

Cycles and Intractability in Social Choice Theory – August 18, 2016

William S. Zwicker^{1,2}

Abstract

We introduce the (j, k) -Kemeny rule – a generalization 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 $\vec{w} = \vec{w}_{cycle} + \vec{w}_{cocycle}$ of [16]. Thus the existence of 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.

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 [11] 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 [6] and [7]), and the Borda mean rule [7]; with an additional restriction one also obtains Borda and plurality voting.³

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 or tournaments⁴ 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 $\vec{w} = \vec{w}_{cycle} + \vec{w}_{cocycle}$ induced by a profile (see [16]; a more complete exposition appears in [7]). In particular, $j \leq 2$ guarantees $\vec{w}_{cycle} = 0$, while $k \leq 2$ guarantees that \vec{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 $w_{cycle} = 0$ (in which case we say the profile is *purely acyclic*, with no *hidden* majority cycles), then the winner can be computed in polynomial time for any values

¹Department of Mathematics, Union College, Schenectady NY 12308, USA. Email: zwickerw@union.edu

²We thank Matthew Anderson, Markus Brill, Dominik Peters, and Alan D. Taylor for their help with content as well as presentation, and thank the referees for an earlier (COMSOC 2016 conference) version, for suggesting some interesting follow-up questions.

³The (j, k) -Kemeny rule discussed here is akin to, but not the same as, the median procedure of [1]; see further discussion in the proof of Proposition 2.

A planned sequel [18] will explain the relationship between these two general aggregation procedures and among various specific rules obtained from them as restrictions.

⁴Weights are assigned to *arcs* (directed edges). A *tournament* is a digraph wherein each pair a, b of vertices is linked by a single edge oriented in one of the two directions possible. For our purposes here, the tournament/digraph distinction is not important; *tournament* will be our default term. See also footnote 7.

of j and k ; one example is that Kemeny = Borda when $\vec{w}_{cycle} = \mathbf{0}$. 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. Note the distinction here between *overt* majority cycles and *hidden* ones; it is common for a profile of strong (or weak) orders to have hidden cycles ($\vec{w}_{cycle} \neq \mathbf{0}$) yet lack any overt ones. Equivalantly, the pairwise majority relation is transitive in the usual sense, but fails to have the stronger property of *quantitative transitivity* (see Definition 5).⁵

Section 2 consists of a very brief, qualitative description of the orthogonal decomposition of a weighted tournament into cyclic and cocyclic components. For readers to whom these ideas are new, we recommend the more detailed exposition in [7], which includes worked examples and diagrams. In section 3 we dissect the relationship between the standard *max-cut* problem for (weighted, undirected) graphs, and our version *max-kOP* for tournaments. 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 4, and observe 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 3 to the winner determination problem for these rules.

Our hardness proof reduces *max-kcut* to *max-kOP*. A first construction induces, from any weighted (undirected) graph, a *purely cyclic* ($\vec{w}_{cocycle} = \mathbf{0}$) weighted digraph with overt cycles. Might overt majority cycles alone suffice to explain intractability in *max-kOP*? If so, there would be no need to appeal to the orthogonal decomposition or hidden cycles to explain why some instances of *max-kOP* are in *P* while others are *NP*-hard. A second, somewhat more complex construction (in the appendix) reduces *max-2cut* to the transitive subcase of *max-3OP*; here the induced weighted tournament always yields a transitive order on the vertices, with no overt cycles. Thus the first construction establishes that the cyclic component taken alone leads to intractability, while the second, by showing that the transitive subcase remains *NP*-hard, establishes the essential role of hidden cycles in explaining the hardness boundaries among various aggregation rules spawned by (j, k) -Kemeny; looking for overt cycles alone would not suffice.⁶

Several of the ideas developed here (but not those related to complexity) first appeared in [7]. Notions of generalized scoring rule implicit in [13] and explicit in [17], [4], [15], and [18] 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 [12] and the comments at the start of the proof of Proposition 2).

2 About the decomposition ...

Given a tournament on a set V of vertices, along with an assignment w of numerical weights to the arcs, these edge-weights can be interpreted as the *flow* of some substance through the channels of a network. For example, a 5 on the $a \rightarrow b$ edge could indicate a flow of 5 amps of electricity in a wire from a to b , or of 5 gallons of water per minute in a pipe, or (in

⁵We are being a bit loose here, as there are multiple versions of (qualitative) transitivity and of acyclicity; one example is that transitivity of the strict majority preference relation $>^\mu$ for weak order ballots (“more ballots rank x strictly above y than rank y strictly above x ”) is equivalent to the absence of cycles of form $a_1 >^\mu a_2 >^\mu \dots >^\mu a_n \geq^\mu a_1$.

⁶We are indebted to one of the referees of an earlier *COMSOC* conference version of this paper, who pointed us to this issue. The referee observed that if we assume transitivity of the majority preference relation, identifying the winning order for the standard Kemeny rule (for aggregating linear orders into a linear order) is computationally easy, and asked whether the same assumption might suffice to render the $(3, 3)$ -Kemeny winner problem easy. The second reduction, in the appendix, shows the answer to be “No.”

the case of interest for us) a flow of *net preference* for a over b : given a *profile* $\Pi = \{\geq_i\}_{i \in N}$ of weak order ballots (see Section 4) the 5 would then indicate the margin by which voters $i \in N$ who rank a over b outnumber those who rank b over a . Sticking with the electricity metaphor for the moment, a weight of *negative* 5 on the $a \rightarrow b$ edge tells us that the current is actually flowing from b to a .⁷ An assignment of weight 1 to each of the edges in a cycle

$$a_1 \rightarrow a_2 \rightarrow \cdots \rightarrow a_k \rightarrow a_1^8$$

and of weight 0 to each edge off the cycle, is called a *basic cycle* (or *loop current* in electrical engineering speak); a *basic cocycle* (or *source*) at vertex a assigns weight 1 to each edge $a \rightarrow x$ from a to another vertex, weight *negative* 1 to each edge $a \leftarrow x$, and weight 0 to each edge not incident to a .

Now let's throw some linear algebra at the situation. Any fixed enumeration of the *arcs* of a tournament allows us to identify each edge-weight assignment w with a point \vec{w} in the vector space $\mathfrak{R}^{\frac{m(m-1)}{2}}$ (where $m = |V|$)⁹, endowed with the standard inner product. With this identification, we obtain subspaces \mathbf{V}_{cycle} and $\mathbf{V}_{cocycle}$ as the linear spans, respectively, of all basic cycles and of all basic cocycles. One can then show that these spaces are orthogonal complements in $\mathfrak{R}^{\frac{m(m-1)}{2}}$, so that every vector \vec{w} in the space has a unique decomposition $\vec{w} = \vec{w}_{cycle} + \vec{w}_{cocycle}$ as a sum in which $\vec{w}_{cycle} \in \mathbf{V}_{cycle}$, $\vec{w}_{cocycle} \in \mathbf{V}_{cocycle}$ (and $\vec{w}_{cycle} \cdot \vec{w}_{cocycle} = 0$).¹⁰ We say that \vec{w} is *purely cyclic* if $\vec{w} = \vec{w}_{cycle}$ (with $\vec{w}_{cocycle} = \mathbf{0}$); \vec{w} is *purely acyclic* if $\vec{w} = \vec{w}_{cocycle}$ (with $\vec{w}_{cycle} = \mathbf{0}$). A purely cyclic \vec{w} “has no hidden cycles.” Pure acyclicity is equivalent to the strong, quantitative form of transitivity of Definition 5.

In electrical engineering, this decomposition serves as the mathematical foundation of *Kirchoff's Laws* of circuit theory, but its roots lie in one-dimensional homology theory (part of algebraic topology – see [5] and [10]). In particular, the orthogonal projection of \vec{w} onto $\mathbf{V}_{cocycle}$ (which yields $\vec{w}_{cocycle}$) coincides with the *boundary map* of homology.

The decomposition was first applied to the flow of net preference in [16], and was later exploited in [14]. Quite recently, it was used in [7] to characterize the *mean rule* for aggregation of dichotomous weak orders, and in [3] to characterize *maximal lotteries*. The relevance of the decomposition to profiles of weak or strict preference ballots can be appreciated from the following points:

- The cyclic component $\vec{w}_{cocycle}$ of \vec{w} assigns, to each edge $a \rightarrow b$ the scaled difference $\frac{a^\beta - b^\beta}{|V|}$ in symmetric Borda scores of the two alternatives ... so we can say that “The Borda count is the boundary map.”
- Thus, for a purely acyclic edge-weighting the pairwise majority relation yields a transitive ranking, which agrees with the ranking induced by Borda scores.
- More generally, these two rankings agree when the cocyclic component is dominant in determining the sign of each edge-weight. However, when the cyclic component becomes large enough to reverse some edge weight signs, without being so large as to introduce Condorcet cycles (“overt” majority cycles), these two rankings differ, and the difference can be attributed to the “hidden” cycles of the cyclic component.

⁷ In the analysis here, we view the initial edge orientation as an arbitrary choice that serves bookkeeping purposes and does not indicate the actual flow direction. A more sophisticated approach includes both arcs $a \rightarrow b$ and $b \rightarrow a$ for each vertex pair, in a “complete” digraph. For a weighted tournament one then requires that the weight assignment w be antisymmetric, satisfying $w(a, b) = -w(b, a)$. In [18] we drop that requirement, allowing for the decomposition of w into (orthogonal) antisymmetric and symmetric components; the symmetric component can then be interpreted as a weight assignment to *undirected* edges.

⁸With *negative* 1 assigned, of course, to any edge $a_j \leftarrow a_{j+1}$ oriented oppositely to the sense of the cycle.

⁹The i^{th} coordinate of \vec{w} is just the weight assigned, by w , to the i^{th} edge in the enumeration.

¹⁰The electrical engineer would say that every flow of a current in a circuit can be written uniquely as a superposition of loop currents added to a superposition of sinks and sources.

3 Max-cut for directed graphs or tournaments

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 or directed graphs $\vec{\mathcal{H}} = (V, E)$ (with $E \subseteq V \times V$), similarly equipped with functions \vec{w} assigning real-number weights (which may be negative) to arcs¹¹, with a linearly ordered partition of the vertices playing 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 $\vec{\mathcal{P}} = \{T > B\}$ is 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 $\vec{\mathcal{H}} = (V, E)$ with edge-weight function \vec{w} , along with an ordered k -partition $\vec{\mathcal{P}}$ corresponding to a k -chotomous weak order \succeq on V , we’ll say that an arc (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 \vec{v} is now defined by:

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

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

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

If we adopt the following convention ...

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

¹¹We put arrows over symbols for directed graphs or tournaments and over ordered partitions, to distinguish the denoted objects from ordinary graphs and unordered partitions. Assuming $\vec{\mathcal{H}}$ 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 $\vec{w} = \vec{w}_{\text{cycle}} + \vec{w}_{\text{cocycle}}$.

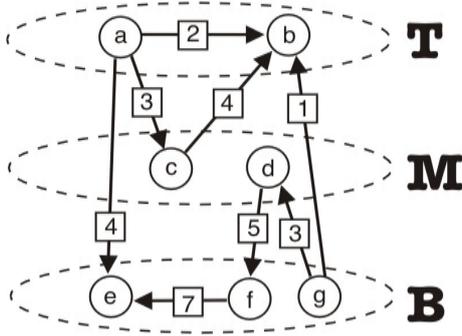


Figure 1: A directed graph and ordered 3-partition

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

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

Definition 2 The *max- k OP* problem takes as inputs a tournament $\vec{\mathcal{H}}$ along with a function \vec{w} that assigns integer weights to $\vec{\mathcal{H}}$'s arcs, 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- k OP* 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 4). A second justification arises from mathematical naturality; *max- k OP* 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 \vec{w}). The equivalence will be discussed in [18].

Our immediate goal is to show that *max- k OP* is *NP*-hard for $k = 3$, but polynomial time both for $k = 2$ and for arbitrary k when $\vec{w}_{\text{cycle}} = 0$; 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- k OP* 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 $\mathcal{O}(|V|^{k-1})$ for *purely acyclic* \vec{w} :

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

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

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 arcs. More precisely:

Definition 3 *Let $\mathcal{G} = (V, E)$ be any complete (finite, undirected) graph and $w: E \rightarrow \mathfrak{R}$ 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 $a \rightarrow d_{ab}$ and $d_{ab} \rightarrow b$ 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 arcs 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

$$a \longrightarrow d_{ab} \longrightarrow b \longrightarrow d_{ba} \longrightarrow a \quad (4)$$

of arcs 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 of line (4) 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 arcs 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.

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

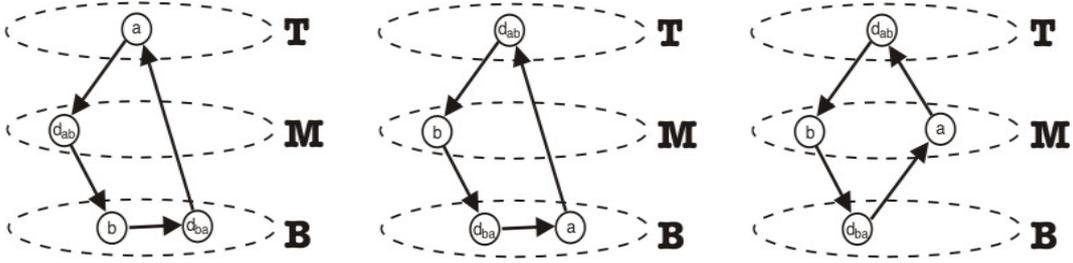


Figure 2: Some possible ways to fit an $\{a, b\}$ 4-cycle into three levels.

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 $\vec{\mathcal{P}}$ of the vertex set \vec{V} of $\vec{\mathcal{H}}_{\mathcal{G}}$ with $\vec{v}(\vec{\mathcal{P}}) \geq k$?”

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 $\{J > K > L\}$, which becomes the ordered partition of $\vec{\mathcal{H}}_{\mathcal{G}}$'s ordinary vertices. 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 $\vec{\mathcal{P}}$ achieves the exact same score: $\vec{v}(\vec{\mathcal{P}}) = v(\mathcal{P}) = j \geq k$.

In the other direction, consider an ordered tripartition $\vec{\mathcal{P}} = \{V_1 \cup D_1 > V_2 \cup D_2 > V_3 \cup D_3\}$ of \vec{V} with $V_1 \cup V_2 \cup V_3 = V$, $D_1 \cup D_2 \cup D_3 = D$, and $\vec{v}(\vec{\mathcal{P}}) \geq k$. 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 $\vec{v}(\vec{\mathcal{P}})$. Thus $v(\mathcal{P}) \geq \vec{v}(\vec{\mathcal{P}}) \geq k$, 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}^* and w^* as follows: add a new vertex \clubsuit along with edges $\{\clubsuit, v\}$ 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}^* of \mathcal{G}^* will place \clubsuit alone in one of the three pieces, while the other two pieces constitute a maximal-score bipartition \mathcal{P} of \mathcal{G} , with $v(\mathcal{P}^*) = |V|\sigma + v(\mathcal{P})$. 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}^* of \mathcal{G}^* with score at least $|V|\sigma + k$. ■

The proofs of Theorems 2 and 3 exploit the decomposition

$$\vec{w} = \vec{w}_{cycle} + \vec{w}_{cocycle} \quad (5)$$

of Section 2 (see also [7] and [16]) and use the following abstract definition of Borda score¹² for a vertex x of a weighted tournament:

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

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

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

$$\vec{w}(x, y) + \vec{w}(y, z) = \vec{w}(x, z) \quad (7)$$

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

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

$$\vec{w}(x, y) = \Gamma(x) - \Gamma(y) \quad (8)$$

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

$$\gamma_1 \leq \gamma_2 \leq \dots \leq \gamma_m \quad (9)$$

enumerating Γ 's values in non-decreasing order.

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

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

Proof: We leave the easy (1) \Leftrightarrow (2) equivalence to the reader. If \vec{w} is purely acyclic, then as an immediate consequence of Observation 11.2 of [7], \vec{w} is difference generated via the function assigning scaled Borda scores:

$$\Gamma: x \mapsto \frac{x^\beta}{|V|}. \quad (10)$$

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

$$\vec{w} \cdot \sigma = [\Gamma(x_1) - \Gamma(x_2)] + [\Gamma(x_2) - \Gamma(x_3)] + \dots + [\Gamma(x_{r-1}) - \Gamma(x_r)] + [\Gamma(x_r) - \Gamma(x_1)] = 0 \quad (11)$$

It follows from linearity of the dot product that $\vec{w} \cdot \tau = 0$ holds for any linear combination of basic cycles – hence $\vec{w} \perp \mathbf{V}_{\text{cycle}}$, and $\vec{w} \in \mathbf{V}_{\text{cocycle}}$. Thus \vec{w} is purely acyclic. (The argument is like that for Proposition 15 in [7].) ■

¹²In section 4 we obtain a tournament $\vec{\mathcal{H}}_\Pi = (A, E)$ and edge-weighting \vec{w}_Π 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 scoring weights $m-1, m-3, \dots, 3-m, 1-m$.

Definition 7 An ordered k -partition $\mathcal{P} = \{P_k \succ' P_{k-1} \succ' \dots \succ' P_1\}$ of a nondecreasing sequence $\gamma_1 \leq \gamma_2 \leq \dots \leq \gamma_m$ of real numbers is monotone if $i < j \Rightarrow \pi(\gamma_i) \leq' \pi(\gamma_j)$, where \leq' 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 :

$$\underline{\gamma_1, \gamma_2, \dots, \gamma_{m_1}} \downarrow_1 \underline{\gamma_{m_1+1}, \dots, \gamma_{m_2}} \downarrow_2 \dots \downarrow_{k-1} \underline{\gamma_{m_{k-1}+1}, \dots, \gamma_m} \quad (12)$$

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

Proof: (of Lemma 3) It is straightforward to show that if some ordered partition $\vec{\mathcal{P}}$ satisfied $i < j$ with $\pi(\gamma_i) \succ' \pi(\gamma_j)$ 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 $\vec{\mathcal{P}}$'s score. A sequence of such swaps converts $\vec{\mathcal{P}}$ into a monotone partition. ■

Proof: (of Theorem 2) Given a purely acyclic instance $\vec{\mathcal{H}} = (V, E)$, \vec{w} of max- k OP with $m = |V|$ vertices, calculate the Borda scores from line (6). Let $W = \max(\{|\vec{w}(x, y)|\}_{x, y \in V})$ and $\Gamma = mW + 1$. Then adding Γ to each of these scores yields strictly positive integers – the *shifted* Borda scores $0 < \gamma_1 \leq \gamma_2 \leq \dots \leq \gamma_m < 2\Gamma$, which we identify with the m vertices. An exhaustive search would then compute $\vec{v}(\vec{\mathcal{P}})$ for each possible monotone ordered k -partition of the γ_j , 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 in $\mathcal{O}(m^{k-1} \log(mW))$ time, but can be improved to $\mathcal{O}(km^2 \log(mW))$ time using standard dynamic programming techniques. ■

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

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. 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, $\vec{v}_{\sigma}(\vec{\mathcal{P}}) = 0$. By linearity, the same holds for any linear combination of basic cycles, and so we conclude that for ordered 2-partitions $\vec{\mathcal{P}}$, $\vec{v}_{\vec{w}_{\text{cycle}}}(\vec{\mathcal{P}}) = 0$. ■

4 The (j, k) -Kemeny rule

Suppose that $V = \{v_1, \dots, v_m\}$ 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 $\vec{\mathcal{H}}_{\Pi} = (V, E)$ and edge-weights \vec{w}_{Π} are as follows:

- $E = \{(v_i, v_j) \mid i < j\}$. Remark: This adds one arc for each two vertices.
- $\vec{w}_{\Pi}(v_i, v_j) = |\{t \in N \mid v_i \geq_t v_j\}| - |\{t \in N \mid v_j \geq_t v_i\}|$. Remark: These weights are the net majorities by which voters favor v_i over v_j .¹³

¹³Loosely, the edge-weighted tournament provides the ‘‘C2’’ information, in Fishburn’s classification, [8]. Note: $w(v_i, v_j)$ is negative when $i < j$ and more voters strictly prefer v_j to v_i than strictly prefer v_i to v_j .

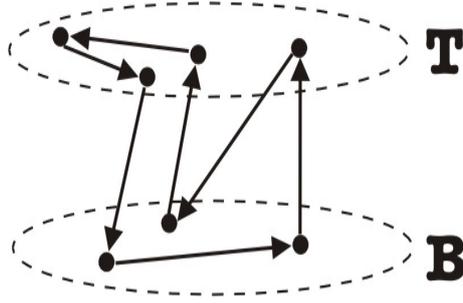


Figure 3: Fitting a cycle into two levels.

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)¹⁴ on V corresponding to the solution(s) of $\max\text{-}k\text{OP}$ for \vec{H}_Π, \vec{w}_Π . 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^* in either position refers to **univalent** dichotomous weak orders – equivalently, ordered 2-partitions $\{T > B\}$ for which $T = \{x\}$ is a singleton.

Proposition 2 Note that a dichotomous weak order $\{T > B\}$ can be interpreted as an approval ballot approving all alternatives in T . A univalent output $\{\{x\} > B\}$ can be interpreted as naming x as winner; as input $\{\{x\} > B\}$ can be interpreted as a plurality ballot for x . With that understanding, special cases of (j, k) -Kemeny include:

1. $(2, 2)$ -Kemeny is the Mean Rule.¹⁵
2. $(2, |V|)$ -Kemeny and $(2, 2^*)$ -Kemeny are approval voting (with outcome the ranking(s) by approval score, and the approval winner(s), respectively).
3. $(2^*, |V|)$ -Kemeny and $(2^*, 2^*)$ -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.¹⁶
5. $(|V|, 2^*)$ -Kemeny is the Borda count voting rule.
6. $(|V|, |V|)$ -Kemeny is the Kemeny voting rule (with rankings as output).

Proof: Some of the Proposition 2 restrictions yield the exact same aggregation rule when applied to the median procedure ([1], [12]) as they do when applied to (j, k) -Kemeny; this happens, for example, whenever $j = |V|$ or $k = |V|$. In particular, the $(2, |V|)$ -median procedure yields, as outcome, the ranking by approval score, and the $(|V|, 2^*)$ -median procedure yields the Borda winner(s). But the $(2, 2)$ -median procedure does not agree with the mean rule (in fact, it seems unlikely that any reasonable version of the mean rule arises

¹⁴Ties 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.

¹⁵The Mean Rule outcome ranks all alternatives with above average approval score over all those with below average score; see [7] for details.

¹⁶Borda Mean Rule acts like the Mean Rule, but with Borda scores replacing approval scores; see [7].

as a median procedure restriction) and the $(2, 2^*)$ -median procedure provably differs from approval voting.

Although the proofs for Proposition 2 are straightforward, explaining the exact relationship to the median procedure would require going beyond the scope of the current paper. For that reason, we are postponing most of these proofs to a planned sequel [18], and will limit ourselves here to a sample, by showing that $(2, 2^*)$ -Kemeny yields, as outcome for a profile of dichotomous weak orders, the alternative(s) having highest approval score for the corresponding profile of approval ballots.

Consider a profile Π^* consisting of a single ballot $\{T > B\}$, corresponding to a single approval ballot of T , with induced tournament $\overrightarrow{\mathcal{H}}_{\Pi^*} = (V, E)$ and edge-weights $\overrightarrow{w}_{\Pi^*}$. It is easy to see that for any two alternatives x and y , the weight $\overrightarrow{w}_{\Pi^*}(x, y)$ on the $x \rightarrow y$ edge is the difference $App_{\Pi^*}(x) - App_{\Pi^*}(y)$ in their approval scores (which will be $+1$, -1 , or 0). Now consider a more general profile Π with a number of ballots. The weight $\overrightarrow{w}_{\Pi}(x, y)$ on the $x \rightarrow y$ edge of $\overrightarrow{\mathcal{H}}_{\Pi}$ is likewise the difference $App_{\Pi}(x) - App_{\Pi}(y)$ in approval scores, because it is a sum of the contributions $App_{\Pi^*}(x) - App_{\Pi^*}(y)$ made by the individual ballots. The score of a univalent ordered partition $\{\{x\} > V \setminus \{x\}\}$ will then be the sum

$$\sum_{y \in V \setminus \{x\}} [App(x) - App(y)],$$

which is maximized when x has a greatest approval score. ■

The results of Theorems 2 and 3 and of Corollary 2 now lift¹⁷ immediately to (j, k) -Kemeny, showing:

Theorem 4 *The problem of determining the winning ordering for a $(\underline{\quad}_1, \underline{\quad}_2)$ -Kemeny election is in P whenever at least one of the blanks contains 2 (or 2^*), 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 $(\underline{\quad}_1, |V|)$ -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 $(\underline{\quad}_1, \underline{\quad}_2)$ -Kemeny, however. To argue for NP -hardness when $\underline{\quad}_1$ 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}_{Π} , 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 $\overrightarrow{w}_{\Pi} = 2\overrightarrow{w}$.¹⁸ To make the argument when $\underline{\quad}_2$ 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- k OP,” and “Max-cut is polynomially reducible to max- k cut,” but these are straightforward generalizations, and we omit the details.

Theorem 5 *The problem of determining the winning ordering for a $(\underline{\quad}_1, \underline{\quad}_2)$ -Kemeny election is NP -hard whenever*

- $\underline{\quad}_1$ is filled with either a fixed $j \geq 3$ or with $|V|$, and

¹⁷Given that the edge weights for (j, k) -Kemeny arise from net majorities, we can replace W (in the Theorem 2 proof) with n , the number of ballots or voters, so that the final conclusion (of $\mathcal{O}(km^2 \log(mW))$ time) becomes $\mathcal{O}(km^2 \log(mn))$ time.

¹⁸ For $V = \{v_1, v_2, v_3, \dots, v_m\}$ consider the following profile Π of two trichotomous weak orders: $\{v_1\} > \{v_2\} > \{v_3, \dots, v_m\}$; $\{v_3, \dots, v_m\} > \{v_1\} > \{v_2\}$. Then $\overrightarrow{w}_{\Pi}(v_1, v_2) = 2$ and \overrightarrow{w}_{Π} assigns weight 0 to every other arc. Thus by combining profiles similar to Π we can build an arbitrary function \overrightarrow{w} taking even integer values.

- $\underline{_2}$ is filled with a fixed $k \geq 3$

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

None of our reasoning here shows *NP*-hardness when $\underline{_2}$ 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.”¹⁹ 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 $\vec{w}_{\Pi_{cycle}} = 0$. As for $(j, |V|)$ -Kemeny, footnote 18 shows that for $j \geq 3$ the induced weights \vec{w}_{Π} 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 with a second interesting question raised by one of our COMSOC referees: what happens to tractability when the cyclic component is simple – what happens, for example, if \vec{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] Brandl, F., Brandt, F., and Seedig, H.G. Consistent Probabilistic Social Choice. To appear in *Econometrica*, 2016.
- [4] 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.
- [5] Croom, F. Basic Concepts of Algebraic Topology. *Springer-Verlag*, New York, 1978.
- [6] Duddy, C., and Piggins, A. Collective approval. *Mathematical Social Sciences* 65, 190-194, 2013.
- [7] Duddy, C., Piggins, A., and Zwicker, W.S. Aggregation of binary evaluations: a Borda-like approach. *Social Choice and Welfare* 46, 301-322, 2016.
- [8] Fishburn, P.C. Condorcet social choice functions. *SIAM Journal on Applied Mathematics* 33, 469-489, 1977.

¹⁹“Max- $|V|$ cut” is in quotes because it is silly. At first, it seemed annoying that our methods did not seem to apply to the original Kemeny rule itself. However, with Theorem 6 (in the appendix), the fundamental nature of this obstacle became clear; the methods we use here establish hardness for the transitive subcase, so they cannot possibly show hardness of Kemeny winner, which is *not* hard for that case.

- [9] Hammer, P.L., Ibaraki, T., and Peled, U.N. Threshold numbers and threshold completions. *Annals of Discrete Math* 11, 125-145, 1981.
- [10] Harary, F. Graph theory and electrical networks. *IRE Trans. Circuit Theory CT-6*, 95-109, 1958.
- [11] Hemaspaandra, E., Spakowski, H., and Vogel, J. The complexity of Kemeny elections. *Theoretical Computer Science* 349, 382-391, 2005.
- [12] 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.
- [13] Myerson, R.B. Axiomatic derivation of scoring rules without the ordering assumption. *Social Choice and Welfare* 12, 59-74, 1995.
- [14] Saari, D.G. *Geometry of Voting*. Springer, 1994.
- [15] 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.
- [16] Zwicker, W.S. The voters' paradox, spin, and the Borda count. *Mathematical Social Sciences* 22, 187-227, 1991.
- [17] 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.
- [18] Zwicker, W.S. Aggregate! *Working paper*, 2016.

5 Appendix: the transitive sub-case of max-3OP

Our goal is to prove:

Theorem 6 *Max-cut is polynomially reducible to the transitive sub-case of max-3OP.*

Transitivity here refers to the ordinary “qualitative” version, as expressed for a weighted tournament²⁰ $\vec{\mathcal{H}} = (V, E), \vec{w}$, by:

$$\text{If } \vec{w}(x, y) > 0 \text{ and } \vec{w}(y, z) > 0 \text{ then } \vec{w}(x, z) > 0, \quad (13)$$

for all $x, y \in V$ with $x \neq y$, and under the *reversal convention*.

Proof: As in the proof of Theorem 1, we specify a polynomial translation that converts a weighted graph $\mathcal{G} = (V, E), w$ containing $|V|$ vertices (which – without loss of generality – is complete and for which $w(e)$ is a nonnegative integer for each edge $e = \{a, b\} \in E$) into a weighted tournament $\vec{\mathcal{F}}_{\mathcal{G}}, \vec{w}$. This time our goals for $\vec{\mathcal{F}}_{\mathcal{G}}, \vec{w}$ are a bit different. Let $C = 1 + \sum_{e \in E} w(e)$ and $\varepsilon = \frac{1}{72|V|^4}$. Then:

²⁰We mean that for a weighted tournament generated from a profile of weak or strict rankings, this condition is equivalent to ordinary transitivity of the strict majority preference relation as defined in footnote 5. However, the construction in the Theorem 6 proof never assigns edge-weights of 0, and in this setting all forms of transitivity of majority preference become equivalent.

1. \vec{w} is transitive in the sense of equation (13), and
2. For each positive integer k the answer to the decision problem “Does there exist a vertex bipartition $\mathcal{P} = \{J, K\}$ of V with $v(\mathcal{P}) \geq k$?” is the same as the answer to “Does there exist an ordered tripartition $\vec{\mathcal{P}}$ of the vertex set \vec{V} of $\vec{\mathcal{F}}_{\mathcal{G}}$ with

$$\vec{v}(\vec{\mathcal{P}}) \geq \langle\langle 3|V|C + k \rangle\rangle? \tag{14}$$

Here $\langle\langle x \rangle\rangle$ rounds x (up or down) to the nearest integer. We leave it to the reader to confirm that for each integer $j \geq 4$ a similar construction substitutes “ordered j -partition” for “ordered tripartition” and “ $(2j - 3)|V|C$ ” for “ $3C$.”

Definition 9 Given $\mathcal{G} = (V, E)$ and $w: E \rightarrow \mathfrak{R}$ as specified above, the tournament $\vec{\mathcal{F}}_{\mathcal{G}} = (\vec{V}, \vec{E})$ and edge-weight function \vec{w} induced by \mathcal{G} and w are defined as follows:

- Choose any reference linear order \triangleright of \mathcal{G} 's vertex set V .
- Each vertex $a \in V$ contributes a quadruple of ordinary vertices a_1, a_2, a_3 , and a_4 to \vec{V} , along with three directed “placement” edges (see figure 4):

$$a_1 \longrightarrow a_2 \longrightarrow a_3 \longrightarrow a_4 \tag{15}$$

weighted as follows: $\vec{w}(a_2, a_3) = 2C$ and $\vec{w}(a_1, a_2) = C = \vec{w}(a_3, a_4)$.

- Each edge $e = \{a, b\}$ of \mathcal{G} having weight w and satisfying $a \triangleright b$ contributes two additional direction vertices d_{ab} and d_{ba} to \vec{V} , along with four directed “adjustment” edges:

$$a_2 \longrightarrow d_{ab} \longrightarrow b_2; \quad b_3 \longrightarrow d_{ba} \longrightarrow a_3 \tag{16}$$

with each such edge assigned weight w (if $w > 0$) or weight ε (if $w = 0$).

- To make $\vec{\mathcal{F}}_{\mathcal{G}}$ into a tournament, we need to add edges between pairs of vertices that are not yet linked in either direction. The direction of each can be chosen carefully (as detailed immediately below) in such a way that no cycles are created. Assign weight ε to each of these “tiny” edges.

Suppose we temporarily erase all weights on the edges of $\vec{\mathcal{F}}_{\mathcal{G}}$, and also omit all the tiny edges, retaining the placement and adjustment edges. The resulting unweighted digraph $\mathcal{E}_{\mathcal{G}}$ is acyclic (meaning there are no cycles that respect edge direction), thanks to the use of the reference linear order \triangleright ; staring at Figure 4 can help explain what’s going on, here. Any such acyclic digraph can be extended to an acyclic tournament by adding new arcs one-at-a-time.²¹ Once these tiny edges are added and assigned positive weight ε , the resulting weighted tournament is transitive, as desired. Moreover, ε is small enough to guarantee that the total contribution of all weight ε edges to the score of any ordered tripartition is less than $\frac{1}{2}$ in absolute value. Thus we are safe when, for the remainder of this proof, we simultaneously ignore the presence of weight ε edges and use of rounding in condition 14.

Any score-maximizing tripartition $\vec{\mathcal{P}} = \{T > M > B\}$ of $\vec{\mathcal{F}}_{\mathcal{G}}$ will also maximize that part of the score contributed by the placement edges (which have weight $2C$ or weight C), because the value of C is large enough to make any nonzero contribution by a single weight- C edge overwhelm the total contribution of all adjustment edges (which have weights from

²¹It is easy to see that if adding edge $x \longrightarrow y$ would introduce a cycle, and adding $y \longrightarrow x$ would also do so, then there must have been a cycle in the original digraph. The argument is essentially the same as that in [9], though the context there – alternating cycles in an undirected pregraph – is somewhat different.

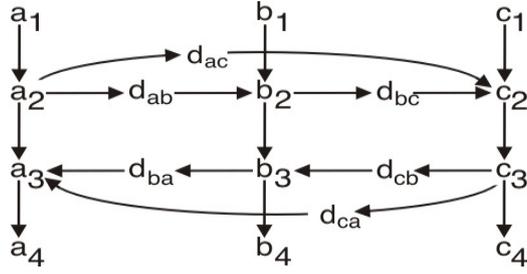


Figure 4: Part of the digraph \mathcal{F}_G showing the (directed) placement and adjustment edges arising from three vertices $a \triangleright b \triangleright c$ (and corresponding undirected edges) of \mathcal{G} . Note the absence of cycles.

the original graph \mathcal{G}). To maximize the placement edge part of the score it is necessary, for each quadruple a_1, a_2, a_3, a_4 of ordinary vertices, either to place $a_1, a_2 \in T$, $a_3 \in M$, and $a_4 \in B$ – in which case we will say that $\vec{\mathcal{P}}$ **places a up** – or $a_1 \in T$, $a_2 \in M$, and $a_3, a_4 \in B$ – in which case $\vec{\mathcal{P}}$ **places a down**. Either such arrangement achieves a contribution of $3C$ from a_1, a_2, a_3, a_4 , and it is easy to see that one can not do better than that. Hence, when seeking to maximize the score of $\vec{\mathcal{P}}$ we may assume that each original vertex of \mathcal{G} is either placed up or down, for a total contribution of $3|V|C$ from placement edges.

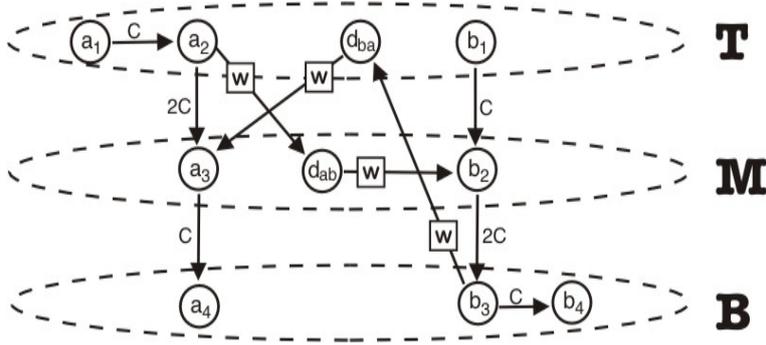


Figure 5: Two quadruples (with a up, b down) and direction vertices (as positioned to achieve a net contribution of w from the adjustment edges).

Assume the weight of some edge $\{a, b\}$ of \mathcal{G} is w , and $a \triangleright b$. Figure 5 shows the corresponding quadruples, placed in an ordered 3-partition $\vec{\mathcal{P}}$ so that a is up and b down. The direction vertices d_{ab} and d_{ba} have been positioned so that the net contribution made to $\vec{v}(\vec{\mathcal{P}})$ by the four corresponding adjustment edges is $w + 0 - w + w = w$. It is straightforward to check that one cannot achieve a contribution greater than w by moving d_{ab} and d_{ba} into different pieces of $\vec{\mathcal{P}}$. The situation is the same when a is down and b is up; one can position d_{ab} and d_{ba} to achieve a net contribution of w from the adjustment edges, and no greater contribution is possible. Finally, when a and b are either both up, or both down, the net contribution made by the four corresponding adjustment edges is 0, regardless of where one positions d_{ab} and d_{ba} .

Thus, given a partition $\mathcal{P} = \{U, D\}$ of \mathcal{G} 's vertices with a score $v(\mathcal{P})$ of at least k one can construct an ordered 3-partition $\vec{\mathcal{P}}$ of $\vec{\mathcal{F}}_G$'s vertices with score at least $3|V|C + k$ by placing all vertices in U up, all in D down, and positioning the direction vertices to maximize

the contribution of adjustment edges, as detailed in the previous paragraph.²² Conversely, given an ordered 3-partition $\vec{\mathcal{P}}$ of $\vec{\mathcal{F}}_{\mathcal{G}}$'s vertices with score at least $3|V|C + k$, we know that each quadruple must be placed in the *up* or *down* position. The considerations of the previous paragraph then imply that by setting $U = \{a \in V \mid a \text{'s quadruple is up}\}$ and $V = \{b \in V \mid b \text{'s quadruple is down}\}$ we obtain a partition $\{U, V\}$ of \mathcal{G} 's vertices with a score of at least k . ■

²²Recall that we are supressing all mention of tiny edges and of ε contributions to $\vec{v}(\vec{\mathcal{P}})$.