Cycles and Intractability in Social Choice Theory

William S. Zwicker[1]

Abstract

We introduce the (j,k)-Kemeny rule – a generalization[a] that aggregates weak orders. Special cases include approval voting, the mean rule and Borda mean rule, as well as the Borda count and plurality voting. Why, then, is the winner problem computationally simple for each of these other rules, but NP-hard for Kemeny? We show that winner intractability for the (j,k)-Kemeny rule first appears at the $j=3$, $k=3$ level. The proof reveals that computational complexity arises from the cyclic component in the fundamental decomposition $\overrightarrow{w} = \overrightarrow{w}_{cycle} + \overrightarrow{w}_{cocycle}$ of [11]. Thus the existence of "underlying" majority cycles – the engine driving both Arrow's impossibility theorem and the Gibbard-Satterthwaite theorem – also serves as a source of computational complexity in social choice.

[a]The (j,k)-Kemeny rule discussed here is akin to, but not the same as, the median procedure of [1].

1 Introduction

In their seminal paper, Bartholdi, Tovey, and Trick [2] showed that the problem of determining the winner of a Kemeny election is NP hard (Hemaspaandra, Spakowski, and Vogel [7] later showed completeness for $P_{||}^{NP}$). We introduce the (j,k)-Kemeny rule, a generalization wherein ballots are weak orders with j indifference classes and the outcome is a weak order with k indifference classes. Different values of j and k yield rules of interest in social choice theory as special cases, including approval voting, the mean rule (see [4] and [5]), and the Borda mean rule [5]; with an additional restriction one also obtains Borda and plurality voting.

Why, then, is the winner problem computationally simple for each of these other rules, but *not* simple for Kemeny? The short answer is that these other rules each satisfy $j \leq 2$ or $k \leq 2$, and we show that winner intractability for the (j,k)-Kemeny rule first appears at the $j=3$, $k=3$ level. This follows from our central result: the well-known NP-complete *max-cut problem* for undirected graphs can be polynomially reduced to *max-3OP*, a version of max-cut for weighted directed graphs in which vertices are partitioned into three pieces rather than two, and these pieces are ordered. The proof reveals that computational complexity arises from the cyclic component in the orthogonal decomposition $\overrightarrow{w} = \overrightarrow{w}_{cycle} + \overrightarrow{w}_{cocycle}$ induced by a profile (see [11]; a more complete exposition appears in [5]). In particular, $j \leq 2$ guarantees $\overrightarrow{w}_{cycle} = 0$, while $k \leq 2$ guarantees that $\overrightarrow{w}_{cycle}$ plays no role in the aggregation; neither guarantee applies when $j, k \geq 3$. However, if the profile *happens* to be one with $w_{cycle} = 0$ (intuitively, the profile has no *hidden* majority cycles), then the winner can be computed in polynomial time for any values of j and k; one example is that Kemeny = Borda when $\overrightarrow{w}_{cycle} = 0$. Thus majority cycles – the engines driving both Arrow's impossibility theorem and the Gibbard-Satterthwaite theorem – also serve as a source of computational complexity in social choice.

[1]This is a conference version, with some proofs omitted. We thank Matthew Anderson and Alan D. Taylor for help with the material and presentation, and the COMSOC referees for catching errors and suggesting some interesting follow-up questions.

In section 2 we dissect the relationship between the standard *max-cut* problem for (weighted, undirected) graphs, and our version *max-kOP* for directed graphs. While standard max-cut is NP-hard even for vertex partitions into two pieces, the directed version is intractable only for partitions into three or more pieces. The cyclic component – a measure of underlying tendency toward majority cycles – accounts for this critical distinction. We introduce the (j,k)-Kemeny rule in section 3, and show that a number of familiar aggregation rules are special cases. Calculating the winner for these rules amounts to solving cases of *max-kOP*; this allows us to transfer complexity results from section 2 to the winner determination problem for these rules.

Parts of this paper extend ideas from [5], which is also the best resource for readers unfamiliar with the orthogonal decomposition of a weighted tournament into cyclic and cocyclic components. Notions of generalized scoring rule implicit in [9] and explicit in [12], [3], and [10] also play an important, behind-the-scenes role. Hudry has several papers considering complexity issues for special cases of the median procedure, including the case of aggregating weak orders (see [8]).[2]

2 Max-cut for directed graphs

In the well known *max-cut problem*, one starts with an undirected graph $\mathcal{G} = (V, E)$ with finite vertex set V. A *vertex cut* is a partition $\mathcal{P} = \{J, K\}$ of V into two pieces, and is assigned a score $v(\mathcal{P})$ equal to the number of edges $\{a, b\} \in E$ whose vertices are "cut" by $\mathcal{P}$ (meaning $a \in J$ and $b \in K$, or $a \in K$ and $b \in J$). The max cut problem takes $\mathcal{G}$ as input, and asks for a vertex cut of maximal score. The corresponding decision problem takes as input the graph $\mathcal{G}$ along with a positive integer k, and asks whether there exists a vertex cut of $\mathcal{G}$ with score at least k. This decision problem, and hence the max-cut problem, are among the best known NP-complete problems. For our purposes, a certain generalization will be useful; the score of a *vertex tripartition* $\mathcal{P} = \{J, K, L\}$ of V will be the number of edges $\{a, b\} \in E$ whose vertices belong to different pieces of $\mathcal{P}$. The *max-tricut problem* and corresponding decision problem are then formulated exactly as one would expect. Our main concern will be with the weighted versions of these problems: each edge e of $\mathcal{G}$ comes equipped with a pre-assigned *edge-weight* $w(e) \geq 0$ and we seek to maximize the sum of the weights assigned to the edges that are cut. These problems are similarly NP-complete. Notice that for the weighted version there is no loss of generality in assuming $\mathcal{G}$ is complete; just add all the missing edges and assign them weight zero.

We consider a version of max cut for tournaments – *directed* graphs $\overrightarrow{\mathcal{H}} = (V, E)$ (with $E \subseteq V \times V$) for which every two distinct vertices a, $b \in V$ satisfy $(a, b) \in E$ or $(b, a) \in E$, but not both – similarly equipped with functions $\overrightarrow{w}$ assigning real-number weights (which may be negative) to directed edges.[3] For tournaments, a linearly ordered partition of the vertices plays the role of a "cut". For example, we might partition V into two disjoint and nonempty pieces, T (for *top*) and B (for *bottom*); the ordered partition $\overrightarrow{\mathcal{P}} = \{T > B\}$ is

[2] There is an important distinction between complexity questions for the (j,k)-Kemeny rule we consider here, and for the corresponding "(j,k)-median procedure." We intend to explore the relationship more thoroughly in a planned extension to this conference version of the paper. Loosely, the difference corresponds to the distinction (discussed in section 2 here) between the well-known max-cut problem for undirected graphs and the corresponding "max-kOP" problem for directed graphs.

[3] We put arrows over symbols for directed graphs and ordered partitions, to distinguish the denoted objects from ordinary graphs and unordered partitions. Assuming $\overrightarrow{\mathcal{H}}$ is a tournament is analogous to assuming completeness for undirected graphs, and similarly does not limit generality of Theorem 1, Proposition 1, or Corollary 1. For the directed problem, allowing negative weights adds no generality; if one reverses an edge while simultaneously reversing the sign of its weight, the effect on the *max-kOP* problem (see Definition 2) is nil. We allow negative weights because they provide notational flexibility needed to develop the decomposition $\overrightarrow{w} = \overrightarrow{w}_{cycle} + \overrightarrow{w}_{cocycle}$.

equivalent to a dichotomous weak order $\succeq$ on V in which two vertices x and y belonging to the same piece satisfy $x \succeq y$ and $y \succeq x$ (in which case we'll write $x \sim y$), but when $x \in T$ and $y \in B$ we get $x \succ y$ (meaning $x \succeq y$ and $y \not\succeq x$). An ordered tripartition $\{T > M > B\}$ similarly corresponds to a *trichotomous* weak order on V (meaning that the equivalence relation $\sim$ has three equivalence classes, rather than two, as in a dichotomous order) while a *linear* order on V is equivalent to an ordered $|V|$-partition, which has as many pieces as there are vertices, so that each piece is a singleton.

Given a tournament or digraph $\overrightarrow{\mathcal{H}} = (V, E)$ with edge-weight function $\overrightarrow{w}$, along with an ordered k-partition $\overrightarrow{\mathcal{P}}$ corresponding to a k-chotomous weak order $\succeq$ on V, we'll say that a directed edge (x, y) goes *down* if $x \succ y$, goes *up* if $y \succ x$, and goes *sideways* if $x \sim y$. For the example in Figure 1, (a, c), (a, e) and (d, f) go down; (g, b), (g, d) and (c, b) go up; and (a, b) and (f, e) go sideways. Score $\overrightarrow{v}$ is now defined by:

$$\overrightarrow{v}_{\overrightarrow{w}}(\overrightarrow{\mathcal{P}}) = \sum_{(x,y)\ goes\ down} \overrightarrow{w}(x, y) \quad - \sum_{(u,v)\ goes\ up} \overrightarrow{w}(u, v), \tag{1}$$

with weights on sideways edges omitted. Thus in Figure 1 we have

$$\overrightarrow{v}_{\overrightarrow{w}}(\overrightarrow{\mathcal{P}}) = [3 + 4 + 5] - [1 + 3 + 4] = 4. \tag{2}$$

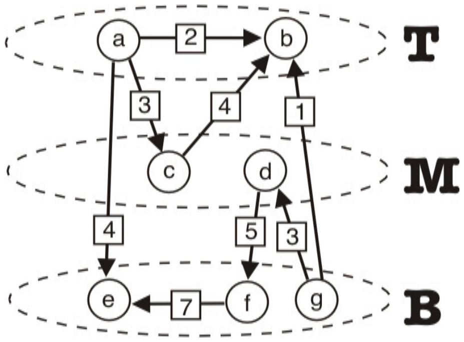

Figure 1: A directed graph and ordered 3-partition

If we adopt the following convention . . .

Definition 1 [Reversal convention] *For an edge-weight assignment $\overrightarrow{w}$ on a tournament $\overrightarrow{\mathcal{H}} = (V, E)$, the reversal convention interprets $\overrightarrow{w}(a, b)$ as $-\overrightarrow{w}(b, a)$ whenever $(a, b) \notin E$.*

. . . then equation (1) can be rewritten as:

$$\overrightarrow{v}_w(\overrightarrow{\mathcal{P}}) = \sum_{x \succ y} w(x, y). \tag{3}$$

Definition 2 *The max-kOP problem takes as inputs a tournament $\overrightarrow{\mathcal{H}}$ along with a function $\overrightarrow{w}$ that assigns real number weights to $\overrightarrow{\mathcal{H}}$'s edges, and seeks an ordered k-partition of maximal score. The corresponding decision problem is defined as one would expect.*

Why propose this particular adaptation of max-cut for tournaments? For one thing, max-*kOP* is implicit in a variety of amalgamation rules known to social choice and judgement aggregation; it is the basis for a generalization of Kemeny voting that yields these known rules as special cases (see section 3). A second justification arises from mathematical naturality; max-*kOP* can be shown equivalent to finding a vector (representing the k-chotomous weak order) that has maximal inner product with a second vector (representing the weight function $\overrightarrow{w}$). The equivalence is not discussed here.

Our immediate goal is to show that max-kOP is NP-hard for $k = 3$, but polynomial time both for $k = 2$ and for arbitrary k when $\overrightarrow{w}_{cycle} = 0$; the argument makes use of the known NP-hardness of max-cut, and is organized in the form of the following five results:

Theorem 1 *Max-tricut is polynomially reducible to max-3OP.*

Proposition 1 *Max-cut is polynomially reducible to max-tricut.*

Thus, we obtain as an immediate corollary:

Corollary 1 *Max-3OP is NP-hard.*

If we attacked max-kOP via brute force search over all ordered k-partitions (of the vertex set V of a weighted tournament), then for any fixed $k \geq 2$ we'd find that the number of such partitions grows exponentially in the number $|V|$ of vertices. The key idea behind the following Theorem 2 and Corollary 2 is that this search space can be reduced to one of size $\leq |V|^{k-1}$ for *purely acyclic* $\overrightarrow{w}$:

Theorem 2 *When restricted to inputs satisfying $\overrightarrow{w}_{\text{cycle}} = 0$ (equivalently, satisfying that $\overrightarrow{w}$ is "purely acyclic," with $\overrightarrow{w} = \overrightarrow{w}_{\text{cocycle}}$), max-$kOP$ is in P.*

Theorem 3 *For any ordered 2-partition $\overrightarrow{\mathcal{P}}$ of a directed graph $\overrightarrow{\mathcal{H}}$ with edge-weight function $\overrightarrow{w}$, $\overrightarrow{v}_{\overrightarrow{w}}(\overrightarrow{\mathcal{P}}) = \overrightarrow{v}_{\overrightarrow{w}_{\text{cocycle}}}(\overrightarrow{\mathcal{P}})$.*

Theorem 3 tells us that max-2OP is equivalent to the restricted version covered by Theorem 2, whence:

Corollary 2 *Max-2OP is in P.*

We turn now to the proof of Theorem 1. The idea is to replace a weighted graph $\mathcal{G}$ with a correspondingly weighted tournament $\overrightarrow{\mathcal{H}_{\mathcal{G}}}$, $\overrightarrow{w}$ in such a way that each tripartition $\mathcal{P}$ of $\mathcal{G}$'s vertices corresponds to an ordered tripartition $\overrightarrow{\mathcal{P}}$ of $\overrightarrow{\mathcal{H}}$'s vertices satisfying $v(\mathcal{P}) = \overrightarrow{v}(\overrightarrow{\mathcal{P}})$. The $\overrightarrow{\mathcal{H}_{\mathcal{G}}}$ construction produces, for each edge $\{a, b\}$ of $\mathcal{G}$, two new vertices (in addition to the original vertices of $\mathcal{G}$) and four directed edges. More precisely:

Definition 3 *Let $\mathcal{G} = (V, E)$ be any complete (finite, undirected) graph and $w : E \to \Re$ be an associated nonnegative edge-weight function. The tournament $\overrightarrow{\mathcal{H}_{\mathcal{G}}}$ and edge-weight function $\overrightarrow{w}$ induced by $\mathcal{G}$ and w are defined as follows:*

- *For each edge $e = \{a, b\} \in E$ of $\mathcal{G}$, construct two* direction vertices *d_{ab} and d_{ba} of $\mathcal{H}_{\mathcal{G}}$. Let $D = \{d_{ab} \mid \{a, b\} \in E\}$ denote the set of direction vertices and assume $D \cap V = \emptyset$.*
- *$\overrightarrow{\mathcal{H}_{\mathcal{G}}}$'s vertex set is $\overrightarrow{V} = D \cup V$, with elements of V referred to as* ordinary vertices.
- *Add all edges of form $a \to d_{ab}$ and $d_{ab} \to b$ to $\overrightarrow{\mathcal{H}_{\mathcal{G}}}$, with $\overrightarrow{w}$ assigning to each the original weight $w(\{a, b\})$ of $\{a, b\}$ in $\mathcal{G}$; then add enough arbitrarily directed edges to make $\overrightarrow{\mathcal{H}_{\mathcal{G}}}$ a tournament, with $\overrightarrow{w}$ assigning weight 0 to each of these.*

Notice that each edge $e = \{a, b\}$ of $\mathcal{G}$ thus contributes an $\{a, b\}$ 4-cycle

$$a \longrightarrow d_{ab} \longrightarrow b \longrightarrow d_{ba} \longrightarrow a \tag{4}$$

of directed edges in $\overrightarrow{\mathcal{H}_{\mathcal{G}}}$, each with weight $w(\{a, b\})$. In particular, $\overrightarrow{w}$ is *purely cyclic.* The combinatorial core of the Theorem 1 proof consists of the following:

Lemma 1 (Fitting a four-cycle into three levels) *Let $\overrightarrow{\mathcal{P}} = \{T > M > B\}$ be any ordered tripartition of the vertex set $\overrightarrow{V}$ of $\overrightarrow{\mathcal{H}_{\mathcal{G}}}$. Then for each weight w edge $e = \{a, b\}$ of $\mathcal{G}$:*

- *if a and b belong to the same piece of $\overrightarrow{\mathcal{P}}$ then the net contribution to the score $\overrightarrow{v}(\overrightarrow{\mathcal{P}})$ made by the edges of the $\{a, b\}$ 4-cycle is zero, and*
- *if a and b belong to any two different pieces of $\overrightarrow{\mathcal{P}}$ then, by appropriately reassigning the direction vertices d_{ab} and d_{ba} among T, M, and B, we can set the net contribution to $\overrightarrow{v}(\overrightarrow{\mathcal{P}})$ made by the edges of the $\{a, b\}$ 4-cycle equal to 0, or w, or $-w$, as we prefer.*

Proof: (Of Lemma 1) Figures 2L, 2C, and 2R (for *Left, Center, Right*) show three possible ways to assign the four vertices a, b, d_{ab} and d_{ba} to membership in the three pieces of $\overrightarrow{\mathcal{P}}$. In 2R ordinary vertices a and b belong to the same piece (here, piece M) of $\overrightarrow{\mathcal{P}}$. Of the four directed edges in the $\{a, b\}$ 4-cycle, two are up edges and two are down edges, so if each edge has weight w the net contribution of these four edges is zero. More generally, whenever $a, b \in M$ it is easy to see that the number of up edges from the $\{a, b\}$ 4-cycle must equal the number of down edges, no matter where d_{ab} and d_{ba} are placed, and that this remains true in case $a, b \in T$ or $a, b \in B$. Thus the net contribution is 0 whenever the ordinary vertices a and b are in the same piece.

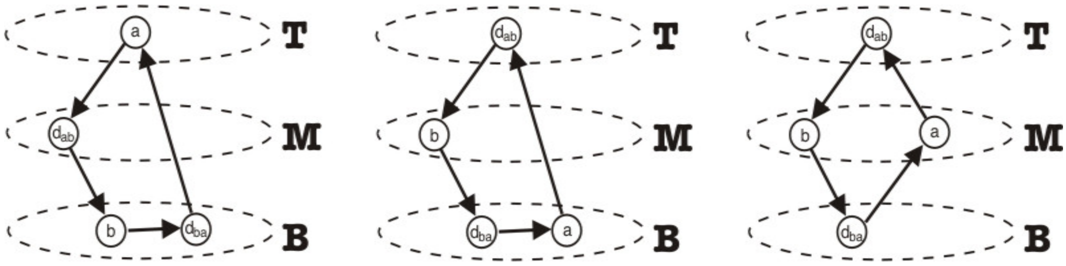

Figure 2: Some possible ways to fit an $\{a, b\}$ 4-cycle into three levels.

In 2L and 2C, ordinary vertices a and b are in different pieces, and we have placed d_{ab} and d_{ba} so that there are two down edges and one up edge. If each edge has weight w then the net contribution of the four edges shown is w. If we exchange the placements of d_{ab} and d_{ba} in 2L (or in 2C), we wind up with two up edges and one down edge for a net contribution of $-w$; if we move d_{ab} and d_{ba} into a common piece, then (as in the previous paragraph) the number of up edges will be equal to the number of down edges, for a net contribution of zero. A moment's thought will convince the reader that for *all* cases in which ordinary vertices a and b belong to different pieces, exactly three possibilities – two up edges + one down, two down edges + one up, or equal numbers of up and down edges – can be achieved by moving d_{ab} and d_{ba} around. This completes the Lemma 1 proof. ∎

Proof: (Of Theorem 1) It suffices to show that given an edge-weighted graph $\mathcal{G}$ and a positive integer k the answer to the decision problem "Does there exist a vertex tripartition $\mathcal{P} = \{J, K, L\}$ of V with $v(\mathcal{P}) \geq k$?" is the same as the answer to "Does there exist an ordered tripartition $\overrightarrow{\mathcal{P}}$ of the vertex set $\overrightarrow{V}$ of $\overrightarrow{\mathcal{H}_{\mathcal{G}}}$ with $\overrightarrow{v}(\overrightarrow{\mathcal{P}}) \geq k$?"

Lemma 1 makes this easy. Given a tripartition $\mathcal{P} = \{J, K, L\}$ of V with $v(\mathcal{P}) = j \geq k$, arbitrarily order $\{J, K, L\}$ as $\{J > K > L\}$. For each weight w edge $\{a, b\}$ of $\mathcal{G}$ cut by $\mathcal{P}$, add each direction vertex d_{ab}, d_{ba} to one of the sets in $\{J > K > L\}$, so as to create two down edges and one up edge from the $\{a, b\}$ 4-cycle; for each original uncut edge $\{a, b\}$ of $\mathcal{G}$

add each vertex d_{ab}, d_{ba} to one of the sets $\{J > K > L\}$ according to the arbitrary dictates of your current mood. It is easy to see that the resulting $\overrightarrow{\mathcal{P}}$ achieves the exact same score: $\overrightarrow{v}(\overrightarrow{\mathcal{P}}) = v(\mathcal{P}) = j \geq k$.

In the other direction, consider an ordered tripartition $\overrightarrow{\mathcal{P}} = \{V_1 \cup D_1 > V_2 \cup D_2 > V_3 \cup D_3\}$ of $\overrightarrow{V}$ with $V_1 \cup V_2 \cup V_3 = V$, $D_1 \cup D_2 \cup D_3 = D$, and $\overrightarrow{v}(\overrightarrow{\mathcal{P}}) \geq k$. Let $\mathcal{P} = \{V_1, V_2, V_3\}$, a tripartition of V. Each weight w edge $\{a, b\}$ of $\mathcal{G}$ cut by $\mathcal{P}$ contributes w to $v(\mathcal{P})$ and contributes w or 0 or $-w$ to $\overrightarrow{v}(\overrightarrow{\mathcal{P}})$. Thus $v(\mathcal{P}) \geq \overrightarrow{v}(\overrightarrow{\mathcal{P}}) \geq k$, as desired. ■

Proof: (of Proposition 1) The reduction is easy and uninteresting, so we omit details. Given an (undirected) graph $\mathcal{G} = (V, E)$ and edge-weight assignment w, create $\mathcal{G}^\star$ and $w^\star$ as follows: add a new vertex ♣ along with edges $\{♣, v\}$ for each $v \in V$, and extend w by assigning weight $\sigma = 1 + \sum_{\{a,b\} \in V} w(\{a, b\})$ to each added edge. Then any maximal-score tripartition $\mathcal{P}^\star$ of $\mathcal{G}^\star$ will place ♣ alone in one of the three pieces, while the other two pieces constitute a maximal-score bipartition $\mathcal{P}$ of $\mathcal{G}$, with $v(\mathcal{P}^\star) = |V|\sigma + v(\mathcal{P})$. Thus, there exists a bipartition $\mathcal{P}$ of $\mathcal{G}$ with score at least k if and only if there exists a bipartition $\mathcal{P}^\star$ of $\mathcal{G}^\star$ with score at least $|V|\sigma + k$. ■

The proofs of Theorems 2 and 3 exploit the decomposition

$$\overrightarrow{w} = \overrightarrow{w}_{cycle} + \overrightarrow{w}_{cocycle} \tag{5}$$

from [11], discussed in greater detail in [5], and use the following abstract definition of Borda score[4] for a vertex x of a weighted tournament:

Definition 4 (Reversal convention from Definition 1 applies) *Given a vertex x of a tournament $\overrightarrow{\mathcal{H}}$ equipped with edge-weight assignment $\overrightarrow{w}$, x's Borda score is given by:*

$$x^\beta = \sum_{y \in V} \overrightarrow{w}(x, y) \tag{6}$$

Definition 5 (Reversal convention applies) *An edge-weight assignment $\overrightarrow{w}$ on a tournament $\overrightarrow{\mathcal{H}} = (V, E)$ satisfies* exact quantitative transitivity *if*

$$\overrightarrow{w}(x, y) + \overrightarrow{w}(y, z) = \overrightarrow{w}(x, z) \tag{7}$$

holds for every three distinct vertices $x, y, z \in V$.

Definition 6 (Reversal convention applies) *An edge-weight assignment $\overrightarrow{w}$ on a tournament $\overrightarrow{\mathcal{H}} = (V, E)$ is* difference generated *if there exists a function $\Gamma : V \to \Re$ such that*

$$\overrightarrow{w}(x, y) = \Gamma(x) - \Gamma(y) \tag{8}$$

holds for every two distinct vertices $x, y \in V$. In this case, we can identify the vertices of $\overrightarrow{\mathcal{H}}$ with a sequence of real numbers

$$\gamma_1 \leq \gamma_2 \leq \cdots \leq \gamma_m \tag{9}$$

enumerating Γ's values in non-decreasing order.

[4] In section 3 we obtain a tournament $\overrightarrow{\mathcal{H}}_\Pi = (A, E)$ and edge-weighting $\overrightarrow{w}_\Pi$ from a profile Π of weak (or linear) orders over a finite set A of m alternatives. The score of a vertex $a \in A$ according to Definition 4 (above) then coincides with the conventional notion of a's Borda score based on Π, as calculated using the "symmetric" Borda weights $m - 1, m - 3, \ldots,\ 3 - m, 1 - m$.

Lemma 2 *Given an edge-weight assignment $\overrightarrow{w}$ on a tournament $\overrightarrow{\mathcal{H}} = (V, E)$, the following are equivalent:*

1. *$\overrightarrow{w}$ satisfies exact quantitative transitivity,*
2. *$\overrightarrow{w}$ is difference generated,*
3. *$\overrightarrow{w}$ is purely acyclic (equivalently, $w_{\text{cycle}} = \mathbf{0}$ in the vector orthogonal decomposition $\overrightarrow{w} = \overrightarrow{w}_{\text{cycle}} + \overrightarrow{w}_{\text{cocycle}}$; equivalently, $\overrightarrow{w} \in \mathbf{V}_{\text{cocycle}}$, the cocycle subspace).*

Proof: We leave the easy (1) $\Leftrightarrow$ (2) equivalence to the reader. If $\overrightarrow{w}$ is purely acyclic, then as an immediate consequence of Observation 11.2 of [5], $\overrightarrow{w}$ is difference generated via the function assigning scaled Borda scores:

$$\Gamma\colon x \mapsto \frac{x^{\beta}}{|V|}. \tag{10}$$

Conversely, assume $\overrightarrow{w}$ is difference generated via Γ, and let $x_1,\ x_2,\ \ldots, x_r,\ x_1$ be any cycle of vertices. The corresponding *basic cycle σ* is an edge-weighting of $\overrightarrow{\mathcal{H}}$ that assigns weight one to each edge $x_i \to x_{i+1}$ or $x_r \to x_1$ from the vertex cycle (under the reversal convention), and weight zero to each other edge. Thus

$$\overrightarrow{w}\cdot\sigma = \big[\Gamma(x_1)-\Gamma(x_2)\big]+\big[\Gamma(x_2)-\Gamma(x_3)\big]+\cdots+\big[\Gamma(x_{r-1})-\Gamma(x_r)\big]+\big[\Gamma(x_r)-\Gamma(x_1)\big] = 0 \tag{11}$$

It follows from linearity of the dot product that $\overrightarrow{w} \cdot \tau = 0$ holds for any linear combination of basic cycles – hence $\overrightarrow{w} \perp \mathbf{V}_{cycle}$, and $\overrightarrow{w} \in \mathbf{V}_{cocycle}$. Thus $\overrightarrow{w}$ is purely acyclic. (The argument is like that for Proposition 15 in [5].) ∎

Definition 7 *An ordered k-partition $\mathcal{P} = \{P_k >' P_{k-1} >' \cdots >' P_1\}$ of a nondecreasing sequence $\gamma_1 \le \gamma_2 \le \cdots \le \gamma_m$ of real numbers is* monotone *if $i < j \Rightarrow \pi(\gamma_i) \le' \pi(\gamma_j)$, where $\le'$ refers to the ordering of $\mathcal{P}$'s pieces, and $\pi(\gamma_i)$ denotes the piece P_s for which $\gamma_i \in P_s$.*

Equivalently, monotone partitions are obtained by "cutting" the γ sequence from line (9) with $k-1$ dividers $\downarrow_i$:

$$\underline{\gamma_1, \gamma_2, \ldots, \gamma_{m_1}}\ \downarrow_1\ \underline{\gamma_{m_1+1}, \ldots, \gamma_{m_2}}\ \downarrow_2\ \cdots\ \downarrow_{k-1}\ \underline{\gamma_{m_{k-1}+1}, \ldots, \gamma_m} \tag{12}$$

Lemma 3 *Given a purely acyclic edge-weight assignment $\overrightarrow{w}$ on a tournament $\overrightarrow{\mathcal{H}} = (V, E)$, there exists a monotone ordered k-partition of $V = \{\gamma_1 \le \gamma_2 \le \cdots \le \gamma_m\}$ that achieves maximal score.*

Proof: (of Lemma 3) It is straightforward to show that if some ordered partition $\overrightarrow{\mathcal{P}}$ satisfied $i < j$ with $\pi(\gamma_i) >' \pi(\gamma_j)$ then swapping γ_i for γ_j (by moving γ_i into the piece to which γ_j initially belonged, and γ_j into γ_i's initial piece) can never decrease $\overrightarrow{\mathcal{P}}$'s score. A sequence of such swaps converts $\overrightarrow{\mathcal{P}}$ into a monotone partition. ∎

Proof: (of Theorem 2) Given an instance of max-kOP with purely acyclic $\overrightarrow{w}$, calculate the scaled Borda scores $\gamma_1 \le \gamma_2 \le \cdots \le \gamma_m$ from line (10) and identify them with the m vertices. Calculate $\overrightarrow{v}(\overrightarrow{\mathcal{P}})$ for each possible monotone ordered k-partition, of which there are at most $(m-1)^{k-1}$, because there are at most $m-1$ options for placing each divider $\downarrow_i$ in line (12); output any optimal partition and its score. This calculation is polynomial in the size of the tournament and maximum edge weight, so acyclic max-kOP is in P. ∎

Proof: (of Theorem 3) For any ordered partition $\overrightarrow{\mathcal{P}}$ of a directed graph $\overrightarrow{\mathcal{H}}$, the score $\overrightarrow{v}(\overrightarrow{\mathcal{P}})$ is a linear functional on the vector space of all possible edge weightings $\overrightarrow{w}$, so that $\overrightarrow{v}_{\overrightarrow{w}}(\overrightarrow{\mathcal{P}}) = \overrightarrow{v}_{\overrightarrow{w}_{cocycle}}(\overrightarrow{\mathcal{P}}) + \overrightarrow{v}_{\overrightarrow{w}_{cycle}}(\overrightarrow{\mathcal{P}})$. Thus, once we demonstrate that for 2-partitions $\overrightarrow{v}_{\overrightarrow{w}_{cycle}}(\overrightarrow{\mathcal{P}}) = 0$, it follows that 2-partitions also satisfy $\overrightarrow{v}_{\overrightarrow{w}}(\overrightarrow{\mathcal{P}}) = \overrightarrow{v}_{\overrightarrow{w}_{cocycle}}(\overrightarrow{\mathcal{P}})$.

Next, observe that the multiple options for fitting a cycle into three levels of an ordered partition (Lemma 1, Figure 2) are severely constrained for ordered partitions having only two levels.

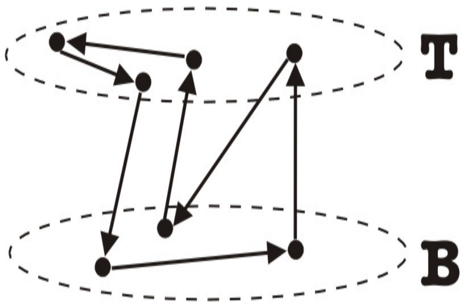

Figure 3: Fitting a cycle into two levels.

As suggested by the example in Figure 3, for 2-partitions the number of down edges will always equal the number of up edges. Thus, for the basic cycle σ that assigns weight 1 to each edge that appears in Figure 3, and weight 0 to every edge not drawn in, $\overrightarrow{v}_{\sigma}(\overrightarrow{\mathcal{P}}) = 0$. By linearity, the same holds for any linear combination of basic cycles, and so we conclude that for ordered 2-parttions $\overrightarrow{\mathcal{P}}$, $\overrightarrow{v}_{\overrightarrow{w}_{cycle}}(\overrightarrow{\mathcal{P}}) = 0$. ∎

3 The (j, k)-Kemeny rule

Suppose that $V = \{v_1, \ldots, v_m\}$ is a (finite) set of m alternatives, and that voters in a finite set N cast weak (or linear) order ballots, resulting in a profile $\Pi = \{\geq_i\}_{i \in N}$. The induced tournament $\overrightarrow{\mathcal{H}_\Pi} = (V, E)$ and edge-weights $\overrightarrow{w}_\Pi$ are as follows:

- $E = \{(v_i, v_j) \mid i < j\}$. Remark: This adds one directed edge for each two vertices.
- $\overrightarrow{w}_\Pi(v_i, v_j) = |\{t \in N \mid v_i \geq_t v_j\}| - |\{t \in N \mid v_j \geq_t v_i\}|$. Remark: These weights are the net majorities by which voters favor v_i over v_j.[5]

Definition 8 *The* (j, k)-Kemeny rule *takes, as input, a profile* Π *of* j*-chotomous weak orders on a finite set* V *of alternatives, and outputs the* k*-chotomous weak order(s)*[6] *on* V *corresponding to the solution(s) of max-*k*OP for* $\overrightarrow{\mathcal{H}_\Pi}$, $\overrightarrow{w}_\Pi$*. The* $(j, |V|)$-, $(|V|, k)$- *and* $(|V|, |V|)$*-Kemeny rules are defined similarly, with linear ordered ballots when* $|V|$ *appears in the* j *position, and linearly ordered outputs when* $|V|$ *appears in the* k *position. A* $2^\star$ *in either position refers to those ordered* 2*-partitions (equivalently, dichotomous weak orders)* $\{T > B\}$ *for which* $T = \{x\}$ *is a singleton (contains a single alternative* x*).*

[5] Loosely, the edge-weighted tournament provides the "C2" information, in Fishburn's classification, [6]. Note that $w(v_i, v_j)$ is negative when $i < j$ and more voters strictly prefer v_j to v_i than strictly prefer v_i to v_j – this compensates for having initially oriented the (v_i, v_j) edge in the direction opposite to the majority. We won't have the option of sticking to positive weights by adjusting the direction of the edge, because the two components of the decomposition can differ as to which direction to choose.

[6] Ties are possible. When the number of ties is large, there may be an exponential blow-up in the number of orders in the output. However for rules (1)-(5) of Proposition 2 the output can be described in a compact language that describes a class of tied orders in terms of ties among individual alternatives.

Proposition 2 *Note that a dichotous weak order $\{T > B\}$ can be interpreted as an approval ballot approving all alternatives in T. When $T = \{x\}$, an output $\{T > B\}$ can be interpreted as naming x as winner, while an input $\{T > B\}$ can be interpreted as a plurality ballot for x. With that understanding, special cases of (j, k)-Kemeny*[7] *include:*

1. $(2, 2)$*-Kemeny is the Mean Rule.*[8]
2. $(2, |V|)$*-Kemeny and* $(2, 2^\star)$*-Kemeny are approval voting (with outcome the ranking(s) by approval score, and the approval winner(s), respectively).*
3. $(2^\star, |V|)$*-Kemeny and* $(2^\star, 2^\star)$*-Kemeny are plurality voting (with outcome the ranking(s) by plurality score, and the plurality winner(s), respectively).*
4. $(|V|, 2)$*-Kemeny is the Borda Mean Rule.*[9]
5. $(|V|, 2^\star)$*-Kemeny is the Borda count voting rule.*
6. $(|V|, |V|)$*-Kemeny is the Kemeny voting rule (with a ranking as output).*

The results of Theorems 2 and 3 and of Corollary 2 now lift immediately to (j, k)-Kemeny, showing:

Theorem 4 *The problem of determining the winning ordering for a* $(\_\__1, \_\__2)$*-Kemeny election is in P whenever at least one of the blanks contains 2 (or* $2^\star$*), and also whenever* $\overrightarrow{w}_{\Pi_{\text{cycle}}} = 0$*. In particular, winner determination is in P for rules (1) – (5) of Proposition 2. Also, for profiles satisfying* $\overrightarrow{w}_{\Pi_{\text{cycle}}} = 0$*, the* $(\_\__1, |V|)$*-Kemeny outcome is the linear ranking induced by Borda scores; in particular, the original Kemeny rule agrees with Borda.*

We need to be a bit more careful when lifting the NP-hardness results from Theorem 1, Proposition 1, and Corollary 1 to the context of $(\_\__1, \_\__2)$-Kemeny, however. To argue for NP-hardness when $\_\__1$ is filled with either a fixed $j \geq 3$ or with $|V|$, we need to know that the specific weighted tournaments $\overrightarrow{\mathcal{H}_\mathcal{G}}$, $\overrightarrow{w}$ constructed in the proof of Theorem 1 are induced as $\overrightarrow{\mathcal{H}_\Pi}$, $\overrightarrow{w}_\Pi$ for some profile Π of j-chotomous orders ($j \geq 3$), and for some profile of linear orders. Actually, it suffices to induce some scalar multiple $C\overrightarrow{w}$ of the Theorem 1 weights as $\overrightarrow{w}_\Pi$, for each of these types of profile. But given an arbitrary integer-valued $\overrightarrow{w}$, for $j \geq 3$ it is straightforward to construct a profile Π of j-chotomous weak orders (or of linear orders) satisfying $\overrightarrow{w}_\Pi = 2\overrightarrow{w}$. To make the argument when $\_\__2$ is filled by a fixed $k \geq 3$ we need versions of Theorem 1 and Proposition 1 asserting "Max-$k$cut is polynomially reducible to max-kOP," and "Max-cut is polynomially reducible to max-$k$cut," but these are straightfoward generalizations, and we omit the details.

Theorem 5 *The problem of determining the winning ordering for a* $(\_\__1, \_\__2)$*-Kemeny election is* NP*-hard whenever*

- $\_\__1$ *is filled with either a fixed* $j \geq 3$ *or with* $|V|$*, and*
- $\_\__2$ *is filled with a fixed* $k \geq 3$

both hold. In particular winner determination is NP*-hard for* $(3, 3)$*-Kemeny.*

[7] When $j = |V|$ or $k = |V|$ (or both), (j, k)-Kemeny coincides with the corresponding (j, k)-median procedure; otherwise, these aggregation rules differ. We'll have more to say about this in a planned expansion of the current conference version of this paper.

[8] The Mean Rule outcome ranks all alternatives with above average approval score over all those with below average score; see [5] for details.

[9] Borda Mean Rule acts like the Mean Rule, but with Borda scores replacing approval scores.

None of our reasoning here shows NP-hardness when $\_\_\__2$ is filled with $|V|$; in particular, Theorem 5 does not allow us to draw hardness conclusions for $(|V|, |V|)$-Kemeny (that is, for the original Kemeny rule itself) or for $(j, |V|)$-Kemeny with $j \geq 3$, because max-cut is *not* polynomially reducible to "max-$|V|$cut."[10] Nonetheless, our argument that computational complexity arises from the cyclic component also applies to the cases missing from Theorem 5. We know from [2] that the original Kemeny rule winner problem is NP-hard, and the last clause of Theorem 4 tells us that Kemeny reduces to a computationally easy rule when $\overrightarrow{w}_{\Pi_{cycle}} = 0$. As for $(j, |V|)$-Kemeny, the comments preceding the Theorem 5 statement reveal that for $j \geq 3$ the induced weights $\overrightarrow{w}_{\Pi}$ from profiles of j-chotomous weak orders are essentially as general as those arising from linear rankings, so winner determination is as hard as for the original Kemeny rule.

We end by mentioning two interesting questions raised by the COMSOC referees. One referee points out that the full power of pure acyclicity (equivalently, of quantitative transitivity) is not needed to make Kemeny's rule easy to compute – ordinary transitivity of majority preference suffices. So, is ordinary transitivity similarly sufficient to make the $(3, 3)$-Kemeny winner problem tractable? A natural initial approach to this question is to reconsider Lemma 3 in this light, and ask what happens if we restrict our search to partitions that are monotone with respect to the linear ordering induced by a transitive majority preference relation – would such a limited search necessarily find an ordered tripartition achieving maximal score? It is not difficult to see that the answer is *no, not necessarily.* About all we can say about the question at this point is that with this obvious approach falling short, we do not know the answer.

A second referee asks what happens to tractability when the cyclic component is simple – what happens, for example, if $\overrightarrow{w}_{cycle}$ can be written as a sum of only one or two simple cycles? We conjecture (but with low confidence) that winner determination for $(3, 3)$-Kemeny would indeed become tractable in this case. In this connection, it seems worth mentioning that the dimension of the cocyclic subspace grows linearly with the number of alternatives, while the dimension of the cyclic space grows quadratically. In a sense, then, the cyclic component is inherently the more complicated one, so that sharply limiting its complexity places a rather strong restriction on the underlying profile.

References

[1] Barthélemy, J.P., and Monjardet, B. The median procedure in cluster analysis and social choice theory. *Mathematical Social Sciences* 1, 235-267, 1981.

[2] Bartholdi, J. III, Tovey, C.A. , and Trick, M.A. Voting Schemes for which It Can Be Difficult to Tell Who Won the Election *Social Choice and Welfare* 6, 157-165, 1989.

[3] Conitzer, V., Rognlie, M., and Xia, L. Preference Functions That Score Rankings and Maximum Likelihood Estimation. In *Proceedings of the Twenty-First International Joint Conference on Artificial Intelligence* (IJCAI-09), 109-115, 2009.

[4] Duddy, C., and Piggins, A. Collective approval, *Mathematical Social Sciences* 65, 190-194, 2013.

[5] Duddy, C., Piggins, A., and Zwicker, W.S. Aggregation of binary evaluations: a Borda-like approach. *Social Choice and Welfare* 46, 301-322, 2016.

[6] Fishburn, P.C. Condorcet social choice functions. *SIAM Journal on Applied Mathematics* 33, 469-489, 1977.

[10] "Max-$|V|$cut" is in quotes because it is silly.

[7] Hemaspaandra, E., Spakowski, H., and Vogel, J. The complexity of Kemeny elections. *Theoretical Computer Science* 349, 382-391, 2005.

[8] Hudry, O. On the computation of median linear orders, of median complete preorders and of median weak orders. *Mathematical Social Sciences* 64 (special issue on Computational Foundations of Social Choice, F. Brandt and W.S. Zwicker, eds.) 2-10, 2012.

[9] Myerson, R.B. Axiomatic derivation of scoring rules without the ordering assumption. *Social Choice and Welfare* 12, 59–74, 1995.

[10] Xia, L. and Conitzer, V. A maximum likelihood approach towards aggregating partial orders. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence* (IJCAI-11), 446-451, 2011.

[11] Zwicker, W.S. The voters' paradox, spin, and the Borda count. *Mathematical Social Sciences* 22, 187-227, 1991.

[12] Zwicker, W.S. Consistency Without Neutrality in Voting Rules: When is a Vote an Average? *Mathematical and Computer Modelling* 48 (special issue on Mathematical Modeling of Voting Systems and Elections: Theory and Applications, A. Belenky, ed.), 1357-1373, 2008.

William S. Zwicker
Department of Mathematics
Union College
Schenectady, NY 12308
USA Email: `zwickerw@union.edu`