Voting by Sequential Elimination with few Voters

Sylvain Bouveret
LIG - Grenoble INP, France
sylvain.bouveret@imag.fr

Yann Chevaleyre
LIPN - Univ. Paris-Nord, France
yann.chevaleyre@lipn.univ-paris13.fr

François Durand
U. Paris-Dauphine, CNRS, PSL, France
fradurand@gmail.com

Jérôme Lang
CNRS, U. Paris-Dauphine, PSL, France
lang@lamsade.dauphine.fr

Abstract

We define a new class of low-communication voting rules, tailored for contexts with few voters and possibly many candidates. These rules are defined by a predefined sequence of voters: at each stage, the designated voter eliminates a candidate, and the last remaining candidate wins. We study deterministic and randomized versions of these rules. We first investigate their axiomatic properties. Then we focus on a subfamily of rules defined by "non-interleaved" sequences, for which we seek a sequence that best approximates the Borda rule under Impartial Culture. Finally, we apply our rules to randomly generated data.

1 Introduction

The communication burden on the agents in collective decision making is often an obstacle to implementing otherwise good voting rules, perhaps even more than computational complexity: while individuals agree to wait a few hours before knowing the outcome of a collective decision mechanism, they will probably not wish to spend a significant amount of time and energy reporting their preferences, especially in low-stake decision contexts, like choosing a collective item to buy, a place for a group dinner, or a film to watch together. In such contexts, the number of alternatives may be high, while the number of agents can be small: it is reasonable, for instance, to consider a group of four individuals choosing a restaurant out of a dozen. The type of problem we have in mind is the design of mechanisms that could be implemented on smartphone applications, and used by agents quickly and with a very low cognitive effort.

A number of works, which we review below, addressed the issue of designing low-communication protocols for computing the exact or approximate outcome of common voting rules. We follow a different path, by *designing new rules that are specifically tailored for coming with low communication protocols*. The family of rules we study in this paper concerns contexts where the number of candidates is larger (possibly much larger) than the number of voters, and works by *having, at each step, a designated agent eliminating one of the remaining alternatives, until only one alternative remains*. We believe that individuals would easily use such rules in low-stake contexts, where simplicity and rapidity of choice matters more than fairness and optimality; as a matter of fact, such rules have been observed in real-life settings by some of us. Although this paper focuses on this precise context, our rules belong to a much larger set, where the number of voters and alternatives can be arbitrary small or large.[1]

Related work can be decomposed into three groups (for a good overview of communication issues in voting see Boutilier and Rosenschein [2015]):

Class 1: Defining low-communication protocols for *computing* voting rules. The communication complexity of voting rules was initiated by Conitzer and Sandholm [2005]: the goal is to identify, for a given rule, lower and upper bounds on the worst-case number of bits to be exchanged so that the outcome can be determined; the upper bound comes with concrete protocols. A different set of results, based on *query complexity*, identifies the worst-case minimal number of queries to determine the winner of some tournament solution, where a query consists in asking whether some edge exists in the majority graph [Balasubramanian *et al.*, 1997; Procaccia, 2008; Dey, 2017]. Finally, Kalech *et al.* [2011] study an interactive protocol where at each round, each voter gives her k-th best alternative; the protocol ends when there is enough information to determine the true winner.

Class 2: Defining low-communication protocols for *approximating* a voting rule. Service and Adams [2012] study the communication complexity of determining approximate winners for rules based on scores (where approximation is relative to the ratio between the score of the outcome and the score of the true winner). Queries in their protocols generally consist in asking voters whether the rank of a candidate lies in a given interval or not. Another series of work that can be located in Class 2 are top-k approximations of voting rules: voters report only their top k alternatives, and a winner is computed from this partial information [Dery *et al.*, 2014; Filmus and Oren, 2014; Lu and Boutilier, 2011].

Class 3: Designing *new rules* based on low-communication

---

[1] For instance, in a context with less alternatives than voters, the designated agent can veto an alternative, and an aternative is eliminated once it obtains a given number of vetos; or else, the designated agent can support an alternative, and the winner is declared once an alternative obtains enough support. Although we have these extensions in mind, we won't discuss them again in the paper.

protocols.[2] Goel and Lee [2012] define a randomized urn-based voting rule where at each stage, a triple of voters (i_1, i_2, i_3) is selected randomly; each of them gives her preferred alternative x_{i_j}, and votes between the other two. If there is a three-way tie, the three voters disappear from the election, otherwise the losing voters are replaced with "copies" of the winning voter. The process is iterated until there remains only one voter (which is guaranteed to eventually occur for a reason we will not explain here). The rule has a low average communication complexity and, when applied to a Euclidean single-peaked profile, achieves a good approximation of the Condorcet winner with high probability. With the *random pairs* rule [Hansen, 2016], each voter is presented a small number of randomly chosen pairs of alternatives and chooses between them; the alternative winning most pairwise contests is declared the winner. Although the rule can violate Pareto-efficiency and even unanimity, such paradoxes are rare in practice; and the required communication is low. Gross *et al.* (2017) define a low-communication randomized rule: voters are selected randomly and asked to report their preferred alternative; as soon as there is an alternative named by two voters, it is declared the winner. The worst-case communication complexity $O(m \log m)$ is the same as for our rules, and the principle definitely has some similarity with ours. However, their mechanism is based only on top-ranked alternatives, while ours uses information "inside" the voters' rankings; related to that, their mechanism is not applicable to settings with few voters, because of the risk that no alternative is top-ranked more than once.

The paper is organized as follows. In Section 2 we define two families of sequential elimination rules. In the first one, the rules are resolute but non-anonymous, as the elimination sequence is fixed. In the second one, the rules are randomized and anonymous: an elimination pattern is fixed, with virtual voters being placeholders, and actual voters are assigned to placeholders with a uniform probability. In Section 3 we conduct an axiomatic study of our rules: they are shown to satisfy Pareto-efficiency, monotonicity, and a weakening of independence of losing alternatives. While the set of satisfied properties may seem light, we emphasize that their communication complexity is extremely low. In Section 4 we compare different elimination sequences according to the expected Borda score of the outcome, where profiles follow the impartial culture distribution. While we do not have a closed-form formula for identifying the best sequence in the general case, we are able to do it for a family of sequences that are especially interesting and convenient: *non-interleaved* sequences, in which the occurrences of each agent in the sequence are contiguous. In Section 5 we study the performance of elimination-based rules on randomly generated data. In Section 6 we suggest further research directions.

2 Elimination-based Voting Rules

Let $N = \{1, \ldots, n\}$ be a set of voters and $C = \{c_1, \ldots, c_m\}$ a set of candidates. A n-voter preference profile over C is a

---

[2]In some sense, class 2 can be interpreted as a subset of class 3, if we see the approximations of voting rules as voting rules *per se*.

collection of votes $V = (V_1, \ldots, V_n)$ where each V_i is a linear order over C. $\mathcal{L}(C)$ is the set of all linears orders over C. $\Delta(C)$ is the set of all probability distributions over C. For $c \in C$, $r(c, V_i)$ denotes the rank of c in V_i (from 1 for the best candidate to m for the worst one). The Borda score of c for V is $S_B(c, V) = \sum_{i=1}^{n} m - r(c, V_i)$. Also, we define the contribution of voter V_i to the Borda score of c as $S_B(c, V_i) = m - r(c, V_i)$. A resolute voting rule is a function from $\mathcal{L}(C)^n$ to C. A randomized voting rule is a function from $\mathcal{L}(C)^n$ to $\Delta(C)$. The Borda rule is the resolute rule that outputs the candidate maximizing $S_B(., V)$ (using a predefined priority relation over C in case of a tie).

An *elimination sequence* for n voters and m candidates is a sequence $\pi = (\pi(1), \ldots, \pi(m-1))$ in N^{m-1}. Given π, the deterministic sequential elimination rule (SER) F_π is defined recursively as follows: for each profile V,

1. $C(0) = C$;
2. For each $i = 1, \ldots, m-1$, $C(i) = C(i-1) \setminus \{\text{last}(C(i-1), V_{\pi(i)})\}$, where $\text{last}(C(i-1), V_{\pi(i)})$ is the least preferred candidate in $C(i-1)$ according to $V_{\pi(i)}$.
3. Output the single element in $C(m)$.

Example 1 *Let $n = 3$, $m = 5$, $V = (abcde, edcba, debca)$ and $\pi = (1, 2, 3, 1)$. Then $C(1) = \{a, b, c, d\}$ since 1 eliminates e, $C(2) = \{b, c, d\}$ since 2 eliminates a, $C(3) = \{b, d\}$ since 3 eliminates c, and finally, $C(4) = \{b\}$ and $F_\pi(V) = b$.*

Given an elimination sequence π and a permutation σ on N, the randomized SER $\hat{F}_\pi$ induced by the pattern π outputs each candidate c with probability $\frac{1}{n!} |\{\sigma \mid F_{\sigma \circ \pi}(V) = c\}|$. Intuitively, the indices in π are not actual voters but placeholders; the randomly chosen permutation σ then assigns actual voters to placeholders.

Example 1 (continued) *Let us apply the randomized pattern $\pi = (1, 2, 3, 1)$ to P. Here are the six possible permutations together with the associated winner: $123 \mapsto 123 : b$; $123 \mapsto 132 : c$; $123 \mapsto 213 : d$; $123 \mapsto 231 : d$; $123 \mapsto 312 : d$; $123 \mapsto 321 : d$. Henceforth, $\hat{F}_\pi(P) = \langle b : \frac{1}{6}, c : \frac{1}{6}, d : \frac{2}{3} \rangle$.*

It is worth noticing that if $\pi = (1, \ldots, 1)$ then $\hat{F}_\pi$ coincides with random dictatorship.

There are two ways of viewing $\hat{F}_\pi$, yielding two ways of computing the winner. In the *offline* view, voters report their preferences without knowing which permutation σ will be chosen; in the *online* view, the permutation σ is randomly chosen first, and voters then proceed to the elimination protocol as in the resolute case. Not only the offline computation of $\hat{F}_\pi$ requires a complete elicitation of the preferences, but it is also hard, as the following result shows.

Proposition 1 *Suppose $m > n$, and let $\pi = (1, 2, \ldots, n, n, n \ldots)$. Checking if a candidate w has a non-zero probability of winning with $\hat{F}_\pi$ is NP-complete.*

Proof sketch. Hardness is proven by the following reduction from 3SAT. Let $x_1, \ldots, x_p$ be the propositional variables, and let $c_1, \ldots, c_q$ the clauses, with $c_i = l_{i,1} \vee l_{i,2} \vee l_{i,3}$, where each $l_{i,j} \in \{x_1 \ldots x_p, \bar{x}_1 \ldots \bar{x}_p\}$ is a literal. The reduction is as follows. We build $m = 3p + 3q + 3$ candidates:

$\{x_i, \bar{x}_i, z_i \mid 1 \leq i \leq p\} \cup \{y_i, \bar{y}_i, c_i \mid 1 \leq i \leq q\} \cup \{a, b, w\}$. We build a profile with $n = 2p + 3q + 1$ voters. Because we are only concerned about the winning probability of w, in each vote V_i, the order of the candidates ranked above w does not matter, so we specify only the order of candidates below w. For each vote V_i, let $\text{below}(V_i, w)$ be the projection of V_i on candidates ranked below w in V_i. For example, $\text{below}(ebdwfa, w) = fa$. The profile V is built as follows:

- For each variable x_j, we have two voters V_{x_j} and $V_{\bar{x}_j}$ such that $\text{below}(V_{x_j}, w) = x_j z_j$ and $\text{below}(V_{\bar{x}_j}, w) = \bar{x}_j z_j$.
- For each clause $c_i = l_{i,1} \vee l_{i,2} \vee l_{i,3}$, we have three voters $V_{i,1}, V_{i,2}, V_{i,3}$ such that $\text{below}(V_{i,j}, w) = c_i l_{i,j} y'_i y_i$ for $j \in \{1, 2, 3\}$.
- We have a last voter V_{last} such that $\text{below}(V_{\text{last}}, w) = x_1 \bar{x}_1 \ldots x_p \bar{x}_p ba$.

The picking sequence is $\pi = (1, 2, \ldots n, n, \ldots n)$. w has a non-zero winning probability if and only if there exists a permutation σ applied to π such that during the $m - 1$ first picking steps, all candidates except w are removed. We can verify that the set of clauses is satisfiable if and only if w has a non-zero winning probability. □

This hardness result has little practical impact: the winner can still be computed efficiently by using the online protocol.

We say that i appears in π if there is a k such that $\pi(k) = i$. Let $\text{nv}(\pi)$ be the number of voters who appear in π. An elimination sequence π is *non-interleaved* if and only if for each $i \in N$, either i does not appear in π or there are two indices $\text{start}(i)$ and $\text{end}(i)$ such that for all $0 \leq k \leq m - 1$, $\pi(k) = i$ if and only if $\text{start}(i) \leq k \leq \text{end}(i)$. Given a non-interleaved sequence π, let ρ be the function yielding the j-th voter participating in the sequence. The *signature* of an non-interleaved sequence is the vector $\tilde{\pi} = (\text{end}(\rho(i)) - \text{start}(\rho(i)) + 1) \mid i = 1, \ldots, \text{nv}(\pi))$. Elimination sequences where voters appear in increasing order are *monotonic*. For instance, with $m = 7$ and $n = 4$, $\pi = (1, 1, 1, 1, 2, 3)$ is monotonic and yields the signature $\tilde{\pi} = (4, 1, 1)$. Without loss of generality, a non-interleaved pattern sequence for a randomized SER can be assumed to be monotonic. Random dictatorship corresponds to the pattern sequence $(1, \ldots, 1)$ with signature $\tilde{\pi} = (m - 1)$.

Observation 1 *The deterministic communication complexity of any resolute elimination rule is in $O(m \log(m))$.*

Indeed, at each step i, voter $\pi(i)$ just has to indicate the candidate she wants to eliminate, which gives $(m - 1)\lceil\log(m)\rceil$ as an immediate upper bound. The exact upper bound is in fact less than that: $\sum_{j=2}^{m}\lceil\log(j)\rceil$, because the voter $m - j + 1$ has to indicate a candidate among the j remaining candidates. This bound is for the set of all sequences; for specific sequences the communication complexity may be lower, especially for non-interleaved sequences. For instance, for the non-interleaved sequence with the signature $(4, 2)$, the first voter needs to communicate $\lceil\log\binom{7}{4}\rceil = 6$ bits, and the second voter $\lceil\log\binom{3}{2}\rceil = 2$ bits: 8 bits in total, whereas $\sum_{j=2}^{m}\lceil\log(j)\rceil = 14$. The minimum is of course obtained for dictatorship.

We now prove that for some sequences, the lower bound

matches the upper bound. Consider the family of sequences $\pi_{\text{simple}} = (1, \ldots, m - 1)$ with $n = m - 1$:

Proposition 2 *The communication complexity of $F_{\pi_{\text{simple}}}$ is in $\Omega(m \log m)$.*

Proof. Let $\mathcal{S}(C)$ be the set of permutations of C. Consider the set of profiles $\mathcal{F} = \{V^\sigma \mid \sigma \in \mathcal{S}(C)\}$, where $V^\sigma = (V_1^\sigma, \ldots, V_n^\sigma)$, and V_i^σ is a vote that ranks $x_{\sigma(i)}$ in the last position and x_m in the second last position (the rest of the vote does not matter). We claim that $\mathcal{F}$ is a fooling set for F_π, that is, for each $V, V' \in \mathcal{F}$, (a) $F(V) = F(V')$, and (b) there is a mixture (V_I, V'_{-I}) of V and V' (coinciding with V on I and with V' on $N \setminus I$) such that $F_\pi(V_I, V'_{-I}) \neq F(V)$. We first verify that for every $\sigma \in \mathcal{S}(C)$ we have $F_\pi(V^\sigma) = x_m$: indeed, at each step $i = 1, \ldots, m - 1$, voter i eliminates her least alternative $x_{\sigma(i)}$, and the remaining one is x_m. Now, we claim that if V^σ and V^τ are two distinct profiles in $\mathcal{F}$, then there is a mixture V' of V^σ and V^τ such that $F_\pi(V') \neq x_m$. Indeed, if $\sigma \neq \tau$ then there exist i, j such that $i \neq j$ and $\sigma(i) = \tau(j)$. Let $V' = (V_1^\sigma, \ldots, V_{j-1}^\sigma, V_j^\tau, V_{j+1}^\sigma, \ldots, V_n^\sigma)$. Because V_i' and V_j' are equal and rank x_i last and x_m second last, the first one whose turn comes will eliminate x_i, and the second one will eliminate x_m; thus, $F_\pi(V') \neq x_m$. Therefore, $\mathcal{F}$ is a fooling set for F_π, and since its cardinality is $(m - 1)!$, the communication complexity of F_π is in $\Omega(\log(m - 1)!)$, i.e, in $\Omega(m \log m)$. □

For randomized SERs with the online view, the communication complexity is exactly the same, if we do not count the communication bits sent by the random number generator to communicate to each agent her role in the sequence — otherwise we have to add $O(n \log(n))$ bits.

3 Axiomatic Properties

We now show that sequential elimination rules enjoy some of the most desirable properties of voting rules.

Clearly, any SER (whether resolute or randomized) is *neutral*: it treats all candidates symmetrically. No resolute SER is *anonymous*: it does not treat voters symmetrically; however, any randomized SER is *ex-ante anonymous*: voters are symmetrical before their assignment to placeholders.

Now, consider Pareto-efficiency: a resolute rule is *Pareto-efficient* if it always elects a Pareto-optimal candidate, where x is Pareto-optimal if no $y \neq x$ is preferred to x by all voters. A randomized rule $\hat{F}$ is *ex-post Pareto-efficient* if for any V, the support of $\hat{F}(V)$ contains only Pareto-optimal candidates.

Observation 2 *A candidate x is Pareto-optimal in profile V if and only if there exists a sequence π such that $F_\pi(V) = x$.*

Proof. Left to right: the sequence π is constructed by choosing at each step a voter who does not rank x last among the remaining candidates. If it was not possible at some step other than the final step, then x would be ranked last in all votes, and thus would be Pareto-dominated. Right to left: if x is Pareto-dominated by y, then whatever π, y will never be eliminated before x, therefore $F_\pi(V) \neq x$. □

This observation leads to the following Proposition.

Proposition 3 *For any π, F_π is Pareto-efficient and $\hat{F}_\pi$ is ex-post Pareto efficient.*

All rules F_π and $\hat{F}_\pi$ satisfy *unanimity*: if there is a unanimously preferred candidate, then she is elected (with probability 1 for randomized rules). We now consider *monotonicity*. An x-improvement of a profile V is a profile V' identical to V except for one vote V'_i, for which x moves one position upward, everything else being unchanged. A resolute rule F is monotonic if for any V with $F(V) = x$, and any x-improvement V' of V, $F(V') = x$. A randomized rule $\hat{F}$ is monotonic if with the same assumptions as above, the probability that x wins cannot decrease when changing V into V'.

Proposition 4 *For any π, F_π and $\hat{F}_\pi$ are monotonic.*

Proof. For the resolute case, this is because if some candidate y is eliminated for V, then for V' it must be eliminated either at the same stage as in V or before. The application of this result to all permutations σ yields the result for $\hat{F}_\pi$. $\square$

We also mention an interesting property, which is a relaxation of Kelly's independence of losing alternatives [Kelly, 1978]. Let $\text{up}(x, V_i) = \{y \mid y \succ_{V_i} x\}$ and $\text{down}(x, V_i) = \{y \mid x \succ_{V_i} y\}$. Given two votes V_i and V'_i, we write $V_i \sim_x V'_i$ if $\text{up}(V'_i, x) = \text{up}(V'_i, x)$ and for each $y, z \in \text{down}(V_i, x)$ we have $y \succ_{V_i} z$ if and only if $y \succ_{V'} z$. Then, all rules F_π satisfy *independence of reshuffling above the winner*: for all profiles V, V' such that $V_i \sim_{F_\pi(V)} V'_i$ for every $i = 1, \ldots, n$, we have $F_\pi(V') = F_\pi(V)$. What makes this property interesting is that it is rather strong, as very few common rules satisfy it (one of the few exceptions being *Antiplurality – a.k.a Veto*).

For lack of space, we only briefly discuss properties that are not satisfied, or satisfied only in specific cases. We summarize all results without proof: (1) if $n \geq 3$ and $m \geq 2$, then no SER is Condorcet-consistent; (2) a resolute (resp. randomized) SER is strategyproof if and only if it is a dictatorship (resp. random dictatorship); (3) A family of randomized SERs, defined for any number of voters and candidates, satisfies SD-participation [Brandl *et al.*, 2015] if and only if it is a random dictatorship for any number of voters and candidates.

4 Borda-optimal Non-interleaved Elimination Sequences

Communication complexity is obviously not the only criterion we may use to choose an elimination sequence (as it would lead us to choose random dictatorship). A way of evaluating the quality of an elimination sequence can be its proximity to a known voting rule, and here we choose the Borda rule. Although this is not the only possible choice, still, it is an natural point to start with, since the Borda score has a clear social welfare interpretation that goes beyond seeing Borda as a voting rule (see *e.g.* d'Aspremont and Gevers [2002]).

For the sake of simplicity, in this section we restrict to non-interleaved elimination sequences. Such a restriction is often natural and desirable: for example, in an online voting setting, voters might want to connect only once to the voting website.

A sequence π will be evaluated by the expected Borda score of the winner $F_\pi(V)$, where profile V is drawn from

a uniform random distribution. We write $IC(n, m)$ the uniform distribution (often called Impartial Culture, or IC) over profiles of n voters and m candidates. First, we estimate the value of $\mathbb{E}_{V \sim IC(n,m)} [S_B (F_\pi(V), V)]$ for a given π. Then, we search for the sequence yielding the best Borda approximation, i.e. maximizing this expected Borda score.

Because we work with IC, all non-interleaved elimination sequences sharing the same signature $\tilde{\pi}$ yield the same expected Borda score. Therefore, without loss of generality, we restrict our study to monotonic elimination sequences. The following proposition gives us a closed formula to compute the expected Borda score, given a non-interleaved elimination sequence.

Proposition 5 *Assume the profile V is drawn from $IC(n, m)$. Let π be a monotonic non-interleaved elimination sequence. Let $l_i = m + 1 - \sum_{j=1}^{i} \tilde{\pi}(j)$ for $i = 1, \ldots n$ and $l_0 = m + 1$. Then, the expected contribution of voter V_i to the Borda score of the winner, as well as the total expected Borda score of the winner are:*

$$\mathbb{E}_{V \sim IC(n,m)} [S_B (F_\pi(V), V_i)] = m - \frac{l_i}{l_{i-1}} \left(\frac{m+1}{2} \right)$$

$$\mathbb{E}_{V \sim IC(n,m)} [S_B (F_\pi(V), V)] = mn - \left(\frac{m+1}{2} \right) \sum_{i=1}^{n} \frac{l_i}{l_{i-1}}$$

Proof. We focus on voter i. If c wins, the contribution of i to its total Borda score is $S_B(c, V_i)$. Let k_i be the number of candidates remaining before i's turn: $k_i = m - \sum_{j=1}^{i-1} \tilde{\pi}(j)$. Let $X_{(1)} \ldots X_{(k_i)}$ be the Borda scores with respect to i of each of these k_i remaining candidates, sorted by increasing scores (so $m - X_{(k_i)}$ is the rank of i's preferred candidate among all candidates still available). Because V_i is drawn from IC, the set $\{X_{(1)}, \ldots, X_{(k_i)}\}$ is uniformly distributed among all k_i-subsets over $\{0, \ldots, m-1\}$. Clearly, i will remove the $\tilde{\pi}(i)$ candidates yielding the lowest Borda scores $S_B(., V_i)$. These Borda scores are $X_{(1)} \ldots X_{(\tilde{\pi}(i))}$. Also, the final winner will be uniformly distributed among the remaining candidates (because of IC). So the expected Borda score of the final winner will be $\mathbb{E}_{j \sim \{\tilde{\pi}(k_i)+1, \ldots, k_i\}} [X_{(j)}]$.

It is a standard result from order statistics that when drawing a uniform subset of k elements from $\{1, \ldots, m\}$, and then ordering them increasingly, the expected value of the ith element is $i \frac{m+1}{k+1}$ (see e.g. Arnold *et al.* [1992, page 56]). Since our Borda scores are drawn from $\{0, \ldots, m-1\}$, we get $\mathbb{E}[X_{(i)}] = i \frac{m+1}{k_i+1} - 1$. For $j \leq k_i$, $1 + \mathbb{E}_{i \sim \{j, \ldots, k\}} [X_{(i)}] = \mathbb{E}\left[\frac{1}{k_i-j+1} \sum_{i=j}^{k} X_{(i)}\right] = \frac{(j+k_i)(k_i-j+1)(m+1)}{2(k_i-j+1)(k_i+1)}$. Finally, $1 + \mathbb{E}_{j \sim \{\tilde{\pi}(k_i)+1, \ldots, k_i\}} [X_{(j)}] = \frac{(\tilde{\pi}(k_i)+1+k_i)(k_i-\tilde{\pi}(k_i))(m+1)}{2(k-\tilde{\pi}(k_i))(k_i+1)} = \frac{(\tilde{\pi}(k_i)+1+k_i)(m+1)}{2(k_i+1)}$. Replacing $k_i + 1$ by l_{i-1}, we get $\mathbb{E}_{j \sim \{\tilde{\pi}(k_i)+1, \ldots, k_i\}} [X_{(j)}] = \frac{(2l_{i-1}-l_i)(m+1)}{2l_{i-1}} - 1 = m - \frac{l_i}{l_{i-1}} \left(\frac{m+1}{2} \right)$. The result follows. $\square$

Now, we look for an optimal sequence, i.e. maximizing the expected total Borda score. We need to minimize:

$$E = \sum_{i=1}^{n} \frac{l_i}{l_{i-1}},$$

with the constraints that l is a non-increasing sequence of integers, $l_0 = m + 1$ and $l_n = 2$.

4.1 Exact Solution with Real Numbers

To forge the intuition, let us relax the assumptions by allowing the (l_i) to be real numbers. For $0 < i < n$, a partial derivation of E with respect to l_i leads to $\frac{l_{i-1}}{l_i} = \frac{l_i}{l_{i+1}}$, hence l must be geometric. We deduce:

$$l_i = (m+1)\left(\frac{2}{m+1}\right)^{\frac{i}{n}}.$$

When the above solution consists of integers, it can be implemented. For example, consider $n = 3$ and $m = 15$. Then $l = (16, 8, 4, 2)$. Remarking that the signature $\tilde{\pi}$ is computed as the differences between consecutive terms in l, we have $\tilde{\pi} = (8, 4, 2)$. More generally, if there exists some integer k such that $m = 2k^n - 1$, then $\tilde{\pi}(i) = 2(k-1)k^{n-i}$.

If the Borda score is viewed as a measure of social welfare, then among the non-interleaved SERs, the solution above is optimal not only for *utilitarian*, but also for *egalitarian* social welfare (ex-ante), since all voters have the same expected utility.

4.2 Properties of the Optimal Sequence

Let us bring back the constraint that (l_i) are integers (still with monotonic non-interleaved elimination sequences).

Proposition 6 *If $m - 1 \leq n$, then the only optimal sequence has signature $\tilde{\pi} = (1, \ldots, 1)$. If $m - 1 \geq n$, then for any optimal sequence, $\mathrm{nv}(\pi) = n$, i.e. all voters participate.*

Proof. Let π be a sequence. Assume that a voter eliminates more than one candidate and that another voter eliminates none. We will then prove that π is not optimal. This immediately implies the two assertions of the proposition.

Let i be the last voter such that $\tilde{\pi}(i) > 1$ and $k = \mathrm{nv}(\pi) + 1$ the first non-participating voter. For $j = i+1, \ldots, k-1$, we have $\tilde{\pi}(j) = 1$. It is easy to check that $l_i = k - i + 1$.

Consider the sequence π' whose signature is identical to $\tilde{\pi}$, except that $\tilde{\pi}'(i) = \tilde{\pi}(i) - 1$ and $\tilde{\pi}'(k) = 1$. Denoting E' the counterpart of E for sequence π', we have: $E' - E = \frac{1}{k-i+1+\tilde{\pi}(i)} + \sum_{j=3}^{k-i+1} \frac{1}{j(j+1)} - \frac{1}{3} = \frac{1}{k-i+1+\tilde{\pi}(i)} - \frac{1}{k-i+2} < 0$. Hence π' is strictly better than π. $\square$

In the following, we often use the fact that if $m - 1 \geq n$, then all voters participate, hence l is strictly decreasing.

Proposition 7 *In an optimal sequence, we have*

$$\frac{l_i(l_i - 1)}{l_{i-1}} \leq l_{i+1} \leq \frac{l_i(l_i + 1)}{l_{i-1}}, \text{ for all } i \leq n - 1. \quad (1)$$

Proof. For $m - 1 \leq n$, this follows from Proposition 6. Assume $m - 1 > n$. Intuitively, we examine a "discrete partial derivative" of E with respect to l_i. Let $\epsilon \in \{-1, +1\}$ and π' a sequence whose signature is identical to $\tilde{\pi}$, except that $\tilde{\pi}'(i) = \tilde{\pi}(i) + \epsilon$ and $\tilde{\pi}'(i+1) = \tilde{\pi}(i+1) - \epsilon$. Comparing E and E', by optimality of π, we have $\frac{l_i}{l_{i-1}} + \frac{l_{i+1}}{l_i} \leq \frac{l_i - \epsilon}{l_{i-1}} + \frac{l_{i+1}}{l_i - \epsilon}$. Some basic algebra leads to $\epsilon l_{i+1} \geq \frac{l_i(\epsilon l_i - 1)}{l_{i-1}}$. Since it is true

for $\epsilon \in \{-1, +1\}$, the result follows. $\square$

Note that the difference between the lower and upper bounds in Equation (1) is $\frac{2l_i}{l_{i-1}} < 2$, hence there are at most two integers in that interval. We will use this fact for our algorithm.

Proposition 8 *The signature $\tilde{\pi}$ of any optimal sequence is non-increasing.*

Proof. Since Proposition 6 proves this assertion for $m - 1 \leq n$, we assume $m - 1 > n$. Let $i \in \{1, \ldots, n-1\}$. Some rewriting of (1) leads to:

$$\left(\tilde{\pi}(i) - 1\right)\frac{l_i}{l_{i-1}} \leq \tilde{\pi}(i+1) \leq \left(\tilde{\pi}(i) + 1\right)\frac{l_i}{l_{i-1}}. \quad (2)$$

Since $l_i < l_{i-1}$, the right inequality implies $\tilde{\pi}(i+1) < \tilde{\pi}(i) + 1$, hence $\tilde{\pi}(i+1) \leq \tilde{\pi}(i)$. $\square$

Finally, we state (omitting its proof due to lack of space) this last property of optimal non-interleaved sequences.

Proposition 9 *The signature $\tilde{\pi}$ of any optimal sequence is approximately convex, in the sense that:*

$$\tilde{\pi}(i) \leq \frac{\tilde{\pi}(i-1) + \tilde{\pi}(i+1) + 1}{2}, \text{ for any } i = 2, \ldots, n-1.$$

For example, for $n = 4$ and $m = 7$, the optimal sequence is $\tilde{\pi} = (2, 2, 1, 1)$ (as shown by our algorithm, see below). It is not convex, because $\tilde{\pi}(2) = 2$ is strictly greater than the average of its two adjacent terms $\tilde{\pi}(1) = 2$ and $\tilde{\pi}(3) = 1$. But it is approximately convex. This example shows that in the above inequality, the "error constant" $\frac{1}{2}$ is tight.

4.3 Computing Borda-optimal Sequences for IC

We present the sketch of an algorithm in $O(nm)$ that computes the signature $\tilde{\pi}$ of an optimal non-interleaved sequence under IC. We will build two tables predecessor and penalty, whose rows are indexed from 0 to n and columns from 2 to m, with the following interpretations. $\mathrm{predecessor}(i-1, j) = k$ means that in an optimal sequence, if $l_i = j$, then $l_{i-1} = k$. $\mathrm{penalty}(i-1, j) = p$ means that in an optimal sequence, if $l_i = j$, then $\frac{l_1}{l_0} + \ldots + \frac{l_i}{l_{i-1}} = p$. For any j, $\mathrm{predecessor}(0, j)$ is initialized to $m + 1$ and $\mathrm{penalty}(0, j)$ is initialized to $\frac{j}{m+1}$. Then, for each i from 1 to $n - 1$, we establish partial optimal sequences until l_{i+1}. To that end, each possible l_i yields two possible values of l_{i+1} thanks to Proposition 7. Since several values of l_i can lead to the same value of l_{i+1}, only the one with lowest penalty is kept as a valid predecessor of l_{i+1}, along with the corresponding penalty. At the end of the algorithm, tables predecessor and penalty are built. We have then $l_n = 2$ and for any $i \in \{1, \ldots, n\}$, $l_{i-1} = \mathrm{predecessor}(i-1, l_i)$. Finally, $\tilde{\pi} = (l_0 - l_1, \ldots, l_n - l_{n-1})$.

To conclude this section, we comment on two other ideas of algorithms: 1) explore only the rounded values of the real solution for sequence l (Section 4.1); or 2) use an integer adaptation of a gradient descent algorithm, until all conditions from Equation (1) are met (i.e. until a local optimum is reached, where incrementing or decrementing any l_i is not profitable). However, consider $n = 3$ and $m = 40$. The real

optimum is $(41.00, 14.98, 5.47, 2.00)$. But (as our algorithm shows) the integer optimum is $(41, 16, 6, 2)$, which does not consist in rounded values of the real solution: this excludes idea 1. Besides, another sequence, $(41, 14, 5, 2)$, also meets all conditions from Equation (1): it is a local optimum, but not the global one, leaving little hope for idea 2.

5 Experimental Study

Now we show to which extent sequential elimination rules provide a good Borda score approximation in practice. To this end we have carried out two sets of experiments. In both sets we have considered random instances drawn from five particular cultures: Impartial Culture (IC), Mallows with a dispersion parameter $\varphi = 0.1$ ($M_{0.1}$), Mallows with $\varphi = 0.5$ ($M_{0.5}$), Single-Peaked Impartial Culture (SPIC), and Single-Peaked with Uniform Peak (SPUP) [Conitzer, 2007].

In the first experiment, the number of voters n varies from 2 to 10 and the number of candidates is $2^{n+1} - 1$. Recall that under IC, optimal non-interleaved sequences in this setting are *geometric* sequences (Section 4.1). For each (n, m) and each culture, we focused on three different sequences: (i) geometric (*e.g.* $\tilde{\pi} = (8, 4, 2)$), (ii) round-robin (*e.g.* $\pi = (1, 2, 3, 1, 2, 3, ...)$) and (iii) random (single) dictator (*e.g.* $\pi = (1, 1, 1, ...)$). We have computed for each one the mean value of the differential ratio $\frac{S_B(F_\pi(V), V) - \min(S_B(., V))}{\max(S_B(., V)) - \min(S_B(., V))}$ over a sample of 10000 profiles generated for each (n, m) and culture. The results are shown in Figure 1 (the result for $M_{0.1}$ is not shown but gives an even better approximation ratio than $M_{0.5}$ for all three sequences).
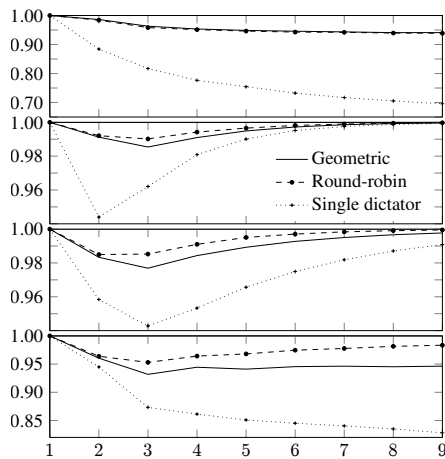

Figure 1: Borda approximation quality of three kinds of sequences for IC, $M_{0.5}$, SPIC, SPUP (from top to bottom).

In the second experiment, the number of voters n varies from 2 to 6 and the number of candidates m from $n+1$ to $2n$. For each (n, m) and each culture, we used an exhaustive search algorithm over the space of all sequences to find the sequence yielding the best expected Borda score (estimated from a sample of 10000 profiles). For all settings, these best sequences, their differential ratios and the standard-deviation of this ratio (both in percentage) are shown in Table 1.

Culture	n	m	opt. sequence	ratio
IC	4	8	$(1, 2, 2, 3, 1, 3, 4)$	$90.5\% \pm 0.08$
IC	5	10	$(1, 1, 2, 3, 2, 1, 3, 4, 5)$	$88.5\% \pm 0.1$
$M_{0.1}$	5	10	$(1, 2, 3, 2, 2, 2, 2, 1, 4)$	$99.5\% \pm 0.005$
SPIC	5	10	$(1, 2, 3, 1, 1, 4, 2, 3, 5)$	$98.2\% \pm 0.01$
SPUP	5	10	$(1, 2, 3, 4, 5, 1, 2, 4, 3)$	$95.1\% \pm 0.04$

Table 1: best sequences, their ratio and standard-deviation

What is remarkable is that the estimated expected value for the differential ratio is almost always over 90% for round-robin, geometric sequences and the optimal sequence and for all the cultures (IC being the worst), and is even close to 100% for more correlated profiles like the Mallows ones. Thus, although the optimal sequence is interleaved, the best non-interleaved sequence (geometric) performs almost as well. Even if it differs from the optimal sequence, round-robin even gives a better approximation of it than the best non-interleaved sequence. Unsurprisingly, random dictatorship performs quite badly in comparison.

Finally, we tested our approach on the real Sushi dataset from PrefLib [Mattei and Walsh, 2013]. To obtain instances with more candidates than voters, we sampled 10000 profiles of 5 voters from the whole set (in this dataset, $m = 10$) and as before computed the best sequence according to the average differential ratio. The sequence obtained is $(1, 1, 1, 2, 2, 2, 3, 4, 5)$, and gives a differential ratio of 93.035%, showing that the method may perform well on realistic datasets as well.

6 Conclusion

Sequential elimination rules have a low communication complexity, while enjoying several important properties. We have defined a framework for identifying optimal sequences for a quality measure, the Borda score; a further research direction consists in identifying optimal sequences for other measures. This work is a first study, in a setting with few voters; we plan to investigate similar low-communication rules, consisting of a sequence of steps where a designated voter vetoes or supports an alternative, applicable in different settings. Moreover, a similar family of protocols in fair division, *picking sequences*, have also been studied under the social welfare point of view [Bouveret and Lang, 2011; Kalinowski *et al.*, 2013; Aziz *et al.*, 2016]. Now, picking sequences can be viewed to some extent as elimination protocols, as each agent, when picking an item, eliminates a set of possible assignments; it is worth exploring this connection further. Finally, as we have seen, this family of voting rules is of course not strategyproof. In this paper, we have left strategical issues aside, but it would be a crucial topic for future research.

Acknowledgements

References

[Arnold *et al.*, 1992] Barry. C. Arnold, N. Balakrishnan, and H. N. Nagaraja. *A First Course in Order Statistics*. Wiley, New York, 1992.

[Aziz *et al.*, 2016] Haris Aziz, Thomas Kalinowski, Toby Walsh, and Lirong Xia. Welfare of sequential allocation mechanisms for indivisible goods. In *ECAI 2016 - 22nd European Conference on Artificial Intelligence, 29 August-2 September 2016, The Hague, The Netherlands - Including Prestigious Applications of Artificial Intelligence (PAIS 2016)*, pages 787–794, 2016.

[Balasubramanian *et al.*, 1997] R. Balasubramanian, Venkatesh Raman, and G. Srinivasaragavan. Finding scores in tournaments. *J. Algorithms*, 24(2):380–394, 1997.

[Boutilier and Rosenschein, 2015] Craig Boutilier and Jeffrey Rosenschein. Incomplete information and communication in voting. In F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. D. Procaccia, editors, *Handbook of Computational Social Choice*, chapter 10. Cambridge University Press, 2015.

[Bouveret and Lang, 2011] Sylvain Bouveret and Jérôme Lang. A general elicitation-free protocol for allocating indivisible goods. In *IJCAI 2011, Proceedings of the 22nd International Joint Conference on Artificial Intelligence*, pages 73–78, 2011.

[Brandl *et al.*, 2015] Florian Brandl, Felix Brandt, and Johannes Hofbauer. Incentives for participation and abstention in probabilistic social choice. In *Proceedings of AAMAS-15*, pages 1411–1419, 2015.

[Conitzer and Sandholm, 2005] Vincent Conitzer and Tuomas Sandholm. Communication complexity of common voting rules. In *Proceedings of ACM-EC 2005*, pages 78–87, 2005.

[Conitzer, 2007] Vincent Conitzer. Eliciting single-peaked preferences using comparison queries. In *Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems*, page 65. ACM, 2007.

[d'Aspremont and Gevers, 2002] Claude d'Aspremont and Louis Gevers. Social welfare functionals and interpersonal comparability. *Handbook of social choice and welfare*, 1:459–541, 2002.

[Dery *et al.*, 2014] Lihi Naamani Dery, Meir Kalech, Lior Rokach, and Bracha Shapira. Reaching a joint decision with minimal elicitation of voter preferences. *Information Sciences*, 278:466–487, 2014.

[Dey, 2017] Palash Dey. Query complexity of tournament solutions. In *AAAI*, pages 2992–2998. AAAI Press, 2017.

[Filmus and Oren, 2014] Yuval Filmus and Joel Oren. Efficient voting via the top-k elicitation scheme: A probabilistic approach. In *ACM-EC 2014*, pages 295–312, 2014.

[Goel and Lee, 2012] Ashish Goel and David Lee. Triadic consensus: A randomized algorithm for voting in a crowd. In *Proceedings of the Eighth Workshop on Internet and Network Economics (WINE-12)*, pages 434–447, 2012.

[Gross *et al.*, 2017] Stephen Gross, Elliot Anshelevich, and Lirong Xia. Vote until two of you agree: Mechanisms with small distortion and sample complexity. In *Proc. of 31st Conference on Artificial Intelligence (AAAI 2017)*, 2017.

[Hansen, 2016] Jeremy Hansen. The random pairs voting rule: Introduction and evaluation with a large dataset. In *Proceedings of COMSOC-16*, 2016.

[Kalech *et al.*, 2011] Meir Kalech, Sarit Kraus, Gal A. Kaminka, and Claudia V. Goldman. Practical voting rules with partial information. *Journal of Autonomous Agents and Multi-Agent Systems*, 22(1):151–182, 2011.

[Kalinowski *et al.*, 2013] Thomas Kalinowski, Nina Narodytska, and Toby Walsh. A social welfare optimal sequential allocation procedure. In *IJCAI 2013, Proceedings of the 23rd International Joint Conference on Artificial Intelligence, Beijing, China, August 3-9, 2013*, pages 227–233, 2013.

[Kelly, 1978] J. S. Kelly. *Arrow Impossibility Theorems*. Academic Press, New York, 1978.

[Lu and Boutilier, 2011] Tyler Lu and Craig Boutilier. Robust approximation and incremental elicitation in voting protocols. In *Proceedings of IJCAI-11*, pages 287–293, 2011.

[Mattei and Walsh, 2013] Nicholas Mattei and Toby Walsh. Preflib: A library of preference data HTTP://PREFLIB.ORG. In *Proceedings of the 3rd International Conference on Algorithmic Decision Theory (ADT 2013)*, Lecture Notes in Artificial Intelligence. Springer, 2013.

[Procaccia, 2008] Ariel Procaccia. A note on the query complexity of the condorcet winner problem. *Information Processing Letters*, 108(6):390–393, 2008.

[Service and Adams, 2012] Travis C. Service and Julie A. Adams. Communication complexity of approximating voting rules. In *Proceedings of AAMAS-12*, pages 593–602, 2012.