# A graphical study of comparative probabilities

Alexander Erreygers[a,*], Enrique Miranda[b,1]

[a] *Ghent University, Belgium*
[b] *University of Oviedo, Spain*

**Abstract**

We consider a set of comparative probability judgements over a finite possibility space and study the structure of the set of probability measures that are compatible with them. We relate the existence of some compatible probability measure to Walley's behavioural theory of imprecise probabilities, and introduce a graphical representation that allows us to bound, and in some cases determine, the extreme points of the set of compatible measures. In doing this, we generalise some earlier work by Miranda and Destercke on elementary comparisons.

*Keywords:* Comparative probabilities, credal sets, lower previsions, sets of desirable gambles, extreme points

## 1. Introduction

The elicitation of probability measures, which is particularly relevant in the context of subjective probability, can be cumbersome in situations of imprecise or ambiguous information. These can arise due to the presence of missing data or contradictory sources of information, or due to the limitations of the measurement devices. To deal with this problem, a number of approaches have been put forward in the literature: (i) we may work with sets of probability measures, or *credal sets* [30]; (ii) we can provide lower and/or upper bounds for the 'true' probability measure, representing the information in terms of non-additive measures [11, 16, 44]; or (iii) we could model the information in terms of its behavioural implications [48]. The different models are sometimes referred to with the common term *imprecise probabilities* [2].

When the available information comes from expert judgements, it may be easier to model it in terms of comparative assessments of the form 'event $A$ is at least as probable as event $B$.' This leads to *comparative probabilities*, that were studied first by de Finetti [10] and later by other authors such as Koopman [28], Good [24] or Savage [41]; see also [8, 20, 26, 29, 46, 49] for some relevant subsequent work. For a recent thorough overview, as well as an extensive philosophical justification and a summary of the most important results, we refer to [27].

In this paper, we consider a collection of comparative probability judgements over a finite possibility space and study the structure of the set of compatible probability measures. Specifically, we shall investigate in which cases this set is non-empty, the number of its extreme points and their features, and the properties of its associated lower probability. While most earlier work on comparative probabilities has mainly focused on the complete case—that is, the case where any two events are compared—ours is not the first study of the incomplete one; in this respect, the most influential works for this paper are those of Walley [48, Section 4.5] and Miranda and Destercke [32]. Our present contribution has the same goals as that of Miranda and Destercke, but our setting is more general: where they exclusively focused on the specific case of comparisons between elementary events, we generalise some of their results to the case of comparisons between arbitrary events.

The paper is organised as follows. We start with a formal introduction of comparative assessments in Section 2, and subsequently discuss the compatibility problem and show that it can be easily tackled using Walley's theory of lower previsions. From Section 3 on, we study the set of extreme points of the associated credal set. To this end, we introduce a graphical representation in Section 4; this representation allows us to determine the number of extreme points in a number of special cases in Section 5, where we also argue that this approach cannot be extended to the general case. We conclude in Section 7 with some additional comments and remarks. In order to ease the reading, proofs have been gathered in an appendix.

We have already reported the results of this contribution in [18]. Two key differences between the present contribution and our previous presentation are that (i) we give a bit more background in Sections 2.1 and 6, and (ii) we provide formal proofs for our results.

## 2. Comparative assessments and compatibility

Consider a finite possibility space $\mathscr{X}$ with cardinality $n$, and a (finite) number $m$ of comparative judgements of the form 'the event $A$ is at least as likely as the event $B$.' For ease of notation, we will represent the $i$-th judgement as a

*Corresponding author
[1] Enrique Miranda's research was supported by project PGC2018-098623-B-I00.

pair $(A_i, B_i)$ of events—that is, subsets of the possibility space $\mathscr{X}$. We collect all $m$ judgements in the comparative assessment

$$\mathscr{C} \coloneqq \{(A_i, B_i) \colon i \in \{1, \dots, m\}, A_i, B_i \subseteq \mathscr{X}\}.$$

Equivalently, the comparative judgements can be represented in terms of a (possibly partial) binary relation $\succeq$ on $2^{\mathscr{X}}$, the power set of the possibility space $\mathscr{X}$, with $A \succeq B$ being equivalent to $(A, B) \in \mathscr{C}$. Miranda and Destercke [32] exclusively dealt with comparative assessments that concern singletons, or equivalently, are a subset of $\{(\{x\}, \{y\}) \colon x, y \in \mathscr{X}\}$. We follow them in calling such comparative assessments *elementary*.

Throughout this contribution we will use a running example to illustrate much of the introduced concepts.

**Running example.** Let $\mathscr{X} \coloneqq \{1, 2, 3, 4\}$ and

$$\mathscr{C} \coloneqq \{(\{1\}, \{2\}), (\{1, 2\}, \{3\}), (\{1, 3\}, \{4\}), (\{1, 2\}, \{4\})\}.$$

Clearly, the corresponding partial binary relation $\succeq$ is given by $\{1\} \succeq \{2\}$, $\{1, 2\} \succeq \{3\}$, $\{1, 3\} \succeq \{4\}$ and $\{1, 2\} \succeq \{4\}$.

Let $\Sigma_{\mathscr{X}}$ denote the set of all probability mass functions on $\mathscr{X}$. We follow the authors of [27, 32, 39, 48] in considering the set

$$\mathscr{M}_{\mathscr{C}} \coloneqq \Big\{p \in \Sigma_{\mathscr{X}} \colon (\forall (A, B) \in \mathscr{C}) \sum_{x \in A} p(x) \geq \sum_{x \in B} p(x)\Big\} \quad (1)$$

of all probability mass functions that are compatible with the comparative judgements. Following Levi [30], we call $\mathscr{M}_{\mathscr{C}}$ the comparative *credal set*.

Given a set $\mathscr{C}$ of comparative judgements, we should first of all determine whether or not there is at least one compatible probability measure—that is, if the comparative credal set $\mathscr{M}_{\mathscr{C}}$ is non-empty. If that is the case, we shall call the comparative assessments *compatible*. In the case of elementary judgements [32], this is trivial because the uniform probability distribution is compatible with any elementary comparative assessment. Unfortunately, when more elaborate judgements are allowed this is no longer the case, as is demonstrated by the following example.

**Example 1.** Consider the possibility space $\mathscr{X} \coloneqq \{1, 2, 3\}$ and the comparative assessment

$$\mathscr{C} \coloneqq \{(\{1\}, \{2, 3\}), (\{2\}, \{1, 3\}), (\{3\}, \{1, 2\})\}.$$

It follows immediately from these judgements that any compatible probability mass function $p$ should satisfy $p(1) \geq 1/2$, $p(2) \geq 1/2$ and $p(3) \geq 1/2$. However, this is clearly impossible, so $\mathscr{M}_{\mathscr{C}} = \emptyset$.

The problem of the existence of a compatible probability measure with a number of comparative assessments on events has been considered quite extensively in the literature [20, 29]. In what follows, we approach this problem by means of the behavioural theory of sets of desirable gambles, for which we give next a succinct introduction.

### *2.1. Connection with sets of desirable gambles*

The existence of a compatible probability measure was characterised in [42, Theorem 4.1] in the case of complete comparative assessments and in [40, Proposition 4] and [39, Section 2] in the case of partial comparative assessments; see also [48, Section 4.5.2]. In this section, we use Walley's result to establish a connection with the theory of *sets of almost-desirable gambles*[2], from which we shall derive a number of additional results. We refer to [48] for a detailed account of the theory.

A *gamble* $f$ is a real-valued map on our finite possibility space $\mathscr{X}$. The set of all gambles on $\mathscr{X}$ is denoted by $\mathcal{L}$, and dominance between gambles is understood pointwise. Within $\mathcal{L}$ we may consider the subset

$$\mathcal{L}^+ \coloneqq \{f \in \mathcal{L} : f \geq 0, f \neq 0\}$$

of non-negative gambles, that in particular includes the *indicator* $\mathbb{I}_A$ of some event $A \in 2^{\mathscr{X}}$, taking value 1 on $A$ and 0 elsewhere. In the context of this paper, non-negative gambles will allow us to encompass trivial comparisons of the type $A \succeq \emptyset$.

It is often convenient to think of a gamble $f$ as an uncertain reward expressed in units of some linear utility scale: in case the outcome of our experiment is $x$, our subject receives the—possibly negative—pay-off $f(x)$. With this interpretation, our subject can specify a set of *almost desirable gambles* $\mathcal{K}$, being some set of gambles—or uncertain rewards—that she considers acceptable. Such a set $\mathcal{K}$ of almost desirable gambles can be extended to include gambles that are implied by rational behaviour; the least-committal of these extensions is the *natural extension* of $\mathcal{K}$, which is defined as $\mathcal{D}_{\mathcal{K}} \coloneqq \overline{\operatorname{posi}(\mathcal{K} \cup \mathcal{L}^+)}$, where we consider the topological closure under the supremum norm and the posi operator is defined for any set of gambles $\mathcal{K}' \subseteq \mathcal{L}$ as

$$\operatorname{posi}(\mathcal{K}') \coloneqq \left\{\sum_{i=1}^{k} \lambda_i f_i \colon k \in \mathbb{N}, \lambda_i > 0, f_i \in \mathcal{K}'\right\},$$

with $\mathbb{N}$ the set of natural numbers—that is, not including zero. We say that a set of almost desirable gambles $\mathcal{K}$ *avoids sure loss* if and only if $\max f \geq 0$ for all $f \in \mathcal{D}_{\mathcal{K}}$, and that it is *coherent* whenever $\mathcal{K} = \mathcal{D}_{\mathcal{K}}$. It turns out that $\mathcal{D}_{\mathcal{K}}$ is coherent if and only if $\mathcal{K}$ avoids sure loss, and that $\mathcal{K}$ avoids sure loss if and only if there exists a probability mass function $p$ such that $\sum_{x \in \mathscr{X}} f(x)p(x) \geq 0$ for every $f \in \mathcal{K}$. As a consequence, the compatibility of $\mathscr{C}$ is equivalent to verifying that

$$\mathcal{K}_{\mathscr{C}} \coloneqq \{\mathbb{I}_A - \mathbb{I}_B \colon (A, B) \in \mathscr{C}\} \quad (2)$$

avoids sure loss, which immediately leads to the following proposition—see also [40, Proposition 4], [39, Section 2] or [48, Lemma 3.3.2].

[2] The reader should not get confused with the more general theory of sets of (really) desirable gambles, that is more informative than credal sets; see [2, Chapter 1] for details.

**Proposition 1.** *The comparative credal set $\mathscr{M}_{\mathscr{C}}$ is non-empty if and only if for all $\lambda_1, \ldots, \lambda_m$ in $\mathbb{N} \cup \{0\}$,*

$$\max \sum_{i=1}^{m} \lambda_i (\mathbb{I}_{A_i} - \mathbb{I}_{B_i}) \geq 0.$$

It is well-known—see for example [48, Section 3.8]—that a coherent set of almost desirable gambles $\mathcal{K}$ has two equivalent representations. The first one is the credal set

$$\mathscr{M}_{\mathcal{K}} := \left\{ p \in \Sigma_{\mathscr{X}} : (\forall f \in \mathcal{K}) \sum_{x \in \mathscr{X}} f(x)p(x) \geq 0 \right\},$$

which is non-empty, closed and convex. Conversely, any non-empty subset $\mathscr{M}$ of $\Sigma_{\mathscr{X}}$ corresponds to a coherent set of almost desirable gambles

$$\mathcal{K}_{\mathscr{M}} := \left\{ f \in \mathcal{L} \colon (\forall p \in \mathscr{M}) \sum_{x \in \mathscr{X}} f(x)p(x) \geq 0 \right\}.$$

Interestingly, $\mathcal{K} = \mathcal{K}_{\mathscr{M}_{\mathcal{K}}}$ if and only if $\mathcal{K}$ is a coherent set of almost desirable gambles, and for any non-empty subset $\mathscr{M}$ of $\Sigma_{\mathscr{X}}$, $\mathscr{M} = \mathscr{M}_{\mathcal{K}_{\mathscr{M}}}$ if and only if $\mathscr{M}$ is closed and convex. The second equivalent representation of the coherent set of almost desirable gambles $\mathcal{K}$ is the *lower prevision* $\underline{P}_{\mathcal{K}}$ on $\mathcal{L}$, defined by

$$\underline{P}_{\mathcal{K}}(f) := \sup\{\mu \in \mathbb{R} : f - \mu \in \mathcal{K}\} \quad \text{for all } f \in \mathcal{L}.$$

Sometimes, it will be convenient to use the conjugate *upper prevision* $\overline{P}_{\mathcal{K}}$, defined by $\overline{P}_{\mathcal{K}}(f) := -\underline{P}_{\mathcal{K}}(-f)$ for all $f \in \mathcal{L}$. Conversely, any real-valued map $\underline{P}$ on $\mathcal{L}$ corresponds to a set of almost desirable gambles

$$\mathcal{K}_{\underline{P}} := \{f \in \mathcal{L} \colon \underline{P}(f) \geq 0\},$$

and we call $\underline{P}$ a coherent lower prevision if and only if $\mathcal{K}_{\underline{P}}$ is a coherent set of almost desirable gambles. Here as well, $\mathcal{K} = \mathcal{K}_{\underline{P}_{\mathcal{K}}}$ if and only if $\mathcal{K}$ is a coherent set of almost desirable gambles, and $\underline{P} = \underline{P}_{\mathcal{K}_{\underline{P}}}$ if and only if $\underline{P}$ is a coherent lower prevision. Clearly, there is also a one-to-one correspondence between coherent lower previsions and non-empty, closed and convex credal sets. More precisely, any coherent lower prevision $\underline{P}$ on $\mathcal{L}$ corresponds to the non-empty, closed and convex credal set

$$\mathscr{M}_{\underline{P}} := \left\{ p \in \Sigma_{\mathscr{X}} : (\forall f \in \mathcal{L}) \sum_{x \in \mathscr{X}} f(x)p(x) \geq \underline{P}(f) \right\},$$

and a non-empty credal set $\mathscr{M}$ corresponds to the coherent lower prevision $\underline{P}_{\mathscr{M}}$ on $\mathcal{L}$ defined by

$$\underline{P}_{\mathscr{M}}(f) := \inf \left\{ \sum_{x \in \mathscr{X}} f(x)p(x) \colon p \in \mathscr{M} \right\} \quad \text{for all } f \in \mathcal{L}.$$

It is clear that $\mathscr{M}_{\mathscr{C}}$ as defined in Eqn. (1) is the credal set corresponding to $\mathcal{K}_{\mathscr{C}}$, and it is not difficult to verify that $\mathscr{M}_{\mathscr{C}} = \mathscr{M}_{\mathcal{D}'}$, with $\mathcal{D}' := \mathcal{D}_{\mathcal{K}_{\mathscr{C}}}$ as defined in Eqn. (2). Throughout this contribution, we let $\underline{P}_{\mathscr{C}}$ and $\overline{P}_{\mathscr{C}}$ denote the lower and upper previsions determined by $\mathcal{D}_{\mathcal{K}_{\mathscr{C}}}$. Recall from Proposition 1 that $\mathscr{M}_{\mathscr{C}}$ is non-empty—meaning that there is a compatible probability mass function—if and only if $\mathcal{K}_{\mathscr{C}}$ avoids sure loss or $\mathcal{D}_{\mathcal{K}_{\mathscr{C}}}$ is coherent. Whenever this is the case, there are three equivalent ways of representing the same information: the closed and convex credal set $\mathscr{M}_{\mathscr{C}}$, the coherent set of almost-desirable gambles $\mathcal{D}_{\mathcal{K}_{\mathscr{C}}}$ and the coherent lower prevision $\underline{P}_{\mathscr{C}}$. The correspondences between these equivalent representations is depicted in Figure 1.

We can use $\underline{P}_{\mathscr{C}}$ to verify whether or not a comparative judgement is saturated and/or redundant. We call a judgement *redundant* if removing it from the assessment does not affect the credal set of compatible probability measures; *saturation* means that there is at least one compatible probability measure that satisfies the constraint with equality. The removal of redundant constraints is mainly interesting from a theoretical point of view. From a practical point of view, it could reduce the computational costs associated with the set of compatible probability measures; however, there is a trade off involved because removing the redundant constraints also has a computational cost. An arguably more relevant practical use for the removal of redundant constraints is that it could also help the elicitator understand better the implications of her assessments.

**Proposition 2.** *Consider an assessment $\mathscr{C}$ such that $\mathscr{M}_{\mathscr{C}}$ is non-empty, and let $\underline{P}_{\mathscr{C}}$ be its associated lower prevision.*

(i) *If there is a comparative judgement $(A, B) \in \mathscr{C}$ such that $\mathbb{I}_A - \mathbb{I}_B \in \operatorname{posi}(\mathcal{L}^+ \cup \mathcal{K}_{\mathscr{C}} \setminus \{\mathbb{I}_A - \mathbb{I}_B\})$, then $\mathscr{M}_{\mathscr{C}} = \mathscr{M}_{\mathscr{C} \setminus \{(A,B)\}}$.*

(ii) *If no $(A, B) \in \mathscr{C}$ satisfies the condition in (i), then $\underline{P}_{\mathscr{C}}(\mathbb{I}_A - \mathbb{I}_B) = 0$ for every $(A, B) \in \mathscr{C}$.*

This means that we should first analyse whether or not each constraint $(A_i, B_i)$ can be expressed as a positive linear combination of the other constraints in $\mathscr{C}$ together with trivial assessments of the type $(A, \emptyset)$ with $\emptyset \neq A \in 2^{\mathscr{X}}$; if this is the case, we can remove $(A_i, B_i)$ from our assessment. Once we have removed all these redundant constraints, any remaining constraint will be saturated by some $p \in \mathscr{M}_{\mathscr{C}}$ when this set is non-empty.

## 3. Bounding the number of extreme points

It follows immediately from the properties of probability mass functions that the comparative credal set $\mathscr{M}_{\mathscr{C}}$ defined in Eqn. (1) is a convex polytope as it is the intersection of $n+m+1$ half spaces, or equivalently, bounded by $n+m+1$ linear (in)equalities. It is well-known—see for instance [5, Section 2]—that whenever a convex polytope is non-empty, it is completely determined by its extreme points, and that any such extreme point must saturate $n = |\mathscr{X}|$ linearly independent constraints.

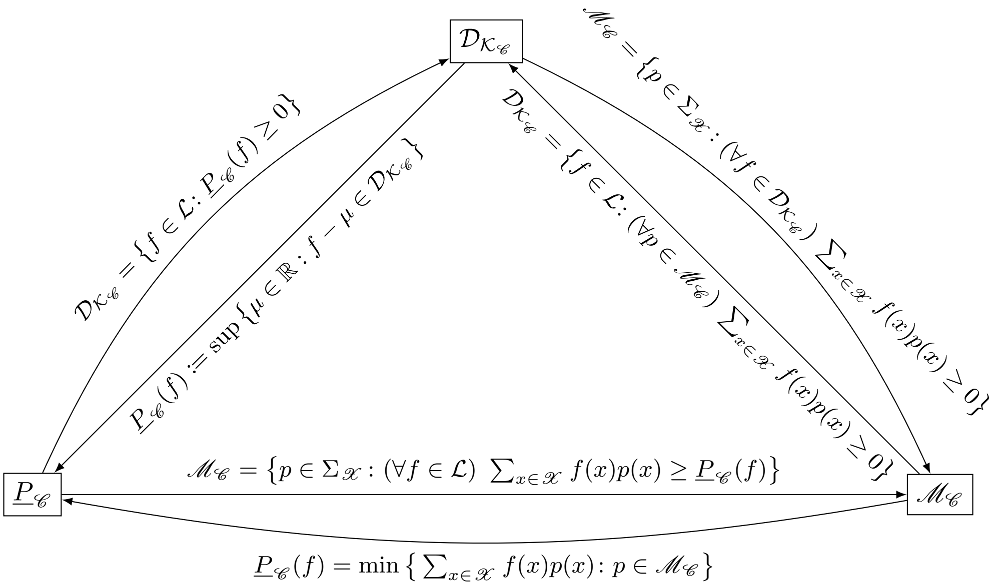


Figure 1: Correspondences between the three representations.

Studying the extreme points is useful because they provide an equivalent—and sometimes more succinct and/or more practical—representation of $\mathscr{M}_\mathscr{C}$. For instance, given the set of extreme points, the lower and upper prevision of some gamble $f$ in $\mathcal{L}$ can then be computed by determining the minimum and maximum of the expectation of $f$ with respect to the extreme points; Miranda and Destercke [32] mention several other cases where the extreme points might be of interest: for example inferences in graphical and statistical models. For this reason, it is interesting that (i) we can bound the number of extreme points a priori; and (ii) there is some procedure to determine them.

A bound on the number of extreme points follows from McMullen's theorem [17]:

$$|\mathrm{ext}(\mathscr{M}_\mathscr{C})| \leq \binom{m+1+\lfloor \frac{n}{2} \rfloor}{m+1} + \binom{m+\lceil \frac{n}{2} \rceil}{m+1}, \quad (3)$$

where $\lfloor n/2 \rfloor$ and $\lceil n/2 \rceil$ denote the largest non-negative integer $k$ such that $k \leq n/2$ and the smallest non-negative integer $\ell$ such that $\ell \geq n/2$, respectively.

It is also possible to establish an upper bound on the number of extreme points that is independent on the number of comparative judgements; its proof is a relatively straightforward modification of the proofs of [12, Theorem 4.4] or [50, Theorem 5.13].

**Proposition 3.** *For any assessment $\mathscr{C}$,*

$$|\mathrm{ext}(\mathscr{M}_\mathscr{C})| \leq n!\,2^n$$

To give a sense of the absolute and relative performance of these bounds, we reconsider our running example.

**Running example.** One can easily verify that the extreme points of the credal set $\mathscr{M}_\mathscr{C}$ are

$$p_1 := (1,0,0,0), \qquad p_2 := (1/2,1/2,0,0),$$
$$p_3 := (1/2,0,1/2,0), \qquad p_4 := (1/2,0,0,1/2),$$
$$p_5 := (1/3,1/3,0,1/3), \qquad p_6 := (1/3,0,1/3,1/3),$$
$$p_7 := (1/4,1/4,1/2,0), \qquad p_8 := (1/5,1/5,1/5,2/5),$$
$$p_9 := (1/6,1/6,1/3,1/3).$$

Hence, $|\mathrm{ext}(\mathscr{M}_\mathscr{C})| = 9$; the upper bounds on the number of extreme points of Eqn. (3) and Proposition 3 are 27 and 384, respectively.

On the other hand, the minimum number of extreme points of a non-empty comparative credal set $\mathscr{M}_\mathscr{C}$, regardless of the cardinality of the possibility space, is 1: if $\mathscr{X} := \{1,\dots,n\}$ and

$$\mathscr{C} := \{(\{i\},\{i+1\}) \colon i = 1,\dots,n-1\} \cup \{(\{n\},\{1\})\},$$

then $\mathscr{M}_\mathscr{C}$ only includes the uniform distribution on $\mathscr{X}$, and as a consequence there is only one extreme point.

Our upper bound on the number of extreme points depends on the cardinality of the space $n$ and the number $m$ of comparative assessments; thus, the bound can be made tighter if we remove constraints that are redundant because they are implied by other constraints and the monotonicity and additivity properties of probability measures. For instance, we may assume without loss of generality that

$$(\forall (A,B) \in \mathscr{C})\ A \neq \mathscr{X}, B \neq \emptyset, A \cap B = \emptyset. \quad (\mathrm{C0})$$

This allows us to bound the cardinality of $\mathscr{C}$:

**Proposition 4.** *If $\mathscr{C}$ satisfies* (C0)*, then*

$$m \leq 3^n - 2^{n+1} + 1.$$

Similarly, we may assume without loss of generality that any $(A, B) \in \mathscr{C}$ cannot be made redundant in the following senses:

(i) $(\nexists(A', B') \in \mathscr{C}, (A', B') \neq (A, B))$ $A' \subseteq A, B' \supseteq B$;
(ii) $(\nexists(A_1, B_1), (A_2, B_2) \in \mathscr{C})$ $A = A_1 \cup A_2, B = B_1 \cup B_2, A_1 \cap A_2 = \emptyset$;
(iii) $(\nexists B_1, A_2 \in 2^{\mathscr{X}}, B_1 \supseteq A_2)$ $(A, B_1) \in \mathscr{C}, (A_2, B) \in \mathscr{C}$.

Nevertheless, it is more fruitful to detect redundant constraints using the theory of coherent lower previsions, as we did in Proposition 2. In this manner, given an initial (finite) set $\mathscr{C}$ of comparative assessments, we may proceed iteratively and remove all the redundant constraints, and then use Eqn. (3) to bound the number of extreme points of the comparative credal set $\mathscr{M}_{\mathscr{C}}$ with this reduced number of constraints.

## 4. A graphical approach

Essential for the results established in [32] is the representation of the elementary comparative assessments as a digraph. In the non-elementary case, such a graphical representation will also be helpful. Throughout this contribution we use the graph theoretic terminology as defined in [21]; we do allow ourselves one difference, though: we prefer to use *nodes* instead of vertices.

### *4.1. Representing the comparative assessment as a graph*

Miranda and Destercke [32] proposed a straightforward but powerful representation of the elementary comparative assessment $\mathscr{C}$ as a digraph: the atoms of the possibility space correspond to the nodes, and a directed edge is added from $x$ to $y$ for every $(\{x\}, \{y\}) \in \mathscr{C}$. The extreme points of the credal set are then obtained through the top subnetworks generated by certain sets of nodes [32, Theorem 1].

Because we do not limit ourselves to elementary comparative judgements, we cannot simply take over their construction. One straightforward generalisation of the aforementioned construction is to add a directed edge from $x$ to $y$ if there is a comparative judgement $(A, B) \in \mathscr{C}$ with $x \in A$ and $y \in B$. However, this approach is not terribly useful because there is loss of information: clearly, the digraph alone does not contain sufficient information to reconstruct the comparative judgements it represents. To overcome this loss of information and to end up with one-to-one correspondence, we borrow a trick from Miranda and Zaffalon [34] and add dummy nodes to our graph.

We represent the assessment $\mathscr{C}$ as a digraph $\mathcal{G}_{\mathscr{C}}$ as follows. First, we add one node for every atom $x$ in the possibility space $\mathscr{X}$. Next, for every comparison $(A_i, B_i)$ in the assessment $\mathscr{C}$, we add an auxiliary node $\xi_i$, and we add a directed edge from every atom $x$ in $A_i$ to this auxiliary node $\xi_i$ and a directed edge from the auxiliary node $\xi_i$ to every atom $y$ in $B_i$. Formally, the set of nodes is $\mathcal{N}_{\mathscr{C}} := \mathscr{X} \cup \{\xi_1, \ldots, \xi_m\}$ and the set of directed edges is

$$\mathcal{E}_{\mathscr{C}} := \bigcup_{i=1}^{m} \{(x, \xi_i) : x \in A_i\} \cup \{(\xi_i, y) : y \in B_i\}.$$

**Running example.** The corresponding digraph $\mathcal{G}_{\mathscr{C}}$ is depicted in Figure 2.

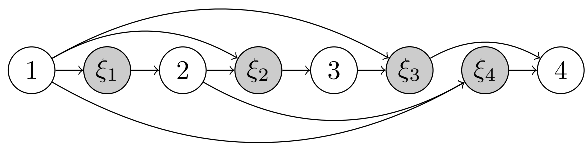


Figure 2: The digraph $\mathcal{G}_{\mathscr{C}}$ that corresponds to the assessment $\mathscr{C}$ or our running example.

Fix some node $\nu$ in the digraph $\mathcal{G}_{\mathscr{C}}$. Following [32], we use $H(\nu)$ to denote the set that consists of the node $\nu$ itself and all of its predecessors, being those nodes $\nu'$ such that there is a directed path from $\nu'$ to $\nu$. Following [9, 32], for any subset $N$ of the set of nodes $\mathcal{N}_{\mathscr{C}}$, we let $H(N) := \cup_{\nu \in N} H(\nu)$ be the so-called *top subnetwork* generated by $N$. We will exclusively be concerned with the restriction of these top subnetworks to non-auxiliary nodes; therefore, we define $H'(x) := H(x) \cap \mathscr{X}$ for any $x$ in $\mathscr{X}$ and $H'(A) := H(A) \cap \mathscr{X} = \cup_{x \in A} H'(x)$ for all $A \in 2^{\mathscr{X}}$.

**Running example.** The top subnetwork of the node 1 is $H(1) = \{1\}$ and that of node 3 is $H(3) = \{1, \xi_1, 2, \xi_2, 3\}$. Hence, $H'(\{1, 3\}) = \{1, 2, 3\}$.

### *4.2. Some basic observations*

The following results are straightforward observations that follow almost immediately from our graphical representation $\mathcal{G}_{\mathscr{C}}$ of the comparative assessment $\mathscr{C}$. The first lemma gives a useful sufficient condition for the existence of a compatible probability measure.

**Lemma 5.** *If the digraph $\mathcal{G}_{\mathscr{C}}$ has a node with indegree zero*[3]*, then $\mathscr{M}_{\mathscr{C}} \neq \emptyset$.*

To facilitate the statement of the following and future results, we introduce some additional notation. For any non-empty event $A \in 2^{\mathscr{X}}$, we denote the *uniform distribution over* $A$ as $p_A$. In the particular case that the event $A$ is the singleton $\{x\}$, we also speak of the *degenerate distribution on* $x$. The second lemma links atoms without predecessors with extreme points that are degenerate distributions.

**Lemma 6.** *The degenerate distribution $p_{\{x\}}$ on $x$ is an extreme point of the comparative credal set $\mathscr{M}_{\mathscr{C}}$ if and only if $x \in \mathscr{X}$ is a node with indegree zero.*

[3]The indegree of a node is the number of directed edges entering into this node, see [21, Section 4.1.6]. Hence, a node with indegree zero is a node without predecessors.

**Running example.** Observe that the node 1 is the only node with indegree zero. Thus, Lemmas 5 and 6 imply that (i) the comparative credal set $\mathscr{M}_{\mathscr{C}}$ is non-empty; and (ii) the degenerate distribution on 1 is an extreme point of $\mathscr{M}_{\mathscr{C}}$.

Our next result uses the well-known fact—see for instance [21, Sections 1.6 and 1.4.1]—that any digraph $\mathcal{H}$ can be uniquely decomposed into its *connected components*: the subdigraphs $\mathcal{H}_1, \dots, \mathcal{H}_k$ such that (i) $\mathcal{H} = \cup_{i=1}^{k} \mathcal{H}_i$, (ii) each subdigraph $\mathcal{H}_i$ is connected, and (iii) $\mathcal{H}_i$ and $\mathcal{H}_j$ are not connected for any $i \neq j$. For elementary comparative assessments, it is shown in [32, Proposition 2] that the extreme points of the comparative credal set can be obtained by determining the extreme points of the (elementary assessments induced by the) connected components separately. Our next result extends this to general comparative assessments.

**Proposition 7.** *Denote the connected components of the digraph $\mathcal{G}_{\mathscr{C}}$ by $\mathcal{G}_1, \dots, \mathcal{G}_k$. For every connected component $\mathcal{G}_i$, we denote its set of non-auxiliary nodes by $\mathscr{X}_i$ and we let $\mathscr{C}_i$ be the comparative assessment with possibility space $\mathscr{X}_i$ that is in one-to-one correspondence with $\mathcal{G}_i$. Then*

$$\operatorname{ext}(\mathscr{M}_{\mathscr{C}}) = \bigcup_{i=1}^{k} \{\operatorname{extend}(p_i) \colon p_i \in \operatorname{ext}(\mathscr{M}_{\mathscr{C}_i})\},$$

*where* $\operatorname{extend}(p_i)$ *is the cylindrical extension of $p_i$ to $\mathscr{X}$ that is obtained by assigning zero mass to $\mathscr{X} \setminus \mathscr{X}_i$.*

Because of this result, without loss of generality we can restrict our attention in the remainder to digraphs $\mathcal{G}_{\mathscr{C}}$ that are connected.

Finally, we establish the following result regarding cycles in the assessment.

**Proposition 8.** *Consider a comparative assessments $\mathscr{C}$, and let*

$$\mathscr{C}' := \mathscr{C} \cup \{(A, B) \colon A, B \in 2^{\mathscr{X}}; A \supset B\}.$$

*If there is a cycle $A_1 \succeq A_2 \succeq A_3 \succeq \cdots \succeq A_k \succeq A_{k+1} = A_1$ in $\mathscr{C}'$, then for any $p \in \mathscr{M}_{\mathscr{C}}$, any $i, j \in \{1, \dots, k\}$ such that $A_j \subset A_i$ and any $x \in A_i \setminus A_j$, $p(x) = 0$.*

In the language of Section 2, this means that $\overline{P}_{\mathscr{C}}(\mathbb{I}_{A_i} - \mathbb{I}_{A_{i+1}}) = 0$ if $A_j \subset A_i$, so any atom in $A_i \setminus A_j$ will always have zero mass. Hence, we can simplify the digraph $\mathcal{G}_{\mathscr{C}}$ by removing nodes that are sure to have zero mass: (i) any atom in $A_i \setminus A_j$ with $A_j \subset A_i$; and (ii) if these removals result in the formation of one or more extra disconnected components, the entirety of those disconnected components that used to be connected exclusively by incoming directed edges from (the direct successors of) the previously removed atoms.

**Example 2.** Consider $\mathscr{X} = \{1, 2, 3, 4, 5, 6\}$ and let

$$\mathscr{C} = \{(\{1\}, \{2, 3\}), (\{2\}, \{4, 5\}), (\{3\}, \{6\}), (\{4\}, \{1\})\}.$$

Then in $\mathscr{C}'$ we can form the cycle

$$\{1\} \succeq \{2, 3\} \succeq \{2\} \succeq \{4, 5\} \succeq \{4\} \succeq \{1\}.$$

Thus, by Proposition 8, any $p \in \mathscr{M}_{\mathscr{C}}$ should satisfy $p(3) = p(5) = 0$, and as a consequence also $p(6) = 0$. From this it follows that $\mathscr{M}_{\mathscr{C}}$ has only one element, namely the probability mass function $(1/3, 1/3, 0, 1/3, 0, 0)$.

**Remark 1.** *Our graphical representation also allows us to simplify somewhat the study of the compatibility problem and the extreme points in the following manner. We define a relationship $R$ between the elements of $\mathscr{X}$ as*

$$xRy \Leftrightarrow \text{ there is a directed cycle going through } x \text{ and } y.$$

*It is easy to see that $R$ is an equivalence relationship. Hence, we may consider the different equivalence classes and the directed edges between them that can be derived from $\mathcal{G}_{\mathscr{C}}$, leading to a new acyclic digraph $\mathcal{G}'_{\mathscr{C}}$ on the equivalence classes. Let $\mathcal{G}'_i$ denote the subdigraph associated with the $i$-th equivalence class and $\mathscr{C}_i$ the corresponding subset of comparative judgements. Observe that*

(i) *the set $\mathscr{M}_{\mathscr{C}}$ is non-empty if and only if there is some subdigraph $\mathcal{G}'_i$ with no predecessors in $\mathcal{G}'_{\mathscr{C}}$ such that $\mathscr{M}_{\mathscr{C}_i}$ is non-empty;*
(ii) *if a subgraph $\mathcal{G}'_i$ is such that $\mathscr{M}_{\mathscr{C}_i}$ is empty, then for each of its successors $\mathcal{G}'_j$ any element of $\mathscr{M}_{\mathscr{C}}$ gives zero probability to the nodes in $\mathcal{G}'_j$.*

*This also allows us to remove redundant parts of the graph.*

*To illustrate this, we consider the following example: let $\mathscr{X} = \{1, 2, 3, 4, 5, 6, 7, 8\}$ and let*

$$\mathscr{C} = \{(\{1\}, \{2, 3\}), (\{2\}, \{1\}), (\{2\}, \{8\}), (\{4\}, \{5, 6\}), (\{5\}, \{4, 6\}), (\{6\}, \{4, 5\}), (\{4\}, \{3\}), (\{3\}, \{7\})\}.$$

*Then the equivalence classes with respect to $R$ are given by $A_1 = \{1, 2\}, A_2 = \{3\}, A_3 = \{4, 5, 6\}, A_4 = \{7\}$ and $A_5 = \{8\}$, and the relationships between them are given by the graph depicted in Figure 3. From the assessment*

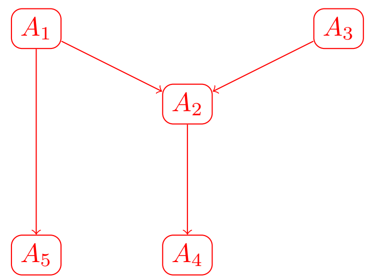

Figure 3: Graph of the equivalence classes in Example 1

$$\mathscr{C}_3 = \{(\{4\}, \{5, 6\}), (\{5\}, \{4, 6\}), (\{6\}, \{4, 5\})\}$$

*associated with the equivalence class $A_3$, it follows that the credal set $\mathscr{M}_{\mathscr{C}_3}$ is empty. This means that we can remove*

*from our set of assessments those involving $\{4,5,6\}$ as well as those associated with the equivalence classes that are their successors in the graph above, namely $A_2$ and $A_4$, because all those elements are bound to have probability zero.*
*The resulting graph is the one associated with the equivalence classes $A_1$ and $A_5$. Since only the first one has no predecessors, it follows that the set $\mathscr{M}_{\mathscr{C}}$ will be non-empty if and only if $\mathscr{M}_{\mathscr{C}_1}$ is non-empty, where $\mathscr{C}_1$ is obtained by making the intersection of $\mathscr{C}$ with the nodes in $A_1 = \{1,2\}$, that is,*

$$\mathscr{C}_1 = \{(\{1\},\{2\}),(\{2\},\{1\})\}.$$

### *4.3. Acyclic digraphs*

If a digraph is free of directed cycles, then we call it *acyclic* [21, Section 4.2]. Any acyclic digraph has at least one node with indegree zero [21, Lemma 4.1]. Therefore, the following result is an immediate corollary of Lemma 5; alternatively, it is also a corollary of Propositions 11 and 13 further on.

**Corollary 9.** *If the digraph $\mathcal{G}_{\mathscr{C}}$ associated with the comparative assessment $\mathscr{C}$ is acyclic, then the associated comparative credal set $\mathscr{M}_{\mathscr{C}}$ is non-empty.*

On the other hand, a digraph is acyclic if and only if it has a topological ordering, sometimes also called an acyclic numbering [21, Proposition 4.1]. This necessary and sufficient condition allows us to establish the following result.

**Proposition 10.** *The digraph $\mathcal{G}_{\mathscr{C}}$ associated with $\mathscr{C}$ is acyclic if and only if there is an ordering $x_1,\dots,x_n$ of the atoms of the possibility space $\mathscr{X}$ such that*

$$(\forall (A,B)\in\mathscr{C})(\exists i\in\{1,\dots,n-1\})\quad A\subseteq\{x_1,\dots,x_i\} \text{ and } B\subseteq\{x_{i+1},\dots,x_n\}.$$

**Running example.** It is easy to verify using Figure 2 that the graph $\mathcal{G}_{\mathscr{C}}$ is acyclic. Thus, by Corollary 9, the comparative credal set is non-empty, which confirms what we have previously seen. Clearly, $1,2,3,4$ is an ordering of $\mathscr{X}$ that satisfies the condition of Proposition 10.

### *4.4. Strict comparative assessments*

Our graphical representation also has implications when we consider a *strict* preference relation, where $A \succ B$ is to be interpreted as 'the event $A$ is more likely than the event $B$.' For a given set $\mathscr{C}$ of comparative judgements, we now consider the set

$$\mathscr{M}_{\mathscr{C}}^{>} := \left\{p\in\Sigma_{\mathscr{X}} : (\forall (A,B)\in\mathscr{C})\ \sum_{x\in A}p(x) > \sum_{y\in B}p(y)\right\}$$

of probability mass functions that are compatible with the strict comparative judgements. Since the set $\mathscr{M}_{\mathscr{C}}$ is a polytope, it follows that it is the closure of $\mathscr{M}_{\mathscr{C}}^{>}$ (in the topology associated with the Euclidean distance), provided that this latter set is non-empty. In our case, we can prove something stronger: $\mathscr{M}_{\mathscr{C}}^{>}$ is the topological interior of $\mathscr{M}_{\mathscr{C}}$.

**Proposition 11.** *For any comparative assessment $\mathscr{C}$, it holds that $\mathscr{M}_{\mathscr{C}}^{>} = \operatorname{int}(\mathscr{M}_{\mathscr{C}})$.*

In our next result, we establish a necessary and sufficient condition for $\mathscr{M}_{\mathscr{C}}^{>}$ to be non-empty; by Proposition 11, this condition ensures that $\mathscr{M}_{\mathscr{C}}$ is non-empty.

**Proposition 12.** *Let $\mathscr{C}$ be a comparative assessment. Then the following are equivalent:*

(i) *$\mathscr{M}_{\mathscr{C}}^{>}$ is non-empty.*
(ii) *Given the set $\mathcal{K}_{\mathscr{C}}$ defined by Eqn. (2), $0\notin \operatorname{posi}(\mathcal{K}_{\mathscr{C}}\cup\mathcal{L}^{+})$.*
(iii) *For every $(A,B)\in\mathscr{C}$, $\overline{P}_{\mathscr{C}}(\mathbb{I}_A-\mathbb{I}_B)>0$.*

In the case of elementary comparisons, it was established in [32, Lemma 1] that $\mathscr{M}_{\mathscr{C}}^{>}$ is non-empty if and only if the digraph $\mathcal{G}_{\mathscr{C}}$ is acyclic. In the general case, the lack of directed cycles turns out to be sufficient as well, leading to a result akin to Corollary 9.

**Proposition 13.** *Let $\mathscr{C}$ be a set of strict comparative assessments. If the associated digraph $\mathcal{G}_{\mathscr{C}}$ is acyclic, then $\mathscr{M}_{\mathscr{C}}^{>}$ is non-empty.*

Quite remarkably and in contrast with the case of elementary probability comparisons, $\mathscr{M}_{\mathscr{C}}^{>}$ can be non-empty even though the digraph $\mathcal{G}_{\mathscr{C}}$ has directed cycles. For example, if $\mathscr{X}=\{1,2,3\}$ and we make the assessments $(\{1,2\},\{3\})$ and $(\{3\},\{1\})$, then the graph has a cycle involving 1 and 3; however, the probability mass function $(0.25,0.45,0.3)$ is compatible with the strict assessments.

On the other hand, a necessary condition for $\mathscr{M}_{\mathscr{C}}^{>}$ to be non-empty is that we cannot derive from $\mathscr{C}$ a cycle of the type $A_1\succ A_2\succ\cdots\succ A_k\succ A_1$. This is equivalent to the graph being acyclic in the case of elementary probability comparisons, and this is what leads to [32, Lemma 1]; however, the two conditions are not equivalent in the general case.

Finally, we may also consider the *mixed* scenario where we consider the union of some set $\mathscr{C}_1$ of strict comparisons—that is, of the type '$P(A)>P(B)$'—and a set $\mathscr{C}_2$ of weak comparisons—as in '$P(A)\geq P(B)$'. It follows from Proposition 11 that in that case the set of compatible probability measures $\mathscr{M}$ is a convex set whose interior is $\mathscr{M}_{\mathscr{C}_1\cup\mathscr{C}_2}^{>}$ and whose closure is $\mathscr{M}_{\mathscr{C}_1\cup\mathscr{C}_2}$. Proposition 13 implies that if the associated digraph $\mathcal{G}_{\mathscr{C}}$ is acyclic, then $\mathscr{M}$ is non-empty. Finally, with a proof similar to that of Proposition 12 we can prove that $\mathscr{M}$ is non-empty when $\overline{P}_{\mathscr{C}_1\cup\mathscr{C}_2}(\mathbb{I}_A-\mathbb{I}_B)>0$ for every $(A,B)\in\mathscr{C}_1$.

## 5. Extreme points of the comparative credal set

As we have often mentioned before, Miranda and Destercke [32] show that in the case of elementary comparative assessments, the extreme points of the comparative credal set can be determined using a graphical representation. More specifically, they show that

E1. all the extreme points of $\mathscr{M}_{\mathscr{C}}$ correspond to uniform probability distributions [32, Lemma 2];
E2. if $C \in 2^{\mathscr{X}}$ is the support of an extreme point, then $C = H'(C)$ [32, Lemma 3];
E3. there are at most $2^{n-1}$ extreme points, and this bound is tight [32, Theorem 4].

Unfortunately, these observations do not hold in the case of non-elementary comparative assessments, as is illustrated by the following example.

**Example 3.** Let $\mathscr{X} := \{1, \dots, 5\}$, and let $\mathscr{C}$ be given by

$$\begin{aligned}\mathscr{C} := \{&(\{1,4\},\{5\}),(\{2,4\},\{1\}),(\{2,5\},\{1\}),\\ &(\{2,3,5\},\{4\}),(\{2,3\},\{1\}),(\{2,4,5\},\{3\}),\\ &(\{1,2,3\},\{4,5\}),(\{3,4\},\{5\}),(\{1,5\},\{3\}),\\ &(\{1,3,4,5\},\{2\}),(\{1,3,5\},\{4\}),(\{3,4,5\},\{1\})\}.\end{aligned}$$

The 34 extreme points of $\mathscr{M}_{\mathscr{C}}$ are reported in Table 1. Note that $34 > 2^5 = 32$. We observe that (E1) does not hold because $p_4$ is not a uniform distribution; (E2) does not hold because the support of $p_1$ is $\{1, 2, 4\}$ but $H'(\{1, 2, 4\}) = \{1, 2, 3, 4, 5\}$; and (E3) does not hold because there are more than $2^{5-1} = 16$ extreme points.

In fact, we see from Example 3 that a comparative credal set can have more than $2^n$ extreme points. Consequently, we cannot use the strategy of [32, Algorithm 1]—that is, construct the possible supports and use the uniform distribution over them—to immediately determine the extreme points of the comparative credal set for some general comparative assessment. This being said, we have nevertheless identified some special cases other than the elementary one in which we *can* generate the extreme points using the digraph $\mathcal{G}_{\mathscr{C}}$.

In the upcoming sections, we will consider a few of these cases by placing conditions on the structure of comparative assessments. As we have mentioned already, [32] deals with the case where all the events are singletons. It is straightforward to show that similar results hold when we consider comparisons on an element of a partition of the possibility space. Our first result takes this idea a bit further.

*5.1. Multi-level partitions of comparative assessments*

As a first special case, we consider a straightforward extension of [32] using a multi-level approach. At the core of this special case are some nested partitions of the possibility space and the restriction that the comparative judgements can only concern events that are on the same level of the nested partitions and belong to the same part of the partition in the previous level. We will here only explain the two-level case in detail; extending the approach to multiple levels is straightforward.

Let $C_1, \dots, C_k$ be a partition of the possibility space $\mathscr{X}$. A comparative assessment $\mathscr{C}$ is *two-level* over this partition if it can be partitioned as

$$\mathscr{C} = \mathscr{C}' \cup \bigcup_{i=1}^{k} \mathscr{C}_i,$$

with $\mathscr{C}' := \mathscr{C} \cap \{(A, B)\colon A, B \in \{C_1, \dots, C_k\}\}$ and $\mathscr{C}_i := \mathscr{C} \cap \{(\{x\}, \{y\})\colon x, y \in C_i\}$ for all $i \in \{1, \dots, k\}$. Observe that if such a decomposition exists, then we can interpret $\mathscr{C}'$ as an elementary comparative assessment with possibility space $\mathscr{X}' := \{C_1, \dots, C_k\}$ and, for all $i \in \{1, \dots, k\}$, we can interpret $\mathscr{C}_i$ as an elementary comparative assessment with possibility space $C_i$. Hence, we may use the algorithm described in [32] to determine the extreme points of the comparative credal sets corresponding to these elementary comparative assessments, which we shall denote by $\mathscr{M}'_{\text{el}}$, $\mathscr{M}_{\text{el},1}$, ..., $\mathscr{M}_{\text{el},k}$, respectively. The following result establishes that we can combine these extreme points to obtain the extreme points of the original comparative credal set $\mathscr{M}_{\mathscr{C}}$.

**Proposition 14.** *Consider a comparative assessment $\mathscr{C}$ that is two-level over the partition $C_1, \dots, C_k$ of the possibility space $\mathscr{X}$. Then $\operatorname{ext}(\mathscr{M}_{\mathscr{C}})$ is given by*

$$\big\{\operatorname{comb}(p, p_1, \dots, p_k)\colon\\ p \in \operatorname{ext}(\mathscr{M}'_{\text{el}}), (\forall i \in \{1, \dots, k\})\ p_i \in \operatorname{ext}(\mathscr{M}_{\text{el},i})\big\},$$

*where $\operatorname{comb}(p, p_1, \dots, p_k)$ is the probability mass function defined for all $i \in \{1, \dots, k\}$ and $x \in C_i$ as*

$$\operatorname{comb}(p, p_1, \dots, p_k)(x) := p(C_i)p_i(x).$$

Furthermore, as a corollary of Proposition 14 and [32, Theorem 4] we obtain the following bound on the number of extreme points.

**Corollary 15.** *Consider a comparative assessment $\mathscr{C}$ that is two-level over some partition. Then $|\operatorname{ext}(\mathscr{M}_{\mathscr{C}})| \le 2^{n-1}$.*

The following example illustrates the result.

**Example 4.** Consider $\mathscr{X} = \{1, 2, 3, 4\}$, and let $\mathscr{C}$ be given by

$$\mathscr{C} = \{(\{1,2\},\{3,4\}),(\{1\},\{2\}),(\{3\},\{4\})\}.$$

This corresponds to a two-level comparative assessment associated with the partition $C_1 := \{1, 2\}$ and $C_2 := \{3, 4\}$. It is not difficult to see that $\operatorname{ext}(\mathscr{M}'_{\text{el}}) = \{(1, 0), (\frac{1}{2}, \frac{1}{2})\}$ (on $\mathscr{X}' = \{C_1, C_2\}$), $\operatorname{ext}(\mathscr{M}_{\text{el},1}) = \{(1, 0), (\frac{1}{2}, \frac{1}{2})\}$ (on $C_1 = \{1, 2\}$) and $\operatorname{ext}(\mathscr{M}_{\text{el},1}) = \{(1, 0), (\frac{1}{2}, \frac{1}{2})\}$ (on $C_2 = \{3, 4\}$). Applying Proposition 14 we conclude that $\operatorname{ext}(\mathscr{M}_{\mathscr{C}})$ is given by

$$\begin{aligned} p_1 &:= (1, 0, 0, 0), & p_2 &:= (1/2, 1/2, 0, 0),\\ p_3 &:= (1/2, 0, 1/2, 0), & p_4 &:= (1/2, 0, 1/4, 1/4),\\ p_5 &:= (1/4, 1/4, 1/2, 0), & p_6 &:= (1/4, 1/4, 1/4, 1/4),\end{aligned}$$

resulting from the combination of the extreme points at the different levels.

Table 1: The extreme points of the comparative assessment in Example 3.

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $p_i(1)$ | $\frac{1}{3}$ | $\frac{1}{3}$ | $\frac{1}{4}$ | $\frac{3}{8}$ | $\frac{1}{7}$ | $\frac{1}{5}$ | $\frac{1}{5}$ | 0 | $\frac{1}{6}$ | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{3}$ | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{3}$ | $\frac{3}{14}$ | $\frac{4}{11}$ |
| $p_i(2)$ | $\frac{1}{3}$ | $\frac{1}{3}$ | 0 | $\frac{2}{8}$ | 0 | 0 | $\frac{1}{5}$ | $\frac{1}{2}$ | 0 | $\frac{1}{4}$ | $\frac{1}{8}$ | $\frac{1}{3}$ | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{3}$ | $\frac{1}{14}$ | $\frac{3}{11}$ |
| $p_i(3)$ | 0 | $\frac{1}{3}$ | $\frac{1}{4}$ | $\frac{1}{8}$ | $\frac{3}{7}$ | $\frac{2}{5}$ | $\frac{2}{5}$ | 0 | $\frac{1}{3}$ | 0 | $\frac{1}{8}$ | 0 | 0 | $\frac{1}{4}$ | $\frac{1}{6}$ | $\frac{3}{14}$ | $\frac{1}{11}$ |
| $p_i(4)$ | $\frac{1}{3}$ | 0 | $\frac{1}{4}$ | $\frac{1}{8}$ | $\frac{1}{7}$ | $\frac{1}{5}$ | 0 | $\frac{1}{4}$ | $\frac{1}{6}$ | $\frac{3}{8}$ | $\frac{3}{8}$ | $\frac{1}{6}$ | $\frac{1}{4}$ | 0 | 0 | $\frac{2}{14}$ | $\frac{1}{11}$ |
| $p_i(5)$ | 0 | 0 | $\frac{1}{4}$ | $\frac{1}{8}$ | $\frac{2}{7}$ | $\frac{1}{5}$ | $\frac{1}{5}$ | $\frac{1}{4}$ | $\frac{1}{3}$ | $\frac{1}{8}$ | $\frac{1}{8}$ | $\frac{1}{6}$ | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{6}$ | $\frac{5}{14}$ | $\frac{2}{11}$ |
| $i$ | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 |
| $p_i(1)$ | $\frac{3}{12}$ | 0 | $\frac{1}{10}$ | $\frac{1}{6}$ | $\frac{1}{5}$ | $\frac{2}{12}$ | 0 | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{6}$ | $\frac{1}{8}$ | $\frac{1}{4}$ | 0 | 0 | $\frac{1}{4}$ | $\frac{1}{7}$ |
| $p_i(2)$ | $\frac{1}{12}$ | $\frac{1}{4}$ | 0 | 0 | $\frac{1}{5}$ | $\frac{1}{12}$ | $\frac{1}{2}$ | $\frac{1}{2}$ | $\frac{1}{2}$ | $\frac{1}{2}$ | $\frac{1}{2}$ | $\frac{1}{2}$ | $\frac{1}{2}$ | $\frac{1}{3}$ | $\frac{1}{2}$ | $\frac{1}{2}$ | $\frac{1}{7}$ |
| $p_i(3)$ | $\frac{2}{12}$ | $\frac{1}{4}$ | $\frac{4}{10}$ | $\frac{1}{3}$ | $\frac{1}{5}$ | $\frac{3}{12}$ | $\frac{1}{6}$ | 0 | 0 | $\frac{1}{4}$ | $\frac{1}{6}$ | $\frac{1}{4}$ | $\frac{1}{8}$ | $\frac{1}{6}$ | $\frac{1}{8}$ | $\frac{1}{8}$ | $\frac{3}{14}$ |
| $p_i(4)$ | $\frac{2}{12}$ | $\frac{1}{4}$ | $\frac{2}{10}$ | $\frac{1}{3}$ | $\frac{2}{5}$ | $\frac{5}{12}$ | $\frac{1}{6}$ | $\frac{1}{4}$ | $\frac{1}{8}$ | 0 | 0 | 0 | 0 | $\frac{1}{3}$ | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{3}{7}$ |
| $p_i(5)$ | $\frac{4}{12}$ | $\frac{1}{4}$ | $\frac{3}{10}$ | $\frac{1}{6}$ | 0 | $\frac{1}{12}$ | $\frac{1}{6}$ | 0 | $\frac{1}{8}$ | 0 | $\frac{1}{6}$ | $\frac{1}{8}$ | $\frac{1}{8}$ | $\frac{1}{6}$ | $\frac{1}{8}$ | 0 | $\frac{1}{14}$ |

**Remark 2.** *While simple, the result in this section allows us to generalise some of the comments on [32, Section 5.2] about imprecise mass functions; see also [1] and [19].*

*Imprecise mass functions arise in the context of evidential theory, where a measure $m$ of imprecise information is allocated into the subsets of $\mathscr{X}$, and then this determines the lower envelope $\underline{P}$ of the set of compatible probability measures by means of*

$$\underline{P}(A) = \sum_{B \subseteq A} m(B).$$

*The lower probability obtained in this manner is a belief function, and is therefore connected with a random set.*

*As discussed in [32], we may have imprecise information about the mass function $m$, and model this information through some comparative assessments of the type $m(A_i) \geq m(A_j)$ for different $A_i, A_j \subseteq \mathscr{X}$. We may consider then the set of compatible mass functions $\mathcal{M}$ that in turn shall determine a convex set of probability measures. Note that this is equivalent to consider second-order probability models, that have been deemed of interest in psychology [14, 23, 31].*

*While in [32] only comparisons between disjoint subsets of $\mathscr{X}$ where considered, we may consider general comparisons, by noticing that (i) these will correspond to elementary subsets of $\mathcal{P}(\mathscr{X})$, and so the results on elementary comparative probabilities are applicable if we place ourselves in that framework; and (ii) that even in $\mathscr{X}$ we could also consider the more general structures tackled in this paper.*

### 5.2. Acyclic digraphs

Recall from Section 4.4 that the absence of cycles simplifies things if we are interested in the compatibility with *strict* comparative judgements. Hence, it does not seem all too far-fetched that determining the (number of) extreme points of the comparative credal set induced by a (non-strict) comparative assessment also simplifies under the absence of cycles. As will become clear in the remainder, this is only certainly so in some special cases.

First, we revisit the three main points of [32] that we recalled at the beginning of this section in the case of acyclic graphs. Our running example shows that also in the acyclic case (E1) does not hold because $p_7$, $p_8$ and $p_9$ are not uniform; (E2) does not hold because $p_3$ has support $C_3 := \{1, 3\}$ but $H'(C_3) = \{1, 2, 3\} \neq C_3$; and (E3) does not hold because there can be more than $2^{4-1} = 8$ extreme points. Furthermore, since different extreme points can have the same support—in our running example, this is the case for $p_7$, $p_8$ and $p_9$—there is no reason why the number of extreme points should be bounded above by $2^n$. Nevertheless, and despite our rather extensive search, we have not succeeded in finding an example of a comparative assessment $\mathscr{C}$ with an acyclic digraph $\mathcal{G}_{\mathscr{C}}$ that has a comparative credal set with more than $2^n$ extreme points. This is in contrast with the cyclic case, as we have shown in Example 3.

While the absence of cycles alone does not seem to allow us to efficiently determine the extreme points, there are two interesting special cases that permit us to do so. Essential to both these special cases is a specific class of subdigraphs of the digraph $\mathcal{G}_{\mathscr{C}}$. To define this class, we first need to introduce two concepts from graph theory. The first concept is that of the *root* of a digraph $\mathcal{H}$: a node $\nu$ such that for any other node $\nu'$, there is a directed path from $\nu$ to $\nu'$. The second concept is that of an *arborescence*: a digraph that has a root and whose underlying graph is a tree. We see for instance that the graph depicted in Figure 2 is not an arborescence, since there is more than one path between the elements 1 and 4, while the digraph depicted in Figure 4 is an arborescence.

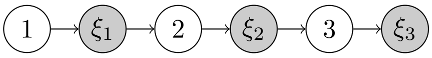


Figure 4: An arborescence that is a subdigraph of the acyclic digraph $\mathcal{G}_{\mathscr{C}}$ depicted in Figure 2.

We now call a subdigraph $\mathcal{G}'$ of the digraph $\mathcal{G}_{\mathscr{C}}$ an *extreme arborescence* if (i) it is an arborescence whose root $x^\star$ has no predecessors in the digraph $\mathcal{G}_{\mathscr{C}}$; and (ii) each of its auxiliary nodes has one direct predecessor and one direct successor.

**Running example.** The digraph depicted in Figure 4 is not an extreme arborescence for $\mathcal{G}_{\mathscr{C}}$, because the auxiliary node $\xi_3$ has no direct successor. Figure 5 depicts some examples of subdigraphs of $\mathcal{G}_{\mathscr{C}}$ that are extreme arborescences.

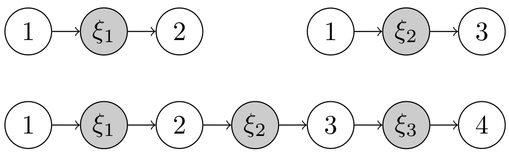


Figure 5: Some extreme arborescences for the digraph $\mathcal{G}_{\mathscr{C}}$ depicted in Figure 2.

Important to note is that all extreme arborescences can be easily procedurally generated. In essence, one needs to (i) select a node $x$ without predecessors in the original digraph $\mathcal{G}_{\mathscr{C}}$; (ii) either stop or, if possible, (a) add one of the outgoing edges of $x$ and the auxiliary node $\xi$ in which it ends, (b) add one of the outgoing edges of $\xi$ and the corresponding successor atom $y$ given that $y$ is not already in the arborescence; (iii) repeat step (ii) but with $x$ being any of the atoms already in the arborescence.

### 5.2.1. *Singular assessments*

The first special case of acyclic digraphs concerns representing digraphs where every (node corresponding to an) atom of has at most one direct predecessor. We call a comparative assessment $\mathscr{C}$ *singular* if

$$(\forall x \in \mathscr{X})\ |\{(A, B) \in \mathscr{C} : x \in B\}| \leq 1.$$

That the asessment is singular means that the events $\{B_i : i = 1, \ldots, m\}$ in our set $\mathscr{C}$ are pairwise disjoint. This could be of interest for instance if we consider partitions of $\mathscr{X}$ and for each element $B$ in the partition we bound above its probability by determining a disjoint subset $A$ such that $P(A) \geq P(B)$; working with partitions could be useful when we first group the elements of $\mathscr{X}$ in clusters and later we model our uncertainty about them using comparative probabilities.

**Remark 3.** *The underlying idea to singular graphs is related to the notion of* sources of contradiction *considered in [34]. Roughly speaking, it was proven there that the problem of the compatibility of a number of (possibly imprecise) conditional probability assessments becomes trivial when the associated coherence graph satisfies two properties: being acyclic and having no nodes with more than one parent. The property of singularity we have introduced accounts for this second condition in our different, and unconditional, context. Like the notion of source of contradiction of coherence graphs, it allows us to simplify the analysis of the structure of the credal set by preventing interactions that may lead to potential conflicts between the assessments.*

**Running example.** We see that the comparative assessment $\mathscr{C}$ is not singular, since 4 appears in both the assessments $(\{1, 3\}, \{4\})$ and $(\{1, 2\}, \{4\})$, while the comparative assessment

$$\mathscr{C}' := \{(\{1\}, \{2\}), (\{1, 2\}, \{3\}), (\{2, 3\}, \{4\})\},$$

whose corresponding digraph is depicted in Figure 6, is.

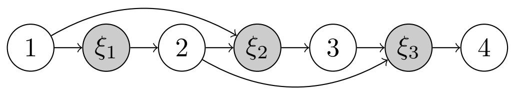


Figure 6: A singular digraph $\mathcal{G}_{\mathscr{C}'}$

The graph associated with a singular assessment need not be acyclic—for example, let $\mathscr{X} = \{1, 2, 3\}$ and consider the comparative judgements $(\{1\}, \{2\})$, $(\{2\}, \{3\})$ and $(\{3\}, \{1\})$. In case it is, we can establish the following two lemmas.

**Lemma 16.** *Consider a singular assessment $\mathscr{C}$ with corresponding digraph $\mathcal{G}_{\mathscr{C}}$ that is acyclic, and let $p$ be an extreme point of the comparative credal set $\mathscr{M}_{\mathscr{C}}$ with support $\mathscr{X}_p := \{x \in \mathscr{X} : p(x) > 0\}$. Then*

(i) *$\mathscr{X}_p$ contains one atom $x^\star$ without predecessors;*
(ii) *there is a subdigraph $\mathcal{G}'$ of $\mathcal{G}_{\mathscr{C}}$ that is an extreme arborescence with root $x^\star$ and atoms $\mathscr{X}_p$.*

**Lemma 17.** *Consider a singular assessment $\mathscr{C}$ such that the corresponding digraph $\mathcal{G}_{\mathscr{C}}$ is acyclic, and fix a non-empty event $C \in 2^{\mathscr{X}}$. If there is a subdigraph $\mathcal{G}'$ of $\mathcal{G}_{\mathscr{C}}$ that is an extreme arborescence with atoms $C$, then there is an extreme point that has $C$ as support.*

These two lemmas allow us to establish the following result.

**Theorem 18.** *Consider a singular assessment $\mathscr{C}$ such that the associated digraph $\mathcal{G}_{\mathscr{C}}$ is acyclic. Then every extreme point $p$ of $\mathscr{M}_{\mathscr{C}}$ corresponds to a unique extreme arborescence $\mathcal{G}' \subseteq \mathcal{G}_{\mathscr{C}}$ and vice versa, in the sense that $p$ is the unique probability mass function that saturates the comparative constraints associated with the auxiliary nodes in $\mathcal{G}'$ and the non-negativity constraints associated with the atoms that are not in $\mathcal{G}'$.*

Because we can procedurally generate all extreme arborescences, it follows that we can use Theorem 18 to generate all extreme points of the comparative credal set. Another consequence of Theorem 18 is that we can establish a lower and upper bound on the number of extreme points in the singular case.

**Theorem 19.** *Consider a singular assessment $\mathscr{C}$ such that the associated digraph $\mathcal{G}_{\mathscr{C}}$ is acyclic. Then*

$$n \leq |\mathrm{ext}(\mathscr{M}_{\mathscr{C}})| \leq 2^{n-1}.$$

These lower and upper bounds are reached, as we can see from [32, Section 4.1].

### 5.2.2. Arborescences

Finally, we consider the case that the corresponding digraph $\mathcal{G}_{\mathscr{C}}$ is an arborescence. Clearly, for this it is necessary that $\mathscr{C}$ is singular and that

$$(\forall (A, B) \in \mathscr{C})\ |A| = 1. \tag{4}$$

We see that the graph depicted in Figure 6 is not an arborescence, while the one associated with the assessment $\mathscr{C} := \{(\{1\}, \{2, 3\}), (\{2\}, \{4\}), (\{3\}, \{5\})\}$, depicted in Figure 7, is. This also illustrates the fact that in the case of singular assessments the sets $B$ must be pairwise disjoint, but they need not be singletons, as was the case with the graph in Figure 6.

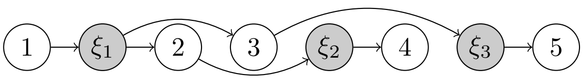


Figure 7: An example of a digraph $\mathcal{G}_{\mathscr{C}}$ that is an arborescence.

As arborescences are special types of acyclic digraphs, we can strengthen Theorem 18 to be—in some sense—similar to [32, Theorem 1]. The key idea here is that with arborescences, we can establish a total order on the possibility space that is compatible with the partial order underlying to the graph, and this allows to somewhat 'partition' our possibility space according to the order, thus making a connection with elementary comparisons. First of all, we establish the following two lemmas.

**Lemma 20.** *Consider an assessment $\mathscr{C}$ such that the associated digraph $\mathcal{G}_{\mathscr{C}}$ is an arborescence. Let $p$ be an extreme point of $\mathscr{M}_{\mathscr{C}}$ with support $\mathscr{X}_p := \{x \in \mathscr{X} : p(x) > 0\}$. Then $\mathscr{X}_p = H'(\mathscr{X}_p)$ and the closest common predecessor of any two $x$ and $y$ in $\mathscr{X}_p$ is a non-auxiliary node.*

**Lemma 21.** *Consider an assessment $\mathscr{C}$ such that the associated digraph $\mathcal{G}_{\mathscr{C}}$ is an arborescence. Let $C$ be a set of states such that for any distinct $x$ and $y$ in $C$, their closest common predecessor is a non-auxiliary node. Then*

$$p \colon \mathscr{X} \to \mathbb{R} \colon x \mapsto p(x) := \begin{cases} \frac{1}{|H'(C)|} & \text{if } x \in H'(C), \\ 0 & \text{otherwise} \end{cases}$$

*is the unique extreme point of $\mathscr{M}_{\mathscr{C}}$ with support $H'(C)$.*

Using these two lemmas, we can derive the following result.

**Theorem 22.** *Consider and assessment $\mathscr{C}$ such that the associated digraph $\mathcal{G}_{\mathscr{C}}$ is an arborescence. Then the set of extreme points of $\mathscr{M}_{\mathscr{C}}$ consists of the uniform distributions on $H'(C)$, where $C$ is any set of atoms such that for all $x, y \in C$, the closest common predecessor of $x$ and $y$ is a non-auxiliary node.*

We also observe that the bound on the number of extreme points established in Theorem 19 is still valid.

**Example 5.** To see that this result does not extend to singular assessments when condition (4) is not satisfied, note that the extreme points of the assessment $\mathscr{C}$ depicted in Figure 6 are

$$\begin{aligned} p_1 &:= (1, 0, 0, 0), & p_2 &:= (1/2, 1/2, 0, 0), \\ p_3 &:= (1/2, 0, 1/2, 0), & p_4 &:= (1/4, 1/4, 1/2, 0), \\ p_5 &:= (1/3, 1/3, 0, 1/3), & p_6 &:= (1/3, 0, 1/3, 1/3), \\ p_7 &:= (1/7, 1/7, 2/7, 3/7), \end{aligned}$$

and that not all of these correspond to uniform distributions.

## 6. Connection with other fields and possible extensions

In this section, we briefly discuss some connections between our results in this paper and other fields as well as some possible extensions of our work.

### 6.1. Ranking sets of objects

The partial orders we have considered in this paper can also be found in the problem of ranking sets of objects [4, 22, 37]. Given a possibility space, a (possibly) partial order $\succeq$ is established on the subsets of $\mathscr{X}$, with the interpretation that $A \succeq B$ means that 'the group formed by the elements of $A$ is preferred to that formed by the elements of $B$'. Then it is analysed under which condition a total ranking can be established over the elements of $\mathscr{X}$, possibly based on a quantitative measure. This has links also with the problem of measuring the strenghts of the coallitions in coallitional game theory [3, 45]. Since the seminal work of Kannai and Peleg [25], a number of impossibility theorems have been established.

The partial orders $\succeq$ we have considered in this paper can thus be embedded into this context, accounting to the case where the ranking is based on some partially known probability measure. Then our characterisation of the extreme points of the set of compatible probabilities allows us to determine the associated lower and upper probabilities of each element, from which we could derive a number of rankings.

Nevertheless, we should mention that, in the context of ranking of sets of objects, there will be interesting scenarios where the order $\succeq$ is not monotone (we may for instance have that $\{x\} \succ \{x, y\}$) and then they could not be modelled by means of comparative probability assessments.

### 6.2. *Social choice theory*

The previous comments are also related to social choice theory, where a number of preference orderings over some space $\mathscr{X}$ must be aggregated into a global order $\mathscr{X}$. In this sense, we could regard our order $\succeq$ as resulting from the union of the partial preferences established by a number of voters, and these partial preferences may be related from comparative probability assessments. In that case, it would first be useful to analyse the compatibility of all these assessments, that is, to analyse whether there are contradictory preferences among the voters, and for this our results in Section 2.1 would be useful.

Moreover, for the connection to be clear, we should also allow for the preferences to be very weak (for instance, a voter may only say that $A$ is preferred to $B$, and establish no further comparison between the subsets). In this respect, it is useful to consider the works on the aggregation of imprecise preferences, as in [15, 38]. As shown in [7], there are interesting connections between belief aggregation [47] and social choice that would be interesting to explore here.

### 6.3. *Comparative imprecise probabilities*

More generally, we could also tackle the problem of comparing *imprecise* probability models. If we consider a family $\mathcal{C}$ of such models (for instance possibility measures, belief functions, or probability intervals), this could be done in two directions: (i) studying if the lower envelope of the set of probability measures compatible with the comparative assessments belongs to $\mathcal{C}$; or (ii) perform the comparison on the elements of $\mathcal{C}$ instead of probability measures, by considering for instance assessments of the type '$\Pi(A) \geq \Pi(B)'$ for a possibility measure $\Pi$.

While much work has to be done, a preliminary analysis leads us to conjecture that the result in (i) will only be achieved by considering approximations of these lower envelopes, considering that in [32] it was proven that for elementary comparisons 2-monotonicity, which is the weakest requirement that may be added to the coherence that holds from being the envelope of a convex set of probability measures, is not satisfied. Nevertheless, the results on outer approximating coherent lower probabilities by more tractable models in [33, 35, 36] would be useful.

Concerning (ii), convexity may be an issue in some cases, such as possibility measures, but not in others, such as belief functions. When that is not an issue, we think that the theoretical results may be extended in a sort of straightforward manner, but that the computational complexity would increase significantly. Thus, it would be interesting to consider particular cases, such as probability intervals [6] or $p$-boxes [13].

## 7. Conclusions

As we have seen, the study of the set of probability measures associated with some compatible comparative judgements becomes significantly more involved when the comparisons involve non-elementary events. Our contributions in this respect have followed two different directions. On the one hand, we have shown that coherent lower previsions and sets of almost-desirable gambles can be used to solve the fundamental problems of compatibility, redundancy and saturation of the constraints. On the other hand, we have shown that a graphical representation can be used to (i) decompose the analysis of the set of compatible probability measures into a set of simpler subproblems; (ii) connect the strict and non-strict probability comparisons; (iii) give a tighter bound on the number of extreme points of the credal sets; and (iv) characterise these extreme points in a number of cases. In addition, we have also shown that these results cannot be extended to the general case.

Although we find the above results promising, there are some open problems that call for additional research, which should help towards making this model more operative for practical purposes. First and foremost, we would like to deepen the study of the acyclic case, and in particular to determine the number and the shape of the extreme points in other particular cases. In addition, a bound on the number of linearly independent constraints, in the manner hinted at in Section 3, should let us get a better bound on the number of extreme points. Finally, we should also look for graph decompositions that allow to work more efficiently with comparative judgements.


## Acknowledgements

We would like to thank the two anonymous reviewers for the careful reading of this paper; thanks to their useful suggestions, we have made a number of improvements to the presentation of our work.


## References


[1] Thomas Augustin. Generalized basic probability assignments. *International Journal of General Systems*, 4, 451–463, 2005.

[2] Thomas Augustin, Frank Coolen, Gert de Cooman, and Matthias Troffaes, editors. *Introduction to Imprecise Probabilities.* Wiley Series in Probability and Statistics. Wiley, 2014.

[3] John F. Banzhaf. Weighted voting does not work: a mathematical analysis. *Rutgers Law Review*, 19, 317–343, 1965.

[4] Salvador Barberà, Walter Bossert and Prasanta Pattanaik. Ranking sets of objects. In *Handbook of Utility Theory* (S. Barberà, P.J. Hammond, C. Seidl, eds.) Boston, MA, pages 893-977, 2004. doi: 10.1007/978-1-4020-7964-1

[5] Dimitris Bertsimas and Jonh N. Tsitsiklis. *Introduction to Linear Optimization.* Athena Scientific, 1997.

[6] Luis. M. de Campos, Juan F. Huete, and Serafín Moral. Probability intervals: a tool for uncertain reasoning. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 2:167–196, 1994.

[7] Arianna Casanova, Enrique Miranda and Marco Zaffalon. Joint desirability foundations of social choice and opinion pooling. *Annals of Mathematics and Artificial Intelligence*, 2021. To appear. doi: 10.1007/s10472-021-09733-7

[8] Michael A. Cohen. Necessary and sufficient conditions for the existence and uniqueness of weak qualitative probability measures. *Journal of Mathematical Psychology*, 35:242–259, 1991. doi: 10.1016/0022-2496(91)90027-Q

[9] Fabio G. Cozman. Credal networks. *Artificial Intelligence*, 120(2):199–233, 2000. ISSN 0004-3702. doi: 10.1016/S0004-3702(00)00029-1.

[10] Bruno de Finetti. Sul significato soggettivo della probabilità. *Fundamenta Mathematicae*, 17:298–329, 1931. doi: 10.4064/fm-17-1-298-329.

[11] Arthur P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, 38: 325–339, 1967.

[12] Jean Derks and Jeroen Kuipers. On the number of extreme points of the core of a transferable utility game. In *Chapters in Game Theory: In honor of Stef Tijs*, pages 83–97. Springer, 2002. doi: 10.1007/0-306-47526-X_4.

[13] Sebastien Destercke, Didier Dubois, and Eric Chojnacki. Unifying practical uncertainty representations: I. Generalized p-boxes. *International Journal of Approximate Reasoning*, 49(3):649–663, 2008.

[14] Stephen H. Dewitt, Norman E. Fenton, Alice Liefgreen and David A. Lagnado. *Propensities and Second Order Uncertainty: A Modified Taxi Cab Problem. Frontiers in Psychology*, October 2020.

[15] Franz Dietrich and Christian List. *Judgement aggregation with consistency alone* METEOR Research Memorandum No. 021, Maastricht University School of Business and Economics, 2007.

[16] Didier Dubois and Henri Prade. *Possibility Theory.* Plenum Press, New York, 1988.

[17] Martin E. Dyer. The complexity of vertex enumeration methods. *Mathematics of Operations Research*, 8(3):381–402, 1983. doi: 10.1287/moor.8.3.381.

[18] Alexander Erreygers and Enrique Miranda. A study of the set of probability measures compatible with comparative judgements. In *Information Processing and Management of Uncertainty in Knowledge-Based Systems (Proceedings of IPMU 2020)*, pages 165–179. Springer, 2020. doi: 10.1007/978-3-030-50143-3_13.

[19] Ronald Fagin and Joseph Halpern. Uncertainty, belief, and probability. *Computational Intelligence*, 6, 160–173, 1991.

[20] Peter Fishburn. The axioms of subjective probability. with discussion. *Statistical Science*, 1:335–358, 1986.

[21] Jean-Claude Fournier. *Graphs Theory and Applications: With Exercises and Problems.* ISTE & Wiley, 2009.

[22] Christian Geist and Ulle Endriss. Automated Search for Impossibility Theorems in Social Choice Theory: Ranking Sets of Objects. *Journal of Artificial Intelligence Research*, 40, 143–174, 2011.

[23] Robert Goldsmith and Niels-Eric Sahlin. The Role of Second-Order Probabilities in Decision Making. *Advances in Psychology*, 14, 455–467, 1983.

[24] Irving John Good. *Probability and the Weighting of Evidence.* Griffin, 1950.

[25] Yakar Kannai and Bezalel Peleg. A Note on the Extension of an Order on a Set to the Power Set. *Journal of Economic Theory*, 32:172–175, 1984.

[26] Michael Kaplan and Terrence Fine. Joint orders in comparative probability. *The Annals of Probability*, 5(2):161–179, 1977.

[27] Jason Konek. Comparative probabilities. In *The Open Handbook of Formal Epistemology*, pages 267–348. PhilPapers, 2019.

[28] Bernard O. Koopman. The axioms and algebra of intuitive probability. *Annals of Mathematics*, 41(2):269–292, 1940. doi: 10.2307/1969003.

[29] Charles Kraft, John Pratt, and A. Seidenberg. Intuitive probability on finite sets. *Annals of Mathematical Statistics*, 30(2): 408–419, 1959.

[30] Isaac Levi. *The Enterprise of Knowledge.* MIT Press, 1980.

[31] Hsin-Hsien Liu and Andrew M. Colman. Ambiguity aversion in the long run: Repeated decisions under riskand uncertainty. *Journal of Economic Psychology*, 30:277–284, 2009.

[32] Enrique Miranda and Sébastien Destercke. Extreme points of the credal sets generated by comparative probabilities. *Journal of Mathematical Psychology*, 64–65:44–57, 2015. doi: 10.1016/j.jmp.2014.11.004.

[33] Enrique Miranda, Ignacio Montes and Paolo Vicig. On the selection of an optimal outer approximation of a coherent lower probability. *Fuzzy Sets and Systems*, to appear. 2021. doi: 10.1016/j.fss.2021.02.011.

[34] Enrique Miranda and Marco Zaffalon. Coherence graphs. *Artificial Intelligence*, 173(1):104–144, 2009. doi: 10.1016/j.artint.2008.09.001.

[35] Ignacio Montes, Enrique Miranda, and Paolo Vicig. 2-monotone outer approximations of coherent lower probabilities. *International Journal of Approximate Reasoning*, 101:181–205, 2018.

[36] Ignacio Montes, Enrique Miranda, and Paolo Vicig. Outer approximating coherent lower probabilities with belief functions. *International Journal of Approximate Reasoning*, 110:1–30, 2019.

[37] Stefano Moretti and Meltem Oztürk. Ordinal power relations and social choice theory. *Sixth International Workshop on Computational Social Choice*, June 2016.

[38] Maria Silvia Pini, Francesca Rossi, Kristen Brent Venable and Toby Walsh. Aggregating partially ordered preferences. *Journal of Logic and Computation*, 19(3): 475–502 (2009)

[39] Giuliana Regoli. Comparative probability and robustness. *Lecture Notes-Monograph Series*, 29:343–352, 1996.

[40] David Ríos Insua. On the foundations of decision making under partial information. *Theory and Decision*, 33:83–100, 1992.

[41] Leonard J. Savage. *The Foundations of Statistics.* Wiley, 1954.

[42] Dana Scott. Measurement structures and linear inequalities. *Journal of Mathematical Psychology*, 1:233–247, 1964.

[43] Teddy Seidenfeld, Mark Schervisch, and Jay Kadane. *Acting and reflecting*, chapter Decisions without ordering, pages 143–170. Springer, 1990.

[44] Glenn Shafer. *A Mathematical Theory of Evidence.* Princeton University Press, Princeton, New Jersey, 1976.

[45] Lloyd S. Shapley. A value for n-person games. *A*nnals of Mathematical Studies, 28, 307–317, 1953.

[46] Patrick Suppes. The measurement of belief. *Journal of the Royal Statistical Society. Series B (Methodological)*, 36:160–191, 1974.

[47] Peter Walley. *The elicitation and aggregation of beliefs.* Statistics research report 23, University of Warwick, 1982.

[48] Peter Walley. *Statistical Reasoning with Imprecise Probabilities.* Chapman and Hall, 1991.

[49] Peter Walley and Terrence Fine. Varieties of modal (classificatory) and comparative probability. *Synthese*, 41:321–374, 1979.

[50] Anton Wallner. Extreme points of coherent probabilities in finite spaces. *International Journal of Approximate Reasoning*, 44(3): 339–357, 2007. doi: 10.1016/j.ijar.2006.07.017.


## Appendix: Proofs

*Proof of Proposition 2 (i).* By the condition of the statement, there are positive real numbers $\lambda_1, \ldots, \lambda_k > 0$, judgements $(A_{j_1}, B_{j_1}), \ldots, (A_{j_k}, B_{j_k}) \in \mathscr{C} \setminus \{(A, B)\}$ and $h \in \mathcal{L}^+$ such that $\mathbb{I}_A - \mathbb{I}_B = h + \sum_{i=1}^{k} \lambda_i(\mathbb{I}_{A_{j_i}} - \mathbb{I}_{B_{j_i}})$.

Then for any $p \in \mathscr{M}_{\mathscr{C}\setminus\{(A,B)\}}$,

$$\begin{aligned}\sum_{x\in A} p(x) - \sum_{x\in B} p(x) \\ &= \sum_{x\in\mathscr{X}} h(x)p(x) + \sum_{i=1}^{k} \lambda_i \left( \sum_{x\in A_{j_i}} p(x) - \sum_{x\in B_{j_i}} p(x) \right) \\ &\geq 0,\end{aligned}$$

and therefore $p \in \mathscr{M}_{\mathscr{C}}$. The converse inclusion is trivial. □

*Proof of Proposition 2 (ii).* $\underline{P}_{\mathscr{C}}$, the lower envelope of $\mathscr{M}_{\mathscr{C}}$, is the natural extension [48, Section 3.1] of the lower prevision $\underline{P}'_{\mathscr{C}}$ on $\mathcal{K}_{\mathscr{C}}$ given by $\underline{P}'_{\mathscr{C}}(\mathbb{I}_A - \mathbb{I}_B) = 0$ for all $(A,B) \in \mathscr{C}$. If there is some $(A,B) \in \mathscr{C}$ such that $\underline{P}_{\mathscr{C}}(\mathbb{I}_A - \mathbb{I}_B) > 0$, this means that $\underline{P}'_{\mathscr{C}}$ is not coherent [48, Theorem 3.1.2], because it does not coincide with its natural extension. Applying [48, Theorem 3.1.1], this means that there are positive real numbers $\mu, \lambda_1, \ldots, \lambda_k > 0$ and judgements $(A_{j_1}, B_{j_1}), \ldots, (A_{j_k}, B_{j_k}) \in \mathscr{C} \setminus \{(A,B)\}$ such that

$$\mathbb{I}_A - \mathbb{I}_B - \mu \geq \sum_{i=1}^{k} \lambda_i (\mathbb{I}_{A_{j_i}} - \mathbb{I}_{B_{j_i}}).$$

However, this implies that $\mathbb{I}_A - \mathbb{I}_B$ belongs to $\operatorname{posi}(\mathcal{L}^+ \cup \mathcal{K}_{\mathscr{C}} \setminus \{\mathbb{I}_A - \mathbb{I}_B\})$, a contradiction. □

*Proof of Proposition 3.* The proof is analogous to that of Wallner [50]. He bounds the number of extreme points of the credal set corresponding to a coherent lower probability $L$. This credal set is bounded by $n^2 - 1$ inequality constraints of the form $\sum_{x\in A} p(x) \geq L(A)$, with $A$ a subset of $\mathscr{X}$, and the normalisation equality constraint. An extreme point of the credal set is determined by $n$ of these inequality constraints, and the corresponding vectors $\mathbb{I}_{A_1}$, …, $\mathbb{I}_{A_n}$ then form a basis for $\mathbb{R}^n$. Different extreme points have incompatible bases—that is, the interior of their convex hulls is disjoint—and he then bounds the number of incompatible bases with $\{0,1\}$-valued basis vectors through the volume of their convex hull: the volume of the convex hull of such a basis is at least $1/n!$, and the volume of the union of the interiors of these bases is bounded above by the volume of $[0,1]^n$—so 1. In our case, we have incompatible bases with $\{-1,0,1\}$-valued basis vectors—that is, one of $p(x) \geq 0$, $\sum_{x\in\mathscr{X}} p(x) = 1$ or $\sum_{x\in A} p(x) \geq \sum_{x\in B} p(x)$—and the volume is bounded by the volume of $[-1,1]^n$ instead of that of $[0,1]^n$, which is $2^n$ instead of 1.

Alternatively, the result also follows from [12, Theorem 4.4], considering that the polytope $Q_A$ defined in that proof satisfies $Q_A \subseteq [-1,1]^n$ and so $\mathcal{V}(Q_A) \leq \mathcal{V}([-1,1]^n) = 2^n$. □

*Proof of Proposition 4.* Note that the judgment $(A,B)$ corresponds to partitioning the possibility space $\mathscr{X}$ in three pairwise disjoint parts, namely $A$, $B$ and $\mathscr{X} \setminus (A \cup B)$, where only the third part can be empty. Clearly, this is equivalent to labelling each atom $x$ of the possibility space with one of the three parts, where we have to ensure that the first two parts are non-empty. Labelling the atoms with one of the three parts can be done in $3^n$ ways. This includes three cases that are invalid: (i) 1 instance of the form $(\emptyset, \emptyset, \mathscr{X})$, (ii) $2^n - 1$ instances of the form $(\emptyset, A, A^c)$ with $A \neq \emptyset$, and (iii) $2^n - 1$ instances of the form $(A, \emptyset, A^c)$ with $A \neq \emptyset$. Therefore,

$$m \leq 3^n - 1 - 2^n + 1 - 2^n + 1 = 3^n - 2^{n+1} + 1. \qquad \square$$

*Proof of Lemma 5.* Let $x$ be a node without predecessors—that is, without any incoming edges. By construction of the digraph $\mathcal{G}_{\mathscr{C}}$, this implies that $x$ is a non-auxiliary node—that is, an atom of the possibility space $\mathscr{X}$—and that $x \notin B_i$ for all $i \in \{1, \ldots, m\}$. One can now immediately verify that the probability mass function $p := \mathbb{I}_x$—that is, the degenerate probability mass function on $x$—satisfies all judgements, and therefore belongs to the comparative credal set $\mathscr{M}_{\mathscr{C}}$. Consequently, $\mathscr{M}_{\mathscr{C}}$ is non-empty. □

*Proof of Lemma 6.* Assume first of all that $x$ is a node without predecessors. From the proof of Lemma 5 that $p$ satisfies all comparisons, so it belongs to the comparative credal set $\mathscr{M}_{\mathscr{C}}$. To verify that $p$ is an extreme point, we observe that (i) $\langle p, \mathbb{I}_{\mathscr{X}} \rangle = 1$ and, for all $y \in \mathscr{X} \setminus \{x\}$, $\langle p, \mathbb{I}_y \rangle = p(y) = 0$; and (ii) $\mathbb{I}_{\mathscr{X}}$ and $\mathbb{I}_y$ with $y \in \mathscr{X} \setminus \{x\}$ are $n$ linearly independent vectors because

$$\mathbb{I}_x = \mathbb{I}_{\mathscr{X}} - \sum_{y\in\mathscr{X}\setminus\{x\}} \mathbb{I}_y.$$

Conversely, if $x$ has a predecessor $y$ in the graph, it follows that $p$ satisfies $p(y) = 0 < p(x) = 1$, and therefore it does not belong to the comparative credal set $\mathscr{M}_{\mathscr{C}}$. As a consequence, it cannot be an extreme point of this set, either. □

*Proof of Proposition 7.* If the digraph $\mathcal{G}_{\mathscr{C}}$ is connected—that is, has a single connected component—then the stated is trivially true. Therefore, we assume that the digraph $\mathcal{G}_{\mathscr{C}}$ has at least two connected components. Our proof is one by contradiction. Fix some extreme point $p$ of $\mathscr{M}_{\mathscr{C}}$, and assume *ex absurdo* that its support $\mathscr{X}_p := \{x \in \mathscr{X} : p(x) > 0\}$ intersects with more than one of the respective supports $\mathscr{X}_1, \ldots, \mathscr{X}_k$ of the connected components, $\mathcal{G}_1, \ldots, \mathcal{G}_k$ with $k \geq 2$. For any $i \in \{1, \ldots, k\}$, we construct a new probability mass function $p_i$ from the extreme point $p$ by setting the mass outside of $\mathscr{X}_i$ to zero and renormalising:

$$p_i \colon \mathscr{X} \to [0,1] \colon x \mapsto p_i(x) := \begin{cases} \dfrac{p(x)}{\alpha_i} & \text{if } x \in \mathscr{X}_i, \\ 0 & \text{otherwise} \end{cases}$$

with $\alpha_i := \sum_{y\in\mathscr{X}_i\cap\mathscr{X}} p(y)$. Observe that, by construction, $p$ is a convex combination of $p_1, \ldots, p_k$ with weights $\alpha_1, \ldots, \alpha_k$. Next, we observe that $p_1, \ldots, p_k$ all satisfy the comparative assessments, so they belong to the comparative credal set $\mathscr{M}_{\mathscr{C}}$. To verify this, it suffices to observe that for any comparative assessment $(A, B) \in \mathscr{C}$, the events $A$ and $B$ are subsets of (the set of nodes of) the same connected component by construction of the digraph $\mathcal{G}_{\mathscr{C}}$. Summarising, we have shown that $p$ is a convex combination of $k \geq 2$ elements of the comparative credal set $\mathscr{M}_{\mathscr{C}}$, and as a consequence it cannot be an extreme point. □

*Proof of Proposition 8.* Fix some $p \in \mathscr{M}_{\mathscr{C}}$. Observe that $A_1 \succeq A_2 \succeq \cdots \succeq A_k \succeq A_{k+1} = A_1$ implies that

$$\sum_{x\in A_1} p(x) \geq \sum_{x\in A_2} p(x) \geq \cdots \geq \sum_{x\in A_k} p(x) \geq \sum_{x\in A_1} p(x),$$

so $\sum_{x\in A_i} p(x) = \sum_{x\in A_j} p(x)$ for all $i, j$ in $\{1, \dots, k\}$. Thus, if $i$ and $j$ are indices in $\{1, \dots, k\}$ such that $A_j \subset A_i$, then clearly

$$\sum_{x\in A_i} p(x) = \sum_{x\in A_j} p(x) + \sum_{x\in A_i\setminus A_j} p(x) = \sum_{x\in A_j} p(x).$$

Because the probability mass function $p$ is non-negative, we infer from the second equality that $p(x) = 0$ for all $x$ in $A_i \setminus A_j$. □

*Proof of Proposition 10.* We first prove the necessity. As the digraph $\mathcal{G}_{\mathscr{C}}$ is acyclic, it is well-known—see for instance [21, Proposition 4.1]—that we can fix an ordering $\nu_1, \dots, \nu_{n+m}$ of its nodes such that the start of any directed edge is ordered before the end of this directed edge. Due to the way the digraph $\mathcal{G}_{\mathscr{C}}$ is constructed, this clearly implies the existence of the ordering in the statement.

Next, we prove the sufficiency. Fix any directed path in the digraph $\mathcal{G}_{\mathscr{C}}$, and some atom $x_i$ that is along this path. Due to the ordering and the condition on the comparisons, it is clear that the next atom $x_j$ that is on the path has a higher index according to the ordering: $j \geq i$. This prevents the path from being a cycle; as the directed path was arbitrary, we infer from this that the graph is acyclic. □

*Proof of Proposition 11.* Since $\overline{\mathscr{M}^{>}_{\mathscr{C}}} = \mathscr{M}_{\mathscr{C}}$, it follows that

$$\mathscr{M}^{>}_{\mathscr{C}} \subseteq \operatorname{int}(\overline{\mathscr{M}^{>}_{\mathscr{C}}}) = \operatorname{int}(\mathscr{M}_{\mathscr{C}}).$$

Conversely, let $p$ be a mass function in $\operatorname{int}(\mathscr{M}_{\mathscr{C}})$. Then there is some $\epsilon > 0$ such that any $q \in \Sigma_{\mathscr{X}}$ at a distance $\epsilon$ from $p$ also belongs to $\mathscr{M}_{\mathscr{C}}$. Assume *ex absurdo* that $p$ does not belong to $\mathscr{M}^{>}_{\mathscr{C}}$. There there must be some $(A, B) \in \mathscr{C}$ such that $\sum_{x\in A} p(x) = \sum_{x\in B} p(x)$. As a consequence, $\sum_{x\in B} p(x) \leq 1/2$, meaning that we can find some $q \in \Sigma_{\mathscr{X}}$ such that

$$\sum_{x\in B} q(x) > \sum_{x\in B} p(x), \ \sum_{x\in A} q(x) \leq \sum_{x\in A} p(x)$$

and

$$(\forall x \in \mathscr{X}) \ |p(x) - q(x)| < \frac{\epsilon}{n}.$$

For example, in case $\sum_{x\in A} p(x) > 0$, we can take any $x_1 \in B$ and $x_2 \in A$ with $p(x_2) > 0$, and define $q$ as

$$q(x) = \begin{cases} p(x) + \frac{\epsilon}{K} & \text{if } x = x_1 \\ p(x) - \frac{\epsilon}{K} & \text{if } x = x_2 \\ p(x) & \text{otherwise} \end{cases}$$

where $K > n$ is large enough so that $p(x_2) - \frac{\epsilon}{K} \geq 0$ and $p(x_1) + \frac{\epsilon}{K} \leq 1$ (i.e., $K > \max\{n, \frac{\epsilon}{p(x_2)}, \frac{\epsilon}{1-p(x_1)}\}$).

As a consequence, we obtain that

$$\sum_{x\in A} q(x) < \sum_{x\in B} q(x),$$

so $q \notin \mathscr{M}_{\mathscr{C}}$. However, because the Euclidean distance between $p$ and $q$ is less than $\epsilon$, $q$ belongs to $\mathscr{M}_{\mathscr{C}}$, which is a contradiction. □

*Proof of Proposition 12.* That (i) and (ii) are equivalent was already stated in [43, Theorem 1]; it can also be derived using the Separation Theorem in [48, Appendix E1], taking into account that our possibility space is finite and that we are considering a finite set $\mathcal{K}$.

To see that (i) implies (iii), note that any probability mass function $p$ in $\mathscr{M}^{>}_{\mathscr{C}}$ is dominated by the upper prevision $\overline{P}$, in the sense that for all $i \in \{1, \dots, m\}$,

$$0 < \sum_{x\in A_i} p(x) - \sum_{x\in B_i} p(x) \leq \overline{P}(\mathbb{I}_{A_i} - \mathbb{I}_{B_i}).$$

Conversely, if $\overline{P}(\mathbb{I}_{A_i} - \mathbb{I}_{B_i}) > 0$ for all $i \in \{1, \dots, m\}$, then it follows from coherence that for every $i \in \{1, \dots, m\}$, there is some $p_i \in \mathscr{M}_{\mathscr{C}}$ such that $\sum_{x\in A_i} p_i(x) - \sum_{x\in B_i} p_i(x) > 0$ and

$$\sum_{x\in A_i} p_i(x) - \sum_{x\in B_i} p(x) \geq 0 \quad \text{for all } j \in \{1, \dots, m\} \setminus \{i\}.$$

If we define $p \coloneqq (p_1 + \cdots + p_m)/m$, then it is a probability mass function that belongs to the convex set $\mathscr{M}_{\mathscr{C}}$ such that for all $i$ in $\{1, \dots, m\}$,

$$\sum_{x\in A_i} p(x) - \sum_{x\in B_i} p(x) = \sum_{j=1}^{m} \frac{1}{m}\Big(\sum_{x\in A_i} p_j(x) - \sum_{x\in B_i} p_j(x)\Big) > 0.$$

From this, we infer that $p \in \mathscr{M}^{>}_{\mathscr{C}}$; consequently, $\mathscr{M}^{>}_{\mathscr{C}}$ is non-empty. □

*Proof of Proposition 13.* Because $\mathcal{G}_{\mathscr{C}}$ has no cycles, it follows from Proposition 10 that we can fix an ordering $x_1, \dots, x_n$ of $\mathscr{X}$ such that for all $i \in \{1, \dots, m\}$, there is a $j \in \{1, \dots, n-1\}$ such that $A_i \subseteq \{x_1, \dots, x_j\}$ and $B_i \subseteq \{x_{j+1}, \dots, x_n\}$. Therefore, any probability mass function $p$ satisfying

$$p(x_i) > \sum_{j=i+1}^{n} p(x_j) \quad \text{for all } i \in \{1, \dots, n-1\}$$

shall belong to $\mathscr{M}_{\mathscr{C}}^{>}$. Thus, we are looking for a probability mass function $p$ such that

$$\begin{aligned}
p(x_1) &> 1 - p(x_1) \\
p(x_2) &> (1 - p(x_1)) - p(x_2) \\
p(x_3) &> (1 - p(x_1) - p(x_2)) - p(x_3) \\
&\vdots \\
p(x_{n-1}) &> (1 - p(x_1) - \cdots - p(x_{n-2})) - p(x_{n-1})
\end{aligned}$$

or equivalently,

$$\begin{aligned}
&2p(x_1) > 1 \\
&2p(x_2) + p(x_1) > 1 \\
&2p(x_3) + p(x_2) + p(x_1) > 1 \\
&\vdots \\
&2p(x_{n-1}) + p(x_{n-2}) + \cdots + p(x_1) > 1
\end{aligned}$$

One probability mass function that satisfies these inequalities is the one defined by $p(x_n) = 0$ and, for all $i \in \{1, \ldots, n-1\}$, by $p(x_i) = \frac{1}{2^i} + \frac{1}{(n-1)2^{n-1}}$. □

*Proof of Proposition 14.* Any $p \in \mathscr{M}_{\mathscr{C}}$ can be expressed as

$$p = \text{comb}(q, q_1, \ldots, q_k),$$

where $q$ is a probability mass function on $\{C_1, \ldots, C_k\}$ defined by $q(C_i) = \sum_{x \in C_j} p(x)$ for all $i \in \{1, \ldots, k\}$, and where for all $i \in \{1, \ldots, k\}$, $q_i$ is the probability mass function on $C_i$ defined by

$$q_i(x) = \begin{cases} \frac{p(x)}{\sum_{y \in C_i} p(y)} & \text{if } \sum_{x \in C_i} p(x) > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Moreover, it follows that in the construction above $q \in \mathscr{M}'_{\text{el}}$ and $q_i \in \mathscr{M}_{\text{el},i}$ for each $i \in \{1, \ldots, k\}$.

Taking this into account, the result is now immediate, since (i) $p$ cannot be an extreme point of $\mathscr{M}_{\mathscr{C}}$ if either $q$ can be expressed as a convex combination of other elements of $\mathscr{M}'_{\text{el}}$ or, for some $i$ such that $q(C_i) > 0$, it holds that $q_i$ is not an extreme point of $\mathscr{M}_{\text{el},i}$; and, conversely, (ii) any combination of extreme points in the respective credal sets will produce a probability mass function that cannot be written as a convex combination of other elements of $\mathscr{M}_{\mathscr{C}}$, and is therefore an element of $\text{ext}(\mathscr{M}_{\mathscr{C}})$. □

*Proof of Corollary 15.* For all $i \in \{1, \ldots, k\}$, we let $n_i$ denote the cardinality of $C_i$. It follows from [32, Theorem 4] that $|\text{ext}(\mathscr{M}'_{\text{el}})| \leq 2^{k-1}$ and, for all $i \in \{1, \ldots, k\}$, $|\text{ext}(\mathscr{M}_{\text{el},i})| \leq 2^{n_i - 1}$. Applying Proposition 14, we deduce that

$$\begin{aligned}
|\text{ext}(\mathscr{M}_{\mathscr{C}})| &\leq |\text{ext}(\mathscr{M}'_{\text{el}})| \prod_{i=1}^{k} |\text{ext}(\mathscr{M}_{\text{el},i})| \\
&\leq 2^{k-1} \prod_{i=1}^{k} 2^{n_i - 1} = 2^{n-1},
\end{aligned}$$

taking into account that $n_1 + \cdots + n_k = n$. □

*Proof of Lemma 16.* Let $n_p := |\mathscr{X}_p|$. We start our proof by collecting all bounds on the credal set—i.e., convex polytope—$\mathscr{M}_{\mathscr{C}}$ that are saturated by the extreme point $p$. Let

$$\mathcal{B}_p := \{\mathbb{I}_x : x \in \mathscr{X}, \langle p, \mathbb{I}_x \rangle = 0\} = \{\mathbb{I}_x : x \in \mathscr{X}_p^{\text{c}}\}$$

and

$$\mathcal{C}_p := \{\mathbb{I}_A - \mathbb{I}_B : (A, B) \in \mathscr{C}, \langle p, \mathbb{I}_A - \mathbb{I}_B \rangle = 0\};$$

these are the vectors associated with the constraints $p(x) = 0$ for $x \in \mathscr{X}$ and $\sum_{x \in A} p(x) = \sum_{x \in B} p(x)$ for $(A, B) \in \mathscr{C}$ that are satisfied by $p$.

Recall that, as $p$ is an extreme point, there are $n-1$ vectors $v_1, \ldots, v_{n-1}$ in $\mathcal{B}_p \cup \mathcal{C}_p$ such that $v_1, \ldots, v_{n-1}, \mathbb{I}_{\mathscr{X}}$ are linearly independent.

To determine which vectors these are, we partition $\mathcal{C}_p$ in two parts:

$$\mathcal{C}_1 := \{\mathbb{I}_A - \mathbb{I}_B \in \mathcal{C}_p : A \cap \mathscr{X}_p = \emptyset = B \cap \mathscr{X}_p\}$$

and

$$\mathcal{C}_2 := \{\mathbb{I}_A - \mathbb{I}_B \in \mathcal{C}_p : A \cap \mathscr{X}_p \neq \emptyset \neq B \cap \mathscr{X}_p\},$$

considering that if $\sum_{x \in A} p(x) = \sum_{x \in B} p(x)$, then either both have strictly positive probability or both have zero probability. The first one represents those constraints that are trivially saturated because both $A, B$ have probability zero, while the second is given by those $(A, B) \in \mathscr{C}$ with strictly positive probability in $p$.

Note that the support of all of the vectors in $\mathcal{B}_p$ and $\mathcal{C}_1$ is a subset of $\mathscr{X}_p^{\text{c}}$, so there are at most $|\mathscr{X}_p^{\text{c}}| = n - n_p$ linearly independent vectors in $\mathcal{B}_p \cup \mathcal{C}_1$. Therefore, at least $n - 1 - (n - n_p) = n_p - 1$ of the vectors $v_1, \ldots, v_{n-1}$ should belong to $\mathcal{C}_2$. However, $\mathcal{C}_2$ can contain at most $n_p - 1$ vectors because (i) the digraph $\mathcal{G}_{\mathscr{C}}$ is acyclic; and (ii) the assessment $\mathscr{C}$ is singular, meaning that for every atom $x$ in the support $\mathscr{X}_p$, there is at most one $(A, B)$ in $\mathscr{C}$ such that $x \in B$. In conclusion, $\mathcal{C}_2$ consists of precisely $n_p - 1$ vectors, and all of these are in the set of $n$ linearly independent vectors determining the extreme point $p$.

Every vector in $\mathcal{C}_2$ corresponds to a unique judgement, and the auxiliary node that corresponds to this judgement clearly has at least one direct predecessor in the support and at least on direct successor in the support. Observe furthermore that for any atom $x$ in the support $\mathscr{X}_p$ that has a direct predecessor, this direct predecessor is an auxiliary node that must also have a direct predecessor that is an atom in the support $\mathscr{X}_p$. Because (a) there are $n_p - 1$ vectors in $\mathcal{C}_2$, (b) there are $n_p$ atoms in the support $\mathscr{X}_p$, (c) every atom has at most one predecessor because the assessment $\mathscr{C}$ is singular, and (d) the digraph $\mathcal{G}_\mathscr{C}$ is acyclic, we infer from this that there is precisely one atom $x^\star$ in the support $\mathscr{X}_p$ that has no predecessors, as required. This proves the first part of the statement.

Next, we prove the second part of the statement. Consider the subdigraph $\mathcal{G}'$ of $\mathcal{G}_\mathscr{C}$ that consists of (the nodes corresponding to the) atoms in the support $\mathscr{X}_p$ and, for each of the $n_p - 1$ vectors in $\mathcal{C}_2$, the corresponding auxiliary node and the two directed edges that connect it to its unique immediate successor and predecessor in the support. Then by construction, this subdigraph $\mathcal{G}'$ is an extreme arborescence with root $x^\star$ and atoms $\mathscr{X}_p$. This proves the second part of the statement. □

*Proof of Lemma 17.* If $C$ is a singleton, then this follows from Lemma 6. Hence, from here on we assume that $n_C := |C| > 1$. Let $x_1, \dots, x_n$ be the order of the atoms that follows from Proposition 10; without loss of generality, we may assume that the root $x^\star$ of the extreme arborescence with atoms $C$ has index 1. Furthermore, we let $\sigma\colon \{1, \dots, n_C\} \to \{1, \dots, n\}$ be the increasing function such that

$$C = \{x_{\sigma(1)}, \dots, x_{\sigma(n_C)}\}.$$

Because $\mathscr{C}$ is singular and the atoms in $C$ are the atoms of an extreme arborescence, it is clear that for all $i \in \{2, \dots, n_C\}$ there is a unique $\mu(i)$ in $\{1, \dots, m\}$ such that $x_{\sigma(i)} \in B_{\mu(i)}$, and $B_{\mu(i)} \cap C = \{x_{\sigma(i)}\}$. Because $\sigma$ is increasing, we have also ensured that

$$(\forall i \in \{2, \dots, n_C\})\ A_{\mu(i)} \cap C \subseteq \{x_{\sigma(1)}, \dots, x_{\sigma(i-1)}\}. \quad (5)$$

We first explicitly construct a probability mass function $p$ that has $C$ as support and saturates the comparative assessments associated with the pairs $(A_{\mu(2)}, B_{\mu(2)})$, …, $(A_{\mu(n_C)}, B_{\mu(n_C)})$. To that end, we define $q\colon \mathscr{X} \to \mathbb{Z}_{\geq 0}$ as $q(x) := 0$ for all $x \in C^c$. Furthermore, we let $q(x_{\sigma(1)}) = 1$ and

$$q(x_{\sigma(i)}) := \sum_{y \in A_{\mu(i)}} q(y) \quad \text{for all } i \in \{2, \dots, n_C\},$$

such that for all $i \in \{2, \dots, n_C\}$,

$$\sum_{x \in A_{\mu(i)}} q(x) = \sum_{y \in B_{\mu(i)}} q(y) = q(x_{\sigma(i)}),$$

because $B_{\mu(i)} \cap C = \{x_{\sigma(i)}\}$. Finally, we let

$$p\colon \mathscr{X} \to [0,1]\colon x \mapsto p(x) := \frac{q(x)}{\sum_{y \in \mathscr{X}} q(y)}.$$

It is easily verified that $p$ is a probability mass function with the desired properties.

Next, we verify that $p$ is an extreme point. Recall from before that this is the case if there are $n$ linearly independent vectors among the constraint vectors that are saturated. It is clear that by construction $p$ saturates the constraints with vectors $\mathbb{I}_\mathscr{X}$, $\mathbb{I}_x$ with $x \in C^c$ and $\mathbb{I}_{A_{\mu(i)}} - \mathbb{I}_{B_{\mu(i)}}$ with $i \in \{1, \dots, n_C\}$. These vectors are linearly independent if and only if for all real numbers $\lambda_1$, …, $\lambda_n$ with at least one non-zero number,

$$\lambda_1 \mathbb{I}_\mathscr{X} + \sum_{i=2}^{n_C} \lambda_i (\mathbb{I}_{A_{\mu(i)}} - \mathbb{I}_{B_{\mu(i)}}) + \sum_{i=n_C+1}^{n} \lambda_i \mathbb{I}_{x_{\sigma'(i)}} \neq 0,$$

where we let $\sigma'$ be any function from $\{n_C + 1, \dots, n\}$ to $\{1, \dots, n\}$ such that $\mathscr{X} \setminus C = \{x_{\sigma'(n_C+1)}, \dots, x_{\sigma'(n)}\}$. In other words, the above sum is not constant on zero, meaning that there is some $x \in \mathscr{X}$ such that

$$\lambda_1 + \sum_{i=2}^{n_C} \lambda_i (\mathbb{I}_{A_{\mu(i)}}(x) - \mathbb{I}_{B_{\mu(i)}}(x)) + \sum_{i=n_C+1}^{n} \lambda_i \mathbb{I}_{x_{\sigma'(i)}}(x) \neq 0.$$

Assume *ex absurdo* that there are real numbers $\lambda_1$, …, $\lambda_n$ with at least one non-zero number such that the sum above is equal to zero on all elements of $\mathscr{X}$. Because $B_{\mu(i)} \cap C = \{x_{\sigma(i)}\}$, this implies in particular that

$$\lambda_1 \mathbb{I}_C + \sum_{i=2}^{n_C} \lambda_i (\mathbb{I}_{A_{\mu(i)} \cap C} - \mathbb{I}_{x_{\sigma(i)}}) = 0. \quad (6)$$

Evaluating the indicators of Eqn. (6) in $x_{\sigma(n_C)}$ and using Eqn. (5), we find that $\lambda_{n_C} = \lambda_1$. Similarly, evaluating the indicators of Eqn. (6) in $x_{\sigma(n_C - 1)}$ and using Eqn. (5), we find that

$$\begin{aligned}\lambda_{n_C - 1} &= \lambda_1 + \lambda_{n_C} \mathbb{I}_{A_{\mu(n_C)} \cap C}(x_{\sigma(n_C - 1)}) \\ &= \lambda_1 + \lambda_1 \mathbb{I}_{A_{\mu(n_C)} \cap C}(x_{\sigma(n_C - 1)}) = k_{n_C - 1} \lambda_1,\end{aligned}$$

where $k_{n_C - 1}$ is a strictly positive integer. Continuing the same reasoning, it is easy to show that for all $i \in \{2, \dots, n_C\}$, $\lambda_i = k_i \lambda_1$ with $k_i$ a strictly positive integer. Hence, we have that

$$\lambda_1 \left( \mathbb{I}_C + \sum_{i=2}^{n_C} k_i (\mathbb{I}_{A_{\mu(i)} \cap C} - \mathbb{I}_{x_{\sigma(i)}}) \right) + \sum_{i=n_C+1}^{n} \lambda_i \mathbb{I}_{x_{\sigma'(i)}} = 0.$$

We evaluate this expression in $x_{(\sigma(1))}$, to yield

$$\lambda_1 \left( 1 + \sum_{i=2}^{n_C} k_i \mathbb{I}_{A_{\mu(i)} \cap C}(x_{\sigma(1)}) \right) = 0.$$

From this, it follows immediately that $\lambda_1 = 0$. Because $k_2$, …, $k_{n_C}$ are all strictly positive, this implies that $\lambda_2 = \dots = \lambda_{n_C} = 0$. Consequently,

$$\begin{aligned}\lambda_1 \mathbb{I}_\mathscr{X} + \sum_{i=2}^{n_C} \lambda_i (\mathbb{I}_{A_{\mu(i)}} - \mathbb{I}_{B_{\mu(i)}}) &+ \sum_{i=n_C+1}^{n} \lambda_i \mathbb{I}_{x_{\sigma'(i)}} \\ &= \sum_{i=n_C+1}^{n} \lambda_i \mathbb{I}_{x_{\sigma'(i)}} = 0,\end{aligned}$$

which can only hold if $\lambda_{n_C+1} = \dots = \lambda_n = 0$. Hence, we find that all $\lambda_i$'s have to be zero, which is the contradiction that we are looking for. □

*Proof of Theorem 18.* Sufficiency follows from Lemma 16, necessity holds due to Lemma 17. □

*Proof of Theorem 19.* Recall from Theorem 18 that the extreme points of the comparative credal set are in one to one correspondence with extreme arborescences. Therefore, any bound on the number of extreme arborescences is a bound on the number of extreme points as well. In the remainder, we will obtain a lower and upper bound on the number of extreme arborescences.

To this end, we recall from Proposition 10 that there is an ordering $x_1, \ldots, x_n$ of the atoms of the possibility space $\mathscr{X}$ such that

$$(\forall (A, B) \in \mathscr{C})(\exists i \in \{1, \ldots, n-1\})$$
$$A \subseteq \{x_1, \ldots, x_i\} \text{ and } B \subseteq \{x_{i+1}, \ldots, x_n\}.$$

For every $k$ in $\{1, \ldots, n\}$, we let $e_k$ denote the number of extreme arborescences such that the non-auxiliary nodes are a subset of $\{x_1, \ldots, x_k\}$. Because $x_1$ is a node without predecessors, it is clear that there is precisely one extreme arborescence such that its non-auxiliary nodes is $\{x_1\}$; thus, $e_1 = 1$. Next, we fix some $k$ in $\{2, \ldots, n\}$. Clearly, $e_k$ is equal to the sum of $e_{k-1}$ and $e'_k$, where $e'_k$ denotes the number of extreme arborescences such that the non-auxiliary nodes are a subset of $\{x_1, \ldots, x_k\}$ and contain $x_k$. If $x_k$ has no predecessors, then clearly $e'_k = 1$. Thus, we focus on the case that $x_k$ does have predecessors. Consider any extreme arborescence whose non-auxiliary nodes are a subset of $\{x_1, \ldots, x_n\}$ and contain $x_k$. Then removing $x_k$ (and the auxiliary node that precedes it), we obtain an extreme arborescence such that its non-auxiliary nodes are included in $\{x_1, \ldots, x_{k-1}\}$. Thus, it is clear that in this case $1 \leq e'_k \leq e_{k-1}$. Because in both cases $1 \leq e'_k \leq e_{k-1}$, we conclude that $k \leq e_k \leq 2e_{k-1}$.

Because $e_k = 1$ and $k \leq e_k \leq 2e_{k-1}$ for every $k \in \{2, \ldots, n\}$, it is clear that $n \leq e_n \leq 2^{n-1}$, as claimed. □

*Proof of Lemma 20.* Because the graph $\mathcal{G}_{\mathscr{C}}$ is an arborescence, it is acyclic and the assessment $\mathscr{C}$ is singular. Hence, it follows from Lemma 16 that (i) the support $\mathscr{X}_p$ contains the root $x^\star$ of the arborescence; and (ii) there is an extreme arborescence $\mathcal{G}'$ with root $x^\star$ and atoms $\mathscr{X}_p$. Because $\mathcal{G}_{\mathscr{C}}$ is furthermore an arborescence itself, we infer that $\mathscr{X}_p = H'(\mathscr{X}_p)$ and that the closest common predecessor of any two $x$ and $y$ in $\mathscr{X}_p$ is a non-auxiliary node. □

*Proof of Lemma 21.* The proof is a straightforward modification of that of Lemma 17. The fact that it is uniform follows from Eqn. (4), as now $A_{\mu(i)}$ is a singleton. □

*Proof of Theorem 22.* Necessity follows from Lemma 20, sufficiency from Lemma 21. □