Multiwinner Rules on Paths From \( k \)-Borda to Chamberlin–Courant\(^*\)

Piotr Faliszewski\(^1\) and Piotr Skowron\(^2\) and Arkadii Slinko\(^3\) and Nimrod Talmon\(^4\)

\(^1\) AGH University, Krakow, Poland
\(^2\) Institut für Softwaretechnik und Theoretische Informatik, TU Berlin, Berlin, Germany
\(^3\) University of Auckland, Auckland, New Zealand
\(^4\) Weizmann Institute of Science, Rehovot, Israel

faliszew@agh.edu.pl, p.k.skowron@gmail.com, a.slinko@auckland.ac.nz, nimrodtalmon77@gmail.com

Abstract

The classical multiwinner rules are designed for particular purposes. For example, variants of \( k \)-Borda are used to find \( k \) best competitors in judging contests while the Chamberlin–Courant rule is used to select a diverse set of \( k \) products. These rules represent two extremes of the multiwinner world. At times, however, one might need to find an appropriate trade-off between these two extremes. We explore continuous transitions from \( k \)-Borda to Chamberlin–Courant and study intermediate rules.

1 Introduction

The multiwinner voting rule \( k \)-Borda [Debord, 1992] and the rule of Chamberlin and Courant [1983] are based on very different ideas and they are used for very different purposes. For example, \( k \)-Borda is a maximum likelihood estimator [Procaccia et al., 2012], which means that it works well when there is an objectively best committee and all the votes are noisy perceptions of the truth. Intuitively speaking, this means that we assume that votes “closer to the truth” are more likely than those “farther from the truth” and so we seek the area of the “highest density of the votes” to derive the winning committee from the votes there (ignoring outliers and voter groups of smaller density). One consequence of this approach is that the members in the winning committee under \( k \)-Borda tend to be very similar to each other (for a visual justification of these intuitive claims, we point the reader to the work of Elkind et al. [2017a] and to the histograms in our experimental section). Thus, \( k \)-Borda is very suitable for applications such as finding finalists of various competitions (based on the rankings provided by the judges), or—generally speaking—for shortlisting tasks where the candidates are evaluated on a single criterion (Barberà and Coelho [2008] and Elkind et al. [2017b] offer discussions regarding shortlisting tasks).

\(^*\)Piotr Faliszewski was supported by the National Science Centre, Poland, under project 2016/21/B/ST6/01509, Piotr Skowron was supported by a Humboldt Research Fellowship for Postdoctoral Researchers, Arkadii Slinko was supported by the Royal Society of NZ Marsden Fund UOA-254.

\(^1\)See, e.g., the work of Gordon and Truchon [2008], where the problem of aggregating the judges’ rankings in figure skating is treated from the maximum likelihood perspective.

The Chamberlin-Courant rule (the CC rule), on the other hand, was designed to determine as diverse set of alternatives as possible (e.g., a set of movies that an airline may offer on its long-range flights [Elkind et al., 2017b]) so that every shade of preference existing in the society is catered for (Elkind et al. [2017a] justify this intuition visually; see also our experimental section). Consequently, the CC rule is very well suited for applications where we select a set of items from which each voter chooses one that he or she desires most (see the works of Lu and Boutilier [2011], Skowron et al. [2016a], and Faliszewski et al. [2016a] for further relevant scenarios in addition to choosing movies). However it is possible that some members of the Chamberlin-Courant elected committee will be liked by a minority of voters and will not score high on the excellence criterion.

To summarize, \( k \)-Borda and Chamberlin–Courant are very different rules and pursue different objectives: excellence versus diversity. These are two opposing objectives and the two rules are as far apart in the ‘space of voting rules’ as possible.

Nonetheless, there are many applications where neither excellence nor diversity are the most appropriate goals to aim for. For example, consider shortlisting candidates for an academic position in computer science. On the one hand, we may wish to shortlist simply the strongest candidates, but, on the other hand, we would also like to secure certain diversity among the selected candidates with respect to areas of specialization, gender balance, etc. Thus, while excellence is still a crucial criterion, we would also like to put significant weight on diversity (this is also related to proportional representation with respect to multiple external attributes [Lang and Skowron, 2016]).

Similarly, in the movie-selection example, while the desire to choose a diverse set of movies is our main driving force, usually we also want to make sure that more popular types of movies appear in greater numbers than the less popular ones (indeed, on a typical long-range flight we see more Hollywood blockbuster movies than, say, documentaries).

This means that many real-life, business-driven applications of multiwinner voting require rules that strike some sort of a compromise between excellence and diversity, sometimes putting more focus on the former and sometimes giving priority to the latter (see also the discussion of Ratliff and Saari [2014]). However, until recently, in the literature we have had a discrete set of voting rules and nothing in be-
between them. In this paper we attempt to rectify this situation by introducing families of rules that, in a certain sense, lie on a “path between $k$-Borda and the Chamberlin–Courant rule”. This is made possible due to the recently introduced class of committee scoring rules [Faliszewski et al., 2016a] and by the fact that both $k$-Borda and the Chamberlin–Courant rule belong to this class. The committee scoring rules work as follows. To each possible committee, each voter assigns a score based on the positions of the committee members in his or her ranking. These scores are summed up and, as a result, each committee obtains a societal score. The committee with the highest societal score wins.

Both $k$-Borda and the Chamberlin-Courant rule use classical Borda scores to evaluate individual committee members. If the total number of alternatives is $m$ and a candidate is ranked in the $i$th position in a voter’s ranking, then this voter’s Borda score of that candidate is $m - i$. For calculating the score of a committee, according to a given voter, under $k$-Borda we sum up all the scores of the committee members while Chamberlin-Courant takes the maximal score. For example, if the Borda scores of the committee members (from an individual voter) are $b_1 > b_2 > \ldots > b_k$, then the $k$-Borda score of this committee (from this voter) would be $b_1 + b_2 + \ldots + b_k$, while the Chamberlin-Courant score (from the same voter) would be $b_1$.

Skowron et al. [2016a] suggested to calculate the committee score as $\lambda_1 b_1 + \ldots + \lambda_k b_k$, where $\Lambda = (\lambda_1, \ldots, \lambda_k)$ is a vector of numbers between 0 and 1, referred to as the ordered weighted average operator (OWA); the rules that use OWA operators in calculating the scores of committees are called OWA-based. In particular, $k$-Borda and Chamberlin–Courant rules are OWA-based relative to the vectors $(1, \ldots, 1)$ and $(1, 0, \ldots, 0)$, respectively. We consider two families of OWA-based rules that connect these two extremes. The first is defined by vectors with $t$ ones followed by $k - t$ zeros (where $t$ ranges between 1 and $k$), and the second one uses vectors of the form $(1, 1/2^p, 1/3^p, \ldots, 1/\kappa^p)$, for $q \in (0, \infty)$.

Another way to connect $k$-Borda and the Chamberlin-Courant rule is to observe that summing up the Borda scores of the committee members—as in $k$-Borda—simply means taking the $\ell_1$-norm of the vector of those Borda scores, whereas taking the maximal score—as in CC—means taking the $\ell_\infty$-norm. Thus, we consider the family of rules between $k$-Borda and Chamberlin–Courant that use the $\ell_p$-norms with $p \in [1, \infty)$ (we refer to these rules as $\ell_p$-Borda rules).

We test our rules on elections from the 2D Euclidean domain, following Elkind et al. [2017a]. The rules behave in a rather unexpected way. For example, as we move from $\ell_1$-Borda to $\ell_\infty$-Borda, we first observe a rapid phase transition from rules that are similar to $k$-Borda to rules that resemble the Bloc rule (a rule that, at first sight, does not seem to have anything to do with $\ell_p$-Borda rules) and then there is a smooth transition to rules that are similar to CC. Rules using our first family of OWA vectors behave similarly (but “more smoothly”, which is quite surprising given that theseOWAs have a much more discrete nature), whereas the rules based on the second family of OWAs follow a very different path.

As a consequence of our studies, we also find more decisive variants of the $k$-Borda and Chamberlin–Courant rules.

## 2 Preliminaries

In this paper we focus on committee scoring rules defined by Elkind et al. [2017b]. In our model, an election $E = (C, V)$ consists of a set $C = \{c_1, \ldots, c_m\}$ of candidates and a collection $V = \{v_1, \ldots, v_n\}$ of voters, where each voter $v_i$ has a preference order $\succ_i$, thereby ranking the candidates from the most desirable to the least desirable one. For a candidate $c$ and a voter $v$, we write $\text{pos}_v(c)$ to denote the position of $c$ in $v$’s preference order (e.g., $\text{pos}_v(c) = 1$ if $v$ ranks $c$ in the top position and $\text{pos}_v(c) = m$ if $v$ ranks $c$ last). We refer to subsets of $C$ as committees and if $S$ is a committee and $v$ is a voter, then by $\text{pos}_v(S)$ we mean the sequence obtained by sorting the set $\{\text{pos}_v(c) \mid c \in S\}$ in the increasing order.

For positive integers $m$ and $k$, we write $[m]$ to denote the set $\{1, \ldots, m\}$ and $[m]_k$ to denote the set of all length-$k$ increasing sequences of elements from $[m]$. In other words, for a preference order over $m$ candidates, $[m]$ is the set of all possible committees of size $k$ can occupy. For two committee positions $I = (i_1, \ldots, i_k)$ and $J = (j_1, \ldots, j_k)$ from $[m]_k$, we say that $I$ dominates $J$, denoted $I \succeq J$, if for each $t \in [k]$, it holds that $i_t \leq j_t$.

A multiwinner rule $R$ is a function that, given an election $E = (C, V)$ and a positive integer $k$, outputs a family $R(E, k)$ of size-$k$ committees that tie as winners of $E$ (we typically expect a single winning committee, but there can be more due to symmetries in the preference orders).

A committee scoring function, for elections with $m$ candidates and committees of size $k$, is a function $f_{m,k} : [m]_k \to \mathbb{R}$ that assigns a numerical score to each committee position, so that if $I, J \in [m]_k$ are two positions and $I \succeq J$, then $f(I) \geq f(J)$. For an election $E = (C, V)$ with $m$ candidates and a size-$k$ committee $S$, we define the $f_{m,k}$-score of $S$ as:

$$\text{score}_{f_{m,k}}(S, E) = \sum_{v \in V} f_{m,k}(\text{pos}_v(S)).$$

**Definition 1.** Let $f = \{f_{m,k}\}_{1 \leq k \leq m}$ be a family of committee scoring functions. A committee scoring rule $R_f$ is a multiwinner rule that, given an election $E = (C, V)$ and committee size $k$, outputs those committees $S$ (of size $k$) for which $\text{score}_{f_{|C|,k}}(S, E)$ is the highest.

For $m$ candidates, the (single-winner) Borda score awarded to a candidate for occupying position $i$ in a ranking of some voter is $\beta_m(i) = m - i$. The $k$-Borda and Chamberlin–Courant committee scoring rules are defined through the following scoring functions:

1. For the $k$-Borda rule, the scoring functions are:

$$f^{k\text{-Borda}}_{m,k}(i_1, \ldots, i_k) = \beta_m(i_1) + \ldots + \beta_m(i_k).$$

To find a $k$-Borda winning committee it suffices to compute the Borda scores separately for each candidate and pick those $k$ candidates with the highest scores. That is, $k$-Borda does not consider any dependencies between the candidates and simply picks the $k$ best ones according to the Borda social welfare function which ranks all candidates by their Borda scores.

2. For the CC rule, the scoring functions are:

$$f^{\text{CC}}_{m,k}(i_1, \ldots, i_k) = \beta_m(i_1).$$
Intuitively, under the CC rule each voter treats his or her highest-ranked committee member as his or her representative and the score of a committee is the sum of the Borda scores that the voters give to their representatives. In our discussion, we will also refer to the Bloc rule:

3. Under Bloc, each voter names his or her \( k \) favorite candidates and the winning committee consists of those mentioned most frequently. In other words, Bloc uses the scoring functions

\[
f_{\text{Bloc}}^{m,k}(i_1, \ldots, i_k) = \sum_{i=1}^{k} \alpha_k(i),
\]

where \( \alpha_k(i) = 1 \) for \( i \leq k \) and \( \alpha_k(i) = 0 \) otherwise.

**Example 1.** Consider an election with the following votes:

\[
\begin{align*}
    v_1 &: a \succ b \succ c \succ d \succ e \succ f \succ g \succ h, \\
    v_2 &: f \succ h \succ e \succ g \succ d \succ c \succ b \succ a, \\
    v_3 &: a \succ b \succ g \succ h \succ e \succ f \succ c \succ d, \\
    v_4 &: a \succ c \succ d \succ g \succ h \succ b \succ f \succ e.
\end{align*}
\]

Under 4-Borda, the unique winning committee of size 4 is \( \{a, b, c, g\} \), under CC all committees that include \( a \) and \( f \) win.

While winning committees for \( k \)-Borda and Bloc can be computed in polynomial time, the CC rule is computationally hard [Procaccia et al., 2008; Lu and Boutilier, 2011; Betzler et al., 2013]. Nonetheless, CC becomes polynomial-time solvable under certain domain restrictions [Betzler et al., 2013; Skowron et al., 2015b; Yu et al., 2013; Peters and Elkind, 2016] and there are approximation algorithms and heuristics for it [Skowron et al., 2015a; Faliszewski et al., 2016c].

Naturally, many other committee scoring rules exist. We point the reader to the works of Faliszewski et al. [2016a; 2016b] for an overview of the internal structure of the class of committee scoring rules, and to the work of Skowron et al. [2016b] for their axiomatic characterization.

### 3 Three Paths Between \( k \)-Borda and CC

Below we describe the three families of rules that we view as paths between \( k \)-Borda and CC. We start with two families based on OWA operators, then we introduce the \( \ell_p \)-Borda family of rules, and finally we discuss convergence of the rules in our families to CC and to \( k \)-Borda, as well as some of their properties.

**The OWA-Based Paths.** A family \( \Lambda = (\Lambda^k)_{k \in \mathbb{N}} \) of OWA operators, \( \Lambda^k = (\lambda_1^k, \ldots, \lambda_k^k) \), with one operator for each dimension, defines a committee scoring rule with the following scoring functions:

\[
f_{m,k}^\Lambda(i_1, \ldots, i_k) = \lambda_1^k \beta_m(i_1) + \ldots + \lambda_k^k \beta_m(i_k).
\]

Multiwinner rules of this type were introduced by Skowron et al. [2016a] in a somewhat more general context (i.e., not focusing only on the Borda scores), and were later studied by Faliszewski et al. [2016a].

We consider the following OWA operators \( \Lambda^k \):

1. For each positive integer \( t \), the \( t \)-Best family of OWA operators has vectors of \( t \) ones followed by \( t-k \) zeros.
2. For each non-negative real number \( p \), the \( p \)-Harmonic OWA operators are of the form \( (1, 1/2^p, 1/3^p, \ldots, 1/k^p) \).

For each \( t \), we set \( t \)-Borda to be the rule defined using Equation (1) and the \( t \)-Best OWA operators (note that this notation is consistent with the name \( k \)-Borda), and for each non-negative real number \( p \), we let \( p \)-HarmonicBorda be defined analogously, but using the \( p \)-Harmonic OWAs. (1-Harmonic OWAs are also used in the definition of the PAV rule [Thiele, 1895; Kilgour and Marschall, 2012], that was shown to have very good axiomatic properties [Aziz et al., 2017].)

**Example 2.** Consider the election from Example 1. The 2-Borda rule, for \( k = 4 \), chooses committees \( \{a, b, f, h\} \) and \( \{a, c, f, h\} \) (both of which are winning under CC), 2-HarmonicBorda chooses only \( \{a, b, f, h\} \), while \( 1/2 \)-HarmonicBorda chooses \( \{a, b, c, h\} \), which differs from the 4-Borda winning committee only by replacing \( g \) with \( h \).

Intuitively, as \( p \) gets larger, \( p \)-HarmonicBorda becomes more similar to CC and \( 1/p \)-HarmonicBorda becomes more similar to \( k \)-Borda. This intuition is almost correct and later we will make it formal.

The rules on our paths (with the exception of \( k \)-Borda) are NP-hard to compute (this follows immediately from the results of Skowron et al. [2016a]).

**Proposition 1.** For each positive integer \( t > 1 \), it is NP-hard to decide if a winning committee under \( t \)-Borda has at least a given score. The same holds for each positive real number \( p \) and \( p \)-HarmonicBorda.

While the above result speaks of values of \( t \) that are constant (and, in particular, do not depend on \( k \)), Skowron et al. also showed that, e.g., \( (k-1) \)-Borda is NP-hard to compute.

**The \( \ell_p \)-Path.** Let \( p \) be a real number, \( p \geq 1 \). By \( \ell_p \)-Borda we mean the committee scoring rule defined by the committee scoring functions:

\[
f_{m,k}^{\ell_p}(i_1, \ldots, i_k) = \sqrt[p]{\beta_m(i_1)^p + \ldots + \beta_m(i_k)^p}.
\]

That is, under \( \ell_p \)-Borda, the score associated with a committee position \( (i_1, \ldots, i_k) \) is the \( \ell_p \)-norm of the vector \( (\beta(i_1), \ldots, \beta(i_k)) \). Naturally, \( \ell_1 \)-Borda is \( k \)-Borda and \( \ell_\infty \)-Borda is CC. We are interested, however, in \( \ell_p \)-Borda rules where \( p \) is not \( \infty \).

**Example 3.** Consider the election from Example 1. \( \ell_2 \)-Borda chooses committee \( \{a, b, c, f\} \), a CC winning committee that is close to the \( k \)-Borda winning committee (the fact that the committee is winning under CC is mostly a coincidence). Note that the rule in this example and all the rules in Examples 1 and 2 give different sets of committees as outcomes.

All the \( \ell_p \)-Borda rules for \( p > 1 \) are NP-hard to compute.

**Theorem 2.** For each rational \( p > 1 \), deciding whether there is a winning committee with at least a given score is NP-hard.

In principle, the idea behind the proof of this theorem is close to that behind the proof of the NP-hardness of CC [Lu and Boutilier, 2011]. However, it requires significant technical work to introduce sufficient (but polynomially-bounded)
paddings that, roughly speaking, ensures that $\ell_p$-Borda rules behave similarly to CC.

**Convergence of the Rules on the Paths.** While the t-Borda family of rules contains k-Borda and CC by definition, the situation is more intricate for the $\ell_p$-Borda and $p$-HarmonicBorda families. In particular, we are interested in convergence of $p$-HarmonicBorda and $\ell_p$-Borda to both k-Borda and CC. (Note that 0-HarmonicBorda and $f_1$-Borda are simply nicknames for k-Borda, but the rules that could be dubbed $\lim_{p \to 0} p$-HarmonicBorda and $\lim_{p \to 1+} \ell_p$-Borda offer some advantages regarding decisiveness; due to limited space, we mostly focus on convergence to CC, however).

We next show how well the CC winning committees approximate the $\ell_p$-Borda and $p$-HarmonicBorda winning ones.

**Definition 2.** Let $R_f$ be a committee scoring rule defined through a family of committee scoring functions $f = (f_{m,k})_{1 \leq k \leq m}$. Let $E = (C,V)$ be an election and $W$ be a winning committee of $E$ under $R_f$. Let $\alpha$ be a real number such that $0 < \alpha < 1$. We say that a committee $S$ is an $\alpha$-approximate winner for the $R_f$-election $E$, if $score_{f(\cdot),k}(S,E) \geq \alpha \cdot score_{f(\cdot),k}(W,E)$.\(^3\)

**Proposition 3.** Let $E$ be an election with $m$ candidates and let $k$ be the target committee size.
1. For each $p \geq 1$, each CC winning committee for $E$ is a $\frac{1}{\sqrt{k}}$-approximate $\ell_p$-Borda winner for $E$.
2. For each $p \geq 0$, each $p$-HarmonicBorda winning committee for $E$ is a $1/(\sum_{i=1}^{n} 1/j^{i})$-approximate $p$-HarmonicBorda winner for $E$.

**Proof.** We consider the case of $\ell_p$-Borda (the case of $p$-HarmonicBorda is analogous and omitted due to restricted space).

Let $W_p$ be an $\ell_p$-Borda winning committee for $E$ and let $W_\infty$ be a CC winning committee. For brevity, we write $f$ to refer to the scoring function $f_{m,k}$. For a given committee $S$ and each voter $v$ we have:

$$f(pos_v(S)) \leq \sqrt{k} \cdot f_{m,k}(pos_v(S)) \leq \sqrt{k} \cdot f(pos_v(S)).$$

Thus it holds that:

$$score_f(W_p,E) = \sum_{v \in V} f(pos_v(W_p)) \leq \sqrt{k} \cdot \sum_{v \in V} f_{m,k}(pos_v(W_p)) \leq \sqrt{k} \cdot \sum_{v \in V} f_{\ell_p}(pos_v(W_\infty))$$

which proves the statement. \(\Box\)

The estimates in Proposition 3 are asymptotically tight.

**Example 4.** Let $m$, $n$, and $k$ be three positive integers such that $m \geq 2k$. Consider an election $E$ with candidates $c_1, \ldots, c_m$ and with $n + (k-1)$ voters, where $k$ is the target committee size. For each $i \in [k-1]$, the $i$th voter ranks $c_{m-i+1}$ first; and each of the following $n$ voters has preference order $c_1 > c_2 > \ldots > c_m$. For large enough $n$, we have that: (a) $c_1, \ldots, c_k$ is the unique winning committee under each $\ell_p$-Borda and each $p$-HarmonicBorda rule; (b) $c_1, c_m, c_{m-1}, \ldots, c_{m-k+2}$ is the unique CC winning committee; (c) this committee is an $O(1/\sqrt{k})$-approximate $\ell_p$-Borda winner for $E$ and an $O(1/(\sum_{i=1}^{k} \frac{1}{j^i}))$-approximate $p$-HarmonicBorda winner for $E$.

For a given committee size, Proposition 3 indicates which values of $p$ may lead to rules very similar to the CC rule.

**Example 5.** Let us consider some election $E$ and the committee size $k = 20$. CC winning committees are 0.54-approximations of $\ell_5$-Borda winning committees (so, intuitively, we should not expect $\ell_5$-Borda and CC results to be similar), while they are 0.94-approximations of $\ell_{50}$-Borda winning committees (so we expect similarity). Similarly, CC winning committees are 0.62-approximations of 2-HarmonicBorda winning committees, but 0.96-approximation of 5-HarmonicBorda ones.

Above, we studied convergence of the scores of the winning committees. Now we move to the convergence of actual winning committees (these issues differ as one could imagine two disjoint committees with nearly identical scores).

**Definition 3.** Let $(R_i)_{i \in \mathbb{N}}$ be a family of multiwinner voting rules. We say that this sequence converges election-wise, if for each election $E$ and each target committee size $k$, there exists a positive integer $i_0$ such that $R_{i_0}(E,k) = R_i(E,k)$ for each $i \geq i_0$.

A sequence $(R_i)_{i \in \mathbb{N}}$ of rules that converges election-wise defines a new multiwinner rule: Given an election $E$ and a target committee size $k$, it outputs the committee $R_i(E,k)$ for sufficiently large $i$ (when the sequence of outputs becomes stable).

A multiwinner rule $R$ is a refinement of a rule $Q$ if for each election $E$ and each committee size $k$, $R(E,k) \subseteq Q(E,k)$.

**Theorem 4.** The sequence of rules $(p$-HarmonicBorda$)_{p\in \mathbb{N}}$ converges, as $p \to 0$, to a refinement of CC. The sequence $(1/p$-HarmonicBorda$)_{p\in \mathbb{N}}$ converges to a refinement of k-Borda.

**Proof.** Let us consider the sequence $(p$-HarmonicBorda$)_{p\in \mathbb{N}}$. We may assume that $p$ changes continuously, i.e., that $p$ is a real number. We fix an election $E$ and the size of committee $k$. For a committee $Z$, we write $score_p(Z)$ to denote its $p$-HarmonicBorda score in $E$. It suffices to show that for each two distinct committees $X$ and $Y$, starting from some $p_0$ and for all $p > p_0$, we have either $score_p(X) > score_p(Y)$ or $score_p(X) = score_p(Y)$ or $score_p(X) < score_p(Y)$. Suppose $pos_v(X) = \{e_{i_1}^{(1)}, \ldots, e_{i_k}^{(1)}\}$. Then:

$$score_p(X) = \sum_{i=1}^{n} \sum_{j=1}^{k} \frac{1}{p^{\beta_m(e_{i_j}^{(1)})}} = \sum_{j=1}^{k} \frac{1}{p^{\beta_m(e_{i_j}^{(1)})}} = \sum_{j=1}^{k} \frac{1}{p^{\beta_m(e_{i_j}^{(1)})}}$$

where $a_j = \sum_{i=1}^{n} \beta_m(e_{i_j}^{(1)})$. Similarly, $score_p(Y) = \sum_{j=1}^{k} \frac{1}{p^{b_j}}$ for some $b_1, \ldots, b_k$. If $a_j = b_j$ for all $j \in [k]$, then the scores of $X$ and $Y$ are always the same, irrespective of the value of $p$. Suppose this does not happen.
As \( p \) grows, the relation between \( \text{score}_p(X) \) and \( \text{score}_p(Y) \) may change several times, but we claim that this number of changes is smaller than \( k \). We will show that the difference

\[
\text{score}_p(X) - \text{score}_p(Y) = \sum_{j=1}^{k} \frac{1}{p_j} (a_j - b_j)
\]

cannot be equal to zero for \( k \) distinct values of \( p \). Assume the contrary, i.e., that it is equal to zero for \( p_1, \ldots, p_k \). Then:

\[
\begin{bmatrix}
1 & (1/2)p_1 & \cdots & (1/k)p_1 \\
1 & (1/2)p_2 & \cdots & (1/k)p_2 \\
\vdots & \vdots & \ddots & \vdots \\
1 & (1/2)p_k & \cdots & (1/k)p_k \\
\end{bmatrix}
\begin{bmatrix}
a_1 - b_1 \\
a_2 - b_2 \\
\vdots \\
a_k - b_k \\
\end{bmatrix}
= \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.
\]

Since \( a_j - b_j \neq 0 \) for at least one \( j \in [k] \), the determinant of the matrix on the left-hand side must be zero. However, this is a generalized Vandermonde matrix which is known to have a non-zero determinant [Wen and Zhang, 2006]. This completes the proof of our claim that the relations between the scores of committees are fixed for sufficiently large \( p \).

As \( p \) approaches \( \infty \), for any committee \( X \) we have \( \text{score}_p(X) \rightarrow \text{score}_\infty(X) \); thus, the winning committees of the limiting voting rule have to be CC winning committees and it is a refinement of CC. The argument for \( k\)-Borda \(( p \rightarrow 0 \) is similar.

The situation with \( \ell_p\)-Borda rules is much more difficult and we only obtain a weaker result.

**Proposition 5.** For each election \( E \) and each committee size \( k \), there is a value \( p_0 \) such that for each \( p > p_0 \), \( \ell_p\)-Borda\((E,k) \) is a set of CC winning committees.

**Decisiveness of the Rules on the Paths.** Consider the next example.

**Example 6.** In Examples 2 and 3 we have already seen that \( p\)-HarmonicBorda and \( \ell_p\)-Borda rules can provide results that refine the CC rule; the same happens for the \( k\)-Borda rule. Consider an election with \( m \geq 2k \) candidates and two votes: \( c_1 > c_2 > \cdots > c_m \) and its reverse. For the target committee size \( 2k \), the \( 2k\)-Borda rule outputs all committees of size \( 2k \). On the other hand, each \( \ell_p\)-Borda rule (with \( p > 1 \)) and each \( p\)-HarmonicBorda rule (with \( p > 0 \)) output a unique winning committee \( \{c_1, \ldots, c_k, c_m, \ldots, c_{m-k+1}\} \).

Indeed, our new rules can quite substantially improve the decisiveness of \( k\)-Borda and CC. We have generated 30,000 elections with 10 candidates and 10 voters each (using the Impartial Culture assumption, i.e., we chose each voter's preference order uniformly at random) and computed all winning committees of size \( k = 4 \). Table 1 shows the results that we have obtained (we used brute-force search to compute all winning committees). Indeed, it seems that (the limiting variants of) the \( \ell_p\)-Borda and \( p\)-HarmonicBorda rules can be used as natural, more decisive, variants of \( k\)-Borda and CC.

<table>
<thead>
<tr>
<th>rule</th>
<th>avg. num. of win. committees</th>
<th>max. num. of win. committees</th>
</tr>
</thead>
<tbody>
<tr>
<td>( k)-Borda</td>
<td>1.206</td>
<td>10</td>
</tr>
<tr>
<td>( \ell_{10})-HarmonicBorda</td>
<td>1.00003</td>
<td>2</td>
</tr>
<tr>
<td>( \ell_{0.01})-Borda</td>
<td>1.00</td>
<td>1</td>
</tr>
<tr>
<td>CC</td>
<td>2.119</td>
<td>28</td>
</tr>
<tr>
<td>5-HarmonicBorda</td>
<td>1.00016</td>
<td>2</td>
</tr>
<tr>
<td>( \ell_{0.01})-Borda</td>
<td>1.00007</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 1: Average number of winning committees and the highest number of winning committees observed; under Impartial Culture assumption with \( n = m = 10 \) and \( k = 4 \).

In the 2D Euclidean model of preferences, voters and candidates are represented as points on a plane (often referred to as the ideal points of the respective voters/candidates). Each voter forms a preference order by sorting the candidates with respect to the distance of their ideal points from his or hers.

In our experiments—following Elkind et al. [2017a]—we consider elections with \( m = 200 \) candidates, \( n = 200 \) voters, and with committee size \( k = 20 \). These numbers are chosen since they are big enough to be quite representative and robust, but small enough to make experimentation feasible. All our elections are generated using the 2D Euclidean model, where both the candidates’ and voters’ ideal points are distributed uniformly on a \([-3, 3] \times [-3, 3]\) square. Elkind et al. [2017a] have considered other distributions and other election sizes, showing that the graphical results are robust to such variations (unless we consider, e.g., very small committees).

**Computing Winning Committees.** Since our rules are NP-hard, we used the following heuristic based on simulated annealing. We begin by sampling a random committee \( S_0 \). Then, in each iteration, we take the committee \( S_{i-1} \) and form a temporary committee \( S_i \) by randomly replacing one member in \( S_{i-1} \). If the score of \( S_i \) is greater than that of \( S_{i-1} \), then we set \( S_{i-1} = S_i \). Otherwise, we draw a random number between 0 and 1; if it is below \( pq^i \) (where \( p \) and \( q \) are two parameters, we used \( p = 0.2 \) and \( q = 0.999 \), then we set \( S_{i-1} = S_i \); otherwise, we set \( S_i = S_{i-1} \). We execute 2000 iterations and output the highest-scoring committee encountered.

We compared the results obtained from using our heuristic to the optimal ones (for some of our OWA-based rules) and they turned out to be very close (both in terms of the approximation ratio and similarity of the histograms; see below).

**The Experiment.** We have generated 5,000 elections for each of our rules and computed their results using our heuristic (we considered \( t\)-Borda rules for \( t \in \{1, 2, \ldots, 20\} \), \( p\)-HarmonicBorda for \( p \in \{0.1, 0.2, \ldots, 5.0\} \), and \( \ell_p\)-Borda rules for \( p \in \{1, 2, \ldots, 50\} \)). For each of the rules, we have computed a histogram showing how frequently candidates in various areas of the \([-3, 3] \times [-3, 3]\) square are selected. Specifically, we partitioned the square into 120 \( \times \) 120 small squares (called cells) and computed how many winners from the generated elections fall into each. We present (some of) the resulting histograms in Figures 1–3 (the darker a given cell, the more winners fall there; Elkind et al. give more details on converting numerical values to colors). For comparison, Figure 4 contains histograms for \( k\)-Borda, Bloc, and CC.
Finally, for each of the generated histograms, we computed its earth-mover distance from the histograms for \( k \)-Borda, Bloc, and CC. The plots showing how these distances change depending on our rules’ parameters are shown in Figures 1–3.

**Earth-Mover Distance.** The earth-mover distance (EMD) is a widely-used measure of similarity between pictures [Pele leg et al., 1989]. Figuratively speaking, it views each cell in a histogram as containing some grains of sand (in proportion to the value of the cell), and the distance between two histograms is the minimum sum of Euclidean distances that all the grains of sand need to travel to transform one histogram into the other. To compute earth-mover distances, we used its standard formulation as an integer linear program (ILP) and used an ILP solver. The smaller the earth-mover distance between two histograms, the more similar these histograms are. The EMDs computed for our histograms are all normalized in the same way, so the results are comparable between rules. To assess which values of the distance are “small” and which are “large,” one may compare presented example histograms.

**Analysis.** Perhaps somewhat surprisingly, the graphical results that we have obtained for the \( t \)-Borda and \( \ell_p \)-Borda rules progress from \( k \)-Borda to CC through histograms very similar to those for Bloc (however, there is a big difference between Bloc and these rules; for each single election, Bloc typically chooses candidates concentrated in some parts of the square—leading to the histogram from Figure 4 only in the aggregate)—whereas winners under \( t \)-Borda and \( \ell_p \)-Borda rules are distributed very similarly to what their histograms suggest; the same good property holds for the \( p \)-HarmonicBorda rules). Further, for \( \ell_p \)-Borda we observe a rather abrupt transition from histograms very similar to those for \( k \)-Borda (for \( p \in \{1, \ldots, 4\} \)) to those similar to Bloc (for \( p \in \{5, \ldots, 17\} \)). The further transition, however, is smooth.

On the other hand, the histograms for \( p \)-HarmonicBorda rules do not pass though areas of similarity to Bloc (even though the histograms are more similar to those for CC than those for Bloc only for \( p > 2 \)). Indeed, while the histograms for \( t \)-Borda and \( \ell_p \)-Borda have areas of “lower density in the center,” this never happens for \( p \)-HarmonicBorda.

**5 Conclusions**

We analyzed three families of committee scoring rules that form continuous paths between the \( k \)-Borda and CC rules. We have shown that the rules on these paths can lead to more decisive variants of \( k \)-Borda and CC. Further, these rules may be useful for tasks that need some compromise between excellence and diversity, provided by \( k \)-Borda and CC, respectively. Quite surprisingly, we have found that two of our paths seem to pass through rules with unexpected features.

For future research, we believe it would be interesting to consider paths between other multiwinner rules (e.g., between approval-based rules).