



Extended formulations for order polytopes through network flows

Clinton P. Davis-Stober^a, Jean-Paul Doignon^{b,*}, Samuel Fiorini^b, François Glineur^c, Michel Regenwetter^d

^a University of Missouri, Columbia, United States

^b Université Libre de Bruxelles, Belgium

^c Université catholique de Louvain, Belgium

^d University of Illinois at Urbana–Champaign, United States

HIGHLIGHTS

- Polyhedral combinatorics deliver information on random preference models.
- New extended formulations break through combinatorial explosion of some of these models.
- New formulations enable study of random models with more alternatives for linear, weak, interval and semiorders.
- For three models, the size of the description achieves asymptotic optimality.

ARTICLE INFO

Article history:

Received 5 October 2017

Received in revised form 27 February 2018

Keywords:

Order polytopes

Extended formulations

Network flows

Probabilistic choice

Distribution free random utility

ABSTRACT

Mathematical psychology has a long tradition of modeling probabilistic choice via distribution-free random utility models and associated random preference models. For such models, the predicted choice probabilities often form a bounded and convex polyhedral set, or polytope. Polyhedral combinatorics have thus played a key role in studying the mathematical structure of these models. However, standard methods for characterizing the polytopes of such models are subject to a combinatorial explosion in complexity as the number of choice alternatives increases. Specifically, this is the case for random preference models based on linear, weak, semi- and interval orders. For these, a complete, linear description of the polytope is currently known only for, at most, 5–8 choice alternatives. We leverage the method of extended formulations to break through those boundaries. For each of the four types of preferences, we build an appropriate network, and show that the associated network flow polytope provides an extended formulation of the polytope of the choice model. This extended formulation has a simple linear description that is more parsimonious than descriptions obtained by standard methods for large numbers of choice alternatives. The result is a computationally less demanding way of testing the probabilistic choice model on data. We sketch how the latter interfaces with recent developments in contemporary statistics.

© 2018 Elsevier Inc. All rights reserved.

1. Introduction

Much of the literature on choice behavior in the social sciences centers on modeling the unobservable hypothetical preferences or cognitive processes underlying observable choice behavior. This endeavor encounters a huge hurdle in empirical applications: How can models accommodate heterogeneity in choice behavior across individuals as well as within any given individual? To this end, researchers have employed a variety of probabilistic modeling

approaches. Prominent examples include computational, stochastic process models that mimic hypothetical cognitive processes, such as the well-known diffusion model (e.g., Ratcliff & Rouder, 1998; Ratcliff & Smith, 2004) and the linear ballistic accumulator model (Brown & Heathcote, 2008; Trueblood, Brown, & Heathcote, 2014). Other approaches include multinomial processing tree models (see, e.g., Erdfelder et al., 2009, for a review). Rather than model latent cognitive processes, we concentrate on the more abstract notion of latent preferences. In so doing, we consider models that use a minimum of mathematical assumptions and that delineate large classes of theories of choice behavior. This approach is grounded in a long tradition of axiomatization and axiom testing in mathematical psychology (see, e.g., Luce, 2000) and naturally interfaces with contemporary statistical methods (Davis-Stober & Brown, 2011; Davis-Stober, Brown, & Cavagnaro, 2015).

* Correspondence to: U.L.B. c.p. 216, Bd du Triomphe, B-1050 Bruxelles, Belgium.

E-mail addresses: stoberc@missouri.edu (C.P. Davis-Stober), doignon@ulb.ac.be (J.-P. Doignon), sfiorini@ulb.ac.be (S. Fiorini), francois.glineur@uclouvain.be (F. Glineur), regenwet@illinois.edu (M. Regenwetter).

Evaluating choice data against such models allows for strong inferences regarding the latent preferences that give rise to choice behavior. Here, we tackle challenges with the mathematical characterization of such models.

Specifically, we consider classes of choice theories based upon four types of transitive binary preference relations: linear orders (rankings without ties), weak orders (rankings with or without ties), semiorders (partial rankings up to a constant threshold of discrimination), and interval orders (partial rankings up to a stimulus-dependent threshold of discrimination). For each of these algebraic representations of preferences, we consider probabilistic choice induced by an (unspecified) probability distribution over permissible preference states, i.e., over relations of the selected type. These models are also known in the literature as *random preference models*, *mixture models*, or *random relation models*. Because each of these models has been shown to have a random utility representation, our results also directly apply to characterizing various kinds of distribution-free random utility models, as well as random function models (see, e.g., [Regenwetter & Marley, 2001](#)).

Among the most important applications of these models are studies aimed at testing whether decision makers have transitive preferences (e.g., because they employ compensatory decision strategies by which they coherently trade-off between competing decision attributes) or whether they violate transitivity (say, because they employ simple heuristics that cause ‘incoherent’ preferences). It is well-known ([Roberts, 1979](#)) that, for finitely many choice options, complete asymmetric binary preferences are transitive if and only if they are linear orders. Dropping the requirement of complete preference, asymmetric binary preferences on a finite set of choice options are negatively transitive (hence also transitive) if and only if they are strict weak orders ([Roberts, 1979](#)). Testing random preference models on laboratory data faces two main hurdles: (i) Because the prediction ranges of these models form convex polytopes (that is, bounded convex polyhedral sets as in [Ziegler, 1998](#)), they require order-constrained statistical methods; (ii) The order constraints (geometrically, the linear inequalities defining facets of the polytope) are fully known only when the number of choice alternatives is small. Recent developments have essentially solved the first challenge (see [Davis-Stober, 2009](#); [Karabatsos, 2006](#); [Klugkist & Hoijtink, 2007](#); [Myung, Karabatsos, & Iverson, 2005](#); [Silvapulle & Sen, 2005](#)).

Using some of these approaches, [Regenwetter, Dana, and Davis-Stober \(2011\)](#) provided the first empirical and statistically adequate test of the “linear ordering model” in the literature. Following [Tversky \(1969\)](#), they used five choice alternatives per stimulus set. [Regenwetter and Davis-Stober \(2012\)](#), who reported the first empirical test of the “weak order model”, also used five choice alternatives per stimulus set, in this case because the mathematical structure of the weak order polytope is not yet fully understood for larger numbers of choice alternatives. [Regenwetter and Davis-Stober \(2011\)](#) used four choice alternatives to test the “semiorder model” and the “interval order model” because they did not know a full description of the corresponding polytopes for more than four options.

In this article, we leverage extended formulations. We study network flow polytopes to characterize random preference models of linear, weak, semi-, and interval orders in novel ways. This approach generates a simpler, more parsimonious description of the corresponding random preference model and allows empirical researchers to study these models for larger numbers of choice alternatives than previously possible. While in [Fiorini \(2004\)](#) consideration of network flow polytopes led to a new proof of the main theorem of [Falmagne \(1978\)](#), here we show that it is crucial for generating more efficient ways to test four random utility models.

The paper proceeds as follows. First, we review key facts about linear, weak, semi-, and interval orders. Then, we discuss the

convex polytopes associated with the random preference models based on each of these four types of preference relations and the obstacles faced by traditional mathematical methods for characterizing their model ranges. The following two sections overcome these obstacles and investigate extended formulations, especially those based on network flows, and associated network flow polytopes. We then discuss how these network flow polytopes interface with some contemporary developments in statistics. We end with a summary and detail future directions of this work.

2. Models of pairwise preferences and their numerical representations

Throughout the paper, S denotes a finite set of n choice alternatives (thus, $|S| = n$). The preferences of a decision maker among the alternatives are cast as a relation R on S , with iRj meaning that the decision maker likes i strictly less than j (hence strictly prefers j over i). In a simple case, the relation R is a *linear order*, that is an irreflexive, transitive and complete relation (for a definition of these and similar terms, see for instance [Fishburn, 1985, Chapter 1](#), or [Pirlot & Vincke, 1997, Chapter 3](#)). As is well-known, such an ideal situation occurs when preferences reflect comparisons of utility values and no two choice alternatives have the same utility.

Proposition 1 (Linear Order). *The relation R on the set S of alternatives is a linear order if and only if there exists an injective mapping $u : S \rightarrow \mathbb{R}$ such that, for all i, j in S :*

$$iRj \iff u(i) < u(j). \quad (1)$$

If we drop the injectivity condition in [Proposition 1](#), thus making room for alternatives having equal utility, we get a (strict) *weak order* (an asymmetric and negatively transitive relation).

Proposition 2 (Weak Order). *The relation R on the set S of alternatives is a weak order if and only if there exists a mapping $u : S \rightarrow \mathbb{R}$ such that, for all i, j in S :*

$$iRj \iff u(i) < u(j). \quad (2)$$

In many situations, comparing utility values defined as real numbers is not realistic. According to an extended viewpoint, any alternative i from S is assigned a range of values taking the form of a real interval $[\ell(i), h(i)]$ in such a way that iRj holds exactly if the interval $[\ell(i), h(i)]$ lies entirely before the interval $[\ell(j), h(j)]$ (see [Fig. 1](#) for a geometric illustration). The relation R is then called an *interval order* on S . Here, one may interpret $\ell(i)$ as the lower utility value of i , $h(i)$ as the upper utility value, and $h(i) - \ell(i)$ as a perceptual *threshold* for utilities. We call such a pair of mappings (ℓ, h) a *representation* of R .

A combinatorial characterization of interval orders, due independently to [Fishburn \(1970\)](#) and [Mirkin \(1972\)](#) (but also included in the result on ‘bi-quasi-series’ in [Ducamp & Falmagne, 1969](#)), is as follows.

Proposition 3 (Interval Order). *The relation R on the set S of alternatives satisfies the following two conditions, for all i, j, i', j' in S :*

$$\text{not } (iRi); \quad (3)$$

$$(iRj \text{ and } i'Rj') \implies (iRj' \text{ or } i'Rj) \quad (4)$$

if and only if there exist two mappings $\ell, h : S \rightarrow \mathbb{R}$ such that the following two conditions hold, for all i, j in S :

$$\ell(i) \leq h(i); \quad (5)$$

$$iRj \iff h(i) < \ell(j). \quad (6)$$

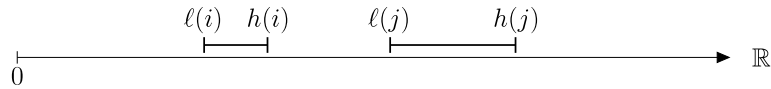


Fig. 1. Geometric representation of iRj in the case of an interval order R , thus: i is less preferable than j if and only if $h(i) < l(j)$.

When the threshold $h(i) - \ell(i)$ in Proposition 3 is required to be independent of the alternative i , we get the semiorder model, and say that the relation R is a *semiorder* on S . The latter model appears in Wiener (1915), Armstrong (1939), and Luce (1956) and was given its name by the last author. The following characterization is due to Scott and Suppes (1958).

Proposition 4 (Semiorder). *The relation R on the set S of alternatives satisfies the following three conditions, for all i, j, i', j', i'', j'' in S :*

not (iRi) ;

$(iRj \text{ and } i'Rj') \implies (iRj' \text{ or } i'Rj)$;

$(iRi' \text{ and } i'Ri'') \implies (iRj \text{ or } jRi'')$

if and only if there exist two mappings $\ell, h : S \rightarrow \mathbb{R}$ and a nonnegative real number r such that the following two conditions hold, for all i, j in S :

$$iRj \iff h(i) < \ell(j); \quad (7)$$

$$h(i) - \ell(i) = r. \quad (8)$$

3. Random preference models and their associated polytopes

Relations like those of Propositions 1–4 can be used to model the decision maker's preferences at a particular moment. In general, the decision maker may be uncertain about her preferences and may probabilistically sample a preference from some collection of preference states, when required to make decisions. We assume all permissible preference states obey the model selected. Thus, at any time we may interview the decision maker about any pair (i, j) of alternatives, and collect an answer reflecting her currently sampled preference. As a consequence, repeated interviews will lead to frequencies of having iRj (or not). On the theoretical side, this leads us to posit a probability distribution P on the set $\mathcal{LO}_S$ of all linear orders on S —to take this model as an example; similar considerations of course hold in the other three models, for the collection $\mathcal{WO}_S$ of all weak orders, $\mathcal{IO}_S$ of all interval orders, or $\mathcal{SO}_S$ of all semiorders. With $P(R)$ denoting the probability that the decision maker's (currently sampled) relation is R , we assume that the probability that she prefers alternative j over alternative i is given by the following formula:

$$p_{(i,j)} = \sum \{P(R) \mid R \in \mathcal{LO}_S \text{ with } iRj\}. \quad (9)$$

Not all collections of real values $\{p_{(i,j)} \mid (i,j) \in S \times S\}$ can be produced in this way, for a probability distribution P on $\mathcal{LO}_S$. For instance, Eq. (9) implies $p_{(i,j)} \geq 0$ and, for any $i \neq j$, $p_{(i,j)} + p_{(j,i)} = 1$. Other necessary conditions are known, but necessary and sufficient conditions are known only for small values of n . There are doubts that a tractable set of necessary and sufficient conditions could be found that is valid for all n (because the 'linear ordering problem' is NP-hard, see Garey & Johnson, 1979). To get a better understanding of the collections $\{p_{(i,j)} \mid (i,j) \in S \times S\}$ that satisfy Eq. (9) for at least one probability distribution P on $\mathcal{LO}_S$, we now develop a geometric interpretation and we do the same later for the other three models.

Each linear order on the set S , being a collection of pairs of alternatives of S , is a subset of $S \times S$. Because linear orders (as well as the other orders we consider in this paper) are irreflexive, they never contain a pair of the form (i, i) , with $i \in S$. This prompts us

to discard all pairs (i, i) , for $i \in S$. Let $S \star S := (S \times S) \setminus \{(i, i) \mid i \in S\}$. Consider the space $\mathbb{R}^{S \star S}$, in which a point x has a coordinate $x_{(i,j)}$ for each pair (i, j) of distinct alternatives from S . To any linear order R on S , associate the point x^R with coordinates given by

$$(x^R)_{(i,j)} := \begin{cases} 1 & \text{if } iRj, \\ 0 & \text{otherwise.} \end{cases} \quad (10)$$

In other words, x^R is the *characteristic vector* of the linear order R . The convex hull in $\mathbb{R}^{S \star S}$ of the characteristic vectors of all linear orders on S , that is the set of all convex combinations of those vectors (linear combinations with nonnegative coefficients summing to one), is the *linear order polytope* of S , which we denote P_{LO}^S :

$$P_{LO}^S := \text{conv} \{x^R \mid R \in \mathcal{LO}_S\}. \quad (11)$$

A collection $\{p_{(i,j)} \mid (i,j) \in S \star S\}$ satisfies Eq. (9) for at least one probability distribution P on $\mathcal{LO}_S$ if and only if the point p lies in the linear order polytope P_{LO}^S . Thus, to characterize such collections, we would like a criterion for a point x from $\mathbb{R}^{S \star S}$ to lie in P_{LO}^S . A classical result (see, for instance, Ziegler, 1998) states that any polytope admits a linear description, that is a finite system of linear equations and inequalities whose solution set is exactly the polytope. Even if we insist that the system has a minimum number of equations and inequalities, it is generally not unique; however, the number of linear equations is then the codimension of the polytope, and there is exactly one inequality per facet of the polytope. In the case of the linear order polytope P_{LO}^S , the number of equations on $\mathbb{R}^{S \star S}$ equals $\binom{n}{2} = n(n-1)/2$ (those being the already mentioned $p_{(i,j)} + p_{(j,i)} = 1$), but no precise estimation seems to be available on the number of inequalities. As far as we know, a complete linear description of P_{LO}^S is lacking when $|S| = n > 7$. For $n = 8$, the number of vertices is $8! = 40,320$ but only a lower bound on the number of facets is available. This lower bound is huge: the polytope has at least 488,602,996 facets (Christof & Reinelt, 1996). For other results on the linear order polytope, we refer the reader to Boyd and Puleyblank (2009), Doignon, Fiorini, and Joret (2006), Doignon, Fiorini, and Joret (2009), Fiorini (2006a, b), Martí and Reinelt (2011), Oswald, Reinelt, and Seitz (2009) and their references. For an application of the linear order model to test transitivity of preferences on empirical data, see Regenwetter, Dana, and Davis-Stober (2010), Regenwetter et al. (2011) and Regenwetter, Dana, Davis-Stober, and Guo (2011).

The *weak order polytope* of S , which we denote by P_{WO}^S , is similarly defined in $\mathbb{R}^{S \star S}$:

$$P_{WO}^S := \text{conv} \{x^R \mid R \in \mathcal{WO}_S\}. \quad (12)$$

Note in passing that several publications (e.g., Doignon & Fiorini, 2002; Fiorini & Fishburn, 2004) consider 'complete weak orders.' The latter are transitive and complete relations, in other words the inverses of the complements of weak orders. The resulting 'complete weak order polytope' is the image of our polytope P_{WO}^S by the affine transformation $\mathbb{R}^{S \star S} \rightarrow \mathbb{R}^{S \star S} : x \mapsto x'$ with $x'_{(i,j)} = 1 - x_{(j,i)}$. Hence, all the results about one of these polytopes have counterparts for the other.

The Ph.D. dissertation of Fiorini (2001) contains an overview of results about the two order polytopes we just introduced, including a complete description of P_{WO}^S for $|S| = n = 4$. To our knowledge, a complete description of P_{WO}^S is known only up to $n = 5$. For $n = 5$, the weak order polytope P_{WO}^S has 75,834 facets and 541 vertices (Regenwetter & Davis-Stober, 2012).

Two other polytopes play a role in our study. They are the *interval order polytope* of S

$$P_{IO}^S := \text{conv} \{x^R \mid R \in \mathcal{IO}_S\} \quad (13)$$

and the *semiorder polytope* of S

$$P_{SO}^S := \text{conv} \{x^R \mid R \in \mathcal{SO}_S\}. \quad (14)$$

The interval order polytope is investigated by Müller and Schulz (1995), the semiorder polytope by Suck (1995) and both of them by Doignon and Rexhep (2016). For $n = 4$, Regenwetter and Davis-Stober (2011) report that the interval order polytope has 207 vertices and 191 facets, while the numbers for the semiorder polytope are 183 and 563.

Notice that the three polytopes P_{WO}^S , P_{IO}^S and P_{SO}^S are full dimensional, while P_{LO}^S is not. However, this difference has no impact on the construction of extended formulations through network flows which we now review.

4. Network flow representations of order polytopes

We now move to a third way of investigating our four order polytopes. Many applications in psychology and economics aim to evaluate random preference models on human subjects choice data from laboratory studies. This typically involves maximizing concave functions of choice probabilities (e.g., a log-likelihood function) subject to the constraint that the choice probabilities belong to the polytope in question. Such optimizations are straightforward when the geometric structure of the polytope is fully known. The extended formulations that we now consider make it possible to solve some of these optimization goals in cases where obtaining a complete geometric representation of the polytope in question is prohibitively expensive.

4.1. The concept of an extended formulation

In the previous section, we described how testing whether a collection of choice probabilities satisfies the linear order model amounts to checking whether a given point lies in the linear order polytope P_{LO}^S (we again base our exposition on this model), which can be done by checking the validity of each linear equation and inequality in its description. Unfortunately, unless we limit ourselves to small enough n the latter task is made difficult or even impossible by the inherent intricacy, or sheer unavailability, of a linear description of P_{LO}^S . In this section, we sketch the main steps to overcome this difficulty. We provide technical constructions and proofs in the next section.

The main idea is to work with another, ad hoc polytope that projects in some specific way on P_{LO}^S (more precisely: there is an affine transformation mapping the ad hoc polytope onto P_{LO}^S). Any linear description of such an ad hoc polytope is called an *extended formulation* of P_{LO}^S . For a survey of the notion of an extended formulation, see Conforti, Cornuéjols, and Zambelli (2010), Kaibel (2011) and Wolsey (2011). For the four order polytopes we study in this paper, the ad hoc polytope happens to be a “flow polytope”, so we first define the latter notion and review some known facts.

4.2. The concept of a network flow polytope

Let $D = (N, A)$ be a network with node set N and arc set A (for network terminology, see, for instance, Korte & Vygen, 2008). In N , we designate two special nodes: the *source* node s and the *sink* node t .

For a node v , we denote the sets of arcs leaving and entering v by $\delta^+(v)$ and $\delta^-(v)$, respectively. Formally,

$$\begin{aligned} \delta^+(v) &:= \{a \in A \mid \exists w \in N : a = (v, w)\}, \\ \delta^-(v) &:= \{a \in A \mid \exists w \in N : a = (w, v)\}. \end{aligned}$$

Consider a set B of arcs in A . We encode B by its characteristic vector $\chi^B \in \mathbb{R}^A$, defined by letting $\chi_a^B := 1$ if $a \in B$ and $\chi_a^B := 0$ if $a \in A \setminus B$ (we notate the characteristic vectors x^R and χ^B differently to emphasize that they refer to the distinct master sets $S \star S$ and A , respectively). For a vector Φ in $\mathbb{R}^A$, with components Φ_a for $a \in A$, we define the number

$$\Phi(B) := \sum_{a \in B} \Phi_a. \quad (15)$$

A *flow* in D is a vector Φ from $\mathbb{R}^A$, associating a nonnegative number Φ_a to each arc a of the network, such that the out-flow $\Phi(\delta^+(v))$ equals the in-flow $\Phi(\delta^-(v))$ at each node v distinct from the source node s and the sink node t . The *value* of a flow Φ equals $\Phi(\delta^+(s)) - \Phi(\delta^-(s))$, that is, the net out-flow at the source node. A flow Φ is said to be *integral* if Φ_a is an integer, for all arcs a .

In the context of this work, we assume that the network D is acyclic (this simplifying assumption is met in all our examples). We define the *flow polytope* $F = F(D)$ of the network D as the polytope whose vertices are the characteristic vectors of the sets of arcs of all the s – t (directed) paths in D . Any such characteristic vector is an integral s – t flow in D of value 1. Because we assume that D is acyclic, the converse also holds: any vector Φ from $\mathbb{R}^A$ which is an integral s – t flow in D of value 1 is the characteristic vector of an s – t path in D . A complete linear description of $F(D)$ is as follows (see for instance Korte & Vygen, 2008, Theorem 8.8)—we call it the *canonical description* of the flow polytope $F(D)$:

$$F(D) = \left\{ \Phi \in \mathbb{R}^A \mid \begin{array}{ll} \Phi(\delta^+(v)) - \Phi(\delta^-(v)) &= 0, \quad \forall v \in N \setminus \{s, t\}, \\ \Phi(\delta^+(s)) - \Phi(\delta^-(s)) &= 1, \\ \Phi_a &\geq 0, \quad \forall a \in A \end{array} \right\}. \quad (16)$$

Following the current practice in extended formulations, we define the *size* of a linear description as the total number of inequalities in the description, thus disregarding the number of variables and equalities, as well as the bit complexity of the coefficients. Hence, the size of the canonical description of $F(D)$ is $|A|$.

Again, let S denote the set of alternatives, with $|S| = n$. The conventional notations Ω , O and Θ refer, as usual, to asymptotic comparisons of functions, see Cormen, Leiserson, Rivest, and Stein (2003) for their definitions. We now show that three of our order polytopes, namely the linear order, weak order and interval order polytopes, share the following property for some constant c associated with each family of polytopes: For each of these polytopes, there exist a flow polytope F with a linear description of size $O(c^n)$ and a projection from $\mathbb{R}^A$ to $\mathbb{R}^{S \star S}$ that maps the flow polytope F to the considered order polytope. We explicitly build this flow polytope, which is thus an extended formulation. We also build a flow polytope providing an extended formulation for the semiorder polytope with, surprisingly, canonical description size in $\Omega(n!)$, which is not $O(c^n)$ for any constant c .

As we discuss later in Section 7, it follows from recent results that all our extended formulations are size-optimal, except perhaps that for the semiorder polytope. More precisely, any extended formulation for any of our order polytopes for a set of alternatives of size n has size at least $(3/2)^{\lfloor n/2 \rfloor}$.

In order to specify the networks, the key idea is to use ordinal representations of the relations involved. We start with weak orders.

4.3. The case of weak orders

Remember from Proposition 2 that a relation R on S is a weak order exactly if there exists a mapping u from S to $\mathbb{R}$ such that

$$i R j \iff u(i) < u(j). \quad (17)$$

For each τ in $\mathbb{R}$, we let

$$X(\tau) := \{i \in S \mid u(i) < \tau\}. \quad (18)$$

Now, imagine increasing the level τ continuously from $\tau = \min u(S)$ to $\tau = \max u(S) + 1$. Then the successive values of $X(\tau)$ form a finite sequence of distinct sets $X_0, X_1, \dots, X_m$, starting with the empty set $\emptyset$ and ending with the entire set S . We call this sequence the *profile of the weak order* R (notice that the profile of the weak order R does not depend on the representation of R , and also that it determines R). More formally, let $u(S) = \{\tau_0, \tau_1, \dots, \tau_{m-1}\}$ with $\tau_0 < \tau_1 < \dots < \tau_{m-1}$. For every k in $\{0, 1, \dots, m-1\}$, set $X_k := \{i \in S \mid u(i) < \tau_k\}$, and moreover $X_m := S$. The profile of the weak order R forms a chain

$$\emptyset =: X_0 \subset X_1 \subset \dots \subset X_m := S. \quad (19)$$

(Here and throughout, $\subset$ indicates strict inclusion.)

Now form the network whose nodes are the subsets of S , letting $s := \emptyset$ and $t := S$, and whose arcs are the pairs (X, Z) of subsets of S such that $X \subset Z$. We see that the profile of each weak order R determines an s – t path in this network. Conversely, any s – t path in this network determines a unique weak order. Because the two constructions correspond to mappings that are each other's inverses, we get a bijection from the set of s – t paths to the set of weak orders on S . Later (in Section 5.2), we prove that this bijection yields a projection from the flow polytope of the network we just built to the weak order polytope of S .

4.4. The case of linear orders

We now turn to the linear orders on S . In the above network (for weak orders), they correspond to the s – t paths that use only arcs (X, Z) where $X \subset Z$ with the restriction that $|X| + 1 = |Z|$. Because the s – t paths describing linear orders only use those arcs, we now delete all the other arcs from the network while keeping all the nodes. In the resulting network, the s – t paths bijectively correspond to the linear orders on S . We later infer a projection from the flow polytope of this network to the linear order polytope of S (see Theorem 1).

4.5. The case of interval orders

To create a network for interval orders, we use Proposition 3 according to which a relation R on S is an interval order when there exist two mappings ℓ and h from S to $\mathbb{R}$ such that $\ell(i) \leq h(i)$ for all $i \in S$ and

$$iRj \iff h(i) < \ell(j). \quad (20)$$

In the following, we always assume that ℓ and h are both one-to-one, and that the images of S under ℓ and h are disjoint. Why can we modify the mappings ℓ and h in order to satisfy these two conditions? Each value of ℓ is only constrained by Eqs. (5) and (6) to lie in a convex subset of $\mathbb{R}$ which is determined by the values of h . Moreover this convex set has more than one point. Hence, we can always adjust values of ℓ in order to make ℓ one-to-one. Similarly, we can next adjust values of h in order to make h one-to-one and at the same time $\ell(S)$ and $h(S)$ disjoint.

Each level τ from $\mathbb{R}$ determines two subsets of S (for a geometric example, see Fig. 2, but, for the moment, ignore the last sentence in the caption):

$$X(\tau) := \{i \in S \mid \ell(i) < \tau\}, \quad (21)$$

$$Y(\tau) := \{i \in S \mid h(i) < \tau\}. \quad (22)$$

Because $\ell(i) \leq h(i)$ holds for all $i \in S$, we have $Y(\tau) \subseteq X(\tau)$ for every level τ . When we continuously increase τ from $\tau =$

$\min \ell(S)$ to $\tau = \max h(S) + 1$, we observe a finite sequence of distinct pairs of sets $(X_0, Y_0), (X_1, Y_1), \dots, (X_m, Y_m)$ that forms the *profile of the interval order* R (notice that the profile of an interval order R technically depends on the chosen representation ℓ, h of R ; however, each such profile determines R). Our assumptions on the mappings ℓ and h imply that any two consecutive pairs of sets $(X_k, Y_k), (X_{k+1}, Y_{k+1})$ satisfy the following three conditions:

- $X_k \subseteq X_{k+1}$,
- $Y_k \subseteq Y_{k+1}$,
- either $\begin{cases} |X_{k+1}| = |X_k| + 1 \\ |Y_{k+1}| = |Y_k| \end{cases}$ or $\begin{cases} |X_{k+1}| = |X_k| \\ |Y_{k+1}| = |Y_k| + 1. \end{cases}$

Using these conditions as an inspiration, we define the network D_{10}^S whose nodes are pairs (X, Y) of subsets of S with $Y \subseteq X$, with source node $s := (\emptyset, \emptyset)$ and sink node $t := (S, S)$ (the details are given below, including the definition of the arcs). Each s – t path in this network encodes an interval order on S . Although several of these paths encode the same interval order, we obtain a projection from the flow polytope of D_{10}^S to the interval order polytope of S .

4.6. The case of semiorders

The case of semiorders is a bit more involved. Semiorders are particular interval orders, so they correspond to profiles of interval orders having an additional property. We would like to build a new network whose s – t paths exactly correspond to the profiles of semiorders. We obtain the nodes of the new network similarly as those of D_{10}^S , but we store more information at each node. The crucial property on which we rely is well known (see for instance Bogart & West, 1999): an interval order is a semiorder if and only if it admits a representation by intervals such that no interval contains another interval unless they share an endpoint. We now state this property formally.

Proposition 5. *A relation R on the set S is a semiorder if and only if there exist two mappings $\ell, h : S \rightarrow \mathbb{R}$ such that, for all i, j in S :*

$$\ell(i) \leq h(i); \quad (23)$$

$$iRj \iff h(i) < \ell(j); \quad (24)$$

$$\ell(i) < \ell(j) \implies h(i) \leq h(j). \quad (25)$$

Proposition 5 can be easily deduced from Proposition 4 (direct proofs also exist, see for instance Bogart & West, 1999).

The network we build for semiorders has as nodes all triples (X, Y, L) with $Y \subseteq X \subseteq S$ and with L a linear order on $X \setminus Y$. (Notice that when $X \setminus Y = \emptyset$, we must have $L = \emptyset$.) The source and sink nodes are $s := (\emptyset, \emptyset, \emptyset)$ and $t := (S, S, \emptyset)$, respectively. To motivate the later specification of arcs, we show how each representation of a specific type of a semiorder generates a sequence of nodes that eventually becomes an s – t path. Given a representation ℓ, h as in Proposition 5 for a semiorder R , first adjust the values of ℓ, h to get a representation where ℓ, h are one-to-one and where $\ell(S) \cap h(S) = \emptyset$. Then, for any real number τ with $\min \ell(S) \leq \tau \leq \max h(S) + 1$, take the node obtained by setting $X := X(\tau)$, $Y := Y(\tau)$ as in Eqs. (21) and (22), and for $i, j \in X \setminus Y$ letting iLj exactly when $\ell(i) < \ell(j)$. An illustration is given in Fig. 2. Varying the value of τ , we thus generate a sequence of distinct nodes. We later specify the arcs so that each such sequence becomes an s – t path, and each s – t path is such a sequence. Furthermore, the resulting flow polytope provides an extended formulation of the semiorder polytope. The details are provided in Section 5.4. In particular, we show that the canonical description of this flow polytope grows faster in size than the extended formulations of our three other order polytopes. In the case of the semiorder polytope, we get an extended formulation of size $2^{\Theta(n \log n)}$, while for the other order polytopes we obtain extended formulations of size $2^{\Theta(n)}$.

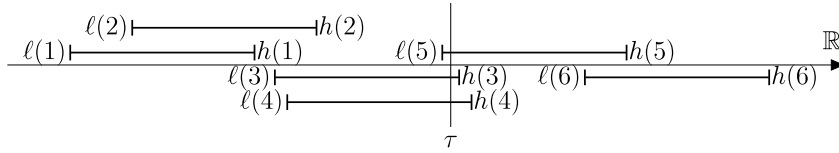


Fig. 2. Illustration for the network node coming from the level τ , in the case of interval orders: $X(\tau) = \{1, 2, 3, 4, 5\}$, $Y(\tau) = \{1, 2\}$ (cf. Eqs. (21) and (22)). In the case of semiordeers, we add the linear order $L = \{(3, 4), (4, 5), (3, 5)\}$, because $\ell(3) < \ell(4) < \ell(5)$.

5. Customizing extended formulations as network flow polytopes

In the next four subsections, we give detailed constructions and proofs of the results announced in Section 4.

5.1. The case of the linear order polytope

The network $D_{LO}^S = (N, A)$ we use for the case of linear orders of S is defined by

$$N := \{X \mid X \subseteq S\},$$

$$A := \{(X, Z) \in N \times N \mid X \subset Z, |Z| = |X| + 1\}$$

(in technical terms, the network D_{LO}^S is the “Hasse diagram” of the “Boolean lattice” of S). Thus, $|N| = 2^n$ and $|A| = n 2^{n-1}$. The source and sink nodes are $s := \emptyset$ and $t := S$, respectively. We now prove that the resulting flow polytope $F(D_{LO}^S)$ projects in a natural way onto the linear order polytope P_{LO}^S .

Theorem 1. Let π be the projection from $\mathbb{R}^A$ to $\mathbb{R}^{S \times S}$, mapping a point $\Phi \in \mathbb{R}^A$ to the point $x \in \mathbb{R}^{S \times S}$ specified for all distinct i, j in S by

$$x_{(i,j)} := \sum \{\Phi_a \mid a = (X, Z) \in A, i \in X \text{ and } j \in Z \setminus X\}. \quad (26)$$

The projection π maps the flow polytope $F(D_{LO}^S)$ onto the linear order polytope P_{LO}^S . As a consequence, the canonical description of the flow polytope $F(D_{LO}^S)$, as in Eq. (16), forms an extended formulation of P_{LO}^S with description size $n 2^{n-1}$.

Proof. It suffices to prove that the image of the vertex set of $F(D_{LO}^S)$ under the projection π is exactly the vertex set of P_{LO}^S .

First, consider a vertex v of P_{LO}^S . Then v is the characteristic vector of a linear order R on S . By considering any ordinal representation $u : S \rightarrow \mathbb{R}$ of R and the corresponding profile, we get an s – t path in D_{LO}^S , say P . The characteristic vector $\chi^{A(P)}$ of the set of arcs of this path is a vertex of $F(D_{LO}^S)$. We establish that $\pi(\chi^{A(P)}) = v$, i.e., $\Phi = \chi^{A(P)}$ in (26) yields $x_{(i,j)} = v_{(i,j)}$ for all distinct alternatives i and j of S .

We evaluate the right-hand side of (26) at $\Phi := \chi^{A(P)}$. The conditions $i \in X$ and $j \in Z \setminus X$ imply $u(i) < u(j)$, and so $i R j$. Moreover, they also imply $Z = X \cup \{j\}$, because here $X \subset Z$ and $|Z| = |X| + 1$. By definition of the characteristic vector $\Phi = \chi^{A(P)}$, the value Φ_a for any arc a is 0 or 1. From what we just proved, $\Phi_a = 1$ can occur in (26) only for the unique arc (X, Z) such that $Z = \{k \in S \mid u(k) \leq u(j)\}$ and $X = Z \setminus \{j\}$. Thus, we have $x_{(i,j)} = 1$ in (26) if and only if $i R j$ (and $x_{(i,j)} = 0$ otherwise). This establishes $\pi(\chi^{A(P)}) = v$.

Second, consider a vertex Φ of $F(D_{LO}^S)$. Then Φ is the characteristic vector of the set of arcs of an s – t path in D_{LO}^S . This path produces a profile that in turn represents a unique linear order R . Letting v denote the characteristic vector of R , we again have $\pi(\Phi) = v$.

The canonical description of the flow polytope $F(D_{LO}^S)$ has $n 2^{n-1}$ inequalities. $\square$

Remark 1. Recall that, for $n = 8$, the linear order polytope has at least 488, 602, 996 facets. In contrast, the size of the canonical description of the corresponding flow polytope F is only 1, 024. Even for $n = 6$ or $n = 7$, the minimal description of P_{LO}^S has size larger than the canonical description of the extended formulation.

5.2. The case of the weak order polytope

For the weak orders of S , we define the network $D_{WO}^S = (N, A)$ by

$$N := \{X \mid X \subseteq S\},$$

$$A := \{(X, Z) \in N \times N \mid X \subset Z\}$$

with the source and sink nodes $s := \emptyset$ and $t := S$, respectively (i.e., the “Boolean lattice” $(\mathcal{P}(S), \subset)$ of S with strict inclusion). Clearly, we have $|N| = 2^n$. Because each arc (X, Z) of D_{WO}^S can be encoded as a word of size n on the alphabet $\{a, b, c\}$ having at least one b (the letters a, b and c correspond to alternatives of $X, Z \setminus X$ and $S \setminus Z$, respectively), we see that $|A| = 3^n - 2^n$.

Theorem 2. Let π be the projection from $\mathbb{R}^A$ to $\mathbb{R}^{S \times S}$, mapping a point Φ from $\mathbb{R}^A$ to the point $x \in \mathbb{R}^{S \times S}$ given by Eq. (26). Then π maps the flow polytope $F(D_{WO}^S)$ to the weak order polytope P_{WO}^S . This yields the extended formulation $F(D_{WO}^S)$ of P_{WO}^S , whose canonical description has size $3^n - 2^n$.

Proof. The proof is parallel to that of Theorem 1, hence we provide only a sketch. Take any vertex v of the weak order polytope P_{WO}^S , with v encoding the weak order R . Then v is the image by π of a well defined vertex $\chi^{A(P)}$ of the flow polytope $F(D_{WO}^S)$. To obtain $\chi^{A(P)}$, take P to be the path in $(\mathcal{P}(S), \subset)$ resulting from the weak order R . Here, for a given weak order R and distinct alternatives i, j in X , the special pair (X, Z) in (26) consists of the smallest set Z in the profile of R that contains j , while X is just the set preceding Z in the profile. $\square$

Remark 2. Recall that, for $n = 5$, the weak order polytope P_{WO}^S has 75,834 facets (Regenwetter & Davis-Stober, 2012). In contrast, the size of the canonical description of the corresponding flow polytope $F(D_{WO}^S)$ is only 211.

5.3. The case of the interval order polytope

The network $D_{IO}^S = (N, A)$ for the interval order polytope of S is defined as follows:

$$N := \{(X, Y) \mid Y \subseteq X \subseteq S\},$$

$$A := \left\{ ((X, Y), (Z, T)) \in N \times N \mid \begin{array}{l} X \subseteq Z, Y \subseteq T, \text{ and} \\ \text{either } |Z| = |X| + 1, \\ \quad |T| = |Y| \\ \text{or } |Z| = |X|, \\ \quad |T| = |Y| + 1 \end{array} \right\}.$$

We choose the source and sink nodes to be $s := (\emptyset, \emptyset)$ and $t := (S, S)$, respectively.

Lemma 1. The numbers of nodes and arcs in the network D_{10}^S are

$$|N| = 3^n \quad \text{and} \quad |A| = 2n3^{n-1}, \quad (27)$$

respectively.

Proof. Let $S = \{s_1, s_2, \dots, s_n\}$. Consider a word $\omega = \omega_1\omega_2 \dots \omega_n$ of size n on the alphabet $\{a, b, c\}$. To each such word ω , we associate a pair of sets (X, Y) such that $Y \subseteq X \subseteq S$ by letting $Y = \{s_i \in S \mid \omega_i = a\}$ and $X = \{s_i \in S \mid \omega_i \in \{a, b\}\}$. We obtain a bijection from the set of words of size n on the alphabet $\{a, b, c\}$ to the node set of the network D_{10}^S . Thus, we have $|N| = 3^n$.

Now, consider a word $\omega = \omega_1\omega_2 \dots \omega_n$ of size n on the alphabet $\{a, b, c, d, e\}$ containing either exactly one d and no e , or exactly one e and no d . We associate an arc $((X, Y), (Z, T))$ of the network D_{10}^S to each such word, as follows. We let

$$X := \{s_i \in S \mid \omega_i \in \{a, b, e\}\},$$

$$Y := \{s_i \in S \mid \omega_i = a\},$$

$$Z := \{s_i \in S \mid \omega_i \in \{a, b, d, e\}\},$$

$$T := \{s_i \in S \mid \omega_i \in \{a, e\}\}.$$

This gives a one-to-one correspondence between the words of size n on the alphabet $\{a, b, c, d, e\}$ containing exactly one of the two letters d and e , and the arcs of D_{10}^S . We conclude that $|A| = 2n3^{n-1}$. $\square$

The last result of this section is as follows.

Theorem 3. Let π be the projection from $\mathbb{R}^A$ to $\mathbb{R}^{S \star S}$, mapping a point $\Phi \in \mathbb{R}^A$ to the point $x \in \mathbb{R}^{S \star S}$ given, for all $(i, j) \in S \star S$, by

$$x_{(i,j)} := \sum \{\Phi_a \mid a = ((X, Y), (Z, T)) \in A, i \in Y, j \in Z \setminus X\}. \quad (28)$$

Then π maps the flow polytope of D_{10}^S onto the interval order polytope of S . Thus, the canonical description of $F(D_{10}^S)$ is an extended formulation of P_{10}^S with description size $|A| = 2n3^{n-1}$.

Proof. The structure of the proof is similar to that of Theorem 1 (see however Remark 3 below. We show that the image of the vertex set of $F(D_{10}^S)$ under the projection π is exactly the vertex set of P_{10}^S .

Consider a vertex $v = x^R$ of P_{10}^S and a vertex $\Phi = x^{A(P)}$ of $F(D_{10}^S)$ such that Φ encodes the profile of some ordinal representation of the interval order R , say by injective mappings ℓ and h from S to $\mathbb{R}$ with $\ell(S) \cap h(S) = \emptyset$. It suffices to prove that $\pi(\Phi) = v$ in this case (notice that each vertex of $F(D_{10}^S)$ comes from some profile of some interval order).

Let i and j be two distinct alternatives of S . The right-hand side of (26) must give $Z = X \cup \{j\}$ and $Y = T$ (by the definition of the arcs in A). In other words, given the profile, the arc $a = ((X, Y), (Z, T))$ is completely determined by i and j . Moreover, such an arc can exist in $A(P)$ only if we have iRj . Hence, the right-hand side of (26) equals 1 if and only if iRj holds (and 0 otherwise), thus $x_{(i,j)} = v_{(i,j)}$. $\square$

Remark 3. In the proofs of Theorems 1 and 2, the projection of the flow polytope on the order polytope induces a bijective mapping between the vertex sets of the two polytopes. This is not the case here: in general, several vertices of the flow polytope are mapped onto the same vertex of P_{10}^S . The reason is that different representations of the same interval order can lead to distinct profiles, and thus to different paths in the network.

Remark 4. Recall that, for $n = 4$, the interval order polytope has 191 facets. This is less than the size of the canonical description of the flow polytope $F(D_{10}^S)$ with $|S| = 4$, which is 216. For larger values of n , we do not know the exact number of facets of the interval order polytope. However, a lower bound on this

number follows from Doignon and Rexhep (2016, Theorem 8 and Corollary 1): the interval order polytope $F(D_{10}^S)$ has at least as many facets as there are ‘PC-graphs’ on S , where a PC-graph is a directed graph in which any node is the tail of at most one arc, and the head of at most one arc (in other words, a PC-graph is a node-disjoint union of paths and cycles plus maybe isolated nodes). For small numbers of alternatives, we provide in Table 1 both the resulting lower bound on the size of any linear description of the interval order polytope P_{10}^S , and the exact size of the canonical description of the flow polytope $F(D_{10}^S)$. The flow polytope definitely has a smaller description size for $n = 6$ (the same could hold for $n = 5$). It is not difficult to see that the same assertion holds for $n \geq 7$ (see Table 1 for a numerical comparison).

5.4. The case of the semiorder polytope

The network $D_{SO}^S = (N, A)$ we build for the semiorder polytope of S is more structured than the one for the interval order polytope of S . The nodes of D_{SO}^S are of the form (X, Y, L) where X and Y are subsets of S such that $Y \subseteq X$, and L is a linear order on $X \setminus Y$ (recall that when $X \setminus Y = \emptyset$ we have $L = \emptyset$). Thus,

$$N = \{(X, Y, L) \mid Y \subseteq X \subseteq S, L \text{ linear order on } X \setminus Y\}.$$

We now indicate when the pair $((X, Y, L), (Z, T, M))$ of nodes from D_{SO}^S forms an arc. There are two cases that we illustrate in Fig. 2 in terms of levels in representations: (α) the threshold moves from $\tau = \ell(5)$ to $\tau = h(3)$; (β) it moves from $\tau = h(3)$ to $\tau = h(4)$. Formally, the two cases are:

(α) for some alternative i in $S \setminus X$:

$$Z = X \cup \{i\},$$

$$T = Y,$$

$$M = L + i,$$

where $L + i$ means that we append i at the end of the linear order L ;

(β) for the alternative j in $X \setminus Y$ which is the first one in L :

$$Z = X,$$

$$T = Y \cup \{j\},$$

$$M = L - j,$$

where $L - j$ means the linear order induced by L on $Z \setminus T = (X \setminus Y) \setminus \{j\}$.

The sink and source nodes are again $s := (\emptyset, \emptyset, \emptyset)$ and $t := (S, S, \emptyset)$, respectively.

Lemma 2. The numbers of nodes and arcs in the network D_{SO}^S are

$$|N| = \sum_{t=0}^n n(n-1) \dots (n-t+1) 2^{n-t}, \quad (29)$$

$$|A| = \sum_{t=0}^n n(n-1) \dots (n-t+1) 2^{n-t-1} (n+t), \quad (30)$$

respectively, and they satisfy

$$n! \leq |N| \leq e^2 n!, \quad n! \leq |A| \leq e^2 n n!. \quad (31)$$

Proof. To obtain a node (X, Y, L) of D_{SO}^S with $t := |X \setminus Y|$, we select first an ordered list of t elements from S . Next we select some subset Y among the $n - t$ remaining elements, and finally we set X equal to Y augmented by the listed elements.

Table 1

Comparison of description sizes in the case of interval orders.

n	Size of minimal description of P_{10}^S	Size of canonical description of $F(D_{10}^S)$
3	17	54
4	191	216
5	≥ 759	810
6	≥ 5557	2,916
7	$\geq 63,839$	10,206

To count the arcs, notice that the number of arcs leaving a node (X, Y, L) equals $n - |Y|$. Assume again $t = |X \setminus Y|$, and let $\ell := |Y|$. Then

$$|A| = \sum_{t=0}^n \left(n(n-1) \dots (n-t+1) \sum_{\ell=0}^{n-t} \binom{n-t}{\ell} (n-\ell) \right). \quad (32)$$

Now

$$\sum_{\ell=0}^{n-t} \binom{n-t}{\ell} (n-\ell) = n 2^{n-t} - (n-t) 2^{n-t-1} = 2^{n-t-1} (n+t). \quad (33)$$

The first bound holds because the last term in Eq. (29) is $n!$. On the other hand, rewriting Eq. (29) as

$$|N| = n! \sum_{t=0}^n \frac{2^{n-t}}{(n-t)!}, \quad (34)$$

we get $|N| \leq n! e^2$. Because the number of arcs leaving any node is at most n , the upper bound on $|A|$ follows from the one on $|N|$. $\square$

Theorem 4. Let π be the projection from $\mathbb{R}^A$ to $\mathbb{R}^{S \times S}$, mapping a point $\Phi \in \mathbb{R}^A$ to the point $x \in \mathbb{R}^{S \times S}$ given for all i, j in S by

$$x_{(i,j)} := \sum_{j \in Z \setminus X} \{\Phi_a \mid a = ((X, Y, L), (Z, T, M)) \in A, i \in Y, j \in Z \setminus X\}. \quad (35)$$

Then π maps the flow polytope of $D_{S_0}^S$ to the semiorder polytope of S . Thus, the flow polytope $F(D_{S_0}^S)$ is an extended formulation of the semiorder polytope $P_{S_0}^S$, with size $|A| \leq e^2 n n!$.

Proof. Once again, we show that π maps the set of vertices of the flow polytope $F(D_{S_0}^S)$ onto the set of vertices of the order polytope, here the semiorder polytope $P_{S_0}^S$ of S . Each s - t path in $D_{S_0}^S$ produces a representation of some interval order R , exactly as in the proof of Theorem 3. Next, taking into account the ordering L in any network node (X, Y, L) , we see that the representation thus obtained has no interval strictly including another one. By Proposition 5, R is then a semiorder. The rest of the proof is similar to the proof of Theorem 3. $\square$

Remark 5. Recall that, for $n = 4$, the semiorder polytope has 563 facets. The description size of the flow polytope $F(D_{S_0}^S)$ with $|S| = 4$ is 520, which is less than the number of facets of the semiorder polytope. It is difficult to prove that this holds for all values of $n \geq 4$, because only a very limited number of facets of semiorder polytopes are known. However we can prove that there exists a natural number n_0 such that, for all $n > n_0$, the description size of the flow polytope $F(D_{S_0}^S)$ will be smaller than the number of facets of the polytope $P_{S_0}^S$. This is due to the fact that there are more facet defining inequalities for $P_{S_0}^S$ than PC-graphs on n elements (Doignon & Rexhep, 2016, Section 9), and more PC-graphs than chain gangs, the latter being the disjoint unions of paths plus maybe isolated nodes. Indeed, the number of chain gangs on n elements equals $n! \sum_{j=0}^{n-1} \binom{n-1}{j} / (j+1)!$ (Sloane, 2016, Sequence A000262), and so it becomes eventually larger¹ than

$e^2 n n!$ for some $n = n_0$, and remains so for all n larger than n_0 . Even if its description size happens to be larger for values of n less than n_0 , our extended formulation has always the advantage that its canonical description is available in a very simple form.

6. Interface between flow polytopes and statistics

In this section, we briefly consider how our network flow polytopes interface with contemporary statistical methods for evaluating random preference models. Typical choice data used in evaluating random preference models are comprised of decision maker responses to a series of repeated paired presentations of choice alternatives (see Regenwetter et al., 2011). We consider the case when pairs of options are offered and the decision maker must choose one (two-alternatives forced choice) or must either choose one or indicate indifference (ternary paired comparisons). Let $p_{(i,j)}$ and $C_{(i,j)}$ denote the probability and the number of times that a decision maker chooses alternative j when offered alternatives i and j . Let $\mathbf{p} := (p_{(i,j)})_{(i,j) \in S \times S}$ and let $\mathbf{C} := (C_{(i,j)})_{(i,j) \in S \times S}$. Let $L(\mathbf{p}|\mathbf{C})$ denote the likelihood of any given data $\mathbf{C}$ as a function of probabilities $\mathbf{p}$, (for examples involving multinomial distributions see e.g., Davis-Stober, 2009; Myung et al., 2005; Regenwetter et al., 2011). In general, in a random preference model, the maximum likelihood estimate does not have a closed form solution and therefore its computation requires some elaborate optimization methods. The maximum likelihood estimate (MLE) for a given random preference model $P_{\mathcal{M}}^S$ that is one of $P_{L_0}^S$, $P_{W_0}^S$, $P_{S_0}^S$, or $P_{I_0}^S$, given choice data $\mathbf{C}$, equals

$$\hat{\mathbf{p}}_{\mathcal{M}}^{\mathbf{C}} := \operatorname{argmax}_{\mathbf{p} \in P_{\mathcal{M}}^S} L(\mathbf{p}|\mathbf{C}). \quad (36)$$

In a network flow representation, the MLE for $P_{\mathcal{M}}^S$ can be calculated via the following optimization program:

$$\begin{aligned} \min_{\mathbf{p} \in \mathbb{R}^{S \times S}, \Phi \in \mathbb{R}^{|A|}} & -\ln L(\mathbf{p}|\mathbf{C}) \\ \text{s.t.} & \Phi \in F(D_{\mathcal{M}}^S) \\ & \mathbf{p} = \pi(\Phi), \end{aligned} \quad (\text{MLE-CP})$$

where $D_{\mathcal{M}}^S$ is one of $D_{L_0}^S$, $D_{W_0}^S$, $D_{S_0}^S$, or $D_{I_0}^S$, accordingly. This is a convex optimization problem because we minimize the negative of the log-likelihood function, which in this case is a convex function (see Davis-Stober, 2009), over a convex feasible region.

Indeed, the two vectors of variables $\mathbf{p}$ and Φ are only constrained by linear equations and inequalities, coming both from the canonical description (16) of the flow polytope $F(D_{\mathcal{M}}^S)$ and the expression that $\mathbf{p}$ is the projection of Φ by π (i.e. its image by an affine mapping).

The solution of this convex program yields the MLE as well as the maximized likelihood value. This convex program can be solved computationally by standard methods in convex optimization such as polynomial-time interior-point methods, see Nesterov and Nemirovskii (1994). Alternatively, the vector of variables $\mathbf{p}$ can be eliminated from the formulation using the projection equalities, leading to

$$\begin{aligned} \min_{\Phi \in \mathbb{R}^{|A|}} & -\ln L(\pi(\Phi)|\mathbf{C}) \\ \text{s.t.} & \Phi \in F(D_{\mathcal{M}}^S). \end{aligned} \quad (\text{MLE-CP}')$$

¹ Computations suggest this happens with $n_0 = 21$.

This is now a convex nonlinear network optimization problem, see e.g. Bertsekas (1998) (note that while the objective function remains convex, it is no longer separable—that is, the function cannot be expressed as the sum of functions of the individual variables).

The maximum likelihood estimate $\hat{p}_{\mathcal{M}}^c$ is obtained by projecting the optimum flow $\hat{\Phi}$ (which is not necessarily unique).

The Bayes factor is a standard method for assessing the relative empirical evidence for/against either of two competing models. It is defined as the ratio of their respective marginal likelihoods (Kass & Raftery, 1995). Following previous approaches in testing random preference models (Davis-Stober et al., 2015), we compare the model $P_{\mathcal{M}}^S$ to an “encompassing” model that places no restrictions on choice probabilities. Assuming suitably chosen prior distributions, the Bayes factor can be re-written as the ratio of two proportions: the proportion of the encompassing prior in agreement with $P_{\mathcal{M}}^S$ and the proportion of the encompassing posterior in agreement with $P_{\mathcal{M}}^S$ (see Klugkist & Hoijtink, 2007). We estimate these proportions by repeatedly drawing from the prior and posterior distributions and counting the samples that satisfy $P_{\mathcal{M}}^S$. Checking whether or not a sampled value p_s of p satisfies a given random preference model, using the extended formulation via a network flow, reduces to solving the following convex optimization program:

$$\begin{aligned} \min_p \quad & \|p - p_s\| \\ \text{s.t.} \quad & \Phi \in F(D_{\mathcal{M}}^S) \\ & p = \pi(\Phi), \end{aligned} \quad (\text{Bayes-CP})$$

where $D_{\mathcal{M}}^S$ is one of D_{LO}^S , D_{WO}^S , D_{SO}^S , or D_{IO}^S , which can be solved in the same ways as (MLE-CP). If the optimal value of p equals p_s , then the sampled point lies inside the model, otherwise it does not. See Davis-Stober et al. (2015) for additional details on using sampling methods to calculate Bayes factors under the weak order polytope.

7. Optimality

Here we discuss the asymptotic optimality of our extended formulations in terms of size. It will be convenient to use the notion of *extension complexity* of a polytope P , defined as the minimum description size of an extended formulation of P . Notice that the extension complexity is affinely invariant, that is, two affinely equivalent polytopes P and Q have the same extension complexity, and that it is monotone in the sense that the extension complexity of a polytope P is at least that of any of its faces F .

Maksimenko (2017) recently proved that the *correlation polytope*

$$P_{\text{COR}}^n := \text{conv}\{xx^T \mid x \in \{0, 1\}^n\} \quad (37)$$

is affinely equivalent to a face of the linear ordering polytope P_{LO}^{2n} (the linear ordering polytope for a set of $2n$ alternatives). This implies that the extension complexity of P_{LO}^{2n} is at least that of P_{COR}^n , which in turn was proved to be at least $(3/2)^n$, see Fiorini, Massar, Pokutta, Tiwary, and Wolf (2015) and Kaibel and Weltge (2014). Therefore, the extension complexity of P_{LO}^{2n} is at least $(3/2)^{\lfloor n/2 \rfloor}$. Moreover, since P_{WO}^n , P_{IO}^n and P_{SO}^n all have P_{LO}^n as a face, we see that their extension complexities are at least that of P_{LO}^n . Since the extended formulations constructed in this paper for P_{LO}^n , P_{WO}^n and P_{IO}^n have size $2^{\Theta(n)}$, this implies that the three corresponding extension complexities are also $2^{\Theta(n)}$, i.e. that these extended formulations are asymptotically optimal (up to exponentiation).

For the semiorder polytope, we only know that the extension complexity of P_{SO}^n is $2^{\Omega(n)}$ and at the same time $2^{O(n \log n)}$. We remark that it is possible that the semiorder polytope admits an extended formulation with size $2^{\Theta(n)}$ but no such extended formulation based on flows. Indeed, flow-based extended formulations

are known to have strong limitations, see Fiorini and Pashkovich (2015).

8. Conclusions and discussion

In this paper, we leverage extended formulations and network flow polytopes to work with random preference models for linear orders, weak orders, semiorders, and interval orders. Our results break through previous barriers in the number of choice alternatives for which parsimonious linear descriptions are known. One fundamental reason is that we provide a complete, linear description of the extended formulation (a polytope which projects on the polytope of choice probabilities permitted by the model). A second reason is that in many cases the extended formulation description is of smaller size than that of the initial polytope. However, in the case of semiorders, the extended formulation relies more heavily on the numerical representations of the relations and consequently entails an excessively large description. One interesting open question is to find a more parsimonious extended formulation in the case of semiorders, or prove that none exists.

A natural application of the extended formulations we have described is to expand empirical studies of fundamental structures (e.g., weakly ordered preferences) for larger numbers of choice alternatives than previously possible (Regenwetter & Davis-Stober, 2012). We leave this for future work.

Acknowledgments

This material is based upon work supported by the Air Force Office of Scientific Research under Award No. FA9550-05-1-0356 (M. Regenwetter, PI), by the National Institutes of Health under Award 774K25AA024182 (C. Davis-Stober, PI), by the National Institute of Mental Health under Award No. PHS 2 T32 MH014257 (M. Regenwetter, PI), by the National Science Foundation under awards SES-1062045, SES-1459699 (M. Regenwetter, PI), and SES-1459866 (C. Davis-Stober, PI). Jean-Paul Doignon, Samuel Fiorini and François Glineur were supported by *Action de Recherche Concertée* grants of the *Fédération Wallonie-Bruxelles* (Belgium). Samuel Fiorini was also supported by ERC Consolidator Grant 615640-ForEFront, European Research Council, Europe. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of their funding agencies or universities.

References

- Armstrong, W. (1939). The determinateness of the utility function. *Economics Journal*, 49, 453–467.
- Bertsekas, D. (1998). *Network optimization: Continuous and discrete models*. Belmont, Massachusetts: Athena Scientific Belmont.
- Bogart, K., & West, D. (1999). A short proof that “proper = unit”. *Discrete Mathematics*, 201, 21–23.
- Boyd, S., & Puleyblank, W. (2009). Facet generating techniques. In *Research trends in combinatorial optimization* (pp. 33–55). Berlin, Germany: Springer.
- Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology*, 57, 153–178.
- Christof, T., & Reinelt, G. (1996). Combinatorial optimization and small polytopes. *TOP*, 4, 1–64.
- Conforti, M., Cornuéjols, G., & Zambelli, G. (2010). Extended formulations in combinatorial optimization. *4OR. A Quarterly Journal of Operations Research*, 8, 1–48.
- Cormen, T., Leiserson, C., Rivest, R., & Stein, C. (2003). *Introduction to algorithms* (2nd ed.). Cambridge, Massachusetts: The M.I.T. Press.
- Davis-Stober, C. P. (2009). Analysis of multinomial models under inequality constraints: Applications to measurement theory. *Journal of Mathematical Psychology*, 53, 1–13.
- Davis-Stober, C. P., & Brown, N. (2011). A shift in strategy or “error”? Strategy classification over multiple stochastic specifications. *Judgment and Decision Making*, 6, 800–813.
- Davis-Stober, C. P., Brown, N., & Cavagnaro, D. R. (2015). Individual differences in the algebraic structure of preferences. *Journal of Mathematical Psychology*, 66, 70–82.

- Doignon, J., & Fiorini, S. (2002). Facets of the weak order polytope derived from the induced partition projection. *SIAM Journal of Discrete Mathematics*, 15, 112–121.
- Doignon, J.-P., Fiorini, S., & Joret, G. (2006). Facets of the linear ordering polytope: a unification for the fence family through weighted graphs. *Journal of Mathematical Psychology*, 50, 251–262.
- Doignon, J.-P., Fiorini, S., & Joret, G. (2009). Weighted graphs defining facets: a connection between stable set and linear ordering polytopes. *Discrete Optimization*, 6, 1–9.
- Doignon, J.-P., & Rexhep, S. (2016). Primary facets of order polytopes. *Journal of Mathematical Psychology*, 75, 231–245.
- Ducamp, A., & Falmagne, J. C. (1969). Composite measurement. *Journal of Mathematical Psychology*, 6, 359–390.
- Erdfelder, E., Auer, T.-S., Hilbig, B. E., Aßfalg, A., Moshagen, M., & Nadarevic, L. (2009). Multinomial processing tree models: A review of the literature. *Zeitschrift für*, 217, 108–124.
- Falmagne, J.-C. (1978). A representation theorem for finite random scale systems. *Journal of Mathematical Psychology*, 18, 52–72.
- Fiorini, S. (2001). *Polyhedral combinatorics of order polytopes* (Ph.D. thesis), Brussels, Belgium: Université Libre de Bruxelles, Département de Mathématiques.
- Fiorini, S. (2004). A short proof of a theorem of Falmagne. *Journal of Mathematical Psychology*, 48, 80–82.
- Fiorini, S. (2006a). $\{0, 1/2\}$ -cuts and the linear ordering problem: Surfaces that define facets. *SIAM Journal of Discrete Mathematics*, 20, 893–912.
- Fiorini, S. (2006b). How to recycle your facets. *Discrete Optimization*, 3, 136–153.
- Fiorini, S., & Fishburn, P. C. (2004). Weak order polytopes. *Discrete Mathematics*, 275, 111–127.
- Fiorini, S., Massar, S., Pokutta, S., Tiwary, H., & Wolf, R. d. (2015). Exponential lower bounds for polytopes in combinatorial optimization. *Journal of the Association for Computing Machinery*, 62, 17.1–17.23.
- Fiorini, S., & Pashkovich, K. (2015). Uncapacitated flow-based extended formulations. *Mathematical Programming*, 153, 117–131.
- Fishburn, P. (1970). Intransitive indifference with unequal indifference intervals. *Journal of Mathematical Psychology*, 7, 144–149.
- Fishburn, P. C. (1985). *Interval orders and interval graphs*. New York, New York: Wiley.
- Garey, M., & Johnson, D. (1979). *Computers and intractability. A guide to the theory of NP-completeness*. New York, New York: Freeman.
- Kaibel, V. (2011). Extended formulations in combinatorial optimization. *Optima*, 85, 2–7.
- Kaibel, V., & Weltge, S. (2014). A short proof that the extension complexity of the correlation polytope grows exponentially. *Discrete & Computational Geometry*, 53, 397–401.
- Karabatsos, G. (2006). Bayesian nonparametric model selection and model testing. *Journal of Mathematical Psychology*, 50, 123–148.
- Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90, 773–795.
- Klugkist, I., & Hoijtink, H. (2007). The Bayes factor for inequality and about equality constrained models. *Computational Statistics & Data Analysis*, 51, 6367–6379.
- Korte, B., & Vygen, J. (2008). *Combinatorial optimization*. Berlin, Germany: Springer.
- Luce, R. D. (1956). Semiorders and a theory of utility discrimination. *Econometrica*, 26, 178–191.
- Luce, R. D. (2000). *Utility of gains and losses: Measurement-theoretical and experimental approaches*. Mahwah, New Jersey: Lawrence Erlbaum Association, Psychology Press.
- Maksimenko, A. N. (2017). A Boolean quadratic polytope is the face of a linear-order polytope. *Sibirskie Elektronnye Matematicheskie Izvestiya*, 14, 640–646 (in Russian).
- Martí, R., & Reinelt, G. (2011). *The linear ordering problem: Exact and heuristic methods in combinatorial optimization*. Applied Mathematical Sciences: vol. 175. Berlin, Germany: Springer.
- Mirkin, B. G. (1972). Description of some relations on the set of real-line intervals. *Journal of Mathematical Psychology*, 9, 243–252.
- Müller, R., & Schulz, A. (1995). The interval order polytope of a digraph. In E. Balas, & J. Clausen (Eds.), *Integer programming and combinatorial optimization. Lecture Notes in Computer Science*: vol. 920 (pp. 50–64). Proceedings of the 4th International IPCO Conference.
- Myung, J. I., Karabatsos, G., & Iverson, G. J. (2005). A Bayesian approach to testing decision making axioms. *Journal of Mathematical Psychology*, 49, 205–225.
- Nesterov, Y., & Nemirovskii, A. (1994). *Interior-point polynomial algorithms in convex programming*. SIAM Studies in Applied Mathematics: vol. 13. Philadelphia, Pennsylvania: Society for Industrial and Applied Mathematics.
- Oswald, M., Reinelt, G., & Seitz, H. (2009). Applying mod-k-cuts for solving linear ordering problems. *TOP*, 17, 158–170.
- Pirlot, M., & Vincke, P. (1997). *Semiorders: Properties, representations, applications*. Dordrecht, The Netherlands: Kluwer.
- Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, 9, 347–356.
- Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, 111, 333–367.
- Regenwetter, M., Dana, J., & Davis-Stober, C. P. (2010). Testing transitivity of preferences on two-alternative forced choice data. *Frontiers in Quantitative Psychology and Measurement*, 1.
- Regenwetter, M., Dana, J., & Davis-Stober, C. P. (2011). Transitivity of preferences. *Psychological Review*, 118, 42–56.
- Regenwetter, M., Dana, J., Davis-Stober, C. P., & Guo, Y. (2011). Parsimonious testing of transitive or intransitive preferences: Reply to Birnbaum (2011). *Psychological Review*, 118, 684–688.
- Regenwetter, M., & Davis-Stober, C. (2011). Ternary paired comparisons induced by semi- or interval order preferences. In E. Dzhaferov, & L. Perry (Eds.), *Descriptive and normative approaches to human behavior. Advanced Series on Mathematical Psychology*: vol. 3 (pp. 225–248). Singapore: World Scientific.
- Regenwetter, M., & Davis-Stober, C. (2012). Behavioral variability of choices versus structural inconsistency of preferences. *Psychological Review*, 119, 408–416.
- Regenwetter, M., & Marley, A. A. J. (2001). Random relations, random utilities, and random functions. *Journal of Mathematical Psychology*, 45, 864–912.
- Roberts, F. S. (1979). *Measurement theory*. London, United Kingdom: Addison-Wesley.
- Scott, D., & Suppes, P. (1958). Foundational aspects of theories of measurement. *Journal of Symbolic Logic*, 23, 113–128.
- Silvapulle, M., & Sen, P. (2005). *Constrained statistical inference. Wiley Series in Probability and Statistics*. Hoboken, New Jersey: Wiley-Interscience [John Wiley & Sons].
- Sloane, N. (2016). The On-Line Encyclopedia of Integer Sequences. Published electronically at <https://oeis.org/>.
- Suck, R. (1995). Random utility representations based on semiorders, interval orders, and partial orders. Unpublished manuscript.
- Trueblood, J. S., Brown, S. D., & Heathcote, A. (2014). The multiattribute linear ballistic accumulator model of context effects in multialternative choice. *Psychological Review*, 121, 179.
- Tversky, A. (1969). Intransitivity of preferences. *Psychological Review*, 76, 31–48.
- Wiener, N. (1915). Studies in synthetic logic. *Proceeding of the Cambridge Philosophical Society*, 18, 14–28.
- Wolsey, L. A. (2011). Using extended formulations in practice. *Optima*, 85, 7–9.
- Ziegler, G. (1998). *Lectures on polytopes* (Revised ed.). Berlin, Germany: Springer-Verlag.