Coalitional Structure of the Muller-Satterthwaite Theorem

Pingzhong Tang and Tuomas Sandholm

Computer Science Department
Carnegie Mellon University
{kenshin,sandholm}@cs.cmu.edu

Abstract. The Muller-Satterthwaite theorem states that social choice functions that satisfy unanimity and monotonicity are also dictatorial. Unlike Arrow's theorem, it does not assume that the function produces a transitive social ordering. Wilson showed that a voting process under Arrow's conditions can be interpreted as a strong and proper *simple game*—as defined by von Neumann and Morgenstern. We show this to be the case also under the Muller-Satterthwaite conditions. Our main theorem, which we prove using two very different approaches—one partially automated and one manual—is that a winning coalition coincides with a blocking coalition under unanimity and monotonicity. This might be of independent interest. We also show that this can be used to generate a short proof of the Muller-Satterthwaite theorem.

1 Introduction

In their seminal book, John von Neumann and Oskar Morgenstern introduced a cooperative game-theoretic concept called a *simple game* [7]. A simple game is uniquely determined by a collection of *winning coalitions*. A winning coalition is a set of agents that are decisive for every pair of outcomes. In other words, if a consensus is reached among the agents in the coalition that one outcome is preferred over another, then the former outcome is preferred over the latter in the social ordering. A simple game is called *proper* if every winning coalition intersects every other winning coalition. This is a natural condition because without it, two disjoint coalitions that prefer different candidates could both be winning, which is essentially a contradiction. A simple game is called *strong* if any coalition that intersects every winning coalition is itself winning.

Wilson [11] recognized a strong resemblance between this formulation of simple games and Arrow's celebrated formulation of a voting process [1]. In particular, Wilson showed that, if a voting process satisfies Arrow's conditions of unanimity (the society is a winning coalition), independence of irrelevant alternatives (the social ordering on any two outcomes is independent of agents' preferences related to other outcomes), and transitivity of the social ordering, then a coalition is winning if and only if it is *blocking*. Blocking is a much weaker notion than that of winning. A coalition is *blocking* if there exists some preference profile where the coalition is the exact set of agents that prefer one outcome to another and the former outcome is preferred over the latter in the social ordering. Wilson's equivalence result not only yields Arrow's famous impossibility theorem as a corollary, but also gives a nice interpretation of Arrow's conditions in terms of simple game: a voting process defined under these conditions is a proper and strong simple game!

One natural question is: what if the requirement of transitivity is relaxed? This question naturally leads us to think about an alternative formulation of a voting process that does not produce a social ordering but only a winner. Such voting processes have been studied extensively; the object of study is a social choice function. The Muller-Satterthwaite theorem states that social choice functions that satisfy unanimity and monotonicity are also dictatorial [6]. Since this impossibility theorem is an analog of Arrow's theorem in this setting, we are curious whether there is also an analog of Wilson's interpretation in this setting.

In this paper, we answer this question affirmatively. We first introduce the corresponding definitions of winning and blocking coalition for the Muller-Satterthwaite setting. We then prove, in this setting, that winning and blocking coalitions coincide. This generalization also enables a short proof of the Muller-Satterthwaite theorem, and gives a coalitional understanding of the Muller-Satterthwaite axioms.

It is not clear, at least to us, how Wilson's proof could be applied to our setting. Instead, we come up with two approaches to prove our results. One adapts a new methodology in computer-aided theorem proving/discovery for social choice to our setting. The methodology was introduced by Tang and Lin [9, 10] for automatically proving and discovering impossibility theorems in social choice and for discovering theorems in two-person games with unique pure Nash equilibrium payoffs. It was later extended by Geist and Endriss [3] for proving and discovering theorems in ranking sets of objects. Our second approach is a manual proof that may be of independent interest.

The rest of the paper is organized as follows. In the next section, we introduce the Muller-Satterthwaite formulation of a voting process and their famous impossibility theorem. We then state our main theorem in Section 3, and prove it using two different approaches. In Section 4, we leverage it to generate a new, short proof of the Muller-Satterthwaite theorem. In Section 5, we connect our main theorem to strong and proper simple games. Finally, we discuss Wilson's interpretation in Section 6 and conclude with some future research ideas in Section 7.

2 The Muller-Satterthwaite formulation and impossibility theorem

A voting model is a tuple (N, O), where N is a finite set of individuals (agents) and O a finite set of outcomes (alternatives). An agent's preference ordering is a linear ordering of O, and a preference profile $>$ of (N, O) is a tuple $(>_1, ..., >_n)$, where $>_i$ is agent i's preference ordering, and $n = |N|$. In what follows, when N is clear from the context, we also call $>$ a preference profile of O. Similarly, when O is clear from the context, we also call $>$ a preference profile of N.

Definition 1. *Given a voting model (N, O), a social choice function is a function $C : L^n \to O$, where L is the set of linear orders on O, and n the number of agents in N.*

Instead of the conditions of unanimity, independence of irrelevant alternatives, and non-dictatorship in Arrow's theorem, Muller and Satterthwaite considered the following three corresponding conditions that are desirable for any social choice function.

The first condition states that an alternative that is dominated by another should never be selected:

Definition 2. *A social choice function C is* weakly unanimous *if for every preference profile $>$, if there is a pair of alternatives a_1, a_2 such that $a_1 >_i a_2$ for every agent i, then $C(>) \neq a_2$.*

The second condition states that if a social choice function selects an outcome for a preference profile, then it will also select this outcome for any other preference profile that does not decrease the relative ranking of this outcome:

Definition 3. *A social choice function C is* monotonic *if, for every preference profile $>$ such that $C(>) = a$, if $>'$ is another profile such that $a >'_i a'$ whenever $a >_i a'$ for every agent i and every alternative a', then $C(>') = a$ as well.*

The third condition states there should not be a dictator:

Definition 4. *An agent i is a dictator in a social choice function C if C always selects i's top choice: for every preference profile $>$, $C(>) = a$ iff for all $a' \in O$ that is different from a, $a >_i a'$. C is* non-dictatorial *if it has no dictator.*

A social choice function is dictatorial if there is a dictator.

The Muller-Satterthwaite theorem states that the three desiderata above cannot coexist in any social choice function:

Theorem 1. ***(Muller-Satterthwaite [6])*** *For any voting model (N, O) such that $|O| \geq 3$, any social choice function that is weakly unanimous and monotonic is also dictatorial.*

3 Our main theorem

In this section, we define the notions of winning coalition and blocking coalition in our setting. We then state and prove our main theorem.

Definition 5. *A coalition S is* winning *if, for any profile $>$ in which every agent in S ranks some alternative a on top of his preference, $C(>) = a$.*

Definition 6. *A coalition S is* blocking *if there exist some profile $>$ such that $C(>) = a$ and a is ranked on the top by all the agents in S and ranked on the bottom by all the agents in $N \setminus S$.*

We are now ready to state our main theorem:

Theorem 2. *For any voting model (N, O) such that $|O| \geq 3$, under weak unanimity and monotonicity, a coalition is winning iff it is blocking.*

Proof. The "only if" part is straightforward. Let S be a winning coalition. Construct a profile such that an alternative a is ranked on the top by all the agents in S and ranked on the bottom by all the agents in $N \setminus S$. According to the definition of winning coalition, we have $C(>) = a$. It follows that S is also blocking.

We now provide two proofs of the "if" part, which is less straightforward. The first one uses Tang and Lin's framework of computer aided proofs for impossibility theorems in social choice theory [9]. The second one is a manual proof of potential independent interest.

For ease of presentation, we use the following notation.

- For any set S, we use S_{-a} to denote $S \setminus \{a\}$, i.e., the set S without element a.
- We extend the above notation to tuples as well: if $t = (t_1, ..., t_n)$, then we use t_{-i} to denote the tuple $(t_1, ..., t_{i-1}, t_{i+1}, ..., t_n)$. Furthermore, we use (t_{-i}, s) to denote the result of replacing ith item in t by s: $(t_{-i}, s) = (t_1, ..., t_{i-1}, s, t_{i+1}, ..., t_n)$. We use $t_{-\{i,j\}}$ to denote $(t_{-i})_{-j}$.
- If $>$ is a linear ordering of O, and $a \in O$, then we let $>_{-a}$ be the restriction of $>$ on O_{-a}: for any $a', a'' \in O_{-a}$, $a' >_{-a} a''$ iff $a' > a''$. On the other hand, if $>$ is a linear ordering of O_{-a} for some $a \in O$, then we let $>^{+a}$ be the extension of $>$ to O such that for any $a' \in O_{-a}$, $a' >^{+a} a$. Thus if $>$ is a linear ordering of O, and $a \in O$, then $>^{+a}_{-a}$ is $(>_{-a})^{+a}$, i.e., the result of moving a to the bottom of the ordering. These notations extend to tuples of orderings. Thus if $>$ is a preference profile of (N, O_{-a}), then

$$>^{+a} = (>_1, ..., >_n)^{+a} = (>^{+a}_1, ..., >^{+a}_n),$$

  which will be a preference profile of (N, O).

The following example explains the last item of our notation above. Given $(N = \{1, 2\}, O = \{a, b, c\})$ and preference profile $>$ as follows:

$$\begin{matrix} a & b \\ c & a \\ b & c \end{matrix}$$

Then, $>_{-a}$ is a profile on $(N = \{1, 2\}, O = \{b, c\})$:

$$\begin{matrix} c & b \\ b & c \end{matrix}$$

Finally, $(>_{-a})^{+a}$ is a profile on $(N = \{1, 2\}, O = \{a, b, c\})$, where a is at the bottom of every agent's preference ordering:

$$\begin{matrix} c & b \\ b & c \\ a & a \end{matrix}$$

3.1 The first proof: A computer-aided approach

In this first proof, we use induction to prove the "if"-part.

The overall proof is a contrapositive argument. We start by arguing that if it is possible for something ("blocking but not winning" in our setting) to happen in the large, then it is possible also in the small. But then we prove that it is impossible in the small. This implies impossibility in the large.

Below, we first discuss the inductive step on the number of candidates and then the inductive step on the number of agents. Finally, we prove the base case where there are two agents and three candidates using a computer program.

The inductive case The inductive case consists of the following two lemmas. The first gives the inductive step on the number of candidates and the second gives the inductive step on the number of agents.

Lemma 1. *If there is a social choice function for n individuals and $m+1$ alternatives that is weakly unanimous, monotonic and there exists a coalition that is blocking but not winning, then there is also a social choice function for n individuals and m alternatives that satisfies these three conditions, for all $n \geq 2, m \geq 3$.*

Proof. Let (N, O) be a voting model such that $|N| = n$ and $|O| = m + 1$, and C a social choice function that satisfies the three conditions in the lemma. For any $a \in O$, we define C_a to be a social choice function that is the "restriction" of C on O_{-a}: for any preference profile $>$ of O_{-a}, $C_a(>) = C(>^{+a})$. In other words, C_a is defined by first adding a to everyone's least preferred position and then calling C on this new profile.

It is straightforward to check that for any candidate $a \in O$, C_a is well-defined and satisfies weak unanimity and monotonicity. We now prove that there exists a candidate d such that there exists a blocking coalition that is not winning in C_d (which operates on the restricted candidate set that does not contain d). Let S be such a coalition for C. By the definition of blocking coalition, there is a profile $>$ on C such that a is ranked on the top by all the agents in S and ranked on the bottom by the remaining agents and $C(>) = a$. By the definition of winning coalition, there is a profile $>'$ on C such that b is ranked on the top by all the agents in S and $C(>') = c \neq b$ (otherwise, C would violate the third condition). Since there are $|O| = m + 1 \geq 4$ candidates, we can find a candidate d distinct from a, b, c. Since $C(>) = a$, by monotonicity, we have $C((>_{-d})^{+d}) = a$. By the definition of C_d, this is equivalent to $C_d(>_{-d}) = a$. Similarly, since $C(>') = c \neq b$, by monotonicity, we have $C((>'_{-d})^{+d}) = c$. Thus, we have $C_d(>'_{-d}) = c \neq b$. In other words, $>_{-d}$ and $>'_{-d}$ are the witnesses of S being a blocking but not winning in C_d. Therefore, we conclude that C_d satisfies all the three conditions in the lemma.

We will now present the inductive step on the number of agents.

Lemma 2. *If there is a social choice function for $n+1$ individuals and m alternatives that is weakly unanimous, monotonic and there exists a coalition that is blocking but not winning, then there is also a social choice function for n individuals and m alternatives that satisfies these three conditions, for all $n \geq 2, m \geq 3$.*

Proof. Let (N, O) be a voting model such that $|N| = n+1$ and $|O| = m$, and C a social choice function that satisfies the three conditions in the lemma. For any pair of agents $i \neq j \in N$, we define $C_{i,j}$ to be the following social choice function for (N_{-i}, O): for any preference profile $>$ of (N, O), $C_{i,j}(>_{-i}) = C(>_{-i}, >_j)$. Again, it can be easily seen that for any pair of agents $i \neq j$, $C_{i,j}$ is weakly unanimous and monotonic. We now prove that there exists two agents, i, j, such that there exist a blocking coalition that is not winning in $C_{i,j}$. Let S be such a coalition for C. Again, we know that there is a profile $>$ on C such that a is ranked on the top by all the agents in S and ranked on the bottom by the remaining agents and $C(>) = a$. Also, there is a profile $>'$ on C such that b is ranked on the top by all the agents in S and $C(>') = c \neq b$. Since there are at least $2+1=3$ agents, we have either $|S| \geq 2$ or $|N \setminus S| \geq 2$. If $|S| \geq 2$, for any two agents $i, j \in S$, by monotonicity, we have $C(>_{-i}, >_j) = C(>) = a$. Similarly, we have $C(>'_{-\{i,j\}}, >''_j, >''_j) = C(>') = c \neq b$, where $>''_j$ is obtained by moving c to the second place (next to b) in $>'_j$ while maintaining the other candidates' positions. In other words, $>_{-i}$ and $(>'_{-\{i,j\}}, >''_j)$ are the witnesses of $S \setminus \{i\}$ being blocking but not winning in $C_{i,j}$. If $|N \setminus S| \geq 2$, we can arbitrarily select two agents, i, j, from $N \setminus S$ and similar arguments go through. Therefore, we conclude that $C_{i,j}$ satisfies all the three conditions in the lemma.

Given the two lemmas above, what remains to be done is to prove the base case, namely that "blocking but not winning" cannot happen in the small.

The base case We now turn to the proof of the base case, and as we mentioned earlier, we use a computer program to do that.

The base case says that when $|N| = 2$ and $|O| = 3$, there is no social choice function on (N, O) that is weakly unanimous, monotonic, and there exists a coalition that is blocking but not winning. The brute-force approach is to generate all possible social choice functions in (N, O) and check all of them one by one for these three conditions. However, there are too many such functions for this to be feasible: there are $3! = 6$ linear orderings of O, resulting in $6 \times 6 = 36$ preference profiles of (N, O), and 3^{36} possible social choice functions.

What we did instead is to generate all functions that satisfy the monotonicity condition, and then check whether any of them satisfy the other two conditions. Similar to Tang and Lin's verification of the base case of Arrow's theorem (cf. [9, p. 1045]), we do so through a constraint satisfaction problem (CSP) formulation [8]. Consider the voting model $(N = \{1, 2\}, O = \{a, b, c\})$ in our base case. We define a CSP for it by introducing 36 variables $x_1, ..., x_{36}$, one for each preference profile of the voting model. The domain of these variables is the set of all possible winners, i.e, $\{a, b, c\}$, and the constraints are the instantiations of the monotonicity condition on the voting model. As can be easily seen, there is a one-to-one correspondence between the social choice functions of the voting model and the assignments of the CSP. Furthermore, a solution to the CSP corresponds to a social choice function that satisfies the monotonicity condition, and vice versa. To solve this CSP, we use a depth-first search that backtracks whenever the current partial assignment violates the monotonicity constraint (see the pseudocode called "Procedure 1" below). Our program returns 17 monotonic functions (2 dictatorial functions, 3 constant functions, and 12 others), none of which satisfy the second and third conditions simultaneously. □

3.2 The second proof: A manual approach

As our second proof of the "if"-part, we give a concise manual proof—in contrast to the computer-aided proof above.

Proof. For the if part, let S be a blocking coalition and $>$ be the corresponding profile where $C(>) = a$ and a is ranked on the top in S and ranked on the bottom in $N \setminus S$.

Procedure 1 Solving CSP By DFS

DFS($assignment$)
if $assignment$ is complete **then**
  return $assignment$
end if
select an unassigned variable x_i
for $l \in \{a, b, c\}$ **do**
  if $x_i = l$ is consistent with every variable assigned so far under monotonicity **then**
    $assignment \leftarrow assignment \cup \{x_i = l\}$
    $result \leftarrow$ DFS($assignment$)
    if $result \neq failure$ **then**
      return $result$
    end if
    $assignment \leftarrow assignment \setminus \{x_i = l\}$ // backtrack
  end if
end for
return $failure$

According to monotonicity, for any profile $>'$ such that a is ranked on the top in S and ranked on the bottom in $N \setminus S$, we have $C(>') = a$. We prove that for any profile $>''$ where any alternative b is ranked on the top in S, $C(>'') = b$.

Consider a profile $>^1$ where b is ranked on the top in S and ranked on the bottom in $N \setminus S$, while a is ranked on the bottom in S and ranked immediately above b in $N \setminus S$. Written out explicitly, $>^1$ is as follows:

$$\begin{array}{cccccc} b & \cdots & b & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & a & \cdots & a \\ a & \cdots & a & b & \cdots & b \end{array}$$

If $C(>^1) = b$ for all such $>^1$ and all such b, we are done since by monotonicity, we can reach our goal above. If $C(>^1) = a$, an immediate contradiction occurs since a must be Pareto dominated by some alternative c since there are more than two alternatives. Thus, the only interesting case to consider is when $C(>^1) = c$ for some $c \notin \{a, b\}$. Again, this is possible since we have at least 3 outcomes. Without loss of generality, let c be ranked second in S after b and ranked on the top in $N \setminus S$.

Consider another profile $>^2$ where b is ranked on the top, followed by a, followed by c in S, while in $N \setminus S$, a is ranked on the bottom, b is ranked immediately above a and c is ranked on the top. $>^2$ is as follows:

$$\begin{array}{cccccc} b & \cdots & b & c & \cdots & c \\ a & \cdots & a & \cdots & \cdots & \cdots \\ c & \cdots & c & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & b & \cdots & b \\ \vdots & \vdots & \vdots & a & \cdots & a \end{array}$$

Clearly, $C(>^2) \neq a$ since a is Pareto dominated by b. Also, $C(>^2) \neq b$, since otherwise it will contradict with $C(>^1) = c$ after applying monotonicity twice (that is, move a to the bottom of S in $>^1$ and move b to the bottom of $N \setminus S$ in $>^2$). Clearly, $C(>^2) \neq d$ for all $d \notin \{a, b, c\}$ since d is Pareto dominated by c. Therefore, we must have $C(>^2) = c$. However, by monotonicity, when moving a above b in S in $>^2$, c must still be the winner, contradicting the fact that a wins in all profiles where a is ranked on the top in S and ranked on the bottom in $N \setminus S$.

We conclude that $C(>^1) = b$ for any such $>^1$ and b. By definition, S is a winning coalition.

4 A short proof of the Muller-Satterthwaite theorem

Our main theorem above (Theorem 2) enables a new, short proof of the Muller-Satterthwaite theorem. See [5, 2] for different proofs that also use coalitional approaches.

Corollary 1. ***(Muller-Satterthwaite theorem)*** *For any voting model (N, O) such that $|O| \geq 3$, under weak unanimity and monotonicity, there is a winning coalition that contains only one agent.*

Proof. Suppose that S is a minimum winning coalition, and it has at least two agents. We can partition N into three disjoint sets: S_1, S_2 and T where $S_1 \cup S_2 = S$ and S_1, S_2 are non-empty. Introduce the following notation: we denote by (abc, def, ghi) a profile where in S_1, a is ranked on the top, b is ranked second, and c is ranked on the bottom; in S_2 d is ranked on the top, e is ranked second, and f is ranked on the bottom; in T (if not empty), g is ranked on the top, h is ranked second, and i is ranked on the bottom. Such preference profiles are as follows:

$$\begin{array}{ccccccccc} a & \cdots & a & d & \cdots & d & g & \cdots & g \\ b & \cdots & b & e & \cdots & e & h & \cdots & h \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ c & \cdots & c & f & \cdots & f & i & \cdots & i \end{array}$$

Now consider a profile $>= (abc, bca, cab)$. $>$ is as follows:

$$\begin{array}{ccccccccc} a & \cdots & a & b & \cdots & b & c & \cdots & c \\ b & \cdots & b & c & \cdots & c & a & \cdots & a \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ c & \cdots & c & a & \cdots & a & b & \cdots & b \end{array}$$

It is easy to show that $C(>) \neq c$ and $C(>) \neq d$, where $d \notin \{a, b, c\}$. Suppose $C(>) = c$; we have $C(>' = (bac, bca, cab)) = c$, contradicting the fact that S is winning. Similarly, suppose $C(>) = d$; we have $C(>' = (bac, bca, cab)) = d$, contradicting the fact that S is winning. We now consider the remaining two cases.

– Case 1. $C(>) = a$. It follows from monotonicity that $C(>^1 = (acb, bca, cab)) = a$. $>^1$ is as follows:

$$\begin{array}{ccccccccc} a & \cdots & a & b & \cdots & b & c & \cdots & c \\ c & \cdots & c & c & \cdots & c & a & \cdots & a \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ b & \cdots & b & a & \cdots & a & b & \cdots & b \end{array}$$

Similarly, it also follows from monotonicity that $C(>^2 = (acb, bca, acb)) = a$. $>^2$ is as follows:

$$\begin{array}{ccccccccc} a & \cdots & a & b & \cdots & b & a & \cdots & a \\ c & \cdots & c & c & \cdots & c & c & \cdots & c \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ b & \cdots & b & a & \cdots & a & b & \cdots & b \end{array}$$

The equation above implies that $S_1 \cup T$ is blocking, and thus by our main theorem, winning. Consider another profile $>^3 = (acb, bca, cba)$. $>^3$ is as follows:

$$\begin{array}{ccccccccc} a & \cdots & a & b & \cdots & b & c & \cdots & c \\ c & \cdots & c & c & \cdots & c & b & \cdots & b \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ b & \cdots & b & a & \cdots & a & a & \cdots & a \end{array}$$

- Subcase 1. $C(>^3) = a$. Here, S_1 is blocking, and thus winning, contradicting our assumption that S is minimum.
- Subcase 2. $C(>^3) = b$. By monotonicity, we have $C(cab, bac, cba) = b$ as well, contradicting our conclusion earlier that $S_1 \cup T$ is winning.
- Subcase 3. $C(>^3) = c$. This immediately contradicts $C(>^1) = a$ by monotonicity.
- Subcase 4. $C(>^3) = d$. We have $C(cab, cba, cba) = d$, contradicting the fact that S is winning.

We conclude that Case 1 is impossible.

– Case 2. $C(>) = b$. It follows from monotonicity that $C(>^4 = (abc, bac, cab)) = b$. $>^4$ is as follows:

$$\begin{array}{ccccccccc} a & \cdots & a & b & \cdots & b & c & \cdots & c \\ b & \cdots & b & a & \cdots & a & a & \cdots & a \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ c & \cdots & c & c & \cdots & c & b & \cdots & b \end{array}$$

Similarly, it also follows from monotonicity that $C(>^5 = (abc, bac, acb)) = b$. $>^5$ is as follows:

$$\begin{array}{ccccccccc} a & \cdots & a & b & \cdots & b & a & \cdots & a \\ b & \cdots & b & a & \cdots & a & c & \cdots & c \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ c & \cdots & c & c & \cdots & c & b & \cdots & b \end{array}$$

This equation ensures that $S_1 \cup T$ is not winning. Consider another profile $>^6=(acb, bac, cab)$. $>^6$ is as follows:

$$\begin{array}{ccccccccc} a & \cdots & a & b & \cdots & b & c & \cdots & c \\ c & \cdots & c & a & \cdots & a & a & \cdots & a \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ b & \cdots & b & c & \cdots & c & b & \cdots & b \end{array}$$

- Subcase 1. $C(>^6) = a$. This immediately contradicts $C(>^4) = b$ by monotonicity.
- Subcase 2. $C(>^6) = b$. This implies that S_2 is blocking, and thus winning as well, contradicting the fact that S is a minimum winning coalition.
- Subcase 3. $C(>^6) = c$. By monotonicity, we have $C(cab, bac, cab) = c$. This implies that $S_1 \cup T$ is blocking, and thus winning. A contradiction.
- Subcase 4. $C(>^6) = d$. We have $C(acb, abc, cba) = d \neq a$, contradicting the fact that S is winning.

We conclude that Case 2 is impossible as well.

Therefore, we have $|S| = 1$.

5 The Muller-Satterthwaite setting can be interpreted as a proper, strong simple game

Recall from the introduction that a simple game is uniquely determined by a collection of winning coalitions. A simple game is called *proper* if every winning coalition intersects every other winning coalition. This is a natural condition because without it, two disjoint coalitions that prefer different candidates could both be winning. A simple game is called *strong* if any coalition that intersects every winning coalition is itself winning.

Wilson showed that a voting process under Arrow's conditions can be interpreted as a strong and proper simple game [11]. As an analog to Wilson's contribution, in this section we show this to be the case also under the Muller-Satterthwaite conditions—which, unlike Arrow's setting, do not assume that the social choice function produces a transitive social ordering.

The following theorem states this. The "if" part proves that it is a proper simple game. The "only if" part proves that it is a strong simple game.

Theorem 3. *Assume a social choice function*

1. *satisfies monotonicity, and*
2. *has the property that a coalition is blocking iff it is winning.*

Then, a coalition is winning iff it intersects every winning coalition.

Proof. The "only if" part is straightforward. Suppose for contradiction that two winning coalitions do not intersect. We can easily construct a profile where one winning coalition unanimously prefers one candidate and the other coalition prefers another candidate. This leads to a conflict in determining the value of the social choice function.

The "if" part is more involved. We are trying to prove that if a coalition S intersects every winning coalition, S is winning. Suppose for contradiction that there exist a profile $>$ such that some candidate a is ranked on the top by S and ranked last by $N \setminus S$, and we have $C(>) = b \neq a$.

Now consider a profile $>^1$ obtained from $>$ by moving b up to the second place in S and to the top in $N \setminus S$, and by moving a to the second place in $N \setminus S$. Clearly, by monotonicity, $C(>^1) = b$. $>^1$ is as follows:

$$\begin{array}{cccccc} a & \cdots & a & b & \cdots & b \\ b & \cdots & b & a & \cdots & a \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{array}$$

Let us consider another profile $>^2$, where a is ranked on the top by all the agents and b is ranked last by S and second by $N \setminus S$, while the remaining candidates are ranked exactly the same as in $>^1$. Clearly, N is winning (start from any profile and move the winner to the top of everyone's preference will maintain its winning position; this new profile is an instance that N is blocking, and thus winning as well). Thus, $C(>^2) = a$. $>^2$ is as follows:

$$\begin{array}{cccccc} a & \cdots & a & a & \cdots & a \\ \vdots & \vdots & \vdots & b & \cdots & b \\ b & b & b & \vdots & \vdots & \vdots \end{array}$$

Now switch the position of a and b in $N \setminus S$, that is, b is now ranked on the top and a is ranked second in $N \setminus S$. Denote the resulting profile by $>^3$. We now show that $C(>^3) = a$. By monotonicity, the only other possibility is $C(>^3) = b$. However, this implies that $N \setminus S$ is a blocking coalition, and thus a winning coalition as well, contradicting the fact that S intersects every winning coalition. $>^3$ is as follows:

$$\begin{array}{cccccc} a & \cdots & a & b & \cdots & b \\ \vdots & \vdots & \vdots & a & \cdots & a \\ b & b & b & \vdots & \vdots & \vdots \end{array}$$

Finally, in $>^3$, we move b to the second place in S and obtain $>^4$. We still have $C(>^4) = a$. $>^4$ is as follows:

$$\begin{array}{cccccc} a & \cdots & a & b & \cdots & b \\ b & \cdots & b & a & \cdots & a \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{array}$$

We observe that $>^4 = >^1$. This contradicts our earlier conclusions $C(>^1) = b \neq a = C(>^4)$.

Naturally, the result in this section also begets an open problem worth further consideration: What are necessary and sufficient conditions that guarantee properness and strongness?

6 Discussion

Wilson [11] defined notions of winning and blocking coalitions in Arrow's setting for social-welfare functions. A winning coalition in that setting is defined to be a set of agents that are decisive on every pair alternatives, for all profiles. A blocking coalition in that setting is a set of agents that are decisive on some pair of alternatives, for some profile. He then proved that, under a set of conditions that are essentially equivalent to Arrow's unanimity, independence of irrelevant alternatives, and transitivity of social ordering, a blocking coalition is also a winning coalition. It is easy to see that a coalition that is winning in Wilson's sense implies it is winning in our setting. Also, a coalition that is blocking in our setting implies it is blocking in Wilson's setting.

One might (therefore) wonder whether Wilson's result could imply our Theorem 2. The answer is negative, since we work in a more general social choice setting, which does not impose transitivity on the social ordering, and thus does not satisfy one of Wilson's assumptions. (Wilson's proof made critical use of the transitivity assumption. It also assumed that there are at least 5 alternatives. It is not clear, at least to us, how to adapt that proof for our purposes. For an interesting characterization of transitivity in proper simple games, see [4].)

7 Conclusions and future research

The Muller-Satterthwaite theorem states that social choice functions that satisfy unanimity and monotonicity are also dictatorial. Unlike Arrow's theorem, it does not assume that the function produces a transitive social ordering. Wilson showed that a voting process under Arrow's conditions can be interpreted as a strong and proper simple game—as defined by von Neumann and Morgenstern. We showed this to be the case also under the Muller-Satterthwaite conditions. Our main theorem, which we proved using two very different approaches—one partially automated and one manual—is that a winning coalition coincides with a blocking coalition under unanimity and monotonicity. This might be of independent interest. We also showed that this can be used to generate a short proof of the Muller-Satterthwaite theorem.

For future research, it would be interesting to think about necessary and sufficient conditions (to replace weak unanimity and monotonicity, which together form a sufficient condition) under which our main theorem holds (a coalition is winning iff it is blocking). We conjecture that those conditions will set the boundary between impossibility and possibility of existence of reasonable voting systems.

References

1. K. Arrow. A difficulty in the concept of social welfare. *Journal of Political Economy*, pages 328–246, 1950.

2. V. Danilov and A. Sotskov. *Social choice mechanisms*. Studies in economic design. Springer, 2002.
3. C. Geist and U. Endriss. Automated search for impossibility theorems in social choice theory: Ranking sets of objects. *J. Artif. Intell. Res. (JAIR)*, 40:143–174, 2011.
4. M. Kaneko. Necessary and sufficient conditions for transitivity in voting theory. *Journal of Economic Theory*, 11(3):385–393, 1975.
5. H. Moulin. *Axioms of Cooperative Decision Making*. Cambridge University Press, 1991.
6. E. Muller and M. A. Satterthwaite. The equivalence of strong positive association and strategy-proofness. *Journal of Economic Theory*, 14(2):412–418, April 1977.
7. J. V. Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, 1944.
8. S. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach*. Prentice-Hall, Englewood Cliffs, NJ, 2nd edition edition, 2003.
9. P. Tang and F. Lin. Computer aided proofs of arrows and other impossibility theorems. *Artificial Intelligence*, 173:1041–1053, 2009.
10. P. Tang and F. Lin. Discovering theorems in game theory: Two-person games with unique pure nash equilibrium payoffs. *Artif. Intell.*, 175(14-15):2010–2020, 2011.
11. R. Wilson. The game-theoretic structure of arrow's general possibility theorem. *Journal of Economic Theory*, 5(1):14–20, August 1972.