\in V_\ell(c)} s^\ell_{\bar v}\,|\Gamma_{x,y,u}| \left(\mathrm{val}_c(u^\ell_{\bar v}) - \ell\right) \\ &= 0. \end{aligned}$$

Similarly, $t_{y,u} = 0$ for all $u \notin \{x, y\}$.

Consider $t_{u,w}$ with neither u nor w in $\{x, y\}$. Then

$$\begin{aligned} c^{\tau\sigma^\ell_{\bar v}}\{u,w\} = u &\iff \\ c^{\sigma^\ell_{\bar v}}\{\tau^{-1}(u), \tau^{-1}(w)\} = \tau^{-1}(u) &\iff \\ c\{u^*, w^*\} = u^*, \end{aligned}$$

where $u^* = (\sigma^\ell_{\bar v})^{-1}(\tau^{-1}(u))$ and $w^* = (\sigma^\ell_{\bar v})^{-1}(\tau^{-1}(w))$. But $\tau^{-1}(u)$ and $\tau^{-1}(w)$ are never x or y, so u^* is never $u^\ell_{\bar v}$ and w^* is never $w^\ell_{\bar v}$. As τ varies, u^* and w^* vary equally over all members of X except $u^\ell_{\bar v}$ and $w^\ell_{\bar v}$, so $c\{u^*, w^*\} = u^*$ just as often as $c\{u^*, w^*\} = w^*$. Thus

$$\sum_{\tau\in\Gamma_{x,y}} W^u_w(c^{\tau\sigma^\ell_{\bar v}}) = 0.$$

Hence

$$t_{u,v} = \frac{1}{|\Gamma_{x,y}|} \sum_{\ell\in\{-1,0,1\}} r_\ell \sum_{\bar v\in V_\ell(c)} s^\ell_{\bar v} \sum_{\tau\in\Gamma_{x,y}} W^u_w(c^{\tau\sigma^\ell_{\bar v}}) = 0.$$

So $d = \mathrm{maj}(\bar t) \in \text{maj-cl}(\mathscr{C})$ has $\mathrm{dom}\, d = \{\{x, y\}\}$; all other pairs have $t_{u,v} = 0$. $\square$

5.9. **Claim.** *If* $\mathscr{C} \subseteq \mathfrak{C}$ *is symmetric and chaotic, then* $\text{maj-cl}(\mathscr{C}) = \mathfrak{C}$.

Proof. Suppose $f \in \mathfrak{C}$.

By Claim 5.7, there is some $c \in \mathscr{C}$ which is valence-imbalanced. So by Claim 5.8, for any $\{x, y\} \in \binom{X}{2}$ there is $d \in \text{maj-cl}(\mathscr{C})$ with $\mathrm{dom}\, d = \{\{x, y\}\}$. If $d\{x, y\} = y$, by symmetry there is $d' \in \text{maj-cl}(\mathscr{C})$ with domain $\{\{x, y\}\}$ and $d'\{x, y\} = x$. So for every $\{x, y\} \in \mathrm{dom}\, f$, let $d_{x,y} \in \text{maj-cl}(\mathscr{C})$ such that $\mathrm{dom}\, d_{x,y} = \{\{x, y\}\}$ and $d_{x,y}\{x, y\} = f\{x, y\}$. Now combine the voter populations which yielded each $d_{x,y}$.

f was arbitrary, so we have shown $\mathfrak{C} \subseteq \text{maj-cl}(\mathscr{C})$. Hence $\text{maj-cl}(\mathscr{C}) = \mathfrak{C}$. $\square$

6. Bounds

We shall consider the loose upper bounds, implied by the preceding proofs, on the necessary number of voters from a symmetric set to yield an arbitrary function in its majority closure, in terms of $\mathbf{n}$. There is no doubt that much tighter bounds could be obtained.

Trivial. 0 functions suffice to obtain no outcome.

Balanced. In order to create a triangular function from a chosen function c, $(\mathbf{n}-3)!$ permutations of the rest of the set were used. These were repeated 3 times, to establish each pair of edges in the triangle.

An arbitrary pseudo-balanced function has each edge in a directed cycle, so one cycle per edge suffices; so we have at most $\binom{\mathbf{n}}{2} = \frac{\mathbf{n}(\mathbf{n}-1)}{2}$ cycles. Each is constructed with two fewer triangles than the number of nodes it contains. This is at most $\mathbf{n}-2$, if some cycle contains every node. So all together we have

$$\frac{1}{2}\mathbf{n}(\mathbf{n}-1)(\mathbf{n}-2)(\mathbf{n}-3)!\,3 = \frac{3\mathbf{n}!}{2}.$$

Partisan. There are at most $\mathbf{n}$ tiers in an arbitrary tiered function, each with at most $\mathbf{n}$ elements (of course, no function meets both conditions.) For each such element x, we take functions for each permutation holding x fixed; there are $(\mathbf{n}-1)!$ such. So we have $\mathbf{n}\,\mathbf{n}(\mathbf{n}-1)! = \mathbf{n}\,\mathbf{n}!$ as an upper bound.

Imbalanced, nonpartisan, but not chaotic. The proof of Claim 4.5 gives the size of the population T_z yielding $c^{a,b,z}$ as $|T_z| = 3(\mathbf{n}-3)!$, with $m = 2(\mathbf{n}-3)!$ votes for the winner of each pair. We chose a partisan function c with l candidates in the winning tier. For each of at most $\binom{\mathbf{n}}{2}$ edges $a \to b$ and each of the $(\mathbf{n}-2)$ $z \in X \setminus \{a,b\}$, we used k copies of T_z, where

$$k = \begin{cases} (\mathbf{n}-2)\binom{\mathbf{n}-3}{l-1} & \text{if } l \le \frac{\mathbf{n}}{2}, \\ (\mathbf{n}-2)\binom{\mathbf{n}-3}{l-2} & \text{if } l > \frac{\mathbf{n}}{2}, \end{cases}$$

and m copies of a population of partisan functions D, where

$$|D| = \begin{cases} (\mathbf{n}-2l)\binom{\mathbf{n}-1}{l} + (\mathbf{n}-2)\binom{\mathbf{n}-2}{l-1} & \text{if } l \le \frac{\mathbf{n}}{2}, \\ (2l-\mathbf{n})\binom{\mathbf{n}-1}{l-1} + (\mathbf{n}-2)\binom{\mathbf{n}-2}{l-1} & \text{if } l > \frac{\mathbf{n}}{2}. \end{cases}$$

In fact, $k' = \frac{k}{\gcd(m,k)}$ copies of T_z, and $m' = \frac{m}{\gcd(m,k)}$ copies of D, would suffice. Certainly $\binom{\mathbf{n}-3}{l-1} = \frac{(\mathbf{n}-3)!}{(l-1)!(\mathbf{n}-l-2)!} \mid (\mathbf{n}-3)!$, and $\binom{\mathbf{n}-3}{l-2} = \frac{(\mathbf{n}-3)!}{(l-2)!(\mathbf{n}-l-1)!} \mid (\mathbf{n}-3)!$, so we may take $k' = (\mathbf{n}-2)$.

If $l \le \frac{\mathbf{n}}{2}$, then we take $m' = 2(l-1)!(\mathbf{n}-l-2)!$, and

$$\begin{aligned} m'\,|D| &= 2(l-1)!(\mathbf{n}-l-2)!\left((\mathbf{n}-2l)\binom{\mathbf{n}-1}{l} + (\mathbf{n}-2)\binom{\mathbf{n}-2}{l-1}\right) \\ &= 2(l-1)!(\mathbf{n}-l-2)!\left(\frac{(\mathbf{n}-2l)(\mathbf{n}-1)!}{l!(\mathbf{n}-l-1)!} + \frac{(\mathbf{n}-2)(\mathbf{n}-2)!}{(l-1)!(\mathbf{n}-l-1)!}\right) \\ &= 2(\mathbf{n}-2)!\left(\frac{(\mathbf{n}-2l)(\mathbf{n}-1)}{l(\mathbf{n}-l-1)} + \frac{\mathbf{n}-2}{\mathbf{n}-l-1}\right). \end{aligned}$$

The last multiplicand is

$$\frac{\mathbf{n}(\mathbf{n}-l-1)}{l(\mathbf{n}-l-1)} = \frac{\mathbf{n}}{l} \le \mathbf{n}.$$

If $l > \frac{\mathbf{n}}{2}$, then we take $m' = 2(l-2)!(\mathbf{n}-l-1)!$, and

$$\begin{aligned} m'\,|D| &= 2(l-2)!(\mathbf{n}-l-1)!\left((2l-\mathbf{n})\binom{\mathbf{n}-1}{l-1} + (\mathbf{n}-2)\binom{\mathbf{n}-2}{l-1}\right) \\ &= 2(l-2)!(\mathbf{n}-l-1)!\left(\frac{(2l-\mathbf{n})(\mathbf{n}-1)!}{(l-1)!(\mathbf{n}-l)!} + \frac{(\mathbf{n}-2)(\mathbf{n}-2)!}{(l-1)!(\mathbf{n}-l-1)!}\right) \\ &= 2(\mathbf{n}-2)!\left(\frac{(2l-\mathbf{n})(\mathbf{n}-1)}{(l-1)(\mathbf{n}-l)} + \frac{\mathbf{n}-2}{l-1}\right). \end{aligned}$$

The last multiplicand is

$$\frac{\mathbf{n}(l-1)}{(l-1)(\mathbf{n}-l)} = \frac{\mathbf{n}}{\mathbf{n}-l} \le \mathbf{n}.$$

So the total number of voters is

$$\begin{aligned} &\binom{\mathbf{n}}{2}\left((\mathbf{n}-2)\,|T_z|\,k' + m'\,|D|\right) \\ &\le \frac{\mathbf{n}(\mathbf{n}-1)}{2}\left((\mathbf{n}-2)3(\mathbf{n}-3)!(\mathbf{n}-2) + 2(\mathbf{n}-2)!\mathbf{n}\right) \\ &= \mathbf{n}(\mathbf{n}-1)(\mathbf{n}-2)!\left(\frac{3}{2}(\mathbf{n}-2)+\mathbf{n}\right) \\ &= \mathbf{n}!\left(\frac{3}{2}\mathbf{n}-3+\mathbf{n}\right) < \frac{5}{2}\mathbf{n}\,\mathbf{n}!. \end{aligned}$$

Chaotic. Given an arbitrary function, we wish to create a "single-edged" $d_{x,y}$, as developed in Section 5, for each edge in the function; there are at most $\binom{\mathbf{n}}{2}$ of these. In creating such a single-edged $d_{x,y}$, we use a collection of symmetric functions, under permutations which take x and y to elements having certain valence combinations. Reading the proof of Claim 5.7 carefully, we see that at most 4 such valence pairs are used. For each of these, all permutations fixing the relevant pair are used, giving $4(\mathbf{n}-2)!$ choice functions used in the creation of the single-edged function. However, these are being combined by some set of rational coefficients, so in fact we must expand each such function taking as many copies as the numerator of its associated coefficient when put over the least common denominator L. The coefficients sum to 1, so the sum of the numbers of copies is L; so $L(\mathbf{n}-2)!$ functions are used for each edge.

Six possible linear combinations of valence pairs yielding $(0,0)$ were considered in Claim 5.7. Using the fact that the original valence pairs are pairs of integers between $-\mathbf{n}$ and $\mathbf{n}$, we can solve for the coefficients by matrix inversion and determine that the least common denominator is at most the determinant of the associated matrix. For instance, in Case 4, suppose there is a point of $y=-x$ strictly between two points of $V_1(c)$ on a line segment not parallel to $y=x$. (This is true in this Case, unless $V_1(c)$ is contained in a line parallel to $y=x$ except for a point on the line $y=-x$; such a situation actually has a smaller common denominator.)

We use Claim 5.6, with $(0,0)$ for $\bar{w}$, and note that either the line from $\bar{p}_0$ through $\bar{w}$ meets the line L, or the line from $\bar{p}_1$ through $\bar{w}$ does. Say $\bar{p}$ is the one which works. Then $(0,0)$ is an imbalanced linear combination of the three valence pairs $\bar{p}$, $\bar{q}_0$, and $\bar{q}_1$. Say $\bar{p}=(a,b)$; then one of the $\bar{q}$ is (b,a) and the other is (c,d). Solving

$$\begin{aligned} r+s+t&=1\\ ar+bs+ct&=0\\ br+as+dt&=0 \end{aligned}$$

amounts to

$$\begin{bmatrix} r\\ s\\ t \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1\\ a & b & c\\ b & a & d \end{bmatrix}^{-1} \begin{bmatrix} 1\\ 0\\ 0 \end{bmatrix}.$$

Since everything on the right hand side is an integer, the only denominator introduced is the determinant of the 3 by 3 matrix, $bd-ac-ad+bc+a^2-b^2$. Since each of a,b,c,d is an integer between $-\mathbf{n}$ and $\mathbf{n}$, this is at most $6\mathbf{n}^2$.

In the cases involving four points, such as Case 1 or Case 5, an additional constraint is needed, which is not a consequence of the other three. In Case 1, use $r_1+r_2+r_3+r_4=1$, $a_1r_1+a_2r_2+a_3r_3+a_4r_4=0$, $b_1r_1+b_2r_2+b_3r_3+b_4r_4=0$, and add $b_3r_3+b_4r_4=-a_3r_3-a_4r_4$, since the points $\bar{v}$ and $(k_0,-k_0)$ were linear combinations of two valence pairs on the line $y=-x$. The lowest common denominator is at most the determinant of

$$\begin{bmatrix} 1 & 1 & 1 & 1\\ a_1 & a_2 & a_3 & a_4\\ b_1 & b_2 & b_3 & b_4\\ 0 & 0 & a_3+b_3 & a_4+b_4 \end{bmatrix},$$

which expands to a 16-term sum of three a_i or b_i each, hence at most $16\mathbf{n}^3$. In Case 5, note that the point $\bar{v}$ between two points of $V_1(c)$ is arbitrary; we can choose any point between the two. Between any two pairs of integers on which do not lie on a line perpendicular to $y=x$, we can choose a point whose projection on $y=x$ is of the form $\left(\frac{l}{4},\frac{l}{4}\right)$ for some integer l. Solving

$$\begin{bmatrix} 1 & 1 & 1 & 1\\ k & a & b & c\\ k & b & a & d\\ 0 & a & b & c \end{bmatrix}^{-1} \begin{bmatrix} 1\\ 0\\ 0\\ \frac{l}{4} \end{bmatrix}$$

gives a determinant $k(2a^2-b^2-ab-ac+bd+bc-ad)$ at most $8\mathbf{n}^3$; we must multiply by 4 because of the fraction in the column vector, giving $32\mathbf{n}^3$. This is the largest denominator of the six possible linear combinations in Claim 5.7.

Thus our upper bound is

$$L(\mathbf{n}-2)!\frac{\mathbf{n}^2-\mathbf{n}}{2} < 16\,\mathbf{n}^3\mathbf{n}(\mathbf{n}-1)(\mathbf{n}-2)! = 16\,\mathbf{n}^3\mathbf{n}!.$$

References

N. Alon. Voting paradoxes and digraphs realizations. *Advances in Appl. Math.*, 29: 126–135, 2002.

J. A. N. d. Condorcet. *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. L'imprimerie royale, Paris, 1785.

P. Erdős and L. Moser. On the representation of directed graphs as unions of orderings. *Magyar Tud. Akad. Mat. Kutats Int. Kvzl.*, 9:125–132, 1964.

D. C. McGarvey. A theorem on the construction of voting paradoxes. *Econometrica*, 21:608–610, 1953.

A. Roslanowski and S. Shelah. Norms on possibilities I: forcing with trees and creatures. *Memoirs of the American Mathematical Society*, 141(671):xii + 167, 1999.

S. Shelah. What majority decisions are possible. *Discrete Mathematics*, 309(8): 2349–2364, 2009.

R. Stearns. The voting problem. *American Mathematical Monthly*, 66:761–763, 1959.

(Paul Larson, Nick Matteo) Department of Mathematics, Miami University, Oxford, OH 45056, USA

Email address: larsonpb@muohio.edu, matteona@muohio.edu

(Saharon Shelah) The Hebrew University of Jerusalem, Einstein Institute of Mathematics, Edmond J. Safra Campus, Givat Ram, Jerusalem 91904, Israel

(Saharon Shelah) Department of Mathematics, Hill Center-Busch Campus, Rutgers, The State University of New Jersey, 110 Frelinghuysen Road, Piscataway, NJ 08854-8019, USA