Compatibility, desirability, and the running intersection property

Enrique Miranda[a], Marco Zaffalon[b,*]

[a] University of Oviedo, Dep. of Statistics and Operations Research. Oviedo (Spain)
[b] Istituto Dalle Molle di Studi sull'Intelligenza Artificiale (IDSIA). Lugano (Switzerland)

Abstract

Compatibility is the problem of checking whether some given probabilistic assessments have a common joint probabilistic model. When the assessments are unconditional, the problem is well established in the literature and finds a solution through the *running intersection property* (RIP). This is not the case of conditional assessments. In this paper, we study the compatibility problem in a very general setting: any possibility space, unrestricted domains, imprecise (and possibly degenerate) probabilities. We extend the unconditional case to our setting, thus generalising most of previous results in the literature. The conditional case turns out to be fundamentally different from the unconditional one. For such a case, we prove that the problem can still be solved in general by RIP but in a more involved way: by constructing a junction tree and propagating information over it. Still, RIP does not allow us to optimally take advantage of sparsity: in fact, conditional compatibility can be simplified further by joining junction trees with *coherence graphs*.

Keywords: Compatibility; coherence; marginal problem; conditional models; probabilistic satisfiability; running intersection property; junction trees; coherence graphs; imprecise probability; coherent sets of desirable gambles.

1. Introduction

What is compatibility?

The marginal problem

Suppose we are given a few marginal probability functions over some variables: e.g., $P_1(X_1, X_2)$, $P_2(X_2, X_3)$, $P_3(X_3, X_4, X_5)$. We wonder whether there is a joint probability $P(X_1, X_2, X_3, X_4, X_5)$ from which we can reproduce P_1, P_2, P_3 by marginalisation.

This is an example of the so-called *marginal problem*: that of the compatibility of a number of marginal assessments with a global model. This problem has received a long-standing interest in the literature, since the seminal works by Boole [14], Hoeffding [44], Fréchet [34], Kellerer [51] and Vorobev [88] (see also [20] and the references therein).

The problem is trivial when the marginal models are defined on disjoint sets of variables: in that case, we could for instance determine a compatible joint model by considering the stochastic product of the marginals. However, when those sets of variables are not disjoint, then the problem is not trivial anymore. More recent work on this problem investigated when some additional constraints are placed on the joint in [76, 80], and has also appeared in other, apparently far, contexts, such as quantum mechanics [35] or coalitional game theory [88]. It has also a very nice application in problems of polynomial optimisation, where it can dramatically reduce the computational complexity of solution algorithms by exploiting sparsity in the problem representation [56].

Obviously, a necessary condition for the compatibility of a number of marginal assessments is their *pairwise compatibility*, that is, the equality of the marginals over common variables; in our example, this requires that

$$P_1(X_2) = P_3(X_2) \text{ and } P_2(X_3) = P_3(X_3). \tag{1}$$

*Corresponding author
Email addresses: mirandaenrique@uniovi.es (Enrique Miranda), zaffalon@idsia.ch (Marco Zaffalon)

This in not enough however. In fact, using the theory of hypergraphs, Beeri et al. [5] (see also [60]) established a necessary and sufficient condition for pairwise compatibility to imply global compatibility: the *running intersection property* (RIP).[1] This requires the existence of a total order on the marginals such that if any two marginals have variables in common, then also all the marginals between them in the order contain those variables too. In our example the natural order P_1, P_2, P_3 makes it. Therefore Eq. (1) being true makes sure that a compatible P exists. There could actually be more than one; the *iterative proportional fitting procedure* (IPFP)[29] yields a sequence of probabilities that converge to the compatible joint that maximises Kullback-Leibler information [19].

The works above investigate the compatibility of probabilities; when the possibility spaces are infinite, they are usually assumed to be countably additive on a suitable σ-field. Another direction of generalisation takes into account the possible partial specification of probabilities: for instance say that P_1, P_2, P_3 in the example are only partly known; this corresponds to replacing each of them with a set of candidate probabilities. The marginal problem then becomes checking whether there is a set of joint probabilities P from which we can recover the marginal (candidate) sets by marginalisation.

Set-based probabilistic modelling goes under the umbrella term of *imprecise probabilities* [4]. They include models of possibility measures [32], belief functions [77] or coherent lower previsions [89], among others. The marginal problem has been investigated for some of these models by Studený [82, 83], Vejnarová [86] and Jirousek [49], using the IPFP; van der Gaag [36] has dealt with it by propagating inequality constraints over a tree.

The compatibility problem

The marginal problem has a generalisation to the conditional case that we shall just call the *compatibility problem*. In this case we have any number of conditional probabilities over a set of variables and the problem is again to verify whether they have a compatible joint.

Instances of the compatibility problem have shown up in Artificial Intelligence in the research concerned with probabilistic logic and probabilistic satisfiability [38, 41, 43, 46, 71]; in these cases the focus is on variables with finite support (or just events) and solutions algorithms are often based on linear programming—yet probabilistic satisfiability is NP-hard [12]. Another approach to satisfiability, originated within de Finetti's school, is based on 'full conditional measures' [17, 31]; this model establishes links between conditional probabilities so as to avoid inconsistencies, and can equivalently be represented as 'zero layers' à la Krauss [54]. This allows in particular to deal with structural constraints (also called structural zeroes) between conditional probabilities via sequences of linear programs. With similar aims and properties, Walley et al. [91] have addressed a generalised version of probabilistic satisfiability that mixes conditional and unconditional information, that allows the assessments to be imprecisely specified, and that is not affected by problems due to zero probabilities.

Note in fact that compatibility needs Bayes' rule to be verified besides the simple use of marginalisation. But Bayes' rule is not applicable in the case of zero-probability events. Neglecting this issue can lead to overlook incompatibilities that 'hide' under these zero probabilities. The problem can eventually yield wrong inferences and it is particularly subtle as it is generally unknown in advance where those zero probabilities happen to be. Cozman and Ianni [18] have recently proposed an approach that builds on Walley et al.'s work and that, as such, correctly deals with these problems.

In a different direction, ten years ago we have observed that the compatibility problem, as well as probabilistic satisfiability, can often be simplified taking sparsity into account through a graphical representation called *coherence graphs* [64, Sections 8.2–8.3].

Compatibility is such a general problem that has a life on its own also in the statistical literature. There we can find some early work by Strassen [81], Okner [72] and Kamakura and Wedel [50], and a great bulk of work made by Arnold et al. [2, 1, 3] that also consider the case of imprecise information. Kuo and Wang [93] have shown that the problem of zero probability is an issue also in the statistical case; in the same year we also have discussed the same question in the statistical literature [65]. In addition, we have proved that there is an iterative procedure that converges to the compatible joint model; this is somewhat similar in spirit to the IPFP, but our procedure works for the more involved conditional case and moreover it yields the entire set of compatible probabilities in the case of imprecision. While most work on compatibility focuses on discrete variables, Wang and Ip [92] are a relevant reference for the continuous case. Kuo et al. [55] provide one of the most recent works on the subject, with many references therein.

[1] In the same year, Lemmer established a condition (a special case of RIP) that, given pairwise compatibility, is sufficient for global compatibility [59, Section 4.2].

So there has been much work about compatibility in the conditional case across different communities (that do not seem to have talked much to each other). However, and to our surprise, we could not find any work making the connection to RIP there, which is even more surprising considered the clear connection that exists with RIP in the unconditional case.

Outline of the paper and main results

Our aim in this paper is to establish a clear connection between RIP and compatibility in the most general possible setting: any possibility space, unrestricted domains (no σ-additivity/measurability problems), imprecise probabilities, conditional and unconditional information, no limitations due to zero probabilities.

To achieve these goals, we base our analysis on the imprecise-probability formalism of *coherent sets of desirable gambles* [89, 94]. As we have recently shown [96, 99], such a formalism is an equivalent reformulation of Bayesian decision theory, once it is freed of the precision constraint, with the advantage that it naturally meets all the requirements listed above. We introduce sets of desirable gambles in Section 2.

In the same section, we define compatibility in the unconditional case for sets of desirable gambles and prove in Theorem 2 that RIP and pairwise compatibility imply compatibility. This result generalises most of the previous work on the marginal problem along the lines discussed at the beginning of this section. We try to clarify this point by first specialising our results to sets of probabilities, and then by commenting on the relation of these results with previous ones.

We move to compatibility for the conditional case in Section 3. First, we give a generalised definition of compatibility (Definition 18). The definition makes us realise that compatibility is nothing else than *strong coherence* in Williams-Walley's theory [98, Definition 25], thus enabling us to exploit established tools in such a theory to pursue our aims. This turns out to be particularly important since we verify that the conditional case cannot be reduced to the unconditional one: in the former, compatibility does not imply pairwise compatibility; pairwise compatibility needs to be replaced by Walley's notion of *avoiding partial loss*. We go on then to specialise some of these notions for sets of probabilities.

In Section 4 we give our main results. We start by recalling the notion of *tree decomposition* related to RIP: i.e., that our probabilistic assessments can be represented graphically so as to eventually organise the variables of our problem in a *junction tree*; in such a tree, nodes are clusters of variables (cliques) that satisfy RIP. We give two procedures, analogous to the standard ones of collect and distribute evidence, for the propagation of desirable gambles over the tree. Then we prove in Theorems 9 and 10 that:

- The first procedure terminates with a coherent set at the root of the tree if and only if our original assessments avoid partial loss. This is a first test of compatibility, because if that is not the case, then the original assessments are not compatible and we can stop.
- Otherwise, the second procedure yields the marginals of the joint compatible set of desirable gambles that extends our original assessments. Then the original assessments are compatible if and only if they coincide with such marginals.

In Appendix A.4 we give also an alternative avenue to the proof of Theorems 9 and 10 based on so-called *valuation algebras* [52, 78]. These are abstract representations of knowledge or information that encode primitive tools for distributed computation on a junction tree. Valuation algebras should provide more accessible proofs of distributed computation to those unfamiliar with desirability; moreover, such an avenue has turned out to be an opportunity for us to discuss more widely the interplay of logic, desirability and algebras.

Irrespectively of the proof method, let us remark that these results, being valid for desirable gambles, hold also for sets of probabilities and in particular for traditional, precise, probability (on any possibility space).

Let us recall that in the unconditional case, RIP is often regarded as the optimal way to exploit sparsity in a problem without loss of information. We show in Section 5 that in the conditional case this is no longer true: there are very common situations where we can immediately tell if compatibility holds without having to build a junction tree and perform a propagation. We systematise this observation by leveraging on our past work on coherence graphs [64]. These simplify the verification of coherence by yielding a partition of the original set of assessments into so-called *superblocks*. Here, we extend past results on coherence graphs to desirable gambles and show in Theorem 12 that in order to check compatibility it is enough to separately check it on superblocks. In addition we give a procedure to compute the compatible joint. The lesson here is that if we want to get the best out of the conditional case, we have to combine coherence graphs with junction trees.

We give our concluding views in Section 6. Appendix A contains additional remarks and observations. All the proofs of the paper have been gathered in Appendix B.

2. Compatibility of unconditional models

2.1. Sets of desirable gambles

The most general model we shall consider in this paper is that of *coherent sets of desirable gambles*. Let us introduce the main notions about this theory; we refer to [4, Chapter 1], [90] and [89, Chapter 3] for further details.

Definition 1 (Gambles). *Consider a possibility space $\mathcal{X}$. A* gamble *on $\mathcal{X}$ is a bounded real-valued function $f : \mathcal{X} \to \mathbb{R}$.*

Gambles are interpreted as uncertain rewards in a linear utility scale. We denote by $\mathcal{L}(\mathcal{X})$ the set of all gambles on $\mathcal{X}$, and by $\mathcal{L}^+(\mathcal{X}) := \{f \in \mathcal{L}(\mathcal{X}) : f \geq 0, f \neq 0\}$ the set of positive gambles. We shall simplify the notation whenever possible by omitting the possibility space $\mathcal{X}$. Thus, we shall write $\mathcal{L}^+$ for the positive gambles and moreover use $f \gneq 0$ in place of $f \geq 0, f \neq 0$.

Definition 2 (Coherence for gambles). *A subset $\mathcal{D} \subseteq \mathcal{L}(\mathcal{X})$ is called* coherent *when it satisfies the following axioms:*

D1. $\mathcal{L}^+ \subseteq \mathcal{D}$ *[Accepting Partial Gains];*

D2. $0 \notin \mathcal{D}$ *[Avoiding Null Gain];*

D3. $f, g \in \mathcal{D} \Rightarrow f + g \in \mathcal{D}$ *[Additivity];*

D4. $f \in \mathcal{D}, \lambda > 0 \Rightarrow \lambda f \in \mathcal{D}$ *[Positive Homogeneity].*

It follows from these axioms that, if f belongs to a coherent set $\mathcal{D}$ and $g \geq f$, then also $g \in \mathcal{D}$.

Whenever a set $\mathcal{D}$ is not coherent, we can try to extend it into a coherent set by means of the following procedure:

Definition 3 (Natural extension for gambles). *Given a set $\mathcal{D} \subseteq \mathcal{L}(\mathcal{X})$, we call*

$$\mathcal{E} := \operatorname{posi}(\mathcal{L}^+ \cup \mathcal{D}), \tag{2}$$

its natural extension*, where* posi *denotes the set of positive linear combinations of the gambles in the argument.*

The natural extension of a set of desirable gambles $\mathcal{D}$ is coherent if and only if it avoids null gain. This motivates the following:

Definition 4 (Avoiding partial loss for gambles). *We say that $\mathcal{D} \subseteq \mathcal{L}(\mathcal{X})$* avoids partial loss *if and only if $0 \notin \mathcal{E}$.*

A set that avoids partial loss can always be extended to a coherent set. The natural extension is just the smallest such set; it can equivalently be represented as the intersection of all the coherent sets that include $\mathcal{D}$.

In this paper, we shall investigate the compatibility of the belief assessments that model our knowledge about different sets of variables [2] . To see how all these different assessments can be embedded into a unified framework, consider non-empty spaces $\mathcal{X}_1, \ldots, \mathcal{X}_n$. Let $N := \{1, \ldots, n\}$. For any subset S of N we shall let $\mathcal{X}_S := \bigtimes_{j \in S} \mathcal{X}_j$ and denote by x_S its generic element. We abuse this notation in two extreme cases to keep it simple: if S is a singleton we shall not write braces, so $\mathcal{X}_{\{j\}}$ will become $\mathcal{X}_j$ (and $x_{\{j\}}$ will become x_j); if $S = N$ we shall just omit it, therefore $\mathcal{X}_S$ will become $\mathcal{X}$ (and x_N will be written as x). The latter is made also to emphasise that $\mathcal{X}_N = \mathcal{X}$ is, from now on, our overall possibility space.[3]

[2] To avoid confusion between our use of the term 'variables' and traditional 'random variables', let us remark that in this paper variables should be understood simply as functions taking values in respective possibility spaces $\mathcal{X}_i$, and are essentially just a mathematical convenience. We do not use random variables in this paper, even though gambles can be thought of as playing their role in the theory of desirability. Thus, if we have two variables X_1, X_2 taking values in respective spaces $\mathcal{X}_1, \mathcal{X}_2$, uncertainty about the joint behaviour of (X_1, X_2) shall be modelled by means of a coherent set of desirable gambles in $\mathcal{L}(\mathcal{X}_1 \times \mathcal{X}_2)$.

[3] We shall thus assume that the underlying variables are *logically independent*, meaning that any value in the Cartesian product of the spaces $\mathcal{X}_1, \ldots, \mathcal{X}_n$ is assumed to be possible. For a discussion of the relevance of this hypothesis in compatibility problems, we refer to [10, Section 3.4] and to [85]. Note that the assumption of logical independence does not preclude the existence of zero probabilities.

Definition 5 (**Projection operator).** *Given spaces $\mathcal{X}_1, \dots, \mathcal{X}_n$ and any subset S of N, we denote by π_S the* projection operator*, given by*

$$\pi_S : \mathcal{X} \to \mathcal{X}_S$$
$$x \hookrightarrow (x_j)_{\{j \in S\}}.$$

Definition 6 (**Measurable gambles).** *We shall say that a gamble f on $\mathcal{X}$ is $\mathcal{X}_S$-measurable if and only if*

$$(\forall x, y \in \mathcal{X} : \pi_S(x) = \pi_S(y))\ f(x) = f(y).$$

We shall denote by $\mathcal{L}_S(\mathcal{X})$ (or simply $\mathcal{L}_S$) the subset of $\mathcal{L}(\mathcal{X})$ given by the $\mathcal{X}_S$-measurable gambles. There exists a one-to-one correspondence between $\mathcal{L}_S(\mathcal{X})$ and $\mathcal{L}(\mathcal{X}_S)$, and we will sometimes abuse the notation by writing $\mathcal{D} \cap \mathcal{L}(\mathcal{X}_S)$ when we mean $\mathcal{D} \cap \mathcal{L}_S(\mathcal{X})$ for a given set of gambles $\mathcal{D} \subseteq \mathcal{L}(\mathcal{X})$. For clarity, we shall use the notation posi_S when the natural extension is applied with respect to the set of S-measurable gambles, and use posi in case the natural extension is taken with respect to $\mathcal{L}(\mathcal{X})$. If we consider $\mathcal{D}_S \subseteq \mathcal{L}_S$, then its natural extension with respect to $\mathcal{L}_S$ is given by

$$\mathrm{posi}_S(\mathcal{L}_S^+ \cup \mathcal{D}_S) = \mathrm{posi}(\mathcal{L}^+ \cup \mathcal{D}_S) \cap \mathcal{L}_S.$$

Definition 7 (**Coherence relative to a set of gambles).** *We shall say that a set $\mathcal{D} \subseteq \mathcal{L}_S(\mathcal{X})$ is* coherent relative to *$\mathcal{L}_S(\mathcal{X})$ when the set $\mathcal{D}_S \subseteq \mathcal{L}(\mathcal{X}_S)$ that we can make a one-to-one correspondence with, is coherent.*

Note that coherence of $\mathcal{D}$ relative to $\mathcal{L}_N(\mathcal{X})$ is just coherence of $\mathcal{D}$, which makes sense given that $\mathcal{L}_N(\mathcal{X}) = \mathcal{L}(\mathcal{X})$. It also follows that if $\mathcal{D}$ is coherent relative to $\mathcal{L}_S$, then it is a *cone*: $\lambda f \in \mathcal{D}$ for every $f \in \mathcal{D}$ and every $\lambda > 0$.

Definition 8 (**Marginal set of gambles).** *Let $\mathcal{D} \subset \mathcal{L}(\mathcal{X})$ be a coherent set of desirable gambles and consider a subset S of N. The S-marginal of $\mathcal{D}$ is the set $\mathcal{D} \cap \mathcal{L}_S$.*

It follows that the S-marginal of a coherent set of desirable gambles is coherent relative to the set $\mathcal{L}_S$.

In this paper, we study the problem of the compatibility of a number of partial assessments into a joint model. We shall assume that these assessments are modelled by coherent sets of desirable gambles. We consider therefore subsets $S_1, \dots, S_r$ of $\{1, \dots, n\}$, and for every $j = 1, \dots, r$ let $\mathcal{D}_j$ be a subset of $\mathcal{L}(\mathcal{X})$ that is coherent with respect to the set $\mathcal{L}_{S_j}(\mathcal{X})$ of $\mathcal{X}_{S_j}$-measurable gambles. Our goal is to find conditions that guarantee the existence of a coherent set of desirable gambles $\mathcal{D}$ that is 'compatible' with $\mathcal{D}_1, \dots, \mathcal{D}_r$. Let us clarify what we mean by compatibility in this context. A more general definition shall be introduced in Section 3.

Definition 9 (**Pairwise compatibility for coherent sets of desirable gambles).** *We say that coherent sets of desirable gables $\mathcal{D}_i, \mathcal{D}_j$, with $i \neq j$ in $\{1, \dots, r\}$, are* pairwise compatible *if and only if*

$$\mathcal{D}_i \cap \mathcal{L}_{S_j}(\mathcal{X}) = \mathcal{D}_j \cap \mathcal{L}_{S_i}(\mathcal{X}).$$

In other words, those gambles on $\mathcal{D}_i$ that are S_j-measurable belong to $\mathcal{D}_j$, and viceversa. If we regard our models as coming from different sources, the interpretation would be that, if two sources provide an assessment about the same gamble f, it cannot be that f is deemed desirable by one of them and not by the other.

Definition 10 (**Compatibility for coherent sets of desirable gambles).** *$\mathcal{D}_1, \dots, \mathcal{D}_r$ are said to be* compatible *if and only if there is a coherent set of desirable gambles $\mathcal{D}$ on $\mathcal{L}(\mathcal{X})$ that is pairwise compatible with each of them, in the sense that $\mathcal{D} \cap \mathcal{L}_{S_j}(\mathcal{X}) = \mathcal{D}_j$ for every $j = 1, \dots, r$. We also say that $\mathcal{D}$* is compatible with *$\mathcal{D}_1, \dots, \mathcal{D}_r$.*

The following result gives an equivalent expression of compatibility in terms of the notion of natural extension from Definition 3:

Proposition 1. *Consider sets of desirable gambles $\mathcal{D}_1, \dots, \mathcal{D}_r$ such that $\mathcal{D}_j$ is coherent relative to $\mathcal{L}_{S_j}$. They are compatible if and only if the natural extension $\mathcal{E}$ of $\cup_{j=1}^r \mathcal{D}_j$ satisfies $\mathcal{E} \cap \mathcal{L}_{S_j} = \mathcal{D}_j$ for $j = 1, \dots, r$.*

From this we deduce that the notion of compatibility in Definition 10 coincides with what we called *strong coherence* in [98, Definition 25]. A result similar to Proposition 1 was established by Studený in [82, Proposition 1], in the particular case when the belief models are possibility measures.

As we mentioned in the Introduction, it was established by Beeri et al. [5] that a number of marginal probability measures that are pairwise compatible are automatically compatible when the sets of variables where they are defined satisfy the running intersection property:

Definition 11 (Running intersection property). *The sets of variables $S_1, \ldots, S_r$ satisfy the* running intersection property *if and only if*

RIP. $(\forall i = 2, \ldots, r)(\exists j^\star < i)\ S_i \cap (\cup_{j<i} S_j) = S_i \cap S_{j^\star}$.

We next extend this result to the case where our belief models are sets of desirable gambles:

Theorem 2. *If $S_1, \ldots, S_r$ satisfy* RIP *and the sets $\mathcal{D}_1, \ldots, \mathcal{D}_r$ are pairwise compatible, then they are compatible.*

This result generalises most of previous work in the literature about compatibility in the unconditional case; we discuss this point at some length in Section 2.3. It is also useful to observe that to verify compatibility according to Theorem 2 we only need to marginalise and compare given sets of desirable gambles (let us call this the 'local' complexity), which means that the computational complexity of this task will be linear in r. Stated differently, such a task will be well solved as long as the local problem will be.

The following example illustrates the result:

Example 1. *Consider $N := \{1, 2, 3, 4\}$, and the sets of variables $S_1 := \{1, 2\}, S_2 := \{1, 3\}, S_3 := \{3, 4\}$. These sets of variables satisfy the running intersection property. Therefore, Theorem 2 tells us that if we model our uncertainty about these variables by means of coherent sets of desirable gambles $\mathcal{D}_{S_1}, \mathcal{D}_{S_2}, \mathcal{D}_{S_3}$, they will be compatible if and only if they are pairwise compatible, which in this case means that*

$$\mathcal{D}_{S_1} \cap \mathcal{L}(\mathcal{X}_1) = \mathcal{D}_{S_2} \cap \mathcal{L}(\mathcal{X}_1) \text{ and } \mathcal{D}_{S_2} \cap \mathcal{L}(\mathcal{X}_3) = \mathcal{D}_{S_3} \cap \mathcal{L}(\mathcal{X}_3).$$

For instance, if we consider binary variables and the sets of desirable gambles

$$\begin{aligned}
\mathcal{D}_{S_1} &:= \{f \in \mathcal{L}(\mathcal{X}_{1,2}) : \min\{f(0,1), f(1,0)\} > 0\} \cup \mathcal{L}^+(\mathcal{X}_{S_1}),\\
\mathcal{D}_{S_2} &:= \{f \in \mathcal{L}(\mathcal{X}_{1,3}) : \min\{f(1,1), f(0,1)\} > 0\} \cup \mathcal{L}^+(\mathcal{X}_{S_2}),\\
\mathcal{D}_{S_3} &:= \{f \in \mathcal{L}(\mathcal{X}_{3,4}) : \min\{f(1,1), f(1,0)\} > 0\} \cup \mathcal{L}^+(\mathcal{X}_{S_3}),
\end{aligned}$$

then pairwise compatibility holds, since we have that

$$\mathcal{D}_{S_1} \cap \mathcal{X}_1 = \mathcal{D}_{S_2} \cap \mathcal{X}_1 = \{f \in \mathcal{X}_1 : \min\{f(0), f(1)\} > 0\} \cup \mathcal{L}^+(\mathcal{X}_1) = \mathcal{L}^+(\mathcal{X}_1)$$

and

$$\mathcal{D}_{S_2} \cap \mathcal{X}_3 = \mathcal{D}_{S_3} \cap \mathcal{X}_3 = \{f \in \mathcal{X}_3 : f(1) > 0\} \cup \mathcal{L}^+(\mathcal{X}_3).$$

This means that they are also globally compatible. One such compatible joint is their natural extension, which gives

$$\mathcal{D} = \{f \in \mathcal{L}(\mathcal{X}_N) : \min\{f(0,1,1,1), f(0,1,1,0), f(1,0,1,1), f(1,0,1,0)\} > 0\} \cup \mathcal{L}^+. \ \lozenge$$

Remark 1. *As suggested by Referee 1, in some cases if our coherent sets of desirable gambles $\mathcal{D}_1, \ldots, \mathcal{D}_r$ represent different pieces of information we may not expect them to carry the same information for the common variables; in other words, we may look for the existence of a coherent superset of $\cup_{i=1}^r \mathcal{D}_i$ without imposing the pairwise compatibility of the sets $\mathcal{D}_1, \ldots, \mathcal{D}_r$.*

The set $\cup_{i=1}^r \mathcal{D}_i$ has a coherent superset if and only if its natural extension $\mathcal{E}$ is coherent, and in that case we obtain the compatibility of the sets $\mathcal{D}'_1, \ldots, \mathcal{D}'_r$, where $\mathcal{D}'_j := \mathcal{E} \cap \mathcal{L}_{S_j}$. We then deduce from Proposition 1 and Theorem 2 that, if $S_1, \ldots, S_r$ satisfy RIP, *then the following are equivalent:*

- *$\mathcal{D}_1 \cup \cdots \cup \mathcal{D}_r$ has a coherent superset,*

- ∘ $\mathcal{E} := \mathrm{posi}(\mathcal{L}^+ \cup \bigcup_{i=1}^r \mathcal{D}_i)$ *is coherent,*
- ∘ $\mathcal{D}'_1, \dots, \mathcal{D}'_r$ *are compatible,*
- ∘ $\mathcal{D}'_1, \dots, \mathcal{D}'_r$ *are pairwise compatible,*

where, for every $j = 1, \dots, r$, $\mathcal{D}'_j := \mathrm{posi}(\mathcal{L}^+ \cup \bigcup_{i=1}^r \mathcal{D}_i) \cap \mathcal{L}_{S_j}$.

This is also relevant for the treatment of compatibility we shall make in the conditional case (see Section 3.1 later on), where we shall verify whether some set of gambles that we can derive from the conditional assessments avoids partial loss. $\Diamond$

2.2. Coherent lower previsions

A slightly more precise model than coherent sets of desirable gambles are coherent lower previsions [89, Chapter 2]. These generalise de Finetti's pioneering work on subjective probability theory [26] to the imprecise case; in fact, as we shall see in Proposition 3 below, the compatibility of different sources is equivalent to de Finetti's notion of coherence, extended by Williams and Walley to the imprecise case.

Definition 12 (Coherent lower and upper previsions). *Let $\mathcal{D}$ be a coherent set of desirable gambles in $\mathcal{L}$. For all $f \in \mathcal{L}$, let*

$$\underline{P}(f) := \sup\{\mu \in \mathbb{R} : f - \mu \in \mathcal{D}\}; \tag{3}$$

it is called the lower prevision *of f. The conjugate value given by $\overline{P}(f) := -\underline{P}(-f)$ is called the* upper prevision *of f. The functionals $\underline{P}, \overline{P} : \mathcal{L} \to \mathbb{R}$ are respectively called a* coherent lower prevision *and a* coherent upper prevision.

A coherent lower prevision satisfies the following conditions for every $f, g \in \mathcal{L}$ and every $\lambda > 0$:

- C1. $\underline{P}(f) \geq \inf f$ [Accepting Sure Gains];
- C2. $\underline{P}(\lambda f) = \lambda \underline{P}(f)$ [Positive Homogeneity];
- C3. $\underline{P}(f + g) \geq \underline{P}(f) + \underline{P}(g)$ [Superlinearity].

These conditions are often taken in the literature as axioms of coherent lower previsions whenever they are used as the primitive models of uncertainty and are defined on $\mathcal{L}$.

Definition 13 (Linear prevision). *Let $\underline{P}, \overline{P}$ be coherent lower and upper previsions on $\mathcal{L}$. If $\underline{P}(f) = \overline{P}(f)$ for some $f \in \mathcal{L}$, then we call the common value the* prevision *of f and we denote it by $P(f)$. If this happens for all $f \in \mathcal{L}$ then we call the functional P a* linear prevision.

Linear previsions correspond to de Finetti's previsions, and their restriction to events are finitely additive probabilities.

A coherent lower prevision $\underline{P}$ has a set of *dominating* linear previsions:

$$\mathcal{M}(\underline{P}) := \{P \text{ linear prevision} : (\forall f \in \mathcal{L})\ P(f) \geq \underline{P}(f)\},$$

which turns out to be closed[4] and convex. Since each linear prevision is in a one-to-one correspondence with a finitely additive probability measure (its restriction to events), we can regard $\mathcal{M}(\underline{P})$ also as a set of probabilities. Moreover, $\underline{P}$ is the lower envelope of the previsions in $\mathcal{M}(\underline{P})$:

$$(\forall f \in \mathcal{L})\ \underline{P}(f) = \min\{P(f) : P \in \mathcal{M}(\underline{P})\}.$$

The coherent upper prevision $\overline{P}$ is the upper envelope of the same set; as a consequence, $\underline{P}(f) \leq \overline{P}(f)$ for all $f \in \mathcal{L}$.

[4] In the weak* topology, which is the smallest topology such that all the evaluation functionals given by $f(P) := P(f)$, where $f \in \mathcal{L}$, are continuous.

Example 2. *If we return to Example 1, we see that the coherent lower previsions associated with the coherent sets of desirable gambles in that example are given by*

$$(\forall f \in \mathcal{L}_{S_1})\ \underline{P}_{S_1}(f) = \min\{f(0,1), f(1,0)\},$$
$$(\forall f \in \mathcal{L}_{S_2})\ \underline{P}_{S_2}(f) = \min\{f(1,1), f(0,1)\},$$
$$(\forall f \in \mathcal{L}_{S_3})\ \underline{P}_{S_3}(f) = \min\{f(1,1), f(1,0)\},$$

which are equivalent to the assessments

$$P(X_1 \neq X_2) = 1 = P(X_3 = 1).$$

Thus, a compatible coherent lower prevision is the lower envelope of the set of probabilities degenerate on the mass functions $\{(0,1,1,0), (0,1,1,1), (1,0,1,0), (1,0,1,1)\}$. *Note here that* $\mathcal{D}_{S_3}$ *is not the only coherent set of desirable gambles that induces* $\underline{P}_{S_3}$*: for instance, we may also use*

$$\mathcal{D}'_{S_3} := \{f : \min\{f(1,1), f(1,0)\} > 0\} \cup \{f : f(1,1) = f(1,0) = 0 < f(0,0)\} \cup \mathcal{L}^+_{S_3}. \ \Diamond \quad (4)$$

More generally speaking, a lower prevision $\underline{P}$ defined on a set of gambles $\mathcal{K} \subseteq \mathcal{L}$ is called coherent if and only if it is the restriction of a coherent lower prevision $\underline{Q}$ on $\mathcal{L}$. The smallest such $\underline{Q}$ is called the *natural extension* of $\underline{P}$, and it is given by

$$\underline{E}(f) := \min\{P(f) : P \text{ linear prevision}, (\forall g \in \mathcal{K})\ P(g) \geq \underline{P}(g)\}.$$

As shown in Example 2, coherent sets of desirable gambles are in general more informative than coherent lower previsions, in the sense that there exist different coherent sets of desirable gambles $\mathcal{D}_1 \neq \mathcal{D}_2$ inducing the same coherent lower prevision by means of Eq. (3); the smallest such set satisfies a property called strict desirability:

Definition 14 (Strict desirability). *A coherent set of gambles* $\mathcal{D}$ *is said to be* strictly desirable *if it satisfies the following condition:*

D0. $f \in \mathcal{D} \setminus \mathcal{L}^+ \Rightarrow (\exists \delta > 0)\ f - \delta \in \mathcal{D}$ *[Openness],*

where addition of a gamble with a constant is meant pointwise.

Strict desirability means that $\mathcal{D} \setminus \mathcal{L}^+$ does not include its topological border. By an abuse of terminology, $\mathcal{D}$ is said to be open too.

There is a one-to-one correspondence between coherent lower previsions and strictly desirable sets: from $\underline{P}$ we can induce the set

$$\mathcal{D}_{\underline{P}} := \{f \in \mathcal{L} : f \gneq 0 \text{ or } \underline{P}(f) > 0\}; \quad (5)$$

$\mathcal{D}_{\underline{P}}$ is coherent and strictly desirable and moreover induces $\underline{P}$ through Eq. (3). Moreover, it is the only coherent and strictly desirable set to do so.

Similarly to Definition 8, given a coherent lower prevision on $\mathcal{L}$ and a subset of variables S, we call its S-marginal the model of the information that $\underline{P}$ encompasses on the variables in S:

Definition 15 (Marginal coherent lower prevision). *Let* $\underline{P}$ *be a coherent lower prevision on* $\mathcal{L}$ *and a non-empty* $S \subseteq N$*. Then the* S-marginal coherent lower prevision *it induces is given by*

$$\underline{P}_S(f) := \underline{P}(f)$$

for all $f \in \mathcal{L}_S$.

The S-marginal is simply the restriction of $\underline{P}$ to $\mathcal{L}_S$.

In terms of coherent lower previsions, the notion of compatibility in Definition 10 means that, given marginal coherent lower previsions $\underline{P}_1, \dots, \underline{P}_r$ with respective domains $\mathcal{L}_{S_1}, \dots, \mathcal{L}_{S_r}$, there exists a coherent lower prevision on $\mathcal{L}$ with these marginals. Pairwise compatibility means that the lower prevision $\underline{P}$ we can define on $\mathcal{K} := \mathcal{L}_{S_1} \cup \dots \cup \mathcal{L}_{S_r}$ by $\underline{P}(f) = \underline{P}_j(f)$ for every $f \in \mathcal{L}_{S_j}$ is well defined.

It is immediate then to show that compatibility is equivalent to the coherence of $\underline{P}$:

Proposition 3. *Let $\underline{P}_1, \ldots, \underline{P}_r$ be coherent lower previsions with respective domains $\mathcal{L}_{S_1}, \ldots, \mathcal{L}_{S_r}$. Assume they are pairwise compatible, and let $\underline{P}$ be the lower prevision they determine on $\mathcal{K} = \cup_{i=1}^r \mathcal{L}_{S_i}$. The following are equivalent:*

(a) $\underline{P}_1, \ldots, \underline{P}_r$ are globally compatible.

(b) $\underline{P}$ is a coherent lower prevision on $\mathcal{K}$.

(c) $\underline{P}_1, \ldots, \underline{P}_r$ are globally compatible with the natural extension $\underline{E}$ of $\underline{P}$.

From Theorem 2 and the correspondence between coherent lower previsions and sets of desirable gambles, it is not difficult to establish the following:[5]

Corollary 4. *Consider subsets $S_1, \ldots, S_r$ of $\{1, \ldots, r\}$ satisfying* RIP *and for every j let $\underline{P}_j$ be a coherent lower prevision on $\mathcal{L}_{S_j}$. The following are equivalent:*

(a) $\underline{P}_1, \ldots, \underline{P}_r$ are pairwise compatible.

(b) There exists a coherent lower prevision $\underline{P}$ on $\mathcal{L}$ with marginals $\underline{P}_1, \ldots, \underline{P}_r$.

2.3. Discussion

As a particular case of Corollary 4, we would obtain the result for linear previsions, that is, expectation operators with respect to a probability. We formalise the case of finite spaces that is the most common in the literature:

Corollary 5. *Consider finite possibility spaces $\mathcal{X}_1, \ldots, \mathcal{X}_n$ and subsets $S_1, \ldots, S_r$ of N. The following are equivalent:*

1. *For any pairwise compatible probability measures $P_1, \ldots, P_r$ on $\mathcal{P}(\mathcal{X}_{S_1}), \ldots, \mathcal{P}(\mathcal{X}_{S_r})$, there exists a probability measure P on $\mathcal{P}(\mathcal{X})$ with marginals $P_1, \ldots, P_r$.*

2. *$S_1, \ldots, S_r$ satisfy the running intersection property.*

In particular, our Corollary 4 can also be applied to possibility measures and belief functions, which were the belief models considered in [82], and that can be regarded as particular cases of coherent upper and lower previsions, respectively. We also cover [86, Proposition 4.2], with one qualification: instead of pairwise compatibility, Vejnarová considers the weaker notion called *projectivity*, which means that the corresponding sets of probability measures have non-empty intersection; this is related to Remark 1.

Nevertheless, it is important to remark that our result in terms of sets of desirable gambles (resp., coherent lower previsions) guarantees the existence of a global set of desirable gambles (resp., coherent lower prevision) whose marginals are the belief models we started with. Although this holds in particular if our set of desirable gambles is associated for instance with a possibility measure, it does not follow immediately that our global model (that we build considering techniques of natural extension) is also associated with a possibility measure; see [82, Example 2] for a counterexample. Therefore if one is interested in achieving a global model *that belongs to the same family as the marginal ones*, they should make additional considerations on top of our results.

Let us finally remark that RIP is necessary for pairwise compatibility to imply compatibility: in fact, Beeri at al. show in [5, Theorem 3.4] that if the sets of variables $S_1, \ldots, S_r$ do not satisfy RIP, then it is possible to find marginal probability measures $P_1, \ldots, P_r$ that are pairwise compatible while not being compatible. This can readily be extended to the case where beliefs are expressed in terms of sets of desirable gambles by using the correspondence in Eq. (5).

3. Compatibility of conditional models

We consider next a more general framework: that where our assessments are possibly of a conditional nature. Thus, given two disjoint subsets O, I of our set of variables N, we assume that we have a belief model about the variables in O, given information about the variables in I. The situation considered in Section 2 corresponds to the particular case where I is empty: then, what we have is marginal information about the variables in O.

[5] As remarked by Referee 1, the key in this next result is that the correspondence between coherent lower previsions and coherent sets of desirable gambles established in (5) is a monomorpishm, where these two belief models are valuation algebras in which the combination operator corresponds to the natural extension of the maximum (resp., union), the focusing operator corresponds to marginalisation and the neutral elements are, respectively, the vacuous coherent lower prevision, $\underline{P}(f) = \inf f \ (\forall f)$ and the set $\mathcal{L}^+$ of non-negative gambles. See [52, Section 3.3.2] for more information.

3.1. Conditional sets of desirable gambles

In this section, we consider the case where our belief models are sets of desirable gambles. We need first to extend the notion of coherence to the conditional case:

Definition 16 (**Separately coherent conditional sets of desirable gambles**). *Consider two disjoint subsets I, O of N with $O \neq \emptyset$. A* separately coherent conditional set of desirable gambles *$\mathcal{D}_O|X_I$ is given by*

$$\mathcal{D}_O|X_I := \cup_{x_I \in \mathcal{X}_I} \mathcal{D}_O|x_I,$$

where, for every $x_I \in \mathcal{X}_I$ $\mathcal{D}_O|x_I$ is defined as

$$\mathcal{D}_O|x_I := \{f \in \mathcal{L}(\mathcal{X}_{O\cup I}) : f = I_{X_I=x_I} f, f(x_I, \cdot) \in \mathcal{D}_O^{x_I}\}$$

for some coherent set of desirable gambles $\mathcal{D}_O^{x_I} \subset \mathcal{L}(\mathcal{X}_O)$ on $\mathcal{X}_O$. In case $I = \emptyset$, $\mathcal{D}_O|X_I$ is a single coherent set of desirable gambles $\mathcal{D}_O$.

Formally, $\mathcal{D}_O|X_I$ is a subset of $\mathcal{L}_{O\cup I}$, but it need not be coherent relative to it: it is only coherent once we focus on each particular element $x_I \in \mathcal{X}_I$. Nevertheless, for the purposes of this paper we can equivalently work with its natural extension on $\mathcal{L}_{O\cup I}$, that is given by

$$\{f \in \mathcal{L}_{O\cup I} : f \neq 0, (\forall x_I \in \mathcal{X}_I) f(x_I, \cdot) \in \mathcal{D}_O|x_I \cup \{0\}\}, \tag{6}$$

and that is indeed coherent relative to $\mathcal{L}_{O\cup I}$.

As one particular instance of separately coherent conditional sets of desirable gambles, we have those induced by unconditional sets:

Definition 17 (**Induced separately coherent conditional set of desirable gambles**). *Let $\mathcal{D}$ be a coherent set of gambles and consider two disjoint subsets I, O of N with $O \neq \emptyset$. The separately coherent conditional set of desirable gambles* induced by *$\mathcal{D}$ is given by*

$$\mathcal{D}_O|X_I := \cup_{x_I \in \mathcal{X}_I} \mathcal{D}_O|x_I, \text{ where } \mathcal{D}_O|x_I := \{f \in \mathcal{D} \cap \mathcal{L}_{O\cup I} : f = \mathbb{I}_{X_I=x_I} f\}. \tag{7}$$

When $I = \emptyset$ Equation (7) reduces to $\mathcal{D}_O := \mathcal{D} \cap \mathcal{L}_O$, i.e., it produces the marginal set of desirable gambles that $\mathcal{D}$ induces on the set of variables O. Thus, Definition 8 is a particular case of this one.

Example 3. *If we return to Example 2 and in particular to Eq.* (4)*, we can see how the coherent sets of desirable gambles $\mathcal{D}_{S_3}$ and $\mathcal{D}'_{S_3}$, which induce the same coherent lower prevision $\underline{P}_{S_3}$, produce different conditional sets of desirable gambles: we obtain*

$$\mathcal{D}_4|(X_3 = 0) = \mathcal{L}_4^+ \text{ while } \mathcal{D}'_4|(X_3 = 0) = \{f \in \mathcal{L}_4 : f(0) > 0\} \cup \mathcal{L}_4^+.$$

This shows that coherent sets of desirable gambles are useful for determining conditional assessments, in particular when the conditioning event has (lower) probability zero. $\Diamond$

Definition 18 (**Compatibility of conditional sets of desirable gambles**). *Consider disjoint subsets O_j, I_j of N, with $O_j \neq \emptyset$, for $j = 1, \dots, r$, and let $\mathcal{D}_{O_j}|X_{I_j}$ be a separately coherent conditional set of desirable gambles for $j = 1, \dots, r$. These sets are said to be* compatible *when there is a coherent set of desirable gambles $\mathcal{D}$ that induces each of them by means of Eq.* (7)*.*

From our comments above, this definition subsumes Definition 10 as a particular case. It is also a generalisation of the notion we called *conformity* in [68, Definition 11] for the particular case where we have one conditional and one unconditional model; the idea is again that there exists a joint model from which we can derive all the assessments. As such the notion of compatibility in Definition 18 is nothing else than what we called 'strong coherence' in [98, Definition 25]: the notion of coherence for a collection of sets of desirable gambles (as opposed to its special case of coherence for a single set, as given in Definition 2).

Remark 2. *Let us remark that, in the context of non-additive measures, which can be regarded as particular cases of coherent sets of desirable gambles, we can find many proposals in the literature to induce a conditional model from an unconditional one; see for instance [22, 30, 33, 39, 63] and the references therein. The notion we consider in Definition 18 for coherent sets of desirable gambles corresponds to Williams-Walley's* generalised Bayes rule *and can be defended based on their behavioural interpretation of desirability. Note that if we apply this procedure to a particular family of non-additive measures, the induced conditional model may not always belong to such a family (this is the same issue we mentioned at the end of Section 2): for this reason, if someone wants to focus on some particular model, such as possibility measures, it would be necessary to consider some alternative proposals, or—probably more sensibly—to approximate the generalised Bayes rule through members of the chosen family.* $\diamondsuit$

One immediate consequence of the above definition is the following result, that is similar to Proposition 1:

Proposition 6. *Consider disjoint subsets O_j, I_j of N, with $O_j \neq \emptyset$, for $j = 1, \ldots, r$, and let $\mathcal{D}_{O_j}|X_{I_j}$ be a separately coherent conditional set of desirable gambles for $j = 1, \ldots, r$.*

1. *If $\mathcal{D}_{O_1}|X_{I_1}, \ldots, \mathcal{D}_{O_r}|X_{I_r}$ are compatible, then $\cup_{j=1}^r \mathcal{D}_{O_j}|X_{I_j}$ avoids partial loss.*
2. *If $\mathcal{D}_{O_1}|X_{I_1}, \ldots, \mathcal{D}_{O_r}|X_{I_r}$ are compatible, the smallest coherent set of desirable gambles that induces $\mathcal{D}_{O_j}|X_{I_j}$ by* (7) *for $j = 1, \ldots, r$ is the natural extension $\mathcal{E}$ of $\cup_{j=1}^r \mathcal{D}_{O_j}|X_{I_j}$.*

Example 4. *Using the notation of Example 1, consider the following two separately coherent sets of desirable gambles:*

$$\mathcal{D}_4|X_3 := \mathcal{L}_3^+ \text{ and } \mathcal{D}_3|X_4 := \mathcal{L}_4^+.$$

These two sets are compatible given that they can both be induced by $\mathcal{D}_{S_3}$ in Example 1 via Eq. (7)*, and as a consequence $\mathcal{L}_3^+ \cup \mathcal{L}_4^+$ avoids partial loss. $\mathcal{D}_{S_3}$ is however not their natural extension since the smallest coherent set that induces them is obviously the vacuous set $\mathcal{L}_{S_3}^+$.* $\diamondsuit$

We deduce from Proposition 6 that the verification of compatibility comprises two parts: the first one is whether our sets of desirable gambles avoid partial loss; if the answer is positive, we should verify next whether the natural extension $\mathcal{E}$ of our assessments induces them by means of (7); note that, for this second part, it suffices to know the marginals $\mathcal{E} \cap \mathcal{L}_{O_j \cup I_j}(\mathcal{X})$ for $j = 1, \ldots, r$.

In this paper, we shall provide two algorithms that will simplify the verification of the condition of avoiding partial loss and the computation of the marginals of the natural extension; but before we tackle this problem, we think it is important to clarify why we cannot express it more simply in terms of unconditional sets of desirable gambles.

Indeed, it follows from the above reasoning that, if we want to compute the natural extension of $\cup_{j=1}^r \mathcal{D}_{O_j}|X_{I_j}$ we may first compute separately the natural extension of each of the sets $\mathcal{D}_{O_j}|X_{I_j}$ for $j = 1, \ldots, r$ by means of (6). If we denote $\mathcal{E}_1, \ldots, \mathcal{E}_r$ these natural extensions, it follows that $\mathcal{E}$ is also the natural extension of $\cup_{j=1}^r \mathcal{E}_j$. Thus, we might be tempted by trying to reduce the problem to that of the compatibility of $\mathcal{E}_1, \ldots, \mathcal{E}_r$, which we have tackled in Section 2, and that can be deduced from pairwise compatibility and RIP.

Unfortunately, such a procedure does not work, because the compatibility of $\mathcal{D}_{O_1}|X_{I_1}, \ldots, \mathcal{D}_{O_r}|X_{I_r}$ does not imply the pairwise compatibility of the sets $\mathcal{E}_1, \ldots, \mathcal{E}_r$. This is discussed in Appendix A.1.

Taking this into account, given a number of coherent sets of desirable gambles $\mathcal{D}_1, \ldots, \mathcal{D}_r$ that gather information on different sets of variables $S_1, \ldots, S_r$, we shall investigate if these sets avoid partial loss, meaning that they have a joint coherent superset; but we are not requiring anymore that $\mathcal{D} \cap \mathcal{L}_{S_j} = \mathcal{D}_j$ for every j. Indeed, if the coherent set $\mathcal{D}_j$ is obtained as the natural extension of a separately coherent conditional set $\mathcal{D}_{O_j}|X_{I_j}$, what we should verify next is whether the coherent superset $\mathcal{D}$ induces $\mathcal{D}_{O_j}|X_{I_j}$ by means of Eq. (7), and not whether $\mathcal{D}_j$ is the marginal of $\mathcal{D}$ on $\mathcal{X}_{O_j \cup I_j}$.

Our first result tells us that if a variable appears only in one of these sets, then our assessments on this variable are not relevant for the compatibility problem:

Proposition 7. *Consider subsets $S_1, \ldots, S_r$ of $\{1, \ldots, n\}$ and sets of desirable gambles $\mathcal{D}_1, \ldots, \mathcal{D}_r$, where $\mathcal{D}_j$ is coherent relative to $\mathcal{L}_{S_j}$. For every $i = 1, \ldots, r$, let $\mathcal{D}_i^* := \mathcal{D}_i \cap \mathcal{L}_{S_i \cap (\cup_{j \neq i} S_j)}$. Then:*

$$\cup_{i=1}^r \mathcal{D}_i \text{ avoids partial loss } \Leftrightarrow \cup_{i=1}^r \mathcal{D}_i^* \text{ avoids partial loss.}$$

This result is actually not surprising: the assessments that are made in only one of our belief models cannot be contradicted by any other, and thus will never cause us to violate compatibility.

3.2. Conditional lower previsions

Similarly to what we did in the unconditional case, from our results on the compatibility of (conditional) sets of desirable gambles we can derive analogous results for (conditional) lower previsions. Let us recall a number of preliminary notions (see [66] for details about the relation of desirable gambles with conditional lower previsions).

Definition 19 (Coherent conditional lower and upper previsions). *Let $\mathcal{D}$ be a coherent set of desirable gambles in $\mathcal{L}$. Consider two disjoint subsets I, O of N, with $O \neq \emptyset$, and $x_I \in \mathcal{X}_I$. For all $f \in \mathcal{L}_{O\cup I}$, let*

$$\underline{P}_O(f|x_I) \coloneqq \sup\{\mu \in \mathbb{R} : \mathbb{I}_{x_I}(f - \mu) \in \mathcal{D}\} \tag{8}$$

be the conditional lower prevision *of f given x_I. The conjugate value given by $\overline{P}_O(f|x_I) \coloneqq -\underline{P}_O(-f|x_I)$ is called the* conditional upper prevision *of f. The functionals $\underline{P}_O(\cdot|x_I), \overline{P}_O(\cdot|x_I) : \mathcal{L}_{O\cup I} \to \mathbb{R}$ are respectively called a* coherent conditional lower prevision *and a* coherent conditional upper prevision*.*

Denote by $\inf_{x_I} f$ the infimum value that f takes on $\{x_I\}$. $\underline{P}_O(\cdot|x_I)$ satisfies the following conditions for all $f \in \mathcal{L}_{O\cup I}$ and all real $\lambda > 0$:

CC1. $\underline{P}_O(f|x_I) \geq \inf_{x_I} f$;

CC2. $\underline{P}_O(\lambda f|x_I) = \lambda \underline{P}_O(f|x_I)$;

CC3. $\underline{P}_O(f + g|x_I) \geq \underline{P}_O(f|x_I) + \underline{P}_O(g|x_I)$.

Again, these conditions can be regarded as axioms of coherent conditional lower previsions.

If we make this procedure for every $x_I \in \mathcal{X}_I$, we obtain the following:

Definition 20 (Separately coherent conditional lower prevision). *Consider two disjoint subsets I, O of N, with $O \neq \emptyset$. For all $x_I \in \mathcal{X}_I$, let $\underline{P}_O(\cdot|x_I)$ be a conditional coherent lower prevision. Then we call*

$$\underline{P}_O(\cdot|X_I) \coloneqq \sum_{x_I \in \mathcal{X}_I} \mathbb{I}_{x_I} \underline{P}_O(\cdot|x_I)$$

a separately coherent conditional lower prevision*.*

For every $f \in \mathcal{L}_{O\cup I}$, $\underline{P}_O(f|X_I)$ is the gamble that takes the value $\underline{P}_O(f|x_I)$ in $x_I \in \mathcal{X}_I$; it is an $\mathcal{X}_I$-measurable gamble: $\underline{P}_O(f|X_I) \in \mathcal{L}_I$.

Now consider a finite number of separately coherent conditional lower previsions $\underline{P}_{O_1}(\cdot|X_{I_1}), \ldots, \underline{P}_{O_r}(\cdot|X_{I_r})$ on respective domains $\mathcal{L}_{O_1\cup I_1}, \ldots, \mathcal{L}_{O_r\cup I_r}$. Their joint coherence is defined very naturally as follows:

Definition 21 (Strong coherence of a collection of separately coherent conditional lower previsions). *Given the collection $\underline{P}_{O_1}(\cdot|X_{I_1}), \ldots, \underline{P}_{O_r}(\cdot|X_{I_r})$, we say that* the conditional lower previsions are (strongly) coherent *if and only if there is a coherent set of desirable gambles $\mathcal{D} \subseteq \mathcal{L}$ such that $\underline{P}_{O_i}(\cdot|X_{I_i})$ can be recovered from $\mathcal{D}$ through* (8)*, for all $i = 1, \ldots, r$.*[6]

Next we consider the consistency condition of avoiding partial loss, which is weaker than strong coherence; it allows us to know when a non-coherent collection of conditional lower previsions can be extended into a coherent one. To this end, first we need to introduce the following notion:

Definition 22 (Dominance of a collection of separately coherent conditional lower previsions). *Given two collections of separately coherent conditional lower previsions $\underline{P}_{O_1}(\cdot|X_{I_1}), \ldots, \underline{P}_{O_r}(\cdot|X_{I_r})$ and $\underline{P}'_{O_1}(\cdot|X_{I_1}),\ldots,\underline{P}'_{O_r}(\cdot|X_{I_r})$, we say that the latter* dominates *the former if and only if*

$$(\forall i = 1, \ldots, r)(\forall f \in \mathcal{L}_{O_i\cup I_i})(\forall x_{I_i} \in \mathcal{X}_{I_i})\ \underline{P}'_{O_i}(f|x_{I_i}) \geq \underline{P}_{O_i}(f|x_{I_i}).$$

[6] This is what Williams originally called coherence [87, 94].

We denote dominance for short also as $\underline{P}'_{O_i}(\cdot|X_{I_i}) \geq \underline{P}_{O_i}(\cdot|X_{I_i})$.

Definition 23 (**Avoiding partial loss of a collection of separately coherent conditional lower previsions**). *Given a collection of separately coherent conditional lower previsions, $\underline{P}_{O_1}(\cdot|X_{I_1})$, ..., $\underline{P}_{O_r}(\cdot|X_{I_r})$, we say that* the collection avoids partial loss *if and only if there is a strongly coherent collection $\underline{P}'_{O_1}(\cdot|X_{I_1})$, ..., $\underline{P}'_{O_r}(\cdot|X_{I_r})$ that dominates it.*

Obviously strong coherence implies that a collection avoids partial loss, but not vice versa. In fact, the condition of avoiding partial loss is tantamount to the possibility of turning a non-coherent collection into a coherent one, by making the assessments more precise. The least-committal way to do so is called the natural extension:

Definition 24 (**Natural extension of a collection of separately coherent conditional lower previsions**). *Given a collection of separately coherent conditional lower previsions that avoids partial loss, $\underline{P}_{O_1}(\cdot|X_{I_1})$, ..., $\underline{P}_{O_r}(\cdot|X_{I_r})$, its* natural extension *is the smallest dominating strongly coherent collection $\underline{E}_{O_1}(\cdot|X_{I_1})$, ..., $\underline{E}_{O_r}(\cdot|X_{I_r})$ (i.e., the one that is dominated by all the dominating ones).*

Similarly to what we mentioned in the unconditional case, two different coherent sets of desirable gambles $\mathcal{D}_1, \mathcal{D}_2$ may determine the same conditional lower prevision by means of Eq. (8). As a consequence, sets of desirable gambles constitute a more general uncertainty model than coherent lower previsions.

On the other hand, and similarly to Eq. (5), if we work with coherent lower previsions as the primary model, we can always make a transformation into sets of desirable gambles: given a separately coherent conditional lower prevision $\underline{P}(X_O|X_I)$ on $\mathcal{L}_{O\cup I}$, the set

$$\mathcal{D}_O|x_I := \{\mathbb{I}_{x_I}(f - \underline{P}(f|x_I) + \varepsilon) : f \in \mathcal{L}_O, \varepsilon > 0\} \cup \{f \in \mathcal{L}^+_{O\cup I} : f = \mathbb{I}_{x_I} f\} \tag{9}$$

is a coherent subset of $\mathcal{L}_{O\cup I}$. Moreover, the union $\mathcal{D}_O|X_I$ induces $\underline{P}(X_O|X_I)$ by means of Eq. (8). Indeed, we have the following:

Proposition 8. *Consider separately coherent conditional lower previsions $\underline{P}_{O_1}(\cdot|X_{I_1})$, ..., $\underline{P}_{O_r}(\cdot|X_{I_r})$ with respective domains $\mathcal{L}_{O_1\cup I_1}, \ldots, \mathcal{L}_{O_r\cup I_r}$. Let $\mathcal{D}_{O_1}|X_{I_1}, \ldots, \mathcal{D}_{O_r}|X_{I_r}$ be the sets of desirable gambles they induce by means of Eq.* (9)*. Define $\mathcal{D} := \cup_{j=1}^r \mathcal{D}_{O_j}|X_{I_j}$ and let $\mathcal{E}$ be its natural extension.*

1. *$\underline{P}_{O_1}(\cdot|X_{I_1})$, ..., $\underline{P}_{O_r}(\cdot|X_{I_r})$ avoid partial loss if and only if $\mathcal{D}$ avoids partial loss.*
2. *$\underline{P}_{O_1}(\cdot|X_{I_1})$, ..., $\underline{P}_{O_r}(\cdot|X_{I_r})$ are strongly coherent if and only if $\mathcal{E}$ induces them by means of Eq.* (8)*.*

In fact, it is proven in [67, Theorem 7(2)] that the natural extension $\mathcal{E}$ of $\mathcal{D}$ induces the natural extensions $\underline{E}_{O_1}(\cdot|X_{I_1}), \ldots, \underline{E}_{O_r}(\cdot|X_{I_r})$ of $\underline{P}_{O_1}(\cdot|X_{I_1}), \ldots, \underline{P}_{O_r}(\cdot|X_{I_r})$.

Therefore, if we consider a number of separately coherent conditional lower previsions, the definition of compatibility that is akin to Definition 18 is that of coherence we have given in Definition 21. In fact, we observed already in [66, Theorem 11] that Definition 21 corresponds to the specialisation of strong coherence for desirability to the case of conditional lower previsions.

As a consequence, in order to verify compatibility, we should check (a) whether the set of desirable gambles determined by the separately coherent conditional lower previsions avoids partial loss; and (b) if its natural extension induces the conditional lower previsions by means of Eq. (8). Thus, the problem reduces to the one we have tackled in Section 3.1.

4. Exploiting the power of tree decomposition

In this section we consider the most general version of the compatibility problem, where we have n variables $X_1, \ldots, X_n$ over which we assess r separately coherent conditional sets of desirable gambles $\mathcal{D}_{O_1}|X_{I_1}, \ldots, \mathcal{D}_{O_r}|X_{I_r}$.

In the following we shall sometimes focus only on the variables involved in a certain set $\mathcal{D}_{O_j}|X_{I_j}$; we denote the qualitative form of their relation by the so-called 'template' $X_{O_j}|X_{I_j}$.

As a running example we consider the following $r = 13$ templates over $n = 15$ variables:

$$\begin{aligned}&X_2|X_1, X_2|X_4, X_3|X_2, X_5|X_4, X_5|X_6, X_{11}|X_5, \{X_9, X_{10}\}|\{X_7, X_8, X_{11}\},\\ &X_7|X_{12}, X_{12}|X_8, X_{13}|X_8, X_{13}|X_{12}, X_{15}|\{X_{13}, X_{14}\}, X_8|X_{15}.\end{aligned} \tag{10}$$

The problem now is how to check the compatibility of $\mathcal{D}_{O_1}|X_{I_1}, \ldots, \mathcal{D}_{O_r}|X_{I_r}$. One issue is that if we let as usual $S_j := O_j \cup I_j$ for all j, RIP will not hold in general. However, it is well-known that we can enable RIP to hold by representing the templates through a graph and then proceeding by a so-called *tree decomposition* [42].

The procedure of tree decomposition has a long history and is related to the possibility of optimally decomposing a problem into smaller ones. The solutions of these smaller problems are then aggregated back to obtain the solution of the original problem, in a dynamic-programming fashion [11]. There is a wealth of applications of tree decomposition in Artificial Intelligence: e.g., in probabilistic inference [27, 48, 58], constraint satisfaction [28, 100], matrix decomposition [13, 73]. We are now going to add our generalised version of the compatibility problem to the list of problems that can be solved by tree decomposition.

Let us then proceed in the traditional way towards a tree decomposition. First, we create the so-called 'domain graph' (we are borrowing some terminology from [47]):

Definition 25 (Domain graph). *Given templates $X_{O_1}|X_{I_1}, \ldots, X_{O_r}|X_{I_r}$ over n variables, the corresponding* domain graph *is an undirected graph with n nodes such that node i is associated with variable X_i, for all $i = 1, \ldots, n$. Two nodes are connected in the domain graph if and only if there is a template j such that both nodes' indexes belong to S_j.*

The domain graph for the running example is shown in Figure 1.

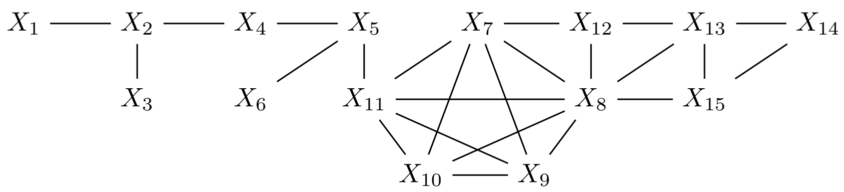

Figure 1: Domain graph for the running example.

The next definition gives an important property that domain graphs may satisfy:

Definition 26 (Triangulated graph). *An undirected graph is* triangulated *if and only if all cycles of length greater than three are cut by a chord (the graph is also called* chordal *in this case).*

It is easy to check that the domain graph of the running example is indeed triangulated (for instance, observe that the cycle $X_8 - X_{12} - X_{13} - X_{15}$, of length four, is cut by cord $X_8 - X_{13}$).

Now we need some additional notion from graph theory:

Definition 27 (Clique). *An undirected graph's* cliques *are its fully connected subgraphs; a clique is said to be* maximal *if and only if it is not contained in any other clique.*

For instance, in the running example the subgraph made of nodes $\{X_7, X_8\}$ is a clique, which, in turn, is contained in the maximal clique $\{X_7, X_8, X_9, X_{10}, X_{11}\}$.

That the graph is chordal implies that the maximal cliques of the domain graph can be arranged in a *join tree* (see, e.g., [47, Theorem 4.4]).

Definition 28 (Join tree). *A* join tree *is an undirected tree whose nodes correspond to the cliques of a domain graph (each node contains the set of variables of the related clique), and with the property that whenever a variable belongs to two nodes, it belongs also to all the nodes in the path between them.*

The latter property is actually the graphical version of RIP, in the sense that if we now let S'_i be equal to the set of variables' indexes in clique i, for $i = 1, \ldots, q \leq r$, then $S'_1, \ldots, S'_q$ satisfy RIP. In other words, the join tree tells us how to optimally aggregate the original variables into clusters (i.e., cliques) so as to make RIP hold over them. Figure 2 shows the join tree for the running example.

The procedure of creating a join tree from a triangulated domain graph is easy, and there are well-known, efficient algorithms to do so (see, e.g., [53, Section 10.4.2]). In case the domain graph is not triangulated, it is always possible to

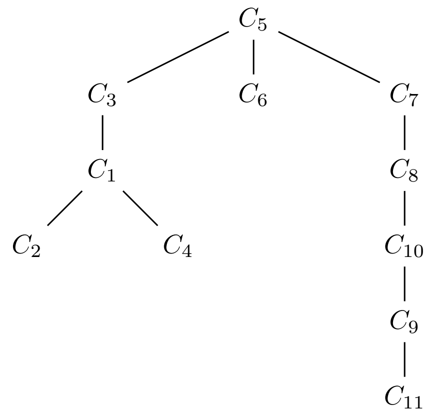

Figure 2: Join tree for the running example. The cliques $C_1, \ldots, C_{11}$ are defined in Table 1.

add edges to the domain graph so as to make it triangulated.[7] Overall, this means that given a collection of templates, we can always assume that there is a triangulated graph associated with it and hence that there is a procedure that outputs the corresponding join tree.

Once the join tree is obtained, and hence we have RIP, we can exploit it to solve the compatibility problem; in particular, we can do this directly on the graph provided that we enrich the join tree by some quantitative information:

Definition 29 (**Junction tree**). *A* junction tree *is obtained from a join tree by (i) assigning the uncertain information about template $X_{O_j}|X_{I_j}$ to a (single) node that contains the variables related to S_j, for all $j = 1, \ldots, r$; (ii) labelling each edge with a so-called* separator *denoting the variables in the intersection of the two connected nodes; and (iii) choosing a 'root' node for the tree in an arbitrary way.*[8,9]

The junction tree for the running example is in Figure 3; Table 1 gives some summary information about it. Note that we have chosen clique $\{X_5, X_{11}\}$ as the root of the tree. Moreover, we take A_i to be the set of indexes of those cliques that are at distance i from this root. Then trivially A_0 consists simply of the index of the root, and if the maximum distance to the root is k, then $\{A_0, A_1, \ldots, A_k\}$ forms a partition of the set of indexes. Labels A_i are displayed in Figure 3 close to the cliques, with $i = 0, \ldots, 5$. We display also the sets of desirable gambles assigned to each clique. Separators between cliques are shown close to the edges connecting them. Note that a node of the junction tree can contain more than one set of desirable gambles.

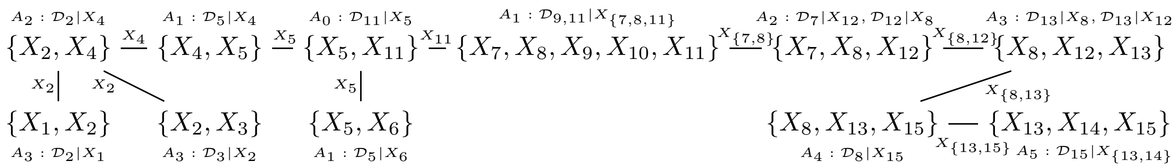

Figure 3: Junction tree for the running example. The cliques are now displayed explicitly through their corresponding sets of variables $S'_1, \ldots, S'_{11}$.

Considered our discussion in Section 3.1, in order to have compatibility we need that our original assessments at least avoid partial loss. For this reason, we can assume that at each node j of the junction tree the associated assessments avoid partial loss: this implies no loss of generality because if they did not, then also the overall set of assessments $\mathcal{D}_{O_1}|X_{I_1}, \ldots, \mathcal{D}_{O_r}|X_{I_r}$ would not avoid partial loss either.

[7]However, this will increase the size of the cliques of the resulting graph and thus may heavily impact the computational complexity of the algorithms that exploit the tree decomposition.

[8]Technically once we choose a root, we should talk of a *rooted* junction tree.

[9]Theorem 10 later on makes sure that the choice of the root node is not relevant.

Clique (C_j)	Variables (S'_j)	Distance to root (A_i)	Desirable gambles ($\mathcal{D}_j$)
C_1	$\{2,4\}$	2	$\mathcal{D}_4\|X_2$
C_2	$\{1,2\}$	3	$\mathcal{D}_2\|X_1$
C_3	$\{4,5\}$	1	$\mathcal{D}_5\|X_4$
C_4	$\{2,3\}$	3	$\mathcal{D}_3\|X_2$
C_5	$\{5,11\}$	0	$\mathcal{D}_{11}\|X_5$
C_6	$\{5,6\}$	1	$\mathcal{D}_5\|X_6$
C_7	$\{7,8,9,10,11\}$	1	$\mathcal{D}_{9,11}\|X_{7,8,11}$
C_8	$\{7,8,12\}$	2	$\mathcal{D}_7\|X_{12} \cup \mathcal{D}_{12}\|X_8$
C_9	$\{8,13,15\}$	4	$\mathcal{D}_8\|X_{15}$
C_{10}	$\{8,12,13\}$	3	$\mathcal{D}_{13}\|X_8 \cup \mathcal{D}_{13}\|X_{12}$
C_{11}	$\{13,14,15\}$	5	$\mathcal{D}_{15}\|X_{13,14}$

Table 1: The clique names, the variables involved in a clique, the distance to the root, as well as the set of desirable gambles associated with each clique. We see for instance that C_{10} is made by the union of two nodes, associated with the assessments $X_{13}|X_8$ and $X_{13}|X_{12}$; if these are modelled by means of the separately coherent conditional sets of desirable gambles $\mathcal{D}_{13}|X_8$ and $\mathcal{D}_{13}|X_{12}$, then the set of desirable gambles associated with C_{10} is given by $\mathcal{D}_{10} := \mathcal{D}_{13}|X_8 \cup \mathcal{D}_{13}|X_{12}$.

At this point we are ready to exploit the tree decomposition. The algorithms that rely on it are usually made of two passes: the first is called *collect evidence* and the second *distribute evidence*. Both require as input the junction tree.

We start by focusing on the first pass of collection of evidence, where all nodes propagate uncertain information towards the root. To simplify the notation, we denote by $\mathcal{D}_j$ the overall set of desirable gambles at node j obtained by taking the union of the assessments in such a node.

Our version of collect evidence is given in Algorithm 1.

Algorithm 1 Collect evidence

```
 1: procedure COLLECTEVIDENCE(a junction tree)
 2:     Let k be the maximum distance of a node from the root;      ▷ Distance 0 is for the root itself.
 3:     let A_i be the set of nodes at distance i from the root, for i = 0, ..., k+1;   ▷ A_{k+1} is always empty.
 4:     for i ← k, 0 do                                              ▷ Focus on distance i.
 5:         for all j ∈ A_i do                                       ▷ Consider the nodes in A_i.
 6:             let A be the set of nodes adjacent to j in A_{i+1};
 7:             let D'_j := posi_{S'_j}(L^+_{S'_j} ∪ D_j ∪ ⋃_{l∈A}(D'_l ∩ L_{S'_j∩S'_l}));   ▷ S'_j ∩ S'_l is the separator of j and l.
 8:         end for
 9:     end for
10:     return the junction tree with the additional information D'_j at each node j = 1, ..., q.
11: end procedure
```

This is essentially the standard form of collect evidence [47, Section 4.4], where we combine uncertain information from a node and some of its neighbours and then marginalise it on the variables of a separator before transmitting it along the related edge. Observe that the combination operator in line 7 is just the natural extension as defined in Eq. (2). The marginalisation operator is denoted, in the same line, by $\mathcal{D}'_l \cap \mathcal{L}_{S'_j \cap S'_l}$, and is the restriction of $\mathcal{D}'_l$ to the set of $S'_j \cap S'_l$-measurable gambles. Note also that the subindex l in $\mathcal{D}'_l$ refers to a node in the junction tree, and that what we obtain is a set of desirable gambles that is coherent relative to $\mathcal{L}(\mathcal{X}_{S'_j})$, where S'_j is the set of variables' indexes in node j. Moreover, the order in which the nodes in the same A_i are used in lines 5–8 in the algorithm is not relevant for the subsequent results, as can be seen from the proofs.

Let us illustrate the procedure with our running example, with the chosen root (clique $\{X_5, X_{11}\}$). Remember that labels $A_i, i = 0, \ldots, 5$ induce a partition of the cliques determined by the distance of each clique from the root, given in the specific case by the following indexes: $A_0 = \{5\}, A_1 = \{3,6,7\}, A_2 = \{1,8\}, A_3 = \{2,4,10\}, A_4 = \{9\}, A_5 = \{11\}$.

Then some instances of the procedure depicted in Algorithm 1 would be as follows:

- In the leaves $j = 2$ and $j = 4$ from A_3, associated with $S'_2 = \{1, 2\}$ and $S'_4 = \{2, 3\}$, respectively, we make $\mathcal{D}'_2 := \mathcal{D}_2$ and $\mathcal{D}'_4 := \mathcal{D}_4$.
- In their neighbour $j = 1 \in A_2$, associated with $S'_1 = \{2, 4\}$, we make $\mathcal{D}'_1 := \mathrm{posi}_{2,4}(\mathcal{L}^+_{2,4} \cup \mathcal{D}_1 \cup (\mathcal{D}'_2 \cap \mathcal{L}_2) \cup (\mathcal{D}'_4 \cap \mathcal{L}_2))$.
- Eventually, we get to the root node $j = 5$, associated with $S'_5 = \{5, 11\}$, and with neighbours $j = 3$ $(S'_3 = \{5, 4\}), j = 6(S'_6 = \{5, 6\}), j = 7(S'_7 = \{7, 8, 9, 10, 11\}$, and we make $\mathcal{D}'_5 := \mathrm{posi}_{5,11}(\mathcal{L}^+_{5,11} \cup \mathcal{D}_5 \cup (\mathcal{D}'_3 \cap \mathcal{L}_5) \cup (\mathcal{D}'_6 \cap \mathcal{L}_5) \cup (\mathcal{D}'_7 \cap \mathcal{L}_{11}))$.

The procedure in Algorithm 1 provides us with the restriction of the natural extension of $\mathcal{D}_1, \dots, \mathcal{D}_q$ to the gambles that depend on the variables from the root node:

Theorem 9. *Let $\mathcal{D}'_0$ denote the set produced by Algorithm 1 in the root node. Then $\mathcal{D}'_0$ is the restriction of the natural extension $\mathcal{E}$ of $\mathcal{D}_1, \dots, \mathcal{D}_q$ to $\mathcal{L}_{S'_0}$.*

We see from this result that

$$\mathcal{D}'_0 = \mathcal{E} \cap \mathcal{L}_{S'_0} = \mathrm{posi}(\cup_l \mathcal{D}_l \cup \mathcal{L}^+) \cap \mathcal{L}_{S'_0},$$

and also that

$$\mathcal{D}_1, \dots, \mathcal{D}_q \text{ avoid partial loss } \Leftrightarrow \mathcal{E} \text{ coherent } \Leftrightarrow \mathcal{D}'_0 \text{ coherent},$$

where the implication "$\mathcal{D}'_0$ coherent $\Rightarrow \mathcal{E}$ coherent" follows because if $\mathcal{E}$ were incoherent it would include the zero gamble and so should do $\mathcal{D}'_0$ then. This means in particular that if $\mathcal{D}'_0$ is not coherent, then the original assessments do not avoid partial loss, and as a consequence they are not compatible. In this case we can stop the procedure here.

Conversely, if $\mathcal{D}'_0$ is coherent we proceed to the reverse procedure of distribute evidence, where the junction tree in input must be the output of collect evidence. Observe that in this case it is not necessary to add the positive gambles, since $\mathcal{L}^+_{S'_j} \subseteq \mathcal{D}'_j$ by construction, and also that, since our graph is a tree, any node has only one immediate neighbour that is closer to the root.

Algorithm 2 Distribute evidence

```
procedure DISTRIBUTEEVIDENCE(a junction tree outputted by Algorithm 1)
    Let k be the maximum distance of a node from the root;        ▷ Distance 0 is for the root itself.
    let A_i be the set of nodes at distance i from the root, for i = −1, 0 ..., k − 1;   ▷ A_{−1} is the empty set.
    for i ← 0, k do                                                ▷ Focus on distance i.
        for all j ∈ A_i do                                         ▷ Consider the nodes in A_i.
            let l be the node adjacent to j in A_{i−1};
            let D''_j := posi_{S'_j}(D'_j ∪ (D''_l ∩ L_{S'_j ∩ S'_l}));   ▷ D''_0 equals D'_0 as it is coherent already.
        end for
    end for
    return the junction tree with the additional information D''_j at each node j = 1, ..., q.
end procedure
```

In order to illustrate the procedure, consider again our running example, depicted in Figure 3. Some instances of the algorithm would be as follows:

- We begin by considering $\mathcal{D}''_5 := \mathcal{D}'_5$ in the root node.
- For clique C_3 associated with $S'_3 = \{5, 4\} \in A_1$, we make $\mathcal{D}''_3 := \mathrm{posi}_{5,4}(\mathcal{D}'_3 \cup (\mathcal{D}''_5 \cap \mathcal{L}_5))$.
- Eventually we get to the leaf $j = 4$ associated with $S'_4 = \{2, 3\} \in A_3$, where we define $\mathcal{D}''_4 := \mathrm{posi}_{2,3}(\mathcal{D}'_4 \cup (\mathcal{D}''_1 \cap \mathcal{L}_2))$.

Let us prove that, for all $j = 1, \dots, q$, the set $\mathcal{D}''_j$ we obtain with this procedure is the restriction of the natural extension of $\mathcal{D}_1, \dots, \mathcal{D}_q$ to the class of $\mathcal{X}_{S'_j}$-measurable gambles. This holds for the root node too, taking into account Theorem 9 and the first step in Algorithm 2.

Theorem 10. *Let $\mathcal{E}$ be the natural extension of $\mathcal{D}_1, \dots, \mathcal{D}_q$. If we follow Algorithm 2, then $\mathcal{D}''_j = \mathcal{E} \cap \mathcal{L}_{S'_j}$ for every $j = 1, \dots, q$.*[10]

After reaching the end of Algorithm 2, it is then a small step to prove whether compatibility holds. For each original assessment $\mathcal{D}_{O_j}|X_{I_j}$, we consider the clique that contains it, and the corresponding set produced by Algorithm 2, say $\mathcal{D}''_i$. From this, using Definition 17, we induce the separately coherent conditional set of desirable gambles $\mathcal{D}''_{O_j}|X_{I_j}$ and verify whether $\mathcal{D}_{O_j}|X_{I_j} = \mathcal{D}''_{O_j}|X_{I_j}$. Compatibility holds if and only if this is the case for all $j = 1, \dots, r$.

With respect to the computational complexity of the procedures of collect and distribute evidence, we should distinguish two cases. In case our probabilistic assessments define a precise compatible joint, then the overall complexity is a linear function of the computation local to the cliques; this is analogous to the traditional procedures that work on junction trees. In the imprecise case, instead, the size of the messages exchanged between cliques may grow exponentially fast with the propagation (e.g., see [61]). This is unavoidable in general, as exact propagation of imprecise information is NP-hard [21].

More generally speaking, the present paper is conceived to lay the foundations of a very general compatibility problem with sets of desirable gambles, and as such we do not go into details of algorithmic implementation. However, since the algorithms require some steps involving marginalisation or natural extension, we would like to briefly mention how these can be practically achieved. In particular, in the case of finite spaces of possibilities, one usually addresses these tasks using linear programming (possibly in a sequence of linear programs). This is detailed for instance in [91]; alternative approaches are described in [17] and the references therein. In the case of infinite spaces, the task is obviously more complicated as one needs to solve non-linear optimisations, or semi-infinite linear programming problems, that are generally intractable. However, when we restrict the attention to the class of polynomial, or piece-wise polynomial, gambles, then approximate solutions to this problem can be obtained by means of Lasserre's *sum-of-squares* hierarchy [57] that are conservative and theoretically sound [8]. Benavoli has released the public software library *PyRational* that implements some of these procedures [7] (see also [6]).

5. Joining coherence graphs and RIP

It is important to realise that RIP or, equivalently, tree decompositions, do not necessarily simplify the compatibility check to the most. Consider for instance a case where the involved assessments define only two templates: $X_1|X_2$ and $X_2|X_3$ (this actually happens in Example 5 in Appendix A); the form of these templates is enough to deduce that the associated numerical assessments, whatever they are (provided that they are separately coherent), are strongly coherent, that is, compatible. In this case, therefore, it would be useless, and inefficient, to construct the junction tree and make the two passes of collect and distribute evidence in order to verify compatibility.

The reason why compatibility immediately holds for templates $X_1|X_2$ and $X_2|X_3$, is that those templates allow for an application of the *marginal extension theorem* (established in [89, Section 6.7] and [62] for coherent lower previsions, and in [69, Proposition 30] for desirable gambles), which, in turn, is the generalisation of the law of total probability to imprecision.[11]

Similar considerations led us in the past to work out the details of the extent to which we can exploit the marginal extension theorem to prove the coherence of some assessments on the sole basis of their templates. The result is the theory of 'coherence graphs', reported in [64]. The coherence graph for the templates in Eq. (10) is represented in Figure 4. It is a straightforward graphical representation where each template is represented by a black circle whose incoming arcs correspond to its conditioning variables and the outgoing arcs to the variables on the left of its conditioning bar.

[10] That the overall procedure of collecting and distributing evidence described above cannot be simplified, is discussed in Appendix A.2.

[11] Referee 1 pointed us to a possible connection between coherence graphs and Kohlas' notion of *kernels* [52, Section 4.5]. Indeed Kohlas' Lemmas 4.17 and 4.18 seem to have a similar aim to the mentioned marginal extension theorem. This prospective relation appears to be worth exploring in future work.

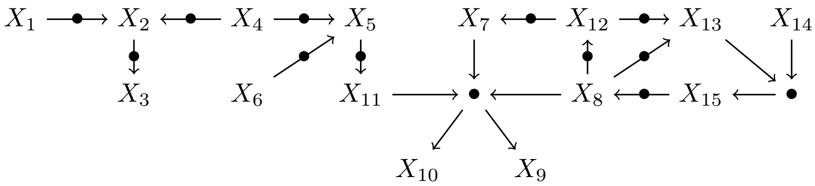

Figure 4: Coherence graph for the running example.

In [64] we showed that in order to verify the coherence of a number of assessments it suffices to do it independently in each of the *superblocks* of their associated coherence graph. These superblocks are built in the following manner:

- Within a coherence graph, we call *source of contradiction* each variable with more than one parent, or that belongs to a cycle. In Figure 4, variables X_2, X_5, X_{13} are sources of contradiction since they have more than one parent; $X_8, X_{12}, X_{13}, X_{15}$ are sources of contradiction as they are involved in cycles.
- The *block* associated with a source of contradiction in made up with all its predecessor circles and related variables (templates) in the coherence graphs. Figure 5 displays the blocks for the running example graph as dashed boxes. On the leftmost part, we can see the two blocks originated by X_2 and X_5, respectively. The remaining box on the rightmost part represents the block that X_{13} originates and that coincides with the block that the variables involved in cycles originate (that is, $X_8, X_{12}, X_{13}, X_{15}$).
- We put together all blocks that have variables in common, thus forming a superblock. In Figure 5 there are two superblocks: the first is given by the union of the two blocks on the left, since they share variable X_4; the second is equal to the single block on the right (or, equivalently, to the union of the two coinciding blocks that share all variables).

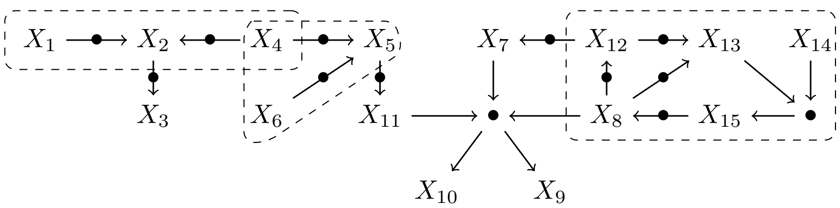

Figure 5: Blocks defining the two superblocks in the coherence graph of the running example.

The structure of the superblocks is equivalent to a partition of our sets of assessments: each superblock makes up for an element of the partition; the assessments not involved in any superblock make up the last element of the partition. It was proven in [64], in the context of coherent lower previsions, that our initial assessments are coherent (avoid partial loss) if those that belong to the same superblock are coherent (avoid partial loss). Similar results hold for sets of desirable gambles:

Proposition 11. *Let us consider a number of templates $X_{O_1}|X_{I_1}, \dots, X_{O_r}|X_{I_r}$ and associated separately coherent conditional sets of desirable gambles $\mathcal{D}_{O_1}|X_{I_1}, \dots, \mathcal{D}_{O_r}|X_{I_r}$. Consider also the associated coherence graph, which induces a partition $\mathcal{B}$ of $\{1, \dots, r\}$. If for each $B \in \mathcal{B}$ it holds that $\cup_{j \in B} \mathcal{D}_{O_j}|X_{I_j}$ avoids partial loss, then $\cup_{j=1}^{r} \mathcal{D}_{O_j}|X_{I_j}$ avoids partial loss.*

In particular, we can also prove that it suffices to verify the compatibility in each superblock separately, since from this we can immediately derive the compatibility overall:

Theorem 12. *Let us consider a number of templates $X_{O_1}|X_{I_1}, \dots, X_{O_r}|X_{I_r}$ and associated separately coherent conditional sets of desirable gambles $\mathcal{D}_{O_1}|X_{I_1}, \dots, \mathcal{D}_{O_r}|X_{I_r}$. Consider also the associated coherence graph, which induces a partition $\mathcal{B}$ of $\{1, \dots, r\}$. If for each $B \in \mathcal{B}$ it holds that $\cup_{j \in B} \mathcal{D}_{O_j}|X_{I_j}$ are compatible, then $\cup_{j=1}^{r} \mathcal{D}_{O_j}|X_{I_j}$ are compatible. Their natural extension is the set $\mathcal{D}_l$ determined by Algorithm 3.*

Algorithm 3 Natural extension

```
procedure NATURAL EXTENSION(a coherence graph; desirability assessments on each node)
    Let B' := {B ∈ B : |B| > 1};                                    ▷ These are associated with the superblocks.
    let D_0 be the natural extension of ∪_{B∈B'} ∪_{j∈B} D_{O_j}|X_{I_j};   ▷ Superblocks have disjoint sets of variables.
    let C := {1, ..., r} \ (∪_{B∈B'} B);                            ▷ These are the remaining indices.
    Consider an order {j_1, ..., j_l} of C so that O_{j_m} ∩_{(m'<m)} (O_{j'_m} ∪ I_{j'_m}) = ∅ ∀m; ▷ It exists by [97, Lemma 1].
    for i ← 1, l do                                                 ▷ We proceed iteratively.
        let D_i be the natural extension of D_{i-1} ∪ D_{O_{j_i}}|X_{I_{j_i}};
    end for
    return D_l.
end procedure
```

In other words, once we are given our conditional sets of desirable gambles on $X_{O_1}|X_{I_1}, \dots, X_{O_r}|X_{I_r}$, we should proceed as follows:

- ◦ We build the coherence graph associated with these sets of variables.
- ◦ On each superblock, we determine the associated junction tree.
- ◦ We verify the compatibility of the subset of the assessments belonging to that junction tree.

In particular, if we consider the assessments in our running example (Eq. (10)), this means that we should only verify the compatibility of:

- ◦ $\mathcal{D}_2|X_1, \mathcal{D}_2|X_4, \mathcal{D}_5|X_4, \mathcal{D}_5|X_6$, on the one hand; and
- ◦ $\mathcal{D}_{13}|X_{12}, \mathcal{D}_{12}|X_8, \mathcal{D}_8|X_{15}, \mathcal{D}_{13}|X_8, \mathcal{D}_{15}|X_{\{13,14\}}$, on the other.

In the first one, we obtain the following junction tree:

$$\overset{\mathcal{D}_2|X_1}{\{X_1, X_2\}} \xrightarrow{X_2} \overset{\mathcal{D}_2|X_4}{\{X_2, X_4\}} \xrightarrow{X_4} \overset{\mathcal{D}_5|X_4}{\{X_4, X_5\}} \xrightarrow{X_5} \overset{\mathcal{D}_5|X_6}{\{X_5, X_6\}}$$

Figure 6: Junction tree for the first superblock.

As a consequence, all we need to do in order to verify the compatibility of the assessments is to compute their natural extension $\mathcal{D}_{1,2,4,5,6}$ (or, more precisely, the intersections $\mathcal{D}_{1,2,4,5,6} \cap \mathcal{L}_{1,2}, \mathcal{D}_{1,2,4,5,6} \cap \mathcal{L}_{2,4}, \mathcal{D}_{1,2,4,5,6} \cap \mathcal{L}_{4,5}$ and $\mathcal{D}_{1,2,4,5,6} \cap \mathcal{L}_{5,6}$) by means of Algorithms 1 and 2 and then check if it induces the original assessments by means of (7).

In the second case, the junction tree is the following:

$$\overset{\mathcal{D}_{13}|X_{12},\, \mathcal{D}_{12}|X_8}{\{X_8, X_{12}, X_{13}\}} \xrightarrow{X_{\{8,13\}}} \overset{\mathcal{D}_8|X_{15},\, \mathcal{D}_{13}|X_8}{\{X_8, X_{13}, X_{15}\}} \xrightarrow{X_{\{13,15\}}} \overset{\mathcal{D}_{15}|X_{\{13,14\}}}{\{X_{13}, X_{14}, X_{15}\}}$$

Figure 7: Junction tree for the second superblock.

Therefore, here we should first of all compute the natural extension of $\mathcal{D}_{13}|X_8, \mathcal{D}_{12}|X_8, \mathcal{D}_{13}|X_{12}, \mathcal{E}_{12,13,8}$; that of $\mathcal{D}_{13}|X_8, \mathcal{D}_8|X_{15}, \mathcal{E}_{8,13,15}$ (note that these two sets are always compatible because of Proposition 13); and then verify the compatibility of $\mathcal{E}_{12,13,8}, \mathcal{E}_{8,13,15}, \mathcal{D}_{15}|X_{14,13}$, by means of Algorithms 1, 2 and Eq. (7).

Thus, the use of coherence graphs allows us to significantly simplify the study of the problem of compatibility.[12]

[12] For a different way to exploit the marginal extension theorem to the extent of checking compatibility, see Appendix A.3.

6. Conclusions

In this paper, we have initially generalised the classical result on the compatibility of a number of marginal probabilities into a global one to the case where our belief models are sets of desirable gambles. This includes as particular cases sets of probability measures and also most models of non-additive measures, such as belief functions or possibility measures. Our generalisation covers also the case of infinite possibility spaces and is not constrained by measurability issues. There are, however, other works on the marginal problem that do not fall into the framework of our Proposition 1: this is for instance the case of Studený's work on ordinal conditional functions and relational databases [82, 83].

We have then considered compatibility in the conditional case and shown that we can solve the problem through junction tree propagation. Apparently, this is the first time that the link between RIP and compatibility is established in the conditional case. We have then shown that the problem can be further simplified joining junction trees and coherence graphs. By these tools, the complexity of checking compatibility may greatly decrease in applications, as it is already known to happen in the unconditional case.

As for future work, the following possibilities seem to be worth considering:

- In this paper we have focused on computing the least-committal joint model compatible with given assessments (i.e., the natural extension). It may be useful to generalise our results so as to make them work with other types of extensions, which satisfy additional requirements. To this end, we think the most promising way would be to expand on our initial connection with information algebras [52] as sketched in Appendix A.4. More generally speaking, and thanks to the motivating comments by Referee 1, we have come to appreciate the power of information algebras, which appear to be very nicely suited to be joined with desirability. We believe there is much to be gained in deepening such a connection.

- At the moment our Algorithm 3, for the computation of the compatible joint in the mixed environment made by junction trees and coherence graphs, does not exploit the form of the coherence graph to decrease the computational complexity. There is certainly room to improve on this, even though the task does not seem immediate to achieve.

- There could be an interesting application of our results to probabilistic satisfiability. The reason is that our framework is general enough to model uniformly both the logical part of the problem (by means of degenerate probabilities) and the probabilistic information on top of it, possibly in an imprecise form. Moreover, it would also be possible to compute the probabilistic implications of the problem on any variables: it would be enough to add those variables to the problem and place a totally uninformative (i.e., vacuous) probability over them, and then let our procedures compute the natural extension (note that this would not be possible using precise probability).

- Computing the natural extension exactly may be costly and it can be necessary to resort to approximate methods. In this light, it would be useful to verify whether our past results on the iterative approximation of the natural extension, in a compatibility context [65, Section 5], can be joined with the current work so as to make a workable algorithm. More generally speaking, the statistical literature has produced a number of algorithms for compatibility that would be useful to merge in some way with our results here.

- Note, from Proposition 6, that if the initial assessments incur partial loss, then compatibility does not hold. One possibility then would be to consider first the approaches to correct incoherent assessments that have been discussed in the literature (e.g., [15, 16, 74] and [98, Section 4.1.1]) and then apply the results in this paper.

- As mentioned in the Introduction, Lasserre has heavily exploited RIP to decrease the complexity of polynomial optimisation problems [56]. Let us recall that Lasserre's work has deep implications on making logic and probability computationally efficient [8]. Since in this paper we relax Lasserre's assumptions (for instance by not relying on σ-additivity and by allowing imprecision) and we enable the conditional case to be treated, in addition to the unconditional one, we expect that our results should be useful to extend his work to other applications.

Finally, we would like to stress that the notion of compatibility we have considered in this paper corresponds to *Williams* (*strong*) *coherence* in imprecise probability. As such, it does not take into account the property of conglomerability (which is relevant to conditioning probabilistic models in the case of infinite spaces, see, e.g., [69]). In fact, Theorem 2 does not extend towards conglomerability, in the following sense: if we consider pairwise compatible and fully conglomerable coherent marginal previsions defined on sets $S_1, \ldots, S_r$ satisfying RIP, their natural extension, while being a coherent compatible joint by Theorem 2, need not be conglomerable. A detailed study of the compatibility problem under conglomerability is one of the main foundational open problems for the future.

Acknowledgements

We acknowledge the financial support of projects PGC2018-098623-B-I00 and GRUPIN/IDI/2018/000176 and of the Swiss NSF grant n. IZKSZ2_162188. We would also like to thank the anonymous Referees for a number of useful remarks.

Appendix A. Additional remarks

Appendix A.1. Conditional compatibility cannot be reduced to the unconditional case

The compatibility of $\mathcal{D}_{O_1}|X_{I_1}, \ldots, \mathcal{D}_{O_r}|X_{I_r}$ does not imply the pairwise compatibility of the sets $\mathcal{E}_1, \ldots, \mathcal{E}_r$. Let us illustrate this question with the following example, where we have three variables X_1, X_2, X_3 and conditional information in terms of $X_1|X_2$ and $X_2|X_3$:

Example 5. *Consider binary spaces $\mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_3$, and let us make the following conditional assessments:*

$$\begin{aligned}
\mathcal{D}_1|(X_2 = 0) &:= \{f \in \mathcal{L}_{12} : f = \mathbb{I}_{X_2=0} f, f(1,0) + f(0,0) > 0\},\\
\mathcal{D}_1|(X_2 = 1) &:= \{f \in \mathcal{L}_{12} : f = \mathbb{I}_{X_2=1} f, f(1,1) + f(0,1) > 0\},\\
\mathcal{D}_2|(X_3 = 0) &:= \{f \in \mathcal{L}_{23} : f = \mathbb{I}_{X_3=0} f, f(1,0) + f(0,0) > 0\},\\
\mathcal{D}_2|(X_3 = 1) &:= \{f \in \mathcal{L}_{23} : f = \mathbb{I}_{X_3=1} f, f(1,1) + f(0,1) > 0\}.
\end{aligned}$$

These four sets of desirable gambles are compatible, in the sense that there is a coherent set of desirable gambles $\mathcal{E} \subseteq \mathcal{L}$ that exactly induces each of them. It is given by:

$$\mathcal{E} := \{f \in \mathcal{L} : \sum_{x \in \mathcal{X}} f(x) > 0\}. \tag{A.1}$$

Indeed, given $f \in \mathcal{L}_{12}$, it holds that

$$\mathbb{I}_{X_2=0} f \in \mathcal{E} \Leftrightarrow f(0,0) + f(1,0) > 0 \Leftrightarrow f \in \mathcal{D}_1|(X_2 = 0),$$

and similarly for the other cases.

Now if we want to consider pairwise compatibility, we need to have coherent sets of desirable gambles on the sets of variables $S_1 := \{1,2\}$ and $S_2 := \{2,3\}$, respectively, which at present we have not. To this end, we need to consider the natural extension of $\mathcal{D}_1|X_2 := \mathcal{D}_1|(X_2 = 0) \cup \mathcal{D}_1|(X_2 = 1)$ to $\mathcal{L}_{12}$, and the natural extension of $\mathcal{D}_2|X_3 := \mathcal{D}_2|(X_3 = 0) \cup \mathcal{D}_2|(X_3 = 1)$ to $\mathcal{L}_{23}$. Using Eq. (6)*, these are respectively given by*

$$\mathcal{E}_{S_1} := \mathrm{posi}_{S_1}(\mathcal{L}^+_{S_1} \cup \mathcal{D}_1|X_2) = \{f \in \mathcal{L}_{12} : f \neq 0, f(1,0) + f(0,0) \geq 0 \textit{ and } f(1,1) + f(0,1) \geq 0\}$$

and

$$\mathcal{E}_{S_2} := \mathrm{posi}_{S_2}(\mathcal{L}^+_{S_2} \cup \mathcal{D}_2|X_3) = \{f \in \mathcal{L}_{23} : f \neq 0, f(1,0) + f(0,0) \geq 0 \textit{ and } f(1,1) + f(0,1) \geq 0\}.$$

However, these two sets are not *pairwise compatible, since*

$$\begin{aligned}
f \in \mathcal{L}(\mathcal{X}_2) \cap \mathcal{E}_{S_1} &\Leftrightarrow f \neq 0, f(0) \geq 0, f(1) \geq 0 \Leftrightarrow f \in \mathcal{L}^+(\mathcal{X}_2) \textit{ while}\\
f \in \mathcal{L}(\mathcal{X}_1) \cap \mathcal{E}_{S_2} &\Leftrightarrow f \neq 0, f(1) + f(0) \geq 0 \Leftrightarrow f(1) + f(0) > 0.
\end{aligned}$$

This implies that the marginals of the joint model $\mathcal{E}$ do not coincide with $\mathcal{E}_{S_1}$ and $\mathcal{E}_{S_2}$. This is the source of the failure of pairwise compatibility.

And yet note that $\mathcal{D}_1|X_2$ and $\mathcal{D}_2|X_3$ jointly avoid partial loss, since they are both included in the coherent set $\mathcal{E}$ given by Eq. (A.1). $\lozenge$

Appendix A.2. The algorithms of collecting and distributing evidence cannot be simplified

Note that the overall procedure of collecting and distributing evidence described above cannot be simplified, in the sense that, for any set of variables A, it does not hold that

$$\text{posi}(\cup_{i=1}^{q}\mathcal{D}_i \cup \mathcal{L}^+) \cap \mathcal{L}_A = \text{posi}(\cup_{S_i \cap A \neq \emptyset}\mathcal{D}_i \cup \mathcal{L}^+) \cap \mathcal{L}_A;$$

that is, even if a set of desirable gambles does not involve any variable in the set A, it could be that it has behavioural implications on A when we propagate information through the tree:

Example 6. *Let X_1, X_2, X_3 be binary variables, and consider the conditional assessments $X_2|X_1$ and $X_3|X_2$ given by*

$$X_1 = 0 \Rightarrow X_2 = 1; X_1 = 1 \Rightarrow X_2 = 1; X_2 = 0 \Rightarrow X_3 = 0; X_2 = 1 \Rightarrow X_3 = 1.$$

These can be modelled by means of the following conditional sets of desirable gambles:

$$\mathcal{D}_{12} := \mathcal{D}_2|X_1 = \{f \in \mathcal{L}_{12} : f(0,1) \geq 0, f(1,1) \geq 0, \max\{f(0,1), f(1,1)\} > 0\};$$
$$\mathcal{D}_{23} := \mathcal{D}_3|X_2 = \{f \in \mathcal{L}_{23} : f(0,0) \geq 0, f(1,1) \geq 0, \max\{f(0,0), f(1,1)\} > 0\}.$$

The gamble $g := \mathbb{I}_{X_3=1} - 2\mathbb{I}_{X_3=0}$ belongs to $\text{posi}(\mathcal{D}_{12} \cup \mathcal{D}_{23} \cup \mathcal{L}^+) \cap \mathcal{L}_3$: to prove this, note that $g \geq f_1 + f_2$, for

$$f_1 := \frac{1}{2}\mathbb{I}_{X_2=1} - 3\mathbb{I}_{X_2=0} \in \mathcal{D}_{12} \text{ and } f_2 := \frac{1}{2}\mathbb{I}_{X_2=X_3} - 3\,\mathbb{I}_{X_2 \neq X_3} \in \mathcal{D}_{23}.$$

However,

$$\text{posi}((\mathcal{D}_{12} \cap \mathcal{L}_3) \cup (\mathcal{D}_{23} \cap \mathcal{L}_3) \cup \mathcal{L}^+) \cap \mathcal{L}_3 = \text{posi}((\mathcal{D}_{23} \cap \mathcal{L}_3) \cup \mathcal{L}^+) \cap \mathcal{L}_3 = \mathcal{L}_3^+$$

and therefore the two sets do not coincide. $\Diamond$

This may be perhaps more easily be seen using a (precise) probabilistic approach: we consider the conditional probabilities $P(X_2|X_1)$ and $P(X_3|X_2)$ such that in the second $X_3 = X_2$ and in the first we effectively obtain that $P(X_2 = 1) = 1$, irrespective of the marginal on X_1, then we deduce that it must be $P(X_3 = 1) = 1$, even if this cannot be obtained from $P(X_3|X_2)$ alone. Note that the example works because we are introducing zero probabilities in the assessment; otherwise it would depend on the unknown marginal distribution of X_1.

Appendix A.3. Compatibility of nested assessments

It is interesting to explicitly characterise the compatibility of the sets of desirable gambles $\mathcal{D}_1, \ldots, \mathcal{D}_r$, understood in terms of avoiding partial loss, in one particular instance of RIP: when the natural order establishes a chain in the pairwise intersections, in the sense that, given $j_1 < j_2 < j_3$, it holds that $S_{j_3} \cap S_{j_1} \subseteq S_{j_3} \cap S_{j_2}$. This may be useful when our assessments are established in an incremental manner, as is for instance the case with sequences of observations, which is a case that would not be treated as effectively with junction trees.

Proposition 13. *Consider sets of variables $S_1, \ldots, S_r$ such that $S_i \cap (\cup_{j<i}S_j) \subseteq S_{i-1}$ for every $i = 2, \ldots, r$, and sets of desirable gambles $\mathcal{D}_1, \ldots, \mathcal{D}_r$, where $\mathcal{D}_j$ is coherent relative to $\mathcal{L}_{S_j}$. Let us define recursively $\mathcal{D}'_1, \ldots, \mathcal{D}'_r$ in the following manner:*

$$\begin{cases} \mathcal{D}'_1 := \mathcal{D}_1, \\ \mathcal{D}'_j := \text{posi}_{S_j}(\mathcal{D}_j \cup (\mathcal{D}'_{j-1} \cap \mathcal{L}_{S_{j-1} \cap S_j})) & \text{if } j > 1. \end{cases}$$

Then

$$\mathcal{D}_1, \ldots, \mathcal{D}_r \text{ avoid partial loss } \Leftrightarrow \mathcal{D}'_r \text{ coherent}.$$

In that case, for every $j = 1, \ldots, r$, $\mathcal{D}'_j = \mathcal{D}'_r \cap \mathcal{L}_{S_j}$.

The procedure in this proposition is a generalisation of the *marginal extension theorem* established in [89, Section 6.7] and [62] for coherent lower previsions, which in turn is an extension of the law of total probability from probability theory. This result also settles the problem of verifying compatibility in case the sets $S_1, \ldots, S_r$ are nested:

- When $\mathcal{D}_1, \ldots, \mathcal{D}_r$ correspond to unconditional assessments, we must check whether $\mathcal{D}'_r$ is a coherent set of desirable gambles and $\mathcal{D}'_j = \mathcal{D}_j$ for every $j = 1, \ldots, r$.

- When $\mathcal{D}_1, \ldots, \mathcal{D}_r$ correspond to conditional assessments, we must check whether $\mathcal{D}'_r$ is a coherent set of desirable gambles and whether $\mathcal{D}'_j$ induces the conditional assessments in $\mathcal{D}_j$ by means of Eq. (7) for every $j = 1, \ldots, r$. $\Diamond$

Appendix A.4. Information and valuation algebras

Valuation algebras are a very general representation of knowledge or information [52, 78]. They abstract away the most important features that appear in nearly every representation, and at such an abstract level, they provide basic operations to make inference. Among these basic operations, valuation algebras provide a very general formulation, as well as a justification, of the junction tree algorithm.

In the original version of this paper, we noticed that there was a natural connection between the present work and valuation algebras, which we were proposing to investigate in future work. However Referee 1 motivated us to start deepening the connection already in this paper. This is what we set out to do in the next section.

Appendix A.4.1. Coherent sets of desirable gambles as valuation algebras

The key observation is that the theory of sets of desirable gambles can be embedded into that of valuation algebras. To prove this, let us start by considering a set of variables V. Each valuation ϕ refers to a finite subset $d(\phi)$ of V, called its *domain*, and it represents some information about these variables. We shall denote by Φ_D the set of all valuations with domain D, and let $\Phi := \cup\{\phi_D : D \subseteq V\}$ be the set of all valuations. The map d is usually called the *labelling* operator. In a valuation algebra, there are two other types of operations: a *combination* operator $\otimes$, which joins the information encoded by two different valuations, and a *marginalisation* operator $\downarrow$, which focuses the knowledge encoded by a valuation onto a smaller domain. Then:

Definition 30 (Valuation algebra). *A system $(\Phi, V, d, \otimes, \downarrow)$ is a* valuation algebra *when it satisfies the following axioms:*

A1. Φ *(resp., Φ_D) is commutative and associative under combination;*

A2. $(\forall \phi, \phi_1, \phi_2 \in \Phi)(\forall D \subseteq d(\phi))\ d(\phi_1 \otimes \phi_2) = d(\phi_1) \cup d(\phi_2)$ *and* $d(\phi^{\downarrow D}) = D$*;*

A3. $(\forall \phi \in \Phi)\ \phi^{\downarrow d(\phi)} = \phi$*;*

A4. $(\forall D \subseteq D' \subseteq d(\phi))(\forall \phi \in \Phi)\ \phi^{\downarrow D} = (\phi^{\downarrow D'})^{\downarrow D}$*;*

A5. *If* $\phi_1, \phi_2 \in \Phi$ *are valuations with* $D_1 := d(\phi_1)$ *and* $D_2 := d(\phi_2)$*, then* $(\phi_1 \otimes \phi_2)^{\downarrow D_1} = \phi_1 \otimes \phi_2^{\downarrow D_1 \cap D_2}$*;*

A6. $(\forall D \subseteq V)(\exists e_D \in \Phi_D)(\forall \phi \in \Phi_D)\ \phi \otimes e_D = e_D \otimes \phi = \phi$*, and moreover* $(\forall D_1, D_2 \subseteq V)\ e_{D_1} \otimes e_{D_2} = e_{D_1 \cup D_2}$.

Let us show that coherent sets of desirable gambles can be embedded into this theory. Using the notation we have employed throughout the paper, given a set of indices $N := \{1, \ldots, n\}$ and possibility spaces $\mathcal{X}_1, \ldots, \mathcal{X}_n$, we let Φ_S be the family of sets of desirable gambles $\mathcal{D} \subseteq \mathcal{L}_S(\mathcal{X})$ that are coherent relative to $\mathcal{L}_S(\mathcal{X})$, in the manner specified in Definition 7. Let $\Phi := \cup_{S \subseteq N} \Phi_S$. The labelling operator is then given by

$$d(\phi) := \cap\{S : \phi \in \Phi_S\}. \tag{A.2}$$

Next, the combination operator we shall consider is related to the natural extension: we let

$$\phi_1 \otimes \phi_2 := \operatorname{posi}(\phi_1 \cup \phi_2 \cup \mathcal{L}^+_{d(\phi_1) \cup d(\phi_2)}); \tag{A.3}$$

and finally, the marginalisation operator is given by

$$\phi^{\downarrow S} := \phi \cap \mathcal{L}_S(\mathcal{X}). \tag{A.4}$$

It is not difficult to establish the following:

Proposition 14. *The set Φ equipped with the operators above is a valuation algebra.*

As a consequence, we can use all the machinery of valuation algebras for coherent sets of desirable gambles. In particular, this means that Theorem 9 follows immediately using [52, Theorem 4.8]; similarly, Theorem 10 is an immediate consequence of [52, Theorem 4.10].

Appendix A.4.2. Desirable gambles, logic, and information algebras

We would like to conclude this detour on algebras by discussing in more general terms the relation between desirable gambles and information algebras. In fact, Kohlas devotes a chapter of his book to *information* algebras that are, loosely speaking, valuation algebras with an additional property of idempotency. It is particularly interesting to focus on Section 6.4 of Kohlas' book, where he describes how a general logical system can be proven to be an information algebra.

In such a context one needs a language and a consequence operator—as originally defined by Tarski [84, Chapter 5]. In our case the language is just the set of all gambles $\mathcal{L}$ and the consequence operator is the natural extension:

$$C(\mathcal{G}) \coloneqq \operatorname{posi}(\mathcal{G} \cup \mathcal{L}^+),$$

which associates with any subset $\mathcal{G} \subseteq \mathcal{L}$ its coherent closure.

It is very easy to prove that $C(\cdot)$ complies with Kohlas' requirements E1 and E2 in [52, Section 6.4], whence satisfying the definition of a consequence operator. Moreover, it is possible to prove that it also satisfies properties C4 and C5 in that same section. From this we get, using Kohlas' results, that desirable gambles are quite a general form of an information algebra. In fact, we could have actually used this path to prove, in the previous section, that desirable gambles make up a valuation algebra; but it was somewhat of an overkill for our aims, whence we dealt with it more directly by showing that desirability satisfies the axioms of valuation algebras.

So why do we think that the question of information algebras is relevant to the discussion here?

We believe it is relevant because traditionally there has been a disconnection between logic and probability, of which we can see many examples in the literature. For example, the original developments about belief revision (e.g., see [37]) were essentially based on, and made for, logic, and only then probability entered the picture in a kind of ad-hoc way; on the philosophical side, we can for instance find Howson that wonders whether probability and logic can be combined [45]; we can see something similar also in Kohlas' book when he devotes the entire Chapter 7 to embedding probability and belief functions in his theory (see also [40, Section 2.2]). For similar reasons, we believe, Kohlas introduces a number of variants of valuation algebras to account for ratios (needed by Bayes' rule) and products (independence).

In our view, these difficulties are originated by one unfortunate, and yet stubborn, choice: that of representing probability in its habitual form—which we regard as the 'primal' representation of probability. This should be contrasted to its 'dual' representation, which is nothing else than desirability. Let us stress that we are actually talking of the *mathematical* dual, which is obtained through linear programming in the finite case or by a separating hyperplane theorem in the infinte case (see [89, Appendix E], [9]). When we move to the dual form we obtain desirability; and desirability, as we have seen, is a pure logical theory. In this form, there is no need to find special ways to accommodate probability in a setting originally conceived for logic, everything becomes straightforward. For example, the embedding of belief functions into algebras on which Kohlas and collaborators have spent much energy, on the wake of Shenoy and Shafer's seminal work [79], becomes a byproduct of the far easier embedding of desirability.[13] Note, in addition, that desirability does not need ratios to define Bayes' rule (Definition 16) and independence does not necessarily need products [24, Definitions 3 and 5]. As a consequence, a bare information algebra is all one needs to live in the most general case.

Overall, we argue that it is the missing duality step that has markedly slowed down the unification of logic and probability, as well as a number of important developments; and we claim it is imprecise probability's merit to have changed perspective, thus allowing for such an alternative avenue. It is a simple step, after all, but one that has not frequently been taken outside the imprecise probability community, not even nowadays, after more than 40 years that desirability has been introduced by Williams [94] and then repeatedly proposed (e.g., [8, 23, 66, 70, 75, 89, 90, 98]).

We hope that this further discussion convinces more people to take up desirability as a very convenient way to work with probability in purely logical terms.

Appendix B. Proofs

Proof of Proposition 1. It suffices to prove the direct implication, the converse being trivial.

[13] Yet, let us remark that those past works adopt a non-probabilistic interpretation of belief functions, unlike desirability.

Let $\mathcal{D}$ be a coherent set of desirable gambles satisfying $\mathcal{D} \cap \mathcal{L}_{S_j} = \mathcal{D}_j$ for every j. Then it holds that $\cup_{j=1}^r \mathcal{D}_j \subseteq \mathcal{D}$, and, since $\mathcal{E}$ is the smallest coherent superset of $\cup_{j=1}^r \mathcal{D}_j$, we deduce that $\cup_{j=1}^r \mathcal{D}_j \subseteq \mathcal{E} \subseteq \mathcal{D}$. As a consequence, $\mathcal{E} \cap \mathcal{L}_{S_j} \subseteq \mathcal{D} \cap \mathcal{L}_{S_j} = \mathcal{D}_j$, and since the inclusion $\mathcal{E} \cap \mathcal{L}_{S_j} \supseteq \mathcal{D}_j$ always holds, we deduce that $\mathcal{E} \cap \mathcal{L}_{S_j} = \mathcal{D}_j$ for $j = 1, \ldots, r$. □

Proof of Theorem 2. Taking Proposition 1 into account, we are going to prove that the natural extension of $\cup_{j=1}^r \mathcal{D}_j$, given by $\text{posi}(\mathcal{L}^+ \cup \bigcup_{i=1}^r \mathcal{D}_i)$ is a coherent set of desirable gambles that is globally compatible with $\mathcal{D}_1, \ldots, \mathcal{D}_r$. We apply induction on r. We begin with the case of $r = 2$.

Let $\mathcal{D} := \text{posi}(\mathcal{L}^+ \cup \mathcal{D}_1 \cup \mathcal{D}_2)$. To prove that this is a coherent set of desirable gambles, it suffices to show that it avoids partial loss. Assume ex-absurdo that $\mathcal{D}$ incurs partial loss. Since $\mathcal{D}_1, \mathcal{D}_2$ are coherent relative to $\mathcal{L}_{S_1}, \mathcal{L}_{S_2}$, respectively, it follows that if $\mathcal{D}$ incurs partial loss there are $f_1 \in \mathcal{D}_1, f_2 \in \mathcal{D}_2$ such that $f_1 + f_2 \leq 0$. Let us define $g_1, g_2 \in \mathcal{L}_{S_1} \cap \mathcal{L}_{S_2}$ by

$$g_1(z) := \sup\{f_1(z') : \pi_{S_1 \cap S_2}(z) = \pi_{S_1 \cap S_2}(z')\} \text{ and } g_2(z) := \sup\{f_2(z') : \pi_{S_1 \cap S_2}(z) = \pi_{S_1 \cap S_2}(z')\} \tag{B.1}$$

for all $z \in \mathcal{X}$. Then by construction $g_1 \geq f_1$ and $g_2 \geq f_2$, whence $g_1 \in \mathcal{D}_1, g_2 \in \mathcal{D}_2$. Since moreover $g_1, g_2 \in \mathcal{L}_{S_1} \cap \mathcal{L}_{S_2}$, we deduce from pairwise compatibility that $g_1, g_2 \in \mathcal{D}_1 \cap \mathcal{D}_2$. Since for instance $\mathcal{D}_1$ is coherent, we deduce that there is some $z \in \mathcal{X}$ such that $(g_1 + g_2)(z) > 0$. By Eq. (B.1), there are z_1, z_2 such that $\pi_{S_1 \cap S_2}(z) = \pi_{S_1 \cap S_2}(z_1) = \pi_{S_1 \cap S_2}(z_2)$ and such that

$$g_1(z) - f_1(z_1) < \frac{g_1(z) + g_2(z)}{2} \text{ and } g_2(z) - f_2(z_2) < \frac{g_1(z) + g_2(z)}{2},$$

whence, by summing the two inequalities, we get that $f_1(z_1) + f_2(z_2) > 0$. Now, considering that $f_1 \in \mathcal{L}_{S_1}$ and that $f_2 \in \mathcal{L}_{S_2}$, we deduce the existence of some $z' \in \mathcal{X}$ such that $\pi_{S_1}(z') = \pi_{S_1}(z_1)$ and $\pi_{S_2}(z') = \pi_{S_2}(z_2)$, taking into account that the projections of z_1, z_2 on $S_1 \cap S_2$ coincide. As a consequence, $f_1(z') + f_2(z') > 0$, a contradiction.

Next, we show that $\mathcal{D} \cap \mathcal{L}_{S_1} = \mathcal{D}_1$; the proof of the equality $\mathcal{D} \cap \mathcal{L}_{S_2} = \mathcal{D}_2$ is analogous. Consider $f \in \mathcal{D} \cap \mathcal{L}_{S_1}$. Then there are $g \in \mathcal{D}_1 \cup \{0\}, h \in \mathcal{D}_2 \cup \{0\}$ such that $f \geq g + h$. Define $h' \in \mathcal{L}_{S_1} \cap \mathcal{L}_{S_2}$ for all $z \in \mathcal{X}$ by $h'(z) := \sup\{h(z') : \pi_{S_1 \cap S_2}(z) = \pi_{S_1 \cap S_2}(z')\}$. Then $h' \geq h$, whence $h' \in \mathcal{D}_2 \cup \{0\}$. Moreover, since $h' \in \mathcal{L}_{S_1} \cap \mathcal{L}_{S_2}$, also $h' \in \mathcal{D}_1 \cup \{0\}$. Besides, since $f \in \mathcal{L}_{S_1}$ and $f \geq g + h$, we deduce that also $f \geq g + h'$:

$$f(z) = \sup_{\pi_{S_1}(z') = \pi_{S_1}(z)} f(z') \geq \sup_{\pi_{S_1}(z') = \pi_{S_1}(z)} (g(z') + h(z')) = g(z) + \sup_{\pi_{S_1}(z') = \pi_{S_1}(z)} h(z') = g(z) + h'(z).$$

Since f is non-zero because it belongs to $\mathcal{D}$, we conclude that it belongs to $\mathcal{D}_1$. The converse inclusion is trivial.

Assume next that the result holds up to $r - 1$. Let us denote $S^{r-1} := \cup_{j=1}^{r-1} S_j$. Then the natural extension $\mathcal{D}^{r-1}$ of $\mathcal{D}_1, \ldots, \mathcal{D}_{r-1}$, given by $\mathcal{D}^{r-1} := \text{posi}_{S^{r-1}}(\mathcal{L}^+_{S^{r-1}} \cup \bigcup_{j=1}^{r-1} \mathcal{D}_j)$ is coherent relative to $\mathcal{L}_{S^{r-1}}$ and moreover it satisfies $\mathcal{D}^{r-1} \cap \mathcal{L}_{S_j} = \mathcal{D}_j$ for $j = 1, \ldots, r-1$.

Let $\mathcal{D}$ be the natural extension of $\mathcal{D}^{r-1} \cup \mathcal{D}_r$, given by $\mathcal{D} := \text{posi}(\mathcal{L}^+ \cup \mathcal{D}^{r-1} \cup \mathcal{D}_r)$. It follows that $\mathcal{D}$ coincides with the natural extension of $\cup_{j=1}^r \mathcal{D}_j$. To prove that it is a coherent set of desirable gambles, it suffices to show that it avoids partial loss. Assume ex-absurdo that $\mathcal{D}$ incurs partial loss. Since $\mathcal{D}^{r-1}, \mathcal{D}_r$ are coherent relative to $\mathcal{L}_{S^{r-1}}, \mathcal{L}_{S_r}$, respectively, this means that there are $f_1 \in \mathcal{D}^{r-1}, f_2 \in \mathcal{D}_r$ such that $f_1 + f_2 \leq 0$. Let us define $g_1, g_2 \in \mathcal{L}_{S^{r-1}} \cap \mathcal{L}_{S_r}$ for all $z \in \mathcal{X}$ by

$$g_1(z) := \sup\{f_1(z') : \pi_{S^{r-1} \cap S_r}(z) = \pi_{S^{r-1} \cap S_r}(z')\} \text{ and } g_2(z) := \sup\{f_2(z') : \pi_{S^{r-1} \cap S_r}(z) = \pi_{S^{r-1} \cap S_r}(z')\}. \tag{B.2}$$

Then by construction $g_1 \geq f_1$ and $g_2 \geq f_2$, whence $g_1 \in \mathcal{D}^{r-1}, g_2 \in \mathcal{D}_r$. Since moreover $g_1, g_2 \in \mathcal{L}_{S^{r-1}} \cap \mathcal{L}_{S_r}$, we conclude that $g_1, g_2 \in \mathcal{D}^{r-1} \cap \mathcal{D}_r$. Since for instance $\mathcal{D}^{r-1}$ is coherent with respect to $\mathcal{L}_{S^{r-1}}$, we deduce that there is some $z \in \mathcal{X}$ such that $(g_1 + g_2)(z) > 0$. By Eq. (B.2), there are z_1, z_2 such that $\pi_{S^{r-1} \cap S_r}(z) = \pi_{S^{r-1} \cap S_r}(z_1) = \pi_{S^{r-1} \cap S_r}(z_2)$ and such that

$$g_1(z) - f_1(z_1) < \frac{g_1(z) + g_2(z)}{2} \text{ and } g_2(z) - f_2(z_2) < \frac{g_1(z) + g_2(z)}{2}.$$

Now, considering that $f_1 \in \mathcal{L}_{S^{r-1}}$ and that $f_2 \in \mathcal{L}_{S_r}$, we deduce the existence of some $z' \in \mathcal{X}$ such that $\pi_{S^{r-1}}(z') = \pi_{S^{r-1}}(z_1)$ and $\pi_{S_r}(z') = \pi_{S_r}(z_2)$, taking into account that the projections of z_1, z_2 on $S^{r-1} \cap S_r$ coincide. As a consequence, $f_1(z') + f_2(z') > 0$, a contradiction.

To conclude, let us show that $\mathcal{D} \cap \mathcal{L}_{S_r} = \mathcal{D}_r$ and $\mathcal{D} \cap \mathcal{L}_{S^{r-1}} = \mathcal{D}^{r-1}$:

- Consider $f \in \mathcal{D} \cap \mathcal{L}_{S_r}$. Then there are $g \in \mathcal{D}^{r-1} \cup \{0\}, h \in \mathcal{D}_r \cup \{0\}$ such that $f \geq g + h$. Assume that $g \neq 0$; otherwise it is immediate that $f \in \mathcal{D}_r$ (it must be $f \neq 0$ because it belongs to $\mathcal{D}$). For any $z \in \mathcal{X}$,

$$f(z) = \sup_{\pi_{S_r}(z')=\pi_{S_r}(z)} f(z') \geq \sup_{\pi_{S_r}(z')=\pi_{S_r}(z)} (g(z') + h(z')) = h(z) + \sup_{\pi_{S_r}(z')=\pi_{S_r}(z)} g(z').$$

  Let g' be given by $g'(z) := \sup_{\pi_{S_r}(z')=\pi_{S_r}(z)} g(z')$. Then $g' \geq g$, whence $g' \in \mathcal{D}^{r-1}$. Moreover, $g' \in \mathcal{L}_{S_r}$. Applying the induction hypothesis, we conclude that $g' \in \mathcal{D}_r$, and since $f \geq g' + h$ we conclude that also $f \in \mathcal{D}_r$. Thus, $\mathcal{D} \cap \mathcal{L}_{S_r} \subseteq \mathcal{D}_r$. The converse inclusion $\mathcal{D}_r \subseteq \mathcal{D} \cap \mathcal{L}_{S_r}$ is trivial.

- Consider $f \in \mathcal{D} \cap \mathcal{L}_{S^{r-1}}$. Then there are $g \in \mathcal{D}^{r-1} \cup \{0\}, h \in \mathcal{D}_r \cup \{0\}$ such that $f \geq g + h$. Assume that $h \neq 0$; otherwise it is immediate that $f \in \mathcal{D}^{r-1}$, given that it is $f \neq 0$ because it belongs to $\mathcal{D}$. For any $z \in \mathcal{X}$,

$$f(z) = \sup_{\pi_{S^{r-1}}(z')=\pi_{S^{r-1}}(z)} f(z') \geq \sup_{\pi_{S^{r-1}}(z')=\pi_{S^{r-1}}(z)} (g(z') + h(z')) = g(z) + \sup_{\pi_{S^{r-1}}(z')=\pi_{S^{r-1}}(z)} h(z').$$

  Let h' be given by $h'(z) := \sup_{\pi_{S^{r-1}}(z')=\pi_{S^{r-1}}(z)} h(z')$. Then $h' \geq h$, whence $h' \in \mathcal{D}_r$. Moreover, by construction $h' \in \mathcal{L}_{S^{r-1}}$, whence it belongs to $\mathcal{L}_{S^{r-1}} \cap \mathcal{L}_{S_r} = \mathcal{L}_{S_j}$ for some $j \in \{1, \ldots, r-1\}$, taking into account that the sets $S_1, \ldots, S_r$ satisfy RIP. Thus, $h' \in \mathcal{D}_r \cap \mathcal{L}_{S_j} = \mathcal{D}_j$ by pairwise compatibility, and as a consequence it also belongs to $\mathcal{D}^{r-1}$. Since $f \geq g + h'$, we conclude that $f \in \mathcal{D}^{r-1}$. The inclusion $\mathcal{D}^{r-1} \subseteq \mathcal{D} \cap \mathcal{L}_{S^{r-1}}$ is trivial.

We deduce that for every $j = 1, \ldots, r-1$, $\mathcal{D} \cap \mathcal{L}_{S_j} = \mathcal{D} \cap \mathcal{D}^{r-1} \cap \mathcal{L}_{S_j} = \mathcal{D}^{r-1} \cap \mathcal{L}_{S_j} = \mathcal{D}_j$. Since we have already proven that $\mathcal{D} \cap \mathcal{L}_{S_r} = \mathcal{D}_r$, we conclude that $\mathcal{D}$ satisfies the desired properties. □

Proof of Corollary 4. We prove the direct implication as the converse is trivial. Assume that $\underline{P}_1, \ldots, \underline{P}_r$ are pairwise compatible, and let $\mathcal{D}_1, \ldots, \mathcal{D}_r$ be their associated sets of strictly desirable gambles, given by Eq. (5). To prove that they are pairwise compatible, take $i \neq j$ in $\{1, \ldots, r\}$ and a gamble $f \in \mathcal{D}_i \cap \mathcal{L}_{S_j}$. Then either $f \gneq 0$, in which case also $f \in \mathcal{D}_j$, or $\underline{P}_i(f) = \underline{P}_j(f) > 0$, whence $f \in \mathcal{D}_j$. In any case, we conclude that $f \in \mathcal{D}_j \cap \mathcal{L}_{S_i}$, and since the converse inclusion is analogous we conclude that $\mathcal{D}_i, \mathcal{D}_j$ are pairwise compatible.

Applying Theorem 2, we conclude that the natural extension $\mathcal{D}$ of $\cup_{i=1}^r \mathcal{D}_i$ is a coherent set of desirable gambles that is compatible with $\mathcal{D}_1, \ldots, \mathcal{D}_r$. Let $\underline{P}$ be the coherent lower prevision it induces by means of (3). Then for $j = 1, \ldots, r$ and any gamble $f \in \mathcal{L}_{S_j}$, it holds that

$$\underline{P}(f) = \sup\{\mu : f - \mu \in \mathcal{D}\} = \sup\{\mu : f - \mu \in \mathcal{D} \cap \mathcal{L}_{S_j}\} = \sup\{\mu : f - \mu \in \mathcal{D}_j\} = \underline{P}_j(f),$$

where the last equality holds because $\mathcal{D}_j$ induces $\underline{P}_j$ by means of (3). Thus, $\underline{P}$ is compatible with $\underline{P}_1, \ldots, \underline{P}_r$. □

Proof of Proposition 6.

1. It follows from Eq. (7) that if $\mathcal{D}$ induces $\mathcal{D}_{O_j}|X_{I_j}$, then it must be $\mathcal{D}_{O_j}|X_{I_j} \subseteq \mathcal{D}$ for every $j = 1, \ldots, r$. As a consequence, $\cup_{j=1}^r \mathcal{D}_{O_j}|X_{I_j}$ has a coherent superset, or, in other words, it avoids partial loss.

2. Assume that $\mathcal{D}$ is a coherent set of desirable gambles that induces $\mathcal{D}_{O_j}|X_{I_j}$ by means of (7) for $j = 1, \ldots, r$. It follows from the first point that it must be $\cup_{j=1}^r \mathcal{D}_{O_j}|X_{I_j} \subseteq \mathcal{E} \subseteq \mathcal{D}$. Take $x_{I_j} \in \mathcal{X}_{I_j}$. Then

$$\begin{aligned}\mathcal{D}_{O_j}|x_{I_j} &\subseteq \{f \in (\cup_{i=1}^r \mathcal{D}_{O_i}|X_{I_i}) \cap \mathcal{L}_{O_j \cup I_j} : f = \mathbb{I}_{X_{I_j}=x_{I_j}} f\} \\ &\subseteq \{f \in \mathcal{E} \cap \mathcal{L}_{O_j \cup I_j} : f = \mathbb{I}_{X_{I_j}=x_{I_j}} f\} \\ &\subseteq \{f \in \mathcal{D} \cap \mathcal{L}_{O_j \cup I_j} : f = \mathbb{I}_{X_{I_j}=x_{I_j}} f\} = \mathcal{D}_{O_j}|x_{I_j},\end{aligned}$$

   whence $\mathcal{E}$ also induces $\mathcal{D}_{O_j}|X_{I_j}$ via Eq. (7). □

Proof of Proposition 7. The direct implication is trivial, so let us prove the converse.

Assume ex-absurdo that $\cup_{i=1}^r \mathcal{D}_i$ incurs partial loss. Since for every i the set $\mathcal{D}_i$ is a cone because of its relative coherence, it follows from Eq. (2) that there are $f_i \in \mathcal{D}_i \cup \{0\}, i = 1, \ldots, r$, not all of them 0, such that $0 \geq f_1 + \cdots + f_r$. Let us define $A := \cup_{i \neq j}(S_i \cap S_j)$. For every i, let $f_i^* \in \mathcal{L}_{S_i}$ be given by

$$f_i^*(x) := \sup\{f_i(y) : \pi_{S_i \cap A}(y) = \pi_{S_i \cap A}(x)\}. \tag{B.3}$$

Then $f_i^* \in \mathcal{L}_{S_i \cap A} = \mathcal{L}_{S_i \cap (\cup_{j \neq i} S_j)}$, and moreover $f_i^* \geq f_i$. Thus, f_i^* belongs to $\mathcal{D}_i \cup \{0\}$ and therefore to $\mathcal{D}_i^* \cup \{0\}$. Now, for every $z \in \mathcal{X}$,

$$(f_1 + \cdots + f_r)(z) = f_1(\pi_{S_1 \cap A}(z), \pi_{S_1 \setminus A}(z)) + \cdots + f_r(\pi_{S_r \cap A}(z), \pi_{S_r \setminus A}(z)) \leq 0. \tag{B.4}$$

Consider $z' \in \mathcal{X}$, and let $\varepsilon > 0$. Then for every $i = 1, \ldots, r$ there exists $z_i \in \mathcal{X}_{S_i}$ such that

$$f_i^*(\pi_{S_i \cap A}(z)) = f_i^*(\pi_{S_i \cap A}(z_i)) \leq f_i(z_i) + \varepsilon.$$

Since the sets $S_1 \setminus A, \ldots, S_r \setminus A$ are pairwise disjoint, we can build $z \in \mathcal{X}$ such that $\pi_{S_i \cap A}(z) = \pi_{S_i \cap A}(z')$ and $\pi_{S_i \setminus A}(z) = \pi_{S_i \setminus A}(z_i)$ for every i. This, together with Eqs. (B.3) and (B.4), implies that

$$(f_1^* + \ldots, f_r^*)(z') \leq f_1(z_1) + \ldots f_r(z_r) + r\varepsilon = (f_1 + \cdots + f_r)(z) + r\varepsilon \leq r\varepsilon.$$

Since this holds for any $\varepsilon > 0$, we deduce that $f_1^* + \cdots + f_r^* \leq 0$. This means that $\cup_{i=1}^r \mathcal{D}_i^*$ incurs partial loss, a contradiction. ☐

Proof of Proposition 8.

1. The direct implication has been established in [67, Theorem 7(1)], and the converse is a consequence of [67, Theorem 8].
2. The direct and converse implications have been established in [67, Proposition 3(2)] and [67, Theorem 8(2)], respectively. ☐

Lemma 15. *Under the notation of Section 4,*

(a) $\operatorname{posi}(\mathcal{D}_0 \cup \bigcup_{i \in A_j, j \geq 1} \mathcal{D}_i' \cup \mathcal{L}^+) = \operatorname{posi}(\mathcal{D}_0 \cup \bigcup_{i \in A_j, j \geq 1} \mathcal{D}_i \cup \mathcal{L}^+)$.

(b) $\operatorname{posi}(\mathcal{D}_0 \cup \bigcup_{i \in A_j, j \geq 1} \mathcal{D}_i' \cup \mathcal{L}^+) \cap \mathcal{L}_{S_0'} = \operatorname{posi}(\mathcal{D}_0 \cup \bigcup_{i \in A_1} \mathcal{D}_i' \cup \mathcal{L}^+) \cap \mathcal{L}_{S_0'} = \operatorname{posi}(\mathcal{D}_0 \cup \bigcup_{i \in A_1} (\mathcal{D}_i' \cap \mathcal{L}_{S_i' \cap S_0'}) \cup \mathcal{L}^+) \cap \mathcal{L}_{S_0'}$.

Proof. (a) By construction, $\mathcal{D}_i \subseteq \mathcal{D}_i'$ for every i. For the converse inclusion, note that, given $i \in A_j$ for $j \neq 0$, any gamble in $\mathcal{D}_i'$ can be expressed as a sum of gambles from $\cup_{l \in A_k, k \geq j} \mathcal{D}_l$.

(b) From the monotonicity of the posi operator with respect to set inclusion, we deduce that

$$\operatorname{posi}(\mathcal{D}_0 \cup \bigcup_{i \in A_j, j \geq 1} \mathcal{D}_i' \cup \mathcal{L}^+) \cap \mathcal{L}_{S_0'} \supseteq \operatorname{posi}(\mathcal{D}_0 \cup \bigcup_{i \in A_1} \mathcal{D}_i' \cup \mathcal{L}^+) \cap \mathcal{L}_{S_0'} \supseteq \operatorname{posi}(\mathcal{D}_0 \cup \bigcup_{i \in A_1} (\mathcal{D}_i' \cap \mathcal{L}_{S_i' \cap S_0'}) \cup \mathcal{L}^+) \cap \mathcal{L}_{S_0'}.$$

Let us prove that the first inclusion is indeed an equality. For this, we shall prove that

$$\operatorname{posi}\left(\mathcal{D}_0 \cup \bigcup_{i \in A_j, j=1,\ldots,k} \mathcal{D}_i' \cup \mathcal{L}^+\right) \cap \mathcal{L}_{S_0'} = \operatorname{posi}\left(\mathcal{D}_0 \cup \bigcup_{i \in A_j, j=1,\ldots,k-1} \mathcal{D}_i' \cup \mathcal{L}^+\right) \cap \mathcal{L}_{S_0'} \tag{B.5}$$

for any $k > 1$. Consider a gamble f on the left-hand side. Then there are $f_0 \in \mathcal{D}_0, f_i \in \mathcal{D}_i' \cup \{0\}$ for every $i \in \cup_{j=1}^k A_j$ such that $f \geq f_0 + \sum_i f_i$. If $f_l = 0$ for every $l \in A_k$, then trivially f belongs to the right-hand side. Assume next that $f_l \neq 0$ for some $l \in A_k$. Then, there exists some adjacent node $l' \in A_{k-1}$ in the path that connects l with the root node. From the RIP condition, it holds that

$$S_l' \cap S_0' \subseteq S_{l'}' \cap S_0' \text{ and for any other variable } j \in S_l' \setminus S_{l'}',\ j \notin \cup\{S_{l''}' : l'' \neq l, f_{l''} \neq 0\};$$

indeed, if $j \in S'_l \cap S'_{l''}$, then RIP implies that j is in all the cliques in the path that connects l and l''. But since $j \notin S'_{l'}$, then l' does not belong to this path. This implies that the unique path that communicates l'' with the root node is the union of the path that joins l'' with l and the path that joins l with the root note. This would mean that l'' is at a greater distance than l from the root node, i.e., at a distance greater than k, a contradiction.

Since f_i is S'_i-measurable for every i, we have that, for every x,

$$\begin{aligned} f(x) = f(\pi_{S'_0}(x)) \geq [f_0 + \sum_i f_i](x) &= f_0(\pi_{S'_0}(x)) + \sum_i f_i(\pi_{S'_i}(x)) \\ &= f_0(\pi_{S'_0}(x)) + \sum_{i \neq l} f_i(\pi_{S'_i}(x)) + f_l(\pi_{S'_l}(x)) \\ &= f_0(\pi_{S'_0}(x)) + \sum_{i \neq l} f_i(\pi_{S'_i}(x)) + f_l(\pi_{S_l \cap S'_{l'}}(x), \pi_{S'_l \setminus S'_{l'}}(x)). \end{aligned}$$

As a consequence,

$$\begin{aligned} f(x) &= \sup\{f(y) : \pi_{(S'_l \setminus S'_{l'})^c}(x) = \pi_{(S'_l \setminus S'_{l'})^c}(y)\} \geq \sup\{[f_0 + \sum_i f_i](y) : \pi_{(S'_l \setminus S'_{l'})^c}(x) = \pi_{(S'_l \setminus S'_{l'})^c}(y)\} \\ &= f_0(\pi_{S'_0}(x)) + \sum_{i \neq l} f_i(\pi_{S'_i}(x)) + \sup\{f_l(y) : \pi_{(S'_l \setminus S'_{l'})^c(x)} = \pi_{(S'_l \setminus S'_{l'})^c}(y)\} \\ &= f_0(\pi_{S'_0}(x) + \sum_{i \neq l} f_i(\pi_{S'_i}(x)) + f_{l'}(\pi_{S'_l \cap S'_{l'}}(x)), \end{aligned}$$

where $f_{l'}$ is the $S'_{l'}$-measurable gamble given by

$$f_{l'}(x) \coloneqq \sup\{f_l(y) : \pi_{S'_l \cap S'_{l'}}(y) = \pi_{S'_l \cap S'_{l'}}(x)\}. \tag{B.6}$$

On the other hand, Eq. (B.6) implies that $f_{l'} \geq f_l$ and it is $S'_l \cap S'_{l'}$-measurable. Thus, $f_{l'} \in \mathcal{D}'_l \cap \mathcal{L}_{S'_l \cap S'_{l'}}$, and therefore also to $\mathcal{D}'_{l'}$ by construction (line 7 in Algorithm 1).

By repeating the process with all the cliques in A_k, we end up with a number of gambles $f'_0 \in \mathcal{D}_0 \cup \{0\}, f'_i \in \mathcal{D}_i \cup \{0\}$, for $i \in \cup_{j=1}^{k-1} A_j$, such that $f \geq f'_0 + \sum_i f'_i$. Therefore, $f \in \text{posi}(\mathcal{D}_0 \cup \bigcup_{i \in A_j, j=1,\dots,k-1} \mathcal{D}'_i \cup \mathcal{L}^+) \cap \mathcal{L}'_{S_0}$ and as a consequence Eq. (B.5) holds. Since we can do this for every $k > 1$, we conclude that

$$\text{posi}\left(\mathcal{D}_0 \cup \bigcup_{i \in A_j, j \geq 1} \mathcal{D}'_i \cup \mathcal{L}^+\right) \cap \mathcal{L}_{S'_0} = \text{posi}\left(\mathcal{D}_0 \cup \bigcup_{i \in A_1} \mathcal{D}'_i \cup \mathcal{L}^+\right) \cap \mathcal{L}_{S'_0}.$$

Let us establish now that the second inclusion is also an equality. Consider a gamble $f \in \text{posi}(\mathcal{D}_0 \cup \bigcup_{i \in A_1} \mathcal{D}'_i \cup \mathcal{L}^+) \cap \mathcal{L}_{S'_0}$. Then there are $f_i \in \mathcal{D}'_i \cup \{0\}$ for $i \in A_1$ and $f_0 \in \mathcal{D}_0$ such that $f \geq f_0 + \sum_{i \in A_1} f_i$. Then, by construction of the tree, the variables that belong to $(\cup_{i \in A_1} S'_i) \setminus S'_0$ appear exactly in one S'_i. As a consequence, if we define $f'_i \in (\mathcal{D}'_i \cup \{0\}) \cap \mathcal{L}_{S'_i \cap S'_0}$ by

$$f'_i(x) \coloneqq \sup\{f_i(y) : \pi_{S'_i \cap S'_0}(y) = \pi_{S'_i \cap S'_0}(x)\}$$

it holds that $f'_i \in \mathcal{L}_{S'_i \cap S'_0}$, and moreover $f'_i \geq f_i$. Thus, $f'_i \in \mathcal{D}'_i \cup \{0\}$. Moreover, we obtain that $f \geq f_0 + \sum_{i \in A_1} f'_i$. Thus, $f \in \text{posi}(\bigcup_{i \in A_1} (\mathcal{D}'_i \cap \mathcal{L}_{S'_i \cap S'_0}) \cup \mathcal{L}^+) \cap \mathcal{L}_{S'_0}$. $\square$

Proof of Theorem 9. It suffices to take into account the following chain of equalities:

$$\begin{aligned}\mathcal{E}\cap\mathcal{L}_{S_0'} &= \mathrm{posi}(\cup_{i=1}^q \mathcal{D}_i \cup \mathcal{L}^+)\cap\mathcal{L}_{S_0'}\\ &= \mathrm{posi}(\mathcal{D}_0 \cup \bigcup_{i\in A_j, j\geq 1} \mathcal{D}_i' \cup \mathcal{L}^+)\cap\mathcal{L}_{S_0'}\\ &= \mathrm{posi}(\mathcal{D}_0 \cup \bigcup_{i\in A_1} \mathcal{D}_i' \cup \mathcal{L}^+)\cap\mathcal{L}_{S_0'}\\ &= \mathrm{posi}(\mathcal{D}_0 \cup \bigcup_{i\in A_1} (\mathcal{D}_i' \cap \mathcal{L}_{S_i'\cap S_0'}) \cup \mathcal{L}^+)\cap\mathcal{L}_{S_0'}\\ &= \mathcal{D}_0'.\end{aligned}$$

Here, the first and last equalities follow by definition; the second, from point (a) in Lemma 15; and the third and fourth, from point (b) in Lemma 15. □

Proof of Theorem 10. First of all, by construction we have that

$$\mathcal{D}_j'' \subseteq \mathrm{posi}_{S_j'}(\cup_{i=1}^q \mathcal{D}_i' \cap \mathcal{L}_{S_j'}) \subseteq \mathrm{posi}(\cup_{i=1}^q \mathcal{D}_i \cup \mathcal{L}^+)\cap\mathcal{L}_{S_j'} = \mathcal{E}\cap\mathcal{L}_{S_j'}.$$

In order to establish the converse we partition the nodes of the graph into four groups:

E_1 : Nodes l in the unique path that connects the root we had before with j (including these two). As an example, consider Figure 3 with the chosen root $\{X_5, X_{11}\}$ and let j correspond to clique $\{X_1, X_2\}$. Then $E_1 = \{\{X_1, X_2\}, \{X_2, X_4\}, \{X_4, X_5\}, \{X_5, X_{11}\}\}$.

E_2 : Nodes l in E_1^c such that the path that connects them with j includes some clique from E_1 different from the root and j, and does not include the root. Using the example in the previous item, we get $E_2 = \{\{X_2, X_3\}\}$.

E_3 : Nodes l in E_1^c such that the path that connects them with j includes the root we had before. In the example: $E_3 = \{\{X_5, X_6\}, \{X_7, X_8, X_9, X_{10}, X_{11}\}, \{X_7, X_8, X_{12}\}, \{X_8, X_{12}, X_{13}\}, \{X_8, X_{13}, X_{15}\}, \{X_{13}, X_{14}, X_{15}\}\}$.

E_4 : Nodes l in E_1^c such that the path that connects them with j does not include any node from E_1, except j. In the case of the example, $E_4 = \emptyset$.

Denote $A_i := \cup_{l\in E_i} S_l'$, for $i = 1,\ldots,4$. If follows from RIP that the sets $A_2 \setminus A_1$, $A_3 \setminus A_1$ and $A_4 \setminus A_1$ are pairwise disjoint. To see this, note that the nodes in E_2 and E_3 are connected via E_1 (and similarly for E_2 and E_4 and for E_3 and E_4). As a consequence, if a variable j belongs to a node in E_2 and to a node in E_3 it should also belong to all the nodes in the path that connects them, and in particular to some node in E_1.

Take $f \in \mathcal{E}\cap\mathcal{L}_{S_j'}$. By Eq. (2), this means that there are $f_i \in \mathcal{D}_i$ for $i = 1,\ldots,q$ such that $f \geq \sum_{i=1}^q f_i$. Since $\{E_1,\ldots,E_4\}$ forms a partition of $\{1,\ldots,q\}$, we can also write

$$f \geq \left(\sum_{i\in E_1} f_i\right) + \left(\sum_{i\in E_2} f_i\right) + \left(\sum_{i\in E_3} f_i\right) + \left(\sum_{i\in E_4} f_i\right).$$

Thus, if we define $g_j := \sum_{i\in E_j} f_i$ for $j = 1,\ldots,4$, we deduce that $g_i \in \mathrm{posi}(\cup_{l\in E_i}\mathcal{D}_l)$ and that $f \geq g_1 + g_2 + g_3 + g_4$. Define

$$g_i'(x) := \sup\{g_i(y) : \pi_{A_1\cap A_i}(y) = \pi_{A_1\cap A_i}(x)\} \text{ for } i = 2, 3, 4.$$

Then $g_i' \geq g_i$, whence $g_i' \in \mathrm{posi}(\cup_{l\in E_i}\mathcal{D}_l \cup \mathcal{L}^+)$ for $i = 2, 3, 4$. Moreover, since $A_2 \setminus A_1$, $A_3 \setminus A_1$ and $A_4 \setminus A_1$ are pairwise disjoint and $f \in \mathcal{L}_{S_j'} \subseteq \mathcal{L}_{A_1}$, it follows that for all x,

$$f(x) = f(\pi_{A_1}(x)) \geq g_1(\pi_{A_1}(x)) + g_2(\pi_{A_1}(x), \pi_{A_2\setminus A_1}(x)) + g_3(\pi_{A_1}(x), \pi_{A_3\setminus A_1}(x)) + g_4(\pi_{A_1}(x), \pi_{A_4\setminus A_1}(x)),$$

whence

$$\begin{aligned} f(x) &\geq g_1(\pi_{A_1}(x)) + \sup_{\pi_{A_1\cap A_2}(y)=\pi_{A_1\cap A_2}(x)} g_2(y) + \sup_{\pi_{A_1\cap A_3}(y)=\pi_{A_1\cap A_3}(x)} g_3(y) + \sup_{\pi_{A_1\cap A_4}(y)=\pi_{A_1\cap A_4}(x)} g_4(y) \\ &= g_1(x) + g_2'(x) + g_3'(x) + g_4'(x). \end{aligned}$$

Now, by construction:

- $g_2' \in \mathcal{L}_{A_2\cap A_1} \subseteq \mathcal{L}_{A_1}$; by Algorithm 1, we deduce that $g_2' \in \text{posi}(\cup_{l\in E_1}\mathcal{D}_l' \cup \mathcal{L}^+)$.
- $g_3' \in \mathcal{L}_{A_3\cap A_1} = \mathcal{L}_{S_0'}$; by Algorithm 1, we deduce that $g_3' \in \mathcal{D}_0'$.
- $g_4' \in \mathcal{L}_{A_4\cap A_1} = \mathcal{L}_{S_j'}$; by Algorithm 1, we deduce that $g_4' \in \mathcal{D}_j'$,

while $g_1 \in \text{posi}(\cup_{l\in E_1}\mathcal{D}_l) \subseteq \text{posi}(\cup_{l\in E_1}\mathcal{D}_l')$.

As a consequence, $f \in \text{posi}(\cup_{l\in E_1}\mathcal{D}_l' \cup \mathcal{L}^+) \cap \mathcal{L}_{S_j'}$. Let us denote the indices in E_1 as $l_0, l_1, \dots, l_k$, where $l_0 \coloneqq j$ and l_i is the unique node in E_1 at a distance i from j and l_k is the root. Then we have $f \geq \sum_{i=0}^k h_i$, where $h_i \in \mathcal{D}_{l_i}' \cup \{0\}$ for every i. Let us prove that

$$f \in \text{posi}(\cup_{i=0}^{k-2}\mathcal{D}_{l_i}' \cup \mathcal{D}_{l_{k-1}}'' \cup \mathcal{L}^+).$$

If $h_k = 0$, this holds simply taking into account that $\mathcal{D}_{l_{k-1}}' \subseteq \mathcal{D}_{l_{k-1}}''$. Assume next that $h_k \neq 0$, and let us define

$$h_k'(x) \coloneqq \sup\{h_k(y) : \pi_{S_{l_{k-1}}'\cap S_{l_k}'}(y) = \pi_{S_{l_{k-1}}'\cap S_{l_k}'}(x)\}.$$

Note that the nodes l_k (the root) and l_{k-1} are adjacent, since the root l_k is at distance k from j and node l_{k-1} is at distance $k-1$. As a consequence, $S_{l_k}' \cap S_{k-1}' \neq \emptyset$.

By definition, $h_k' \geq h_k$, whence $h_k' \in \mathcal{D}_{l_k}'$. Since it also belongs to $\mathcal{L}_{S_{l_{k-1}}'\cap S_{l_k}'}$, we deduce that $h_k' \in \mathcal{D}_{l_{k-1}}''$, by line 7 in Algorithm 2. Now,

$$(\forall x)\ f(x) = f(\pi_{S_{l_0}'}(x)) \geq \sum_{i=0}^{k-1} h_i(\pi_{S_{l_i}'}(x)) + h_k(\pi_{S_{l_k}'}(x)) = \sum_{i=0}^{k-1} h_i(\pi_{S_{l_i}'}(x)) + h_k(\pi_{S_{l_k}'\cap S_{l_{k-1}}'}(x), \pi_{S_{l_k}'\setminus S_{l_{k-1}}'}(x)).$$

If we denote $B \coloneqq (S_{l_k}' \setminus S_{l_{k-1}}')^c = S_{l_k}'^c \cup S_{l_{k-1}}'$, then $S_{l_k}' \cap B = S_{l_k}' \cap S_{l_{k-1}}'$; moreover, $B^c = S_{l_k}' \setminus S_{l_{k-1}}' \subseteq S_{l_i}'^c$ for every $i = 0, \dots, k-1$ (by RIP), or, equivalently, $S_{l_i}' \subseteq B$ for every $i = 0, \dots, k-1$. As a consequence,

$$f(x) \geq \sup_{\pi_B(y)=\pi_B(x)} \sum_{i=0}^{k-1} h_i(\pi_{S_{l_i}'}(y)) + h_k(\pi_{S_{l_k}'\cap S_{l_{k-1}}'}(y), \pi_{S_{l_k}'\setminus S_{l_{k-1}}'}(y)) = \sum_{i=0}^{k-1} h_i(x) + h_k'(x).$$

Thus, $f \in \text{posi}(\cup_{i=0}^{k-1}\mathcal{D}_{l_i}' \cup \mathcal{D}_{l_{k-1}}'' \cup \mathcal{L}^+) = \text{posi}(\cup_{i=0}^{k-2}\mathcal{D}_{l_i}' \cup \mathcal{D}_{l_{k-1}}'' \cup \mathcal{L}^+)$, taking into account that $\mathcal{D}_{l_{k-1}}' \subseteq \mathcal{D}_{l_{k-1}}''$.

With a similar procedure, we can deduce that $f \in \text{posi}(\cup_{i=0}^{k-3}\mathcal{D}_{l_i}' \cup \mathcal{D}_{l_{k-2}}'' \cup \mathcal{L}^+)$, and eventually that $f \in \text{posi}(\mathcal{D}_{l_0}' \cup \mathcal{D}_{l_1}'' \cap \mathcal{L}^+)$. If we now use that $f \in \mathcal{L}_{S_{l_0}'}$, then we are also able to deduce that $f \in \text{posi}(\mathcal{D}_{l_0}' \cup (\mathcal{D}_{l_1}'' \cap \mathcal{L}_{S_{l_0}'}) \cap \mathcal{L}^+) = \mathcal{D}_{l_0}''$, by line 7 in Algorithm 2.

This proves the inclusion $\mathcal{E} \cap \mathcal{L}_{S_{l_0}'} \subseteq \mathcal{D}_{l_0}''$, or, equivalently, $\mathcal{E} \cap \mathcal{L}_{S_j'} \subseteq \mathcal{D}_j''$. As a consequence, we have the equality. □

Lemma 16. *Consider variables $X_1, \dots, X_n$ and separately coherent conditional sets of desirable gambles $\mathcal{D}_{O_j}|X_{I_j}$, $j = 1, \dots, r$, such that $I_1 = \emptyset$ and $O_j \cap (\cup_{k<j} O_k \cup I_k) = \emptyset$ for $j = 2, \dots, r$. Then $\mathcal{D}_{O_j}|X_{I_j}$, $j = 1, \dots, r$ avoid partial loss.*

Proof. Consider gambles $f_i \cup \{0\} \in \mathcal{D}_i$ for $i = 1, \ldots, r$, not all of them equal to zero, and let us prove that there exists some $x \in \mathcal{X}$ such that $(\sum_{i=1}^{r} f_i(x)) > 0$.

Assume for the moment that $f_i \neq 0$ for every $i = 1, \ldots, r$. Then since $\mathcal{D}_{O_1}$ is an unconditional set of desirable gambles, being $I_1 = \emptyset$, there exists some $x_1 \in \mathcal{X}_{O_1}$ such that $f_1(x_1) > 0$. Consider now $y_1 \in \mathcal{X}_{O_1 \cup I_2}$ such that $\pi_{O_1}(y_1) = x_1$. Since $\mathcal{D}_{O_2}|X_{I_2}$ is separately coherent, there must be some $x_2 \in \mathcal{X}_{O_2 \cup I_2 \cup O_1}$ such that $\pi_{O_1 \cup I_2}(x_2) = y_1$ and $f_2(\pi_{O_2}(x_2), \pi_{I_2}(x_2)) > 0$.

Next we consider $y_2 \in \mathcal{X}_{I_3 \cup (\cup_{i=1}^{2} O_i \cup I_i)}$ such that $\pi_{\cup_{i=1}^{2} O_i \cup I_i}(y_2) = x_2$. Since $\mathcal{D}_{O_3}|X_{I_3}$ is separately coherent, there must be some $x_3 \in \mathcal{X}_{\cup_{i=1}^{3} O_i \cup I_i}$ such that $\pi_{\cup_{i=1}^{2} O_i \cup I_i}(x_3) = y_2$ and $f_2(\pi_{O_3}(x_3), \pi_{I_3}(x_3)) > 0$.

If we proceed in this manner, we obtain $x_1, \ldots, x_r$ such that $\pi_{\cup_{k<j} O_k \cup I_k}(x_j) = x_{j-1}$ for $j = 2, \ldots, r$, and such that $f_j(\pi_{O_j \cup I_j}(x_j)) > 0$. As a consequence,

$$\sum_{j=1}^{r} f_j(x_r) = \sum_{j=1}^{r} f_j(\pi_{O_j \cup I_j}(x_r)) = \sum_{j=1}^{r} f_j(\pi_{O_j \cup I_j}(x_j)) > 0.$$

Finally, when there is some $i \in \{1, \ldots, r\}$ such that $f_i = 0$, we consider an arbitrary $x_i \in \mathcal{X}_{\cup_{j=1}^{i} O_j \cup I_j}$ satisfying $\pi_{\cup_{j=1}^{i-1} O_j \cup I_j}(x_i) = x_{i-1}$, and proceed as in the proof above. □

Proof of Proposition 11. Let $\mathcal{B}' := \{B \in \mathcal{B} : |B| > 1\}$ be the set of indices of the templates that belong to some superblock and denote $C := \cup_{B \notin \mathcal{B}'} B$ the remaining indices. For each $B \in \mathcal{B}'$, denote $C_B := \cup_{j \in B}(O_j \cup I_j)$ the indexes of variables in the templates associated with superblock B. Then it follows from the definitions of the superblocks that the sets $\{C_B : B \in \mathcal{B}'\}$ are pairwise disjoint. As a consequence, if $\cup_{j \in B} \mathcal{D}_{O_j}|X_{I_j}$ avoids partial loss for every $B \in \mathcal{B}'$, we trivially obtain that $\cup_{B \in \mathcal{B}'} \cup_{j \in B} \mathcal{D}_{O_j}|X_{I_j}$ avoids partial loss. If we denote $A := \cup_{j \in C^c}(O_j \cup I_j)$, this means that the set

$$\mathcal{D}^* := \text{posi}_A(\cup_{B \in \mathcal{B}'} \cup_{j \in B} \mathcal{D}_{O_j}|X_{I_j} \cup \mathcal{L}^+)$$

is a coherent set of gambles on $\mathcal{L}_A$.

By [64, p. 115, lines 6–11], for each $j \in C$ it holds that $O_j \cap A = \emptyset$. Moreover, [97, Lemma 1] implies the existence of an order $\{j_1, \ldots, j_l\}$ of C so that $O_{j_m} \cap_{(m' < m)} (O_{m'} \cup I_{m'}) = \emptyset$.

This means that the sets $\mathcal{D}^*, \mathcal{D}_{O_{j_1}}|X_{I_{j_1}}, \ldots, \mathcal{D}_{O_{j_l}}|X_{I_{j_l}}$ satisfy the hypotheses of Lemma 16, and as a consequence they avoid partial loss. Since $\mathcal{D}^*$ is a superset of $\cup_{j \in C^c} \mathcal{D}_{O_j}|X_{I_j}$, we deduce that $\cup_{j=1}^{r} \mathcal{D}_{O_j}|X_{I_j}$ avoids partial loss. □

Lemma 17. *Let $\mathcal{D}_1$ be a set of desirable gambles that is coherent with respect to $\mathcal{L}_{O_1}$, and $\mathcal{D}_{O_2}|X_{I_2}$ be a separately coherent conditional set of desirable gambles, where $O_2 \cap (O_1 \cup I_2) = \emptyset$. Then $\mathcal{D}_{O_1}, \mathcal{D}_{O_2}|X_{I_2}$ are compatible.*

Proof. We may assume without loss of generality that $O_1 \cup O_2 \cup I_2 = \{1, \ldots, n\}$.

Consider first of all the case where $O_1 = I_2$. Then it follows from [69, Proposition 29] that the set

$$\mathcal{D} := \{f_1 + f_2 : f_1 \in \mathcal{D}_{O_1} \cup \{0\}, (\forall x_1 \in \mathcal{X}_{O_1})\ f_2(x_1, \cdot) \in \mathcal{D}_{O_2}|x_1 \cup \{0\}\} \setminus \{0\}$$

is a coherent superset of $\mathcal{D}_{O_1}, \mathcal{D}_{O_2}|X_{I_2}$. Let us prove that it induces $\mathcal{D}_{O_1}, \mathcal{D}_{O_2}|X_{I_1}$ by means of marginalization and conditioning:

- Consider $f \in \mathcal{D} \cap \mathcal{L}_{O_1}$. Then there are $f_1 \in \mathcal{D}_{O_1} \cup \{0\}, f_2 \in \mathcal{D}_{O_2}|X_{O_1} \cup \{0\}$ such that $f \geq f_1 + f_2$. If $f_2 = 0$, the result is trivial. Assume then that $f_2 \neq 0$. Define f_2' on $\mathcal{L}_{O_2 \cup O_1}$ by

  $$f_2'(x) := \sup\{f_2(y) : \pi_{O_1}(y) = \pi_{O_1}(x)\}.$$

  Then for every $x_1 \in \mathcal{X}_{O_1}$ such that $f_2(x_1, \cdot) \neq 0$, it follows from the coherence of $\mathcal{D}_{O_2}|x_1$ that $0 \leq \sup_{x_2 \in \mathcal{X}_{O_2}} f_2(x_1, x_2) = f_2'(x_1)$. Thus, $f_2' \in \mathcal{L}_{O_1}^+ \subseteq \mathcal{D}_{O_1}$. Moreover,

  $$(\forall x)\ f(x) = f(\pi_{O_1}(x)) \geq (f_1 + f_2)(x) = f_1(\pi_{O_1}(x)) + f_2(\pi_{O_1}(x), \pi_{O_2}(x)) \Rightarrow f \geq f_1 + f_2'.$$

  As a consequence, $f \in \mathcal{D}_{O_1}$.

- Fix next $x_1 \in \mathcal{X}_{O_1}$, and take $f \in \mathcal{D}$ such that $f = \mathbb{I}_{x_1} f$. Then there must be $f_1 \in \mathcal{D}_{O_1} \cup \{0\}, f_2 \in \mathcal{D}_{O_2}|X_{O_1}$ such that $f \geq f_1 + f_2$. Assume that $f_1 \neq 0$; otherwise $f \geq \mathbb{I}_{x_1} f_2 \in \mathcal{D}_{O_2}|X_{O_1}$. For any $x_1' \neq x_1$, it holds that $\mathbb{I}_{x_1'} f = 0 \geq f_1(x_1') + f_2(x_1', \cdot)$, and, taking into account that $f_2(x_1', \cdot) \in \mathcal{D}_{O_2}|x_1' \cup \{0\}$, this means that $f_1(x_1') \leq 0$ for every $x_1' \neq x_1$. Since $\mathcal{D}_{O_1}$ is coherent, this implies that $f_1(x_1) > 0$. But then $f(x_1, \cdot) \geq f_2(x_1, \cdot)$ and as a consequence $f \in \mathcal{D}_{O_2}|X_{O_1}$.

We consider next the general case. Let us define the conditional set of desirable gambles $\mathcal{D}'_{O_2}|X_{O_1 \cup I_2}$ by

$$\mathcal{D}'_{O_2}|X_{O_1 \cup I_2} := \cup_{x \in \mathcal{X}_{O_1 \cup I_2}} \mathcal{D}'_{O_2}|x, \text{ with } \mathcal{D}'_{O_2}|x := \mathcal{D}_{O_2}|\pi_{I_2}(x).$$

Then by Definition 16 $\mathcal{D}'_{O_2}|X_{O_1 \cup I_2}$ is a separately coherent conditional set of desirable gambles. Consider also

$$\mathcal{D}'_{O_1 \cup I_2} := \operatorname{posi}_{O_1 \cup I_2}(\mathcal{D}_{O_1} \cup \mathcal{L}^+),$$

the natural extension of $\mathcal{D}_{O_1}$ to $\mathcal{L}_{O_1 \cup I_2}$. This is a coherent set of desirable gambles that satisfies $\mathcal{D}'_{O_1 \cup I_2} \cap \mathcal{L}_{O_1} = \mathcal{D}_{O_1}$.

Applying the first part of the proof, $\mathcal{D}'_{O_1 \cup I_2}, \mathcal{D}'_{O_2}|X_{O_1 \cup I_2}$ are compatible with some coherent set of desirable gambles $\mathcal{D}$. It follows that

$$\mathcal{D} \cap \mathcal{L}_{O_1} = \mathcal{D} \cap \mathcal{L}_{O_1 \cup I_2} \cap \mathcal{L}_{O_1} = \mathcal{D}'_{O_1 \cup I_2} \cap \mathcal{L}_{O_1} = \mathcal{D}_{O_1}.$$

Let us prove that $\mathcal{D}$ also induces $\mathcal{D}_{O_2}|X_{I_2}$. Consider $x \in \mathcal{X}_{I_2}$, and take $f \in \mathcal{D} \cap \mathcal{L}_{O_2 \cup I_2}$ satisfying $f = \mathbb{I}_x f$. Then $f \geq g + h$ for $g \in \mathcal{D}'_{O_1 \cup I_2}, h \in \mathcal{D}'_{O_2}|X_{O_1 \cup I_2}$.

For every $y \in \mathcal{X}_{O_1 \cup I_2}$ with $\pi_{I_2}(y) \neq x$, we have that

$$f(y, \cdot) = 0 \geq g(y) + h(y, \cdot),$$

and since $\sup h(y, \cdot) \geq 0$ because $h(y, \cdot) \in \mathcal{D}'_{O_2}|y \cup \{0\}$, it must be $g(y) \leq 0$. Therefore, $g(y) \leq 0$ for every y such that $\pi_{I_2}(y) \neq x$, and since we are assuming that $g \in \mathcal{D}'_{O_1 \cup I_2} \cup \{0\}$, there must be some y with $\pi_{I_2}(y) = x$ and $g(y) \geq 0$.

We obtain that $f(y, \cdot) \geq h(y, \cdot) \in \mathcal{D}'_{O_2}|y = \mathcal{D}_{O_2}|x$. Since we are assuming that $f \in \mathcal{L}_{O_2 \cup I_2}$, then it must be $f(y, \cdot) = f(x, \cdot)$, and then $f(x, \cdot) \geq h(x, \cdot) \in \mathcal{D}_{O_2}|x$. □

Proof of Theorem 12. Let $\mathcal{B}' := \{B \in \mathcal{B} : |B| > 1\}$ be the indices of the superblocks determined by some source of contradiction. For any $B \in \mathcal{B}'$, let $C_B := \cup_{j \in B}(O_j \cup I_j)$. Then it follows from the definitions of the superblocks that the sets $\{C_B : B \in \mathcal{B}'\}$ are pairwise disjoint. As a consequence, if $\cup_{j \in B} \mathcal{D}_{O_j}|X_{I_j}$ are compatible for every $B \in \mathcal{B}'$, we trivially obtain that $\cup_{B \in \mathcal{B}'} \cup_{j \in B} \mathcal{D}_{O_j}|X_{I_j}$ are compatible: the coherent sets of desirable gambles $\{\mathcal{D}_B\}_{B \in \mathcal{B}'}$ that induce them involve disjoint sets of variables, and as a consequence of Theorem 2, their natural extension, given by

$$\mathcal{D}_0 := \{f \neq 0 : f \geq \sum_{B \in \mathcal{B}} f_B \text{ for some} f_B \in \mathcal{D}_B \cup \{0\}, B \in \mathcal{B}\},$$

has marginals $\{\mathcal{D}_B\}_{B \in \mathcal{B}}$ and so induces $\mathcal{D}_{O_j}|X_{I_j}$ for $j \in B \in \mathcal{B}'$.

Let $C := \{1, \ldots, r\} \setminus (\cup_{B \in \mathcal{B}'} B)$ be the remaining indices. Then if we denote $A := \cup_{j \in C^c}(O_j \cup I_j)$, it holds that $\mathcal{D}$ is a coherent set of desirable gambles on $\mathcal{L}_A$.

By [64, p. 115, lines 6–11], for each $j \in C$ it holds that $O_j \cap A = \emptyset$. Moreover, [97, Lemma 1] implies the existence of an order $\{j_1, \ldots, j_l\}$ of C so that $O_{j_m} \cap_{(m' < m)} (O_{j'_m} \cup I_{j'_m}) = \emptyset$.

We now show that the algorithm produces the natural extension in an iterative manner:

- By Lemma 17, the sets $\mathcal{D}$ and $\mathcal{D}_{O_{j_1}}|X_{I_{j_1}}$ are compatible. Let $\mathcal{D}_1$ denote their natural extension. Since $\mathcal{D}_1 \cap \mathcal{L}_A = \mathcal{D}$, it follows that $\mathcal{D}_1$ also induces the sets $\mathcal{D}_{O_j}|X_{I_j}$ for every $j \in B \in \mathcal{B}'$ by means of Eq. (7).

- $\mathcal{D}_1$ is a coherent set of desirable gambles with respect to $S_1 := A \cup (O_{j_1} \cup I_{j_1})$, while $\mathcal{D}_{O_{j_2}}|X_{I_{j_2}}$ is a separately coherent conditional set of desirable gambles such that $O_{j_2} \cap (S_1 \cup I_{j_2}) = \emptyset$. If we now apply Lemma 17 again, we conclude that $\mathcal{D}_1, \mathcal{D}_{O_{j_2}}|X_{I_{j_2}}$ are also compatible, whence their natural extension $\mathcal{D}_2$ also has marginal $\mathcal{D}_1$ (whence it induces $\mathcal{D}_{O_j}|X_{I_j}$ for every $j \in B \in \mathcal{B}'$ and $\mathcal{D}_{O_{j_1}}|X_{I_{j_1}}$) and it also induces the conditional set of desirable gambles $\mathcal{D}_{O_{j_2}}|X_{I_{j_2}}$.

- ∘ By iterating the procedure, we obtain, for any $k = 1, \dots, m$, a coherent set of desirable gambles $\mathcal{D}_j$ that is compatible with $\cup_{j \in B \in \mathcal{B}'} \mathcal{D}_{O_j}|X_{I_j}$ and with $\mathcal{D}_{O_{j_1}}|X_{I_{j_1}}, \dots, \mathcal{D}_{O_{j_k}}|X_{I_{j_k}}$. By considering the case $k = m$ we conclude that $\mathcal{D}_{O_1}|X_{I_1}, \dots, \mathcal{D}_{O_m}|X_{I_m}$ are compatible.

Moreover, note that on each step since $\mathcal{D}_i$ is a compatible superset of $\mathcal{D}_0 \cup \cup_{i'<i} \mathcal{D}_{O_{j_{i'}}}|X_{I_{j_{i'}}}$, it must include the natural extension of $\cup_{j \in B} \mathcal{D}_{O_j}|X_{I_j} \cup \cup_{i'<i} \mathcal{D}_{O_{j_{i'}}}|X_{I_{j_{i'}}}$; but conversely we have that this natural extension must always include the natural extension of $\mathcal{D}_{i-1} \cup \mathcal{D}_{O_{j_i}}|X_{I_{j_i}}$. Thus, the two sets are equal. □

Proof of Proposition 13. Assume first of all that $\mathcal{D}_1, \dots, \mathcal{D}_r$ avoid partial loss, and let us prove that $\mathcal{D}'_j$ is the restriction of the natural extension of $\mathcal{D}_1 \cup \dots \cup \mathcal{D}_j$ to $\mathcal{L}_{S_j}$. This will imply in particular that $\mathcal{D}'_r$ is coherent, since it will be the restriction of a coherent set. We apply induction on j.

We begin by establishing the result for $j = 2$, since it is trivial for $j = 1$. By definition, $\mathcal{D}'_2$ is the natural extension of $\mathcal{D}_2 \cup (\mathcal{D}_1 \cap \mathcal{L}_{S_1 \cap S_2})$. To prove that it coincides with the restriction to $\mathcal{L}_{S_2}$ of $\mathcal{E}_2 := \text{posi}(\mathcal{L}^+ \cup \mathcal{D}_1 \cup \mathcal{D}_2)$, take a gamble $f \in \mathcal{E}_2$. Then there are $g \in \mathcal{D}_1 \cup \{0\}, h \in \mathcal{D}_2 \cup \{0\}$ such that $f \geq g + h$. If $g = 0$, we are done. If $g \neq 0$, we define the S_2-measurable gamble g' by

$$g'(x) := \sup\{g(y) : \pi_{S_2}(y) = \pi_{S_2}(x)\}.$$

Since $f \in \mathcal{L}_{S_2}$, we deduce that $f \geq g' + h$, and since $g' \geq g$, we get that $g' \in \mathcal{D}_1 \cap \mathcal{L}_{S_2}$. Thus, $f \in \text{posi}((\mathcal{D}_1 \cap \mathcal{L}_{S_1 \cap S_2}) \cup \mathcal{D}_2 \cup \mathcal{L}^+) = \mathcal{D}'_2$. The inclusion $\mathcal{D}'_2 \subseteq \mathcal{E}_2$ is trivial.

Assume next that the result holds up to $j - 1$, so that $\mathcal{D}'_{j-1} = \text{posi}(\mathcal{L}^+ \cup \mathcal{D}_1 \cup \dots \cup \mathcal{D}_{j-1}) \cap \mathcal{L}_{S_{j-1}}$. We must prove that

$$\text{posi}(\mathcal{L}^+ \cup \mathcal{D}'_{j-1} \cup \mathcal{D}_j) \cap \mathcal{L}_{S_j} = \text{posi}(\mathcal{L}^+ \cup \mathcal{D}_1 \cup \dots \cup \mathcal{D}_j) \cap \mathcal{L}_{S_j}.$$

($\subseteq$) It suffices to take into account that $\mathcal{D}'_{j-1}$ is included in $\text{posi}(\mathcal{L}^+ \cup \mathcal{D}_1 \cup \dots \cup \mathcal{D}_j)$ by construction.

($\supseteq$) Consider a gamble $f \in \text{posi}(\mathcal{L}^+ \cup \mathcal{D}_1 \cup \dots \cup \mathcal{D}_j) \cap \mathcal{L}_{S_j}$. Then, there are $g_i \in \mathcal{D}_i \cup \{0\}$ for $i = 1, \dots, j$ such that $f \geq g_1 + \dots + g_j$.

Let us define the S_j-measurable gamble g' by

$$g'(x) := \sup\{(g_1 + \dots + g_{j-1})(y) : \pi_{S_j}(y) = \pi_{S_j}(x)\}$$

for every $x \in \mathcal{X}$. Then by construction $g' \geq g_1 + \dots + g_{j-1}$, whence $g' \in \text{posi}(\mathcal{L}^+ \cup \mathcal{D}_1 \cup \dots \cup \mathcal{D}_{j-1})$. On the other hand, $g' \in \mathcal{L}_{S_j}$, and since $g_i \in \mathcal{L}_{S_i}$ for $i = 1, \dots, j-1$, it follows that $g_1 + \dots + g_{j-1} \in \mathcal{L}_{\cup_{i=1}^{j-1} S_i}$, whence $g' \in \mathcal{L}_{(\cup_{i=1}^{j-1} S_i) \cap S_j} = \mathcal{L}_{S_{j-1} \cap S_j}$, where the equality follows by hypothesis. Thus, g' belongs to $\mathcal{L}_{S_{j-1}}$, and therefore $g' \in \mathcal{D}'_{j-1}$, by the induction hypothesis.

Since moreover $f \geq g' + g_j$, because $f \in \mathcal{L}_{S_j}$, we conclude that $f \in \text{posi}(\mathcal{L}^+ \cup \mathcal{D}'_{j-1} \cup \mathcal{D}_j) \cap \mathcal{L}_{S_j}$.

This concludes the first part of the proof. Assume now that $\mathcal{D}'_r$ is coherent, and let us prove that $\mathcal{D}_1, \dots, \mathcal{D}_r$ avoid partial loss. By construction, $\mathcal{D}'_j$ is the restriction to $\mathcal{L}_{S_j}$ of the natural extension of $\mathcal{D}'_{j-1} \cup \mathcal{D}_j$. Thus, if $\mathcal{D}'_j$ is coherent, then so is $\mathcal{D}'_{j-1}$. Therefore, if $\mathcal{D}'_r$ is coherent we deduce that so is $\mathcal{D}'_j$ for every $j = 1, \dots, r-1$.

This means that $0 \notin \text{posi}(\mathcal{L}^+ \cup \mathcal{D}'_{r-1} \cup \mathcal{D}_r)$. If $0 \in \text{posi}(\mathcal{L}^+ \cup \mathcal{D}'_{r-2} \cup \mathcal{D}_{r-1} \cup \mathcal{D}_r)$, then there are $f \in \mathcal{D}'_{r-2} \cup \{0\}, g \in \mathcal{D}_{r-1} \cup \{0\}, h \in \mathcal{D}_r \cup \{0\}$, not all 0, such that $0 \geq f + g + h$. But $f + g \in \text{posi}(\mathcal{L}^+ \cup \mathcal{D}'_{r-2} \cup \mathcal{D}_{r-1})$, and we can define a gamble $f' \in \mathcal{L}_{S_{r-2} \cap S_{r-1}}$ so that $0 \geq f' + g + h$. Since $f' + g \in \text{posi}(\mathcal{L}^+ \cup \mathcal{D}_{r-2} \cup \mathcal{D}_{r-1}) \cap \mathcal{L}_{S_{r-1}} = \mathcal{D}'_{r-1}$, we deduce that $0 \in \text{posi}(\mathcal{L}^+ \cup \mathcal{D}'_{r-1} \cup \mathcal{D}_r)$, a contradiction with the coherence of $\mathcal{D}_r$.

With a similar reasoning, we deduce that

$$0 \notin \text{posi}(\mathcal{L}^+ \cup \mathcal{D}'_{r-3} \cup \mathcal{D}_{r-2} \cup \mathcal{D}_{r-1} \cup \mathcal{D}_r),$$

and, iterating, that

$$0 \notin \text{posi}(\mathcal{L}^+ \cup \mathcal{D}'_1 \cup \mathcal{D}_2 \cup \dots \cup \mathcal{D}_r) = \text{posi}(\mathcal{L}^+ \cup \mathcal{D}_1 \cup \mathcal{D}_2 \cup \dots \cup \mathcal{D}_r),$$

whence $\mathcal{D}_1, \dots, \mathcal{D}_r$ avoid partial loss. □

Proof of Proposition 14. Let us show that the axioms A1–A6 are satisfied:

A1. Commutativity follows trivially from Eq. (A.3). To prove associativity, consider three coherent sets of desirable gambles $\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3$. Then

$$\begin{aligned} f \in \mathcal{D}_1 \otimes \mathcal{D}_2 \otimes \mathcal{D}_3 &\Leftrightarrow f \geq g_1 + g_2 + g_3 \text{ for some } g_i \in \mathcal{D}_i \cup \{0\} \ (i = 1,2,3) \\ &\Leftrightarrow f \geq g + g_3 \text{ for some } g \in \operatorname{posi}(\mathcal{D}_1 \cup \mathcal{D}_2) \cup \{0\}, g_3 \in \mathcal{D}_3 \cup \{0\} \\ &\Leftrightarrow f \in \operatorname{posi}(\operatorname{posi}(\mathcal{D}_1 \cup \mathcal{D}_2) \cup \mathcal{D}_3 \cup \mathcal{L}^+) \\ &\Leftrightarrow f \in \operatorname{posi}(\operatorname{posi}(\mathcal{D}_1 \cup \mathcal{D}_2 \cup \mathcal{L}^+) \cup \mathcal{D}_3 \cup \mathcal{L}^+) \\ &\Leftrightarrow f \in (\mathcal{D}_1 \otimes \mathcal{D}_2) \otimes \mathcal{D}_3. \end{aligned}$$

Thus, the combination operator is associative.

A2. The first property follows immediately from Eq. (A.3), and the second from Eq. (A.2).

A3. This is an immediate consequence of Eq. (A.2).

A4. For any coherent set of desirable gambles $\mathcal{D}$ relative to $\mathcal{L}_S(\mathcal{X})$ and any $D \subseteq D' \subseteq S$, it holds that $\mathcal{D}^{\downarrow D} = \mathcal{D} \cap \mathcal{L}_D(\mathcal{X}) = \mathcal{D} \cap \mathcal{L}_D(\mathcal{X}) \cap \mathcal{L}_{D'}(\mathcal{X}) = (\mathcal{D}^{\downarrow D'})^{\downarrow D}$.

A5. Consider coherent sets of desirable gambles $\mathcal{D}_1, \mathcal{D}_2$ relative to $\mathcal{L}_{S_1}, \mathcal{L}_{S_2}$, respectively. Then

$$\begin{aligned} f \in (\mathcal{D}_1 \otimes \mathcal{D}_2)^{\downarrow S_1} &\Leftrightarrow f \geq g_1 + g_2 \text{ for some } g_1 \in \mathcal{D}_1 \cup \{0\}, g_2 \in \mathcal{D}_2 \cup \{0\}, f \in \mathcal{L}_{S_1} \\ &\Leftrightarrow f(x) \geq g_1(x) + \sup_{y \in \pi^{-1}_{S_2}(x_{S_1 \cap S_2})} g_2(y) \text{ for some } g_1 \in \mathcal{D}_1 \cup \{0\}, g_2 \in \mathcal{D}_2 \cup \{0\} \\ &\Leftrightarrow f \geq g_1 + g_2' \text{ for some } g_1 \in \mathcal{D}_1 \cup \{0\}, g_2' \in (\mathcal{D}_2 \cap \mathcal{L}_{S_1 \cup S_2}) \cup \{0\} \\ &\Leftrightarrow f \in \mathcal{D}_1 \otimes (\mathcal{D}_2^{\downarrow S_1 \cap S_2}). \end{aligned}$$

A6. Given a set of variables $S \subseteq N$, the vacuous set of desirable gambles $e_S := \mathcal{L}^+(\mathcal{X}_S)$ satisfies the properties of the neutral element. □

References

[1] B. C. Arnold, E. Castillo, and J. M. Sarabia. Compatibility of partial or complete conditional probabilities specifications. *Journal of Statistical Planning and Infernce*, 123:133–159, 2004.

[2] B. C. Arnold, E. Castillo, and J.M. Sarabia. *Conditional Specification of Statistical Models*. Springer, 1999.

[3] B. C. Arnold and S. Press. Compatible conditional distributions. *Journal of the American Statistical Association*, 84:152–156, 1989.

[4] T. Augustin, F. Coolen, G. de Cooman, and M. Troffaes, editors. *Introduction to Imprecise Probabilities*. Wiley, 2014.

[5] C. Beeri, R. Fagin, D. Maier, and M. Yannakis. On the desirability of acyclic database schemes. *Journal of the ACM*, 30:479–513, 1983.

[6] A. Benavoli. Dual probabilistic programming. In *PROBPROG 2018: The International Conference on Probabilistic Programming*, 2018. `http://alessiobenavoli.com/wp-content/uploads/2018/09/dual_PP.pdf`.

[7] A. Benavoli. PyRational: A rational choice modelling framework in python. `https://github.com/PyRational/PyRational/`, since 2018.

[8] A. Benavoli, A. Facchini, D. Piga, and M. Zaffalon. Sum-of-squares for bounded rationality. *International Journal of Approximate Reasoning*, 105:130–152, 2019.

[9] A. Benavoli, A. Facchini, M. Zaffalon, and J. Vicente-Pérez. A polarity theory for sets of desirable gambles. In A. Antonucci, G. Corani, I. Couso, and S. Destercke, editors, *ISIPTA '17: Proceedings of the Tenth International Symposium on Imprecise Probability: Theories and Applications*, volume 62 of *PMLR*, pages 37–48. JMLR.org, 2017.

[10] W. Bergsma and T. Rudas. Marginal models for categorical data. *The Annals of Statistics*, 30(1):140–159, 2002.

[11] U. Bertelé and F. Brioschi. *Nonserial Dynamic Programming*. Academic Press, 1972.

[12] V. Biazzo, A. Gilio, T. Lukasiewicz, and G. Sanfilippo. Probabilistic logic under coherence: complexity and algorithms. *Annals of Mathematics and Artificial Intelligence*, 45(1–2):35–81, 2005.

[13] J. Blair and B. Peyton. An introduction to chordal graphs and clique trees. In *Graph theory and sparse matrix computation*, volume 56, pages 1–30. Springer-Verlag, 1993.

[14] G. Boole. *An investigation on the laws of thought, on which are founded the mathematical theories of logic and probabilities*. Walton and Maberley, London, 1854. reprinted by Dover, New York in 1961.

[15] A. Brozzi, A. Capotorti, and B. Vantaggi. Incoherence correction strategies in statistical matching. *International Journal of Approximate Reasoning*, 53(8):1124–1136, 2012.

[16] A. Capotorti, G. Regoli, and F. Vattari. Correction of incoherent conditional probability assessments. *International Journal of Approximate Reasoning*, 51(6):718–727, 2010.

[17] G. Coletti and R. Scozzafava. *Probabilistic Logic in a Coherent Setting*. Kluwer, 2002.

[18] F. G. Cozman and L. F. di Ianni. Probabilistic satisfiability and coherence checking through integer programming. *International Journal of Approximate Reasoning*, 58:57–70, 2015.

[19] I. Csiszár. I-divergence geometry of probability distributions and minimization problems. *The Annals of Probability*, 3(1):146–158, 1975.

[20] C. Cuadras, J. Fortiana, and J.A. Rodriguez-Lallena. *Distributions with Given Marginals and Statistical Modelling*. Kluwer Academic Publishers, 2002.

[21] C. de Campos and F. Cozman. The inferential complexity of Bayesian and credal networks. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pages 1313–1318, Edinburgh, 2005.

[22] G. de Cooman. Possibility theory II: conditional possibility. *International Journal of General Systems*, 25:325–351, 1997.

[23] G. de Cooman. Belief models: an order-theoretic investigation. *Annals of Mathematics and Artificial Intelligence*, 45:5–34, 2005.

[24] G. de Cooman and E. Miranda. Irrelevance and independence for sets of desirable gambles. *Journal of Artificial Intelligence Research*, 45:601–640, 2012.

[25] B. de Finetti. *Teoria delle Probabilità*. Einaudi, Turin, 1970.

[26] B. de Finetti. *Theory of Probability: A Critical Introductory Treatment*. John Wiley & Sons, Chichester, 1974–1975. English translation of [25], two volumes.

[27] R. Dechter. Bucket elimination: A unifying framework for probabilistic inference algorithms. In *Uncertainty in Artificial Intelligence (UAI'96)*, pages 211–219, 1996.

[28] R. Dechter and J. Pearl. Tree-clustering schemes for constraint-processing. In *Proceedings of the Seventh National Conference on Artificial Intelligence*, pages 150–154. AAAI, 1988.

[29] W.E. Deming and F.F. Stephan. On a least square adjustment of a sampled frequency table when the expected marginal totals are known. *Annals of Mathematical Statistics*, 11:427–444, 1940.

[30] D. Denneberg. Conditioning (updating) non-additive measures. *Annals of Operations Research*, 52(1):21–42, 1994.

[31] L. E. Dubins. Finitely additive conditional probabilities, conglomerability and disintegrations. *The Annals of Probability*, 3:88–99, 1975.

[32] D. Dubois and H. Prade. *Possibility Theory*. Plenum Press, New York, 1988.

[33] D. Dubois and H. Prade. Bayesian conditioning in possibility theory. *Fuzzy Sets and Systems*, 92(2):223–240, 1997.

[34] M. Fréchet. Sur les tableaux de correlation dont les marges sont données. *Ann. Univ. Lyon, Section A, Series 3*, 14:53–77, 1951.

[35] T. Fritz and R. Chaves. Entropic inequalities and marginal problems. *IEEE Transactions on Information Theory*, 59(3):803–817, 2013.

[36] L. C. van der Gaag. Computing probability intervals under independency constraints. In *Proceedings of the Sixth Annual Conference on Uncertainty in Artificial Intelligence*, UAI '90, pages 457–466, New York, NY, USA, 1991. Elsevier Science Inc.

[37] P. Gärdenfors, editor. *Belief Revision*. Cambridge University Press, 1992.

[38] G. Georgakopoulos, D. Kavvadias, and C. H. Papadimitriou. Probabilistic satisfiability. *Journal of complexity*, 4(1):1–11, 1988.

[39] I. Gilboa and D. Schmeidler. Updating ambiguous beliefs. *Journal of Economic Theory*, 59(1):33–49, 1993.

[40] R. Haenni. Ordered valuation algebras: a generic framework for approximating inference. *International Journal of Approximate Reasoning*, 37(1):1–41, 2004.

[41] T. Hailperin. Best possible inequalities for the probability of a logical function of events. *The American Mathematical Monthly*, 72(4):343–359, 1965.

[42] R. Halin. S-functions for graphs. *Journal of Geometry*, 8(1–2):171–186, 1976.

[43] P. Hansen, B. Jaumard, M. Poggi De Aragão, F. Chauny, and S. Perron. Probabilistic satisfability with imprecise probabilities. *International Journal of Approximate Reasoning*, 24(2–3):171–189, 2000.

[44] W. Hoeffding. Masstabinvariante korrelationtheorie. *Schriften des Mathematischen Instituts und des Institud für Angewndte Mathematik der Universität Berlin*, 5:181–233, 1940.

[45] C. Howson. Can logic be combined with probability? Probably. *Journal of Applied Logic*, 7:177–187, 2009.

[46] B. Jaumard, H. Hansen, and M. Poggi De Aragão. Column generation methods for probabilistic logic. *ORSA Journal on Computing*, 3:135–148, 1991.

[47] F. V. Jensen and T. Nielsen. *Bayesian Networks and Decision Graphs*. Springer, 2007.

[48] F.V. Jensen, S.L Lauritzen, and K.G. Olesen. Bayesian updating in causal probabilistic networks by local computation. *Computational Statistics Quarterly*, 4:269–292, 1990.

[49] R. Jirousek. Solution of the marginal problem and decomposable distributions. *Kybernetika*, 27(5):403–412, 1991.

[50] W. Kamakura and M. Wedel. Statistical data fusion for cross-tabulation. *Journal of Marketing Research*, 34:485–498, 1997.

[51] H.G. Kellerer. Verteilungsfunktionen mit gegebenen marginalverteilungen. *Z. Wahrscheinlichkeitstheorie*, 3:247–270, 1964.

[52] J. Kohlas. *Information Algebras: Generic Structures for Inference*. 2003.

[53] D. Koller and N. Friedman. *Probabilistic Graphical Models*. The MIT Press, 2009.

[54] P. H. Krauss. Representation of conditional probability measures on boolean algebras. *Acta Mathematica Academiae Scientiarum Hungaricae*, 19:229–241, 1968.

[55] K.-L. Kuo, C.-C. Song, and T. J. Jiang. Exactly and almost compatible joint distributions for high-dimensional discrete conditional distributions. *Journal of Multivariate Analysis*, 157:115–123, 2017.

[56] J. B. Lasserre. Convergent SDP-relaxations in polynomial optimization with sparsity. *SIAM Journal on Optimization*, 17(3):822–843, 2006.

[57] J. B. Lasserre. *Moments, Positive Polynomials and Their Applications*, volume 1. World Scientific, 2009.

[58] S. L. Lauritzen and D. J. Spiegelhalter. Local computations with probabilities on graphical structures and their application to expert systems (with discussion). *Journal of the Royal Statistical Society, Series B*, 50:157–224, 1988.

[59] J. F. Lemmer. Generalized Bayesian updating of incompletely specified distributions. *Large Scale Systems*, 5:51–68, 1983.

[60] F.M. Malvestuto. Existence of extensions and product extensions for discrete probability distributions. *Discrete Mathematics*, 69:61–77, 1988.

[61] D. Mauá, C. P. de Campos, and M. Zaffalon. Updating credal networks is approximable in polynomial time. *International Journal of Approximate Reasoning*, 53(8):1183 – 1199, 2012.

[62] E. Miranda and G. de Cooman. Marginal extension in the theory of coherent lower previsions. *International Journal of Approximate Reasoning*, 46(1):188–225, 2007.

[63] E. Miranda and I. Montes. Coherent updating of non-additive measures. *International Journal of Approximate Reasoning*, 56:159–177, 2015.

[64] E. Miranda and M. Zaffalon. Coherence graphs. *Artificial Intelligence*, 173(1):104–144, 2009.

[65] E. Miranda and M. Zaffalon. Conditional models: coherence and inference through sequences of joint mass functions. *Journal of Statistical Planning and Inference*, 140(7):1805–1833, 2010.

[66] E. Miranda and M. Zaffalon. Notes on desirability and conditional lower previsions. *Annals of Mathematics and Artificial Intelligence*, 60(3–4):251–309, 2010.

[67] E. Miranda and M. Zaffalon. Conglomerable coherence. *International Journal of Approximate Reasoning*, 54(9):1322–1350, 2013.

[68] E. Miranda and M. Zaffalon. Conformity and independence with coherent lower previsions. *International Journal of Approximate Reasoning*, 78:125–137, 2016.

[69] E. Miranda, M. Zaffalon, and G. de Cooman. Conglomerable natural extension. *International Journal of Approximate Reasoning*, 53(8):1200–1227, 2012.

[70] S. Moral and N. Wilson. Revision rules for convex sets of probabilities. In G. Coletti, D. Dubois, and R. Scozzafava, editors, *Mathematical Models for Handling Partial Knowledge in Artificial Intelligence*, pages 113–128. Plenum Press, New York, 1995.

[71] N.J. Nilsson. Probabilistic logic. *Artificial intelligence*, 28:71–87, 1986.

[72] B. Okner. Data matching and merging: an overview. 3(2):347–352, 1974.

[73] A. Pothen and C. Sun. Distributed multifrontal factorization using clique trees. In *Proceedings of the Fifth SIAM Conference on Parallel Processing for Scientific Computing*, pages 34–40. SIAM, 1992.

[74] E. Quaeghebeur. Characterizing coherence, correcting incoherence. *International Journal of Approximate Reasoning*, 56(B):208–223, 2015.

[75] E. Quaeghebeur, G. de Cooman, and F. Hermans. Accept & reject statement-based uncertainty models. *International Journal of Approximate Reasoning*, 57:69–102, 2015.

[76] L. Rüschendorf. Construction of multivariate distributions with given marginals. *Annals of the Institute of Statistical Mathematics*, 37(A):225–233, 1985.

[77] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, NJ, 1976.

[78] P. P. Shenoy. A valuation-based language for expert systems. *International Journal of Approximate Reasoning*, 3(5):383 – 411, 1989.

[79] P. P. Shenoy and G. Shafer. Axioms for probability and belief-function propagation. In R. D. Shachter, T. S. Levitt, L. N. Kanal, and J. F. Lemmer, editors, *Uncertainty in Artificial Intelligence*, volume 9 of *Machine Intelligence and Pattern Recognition*, pages 169–198. North-Holland, 1990.

[80] H. J. Skala. The existence of probability measures with given marginals. *The Annals of Probability*, 21(1):136–142, 1993.

[81] V. Strassen. The existence of probability measures with given marginals. *Annals of Mathematical Statistics*, 36:423–429, 1965.

[82] M. Studený. Marginal problem in different calculi of AI. In *Advances in Intelligent Computing- IPMU'1994*, pages 348–359. Springer, 1994.

[83] M. Studený. Conditional independence and natural conditional functions. *International Journal of Approximate Reasoning*, 12:43–68, 1995.

[84] A. Tarski. *Logic, Semantics, Metamathematics. Papers from 1923 to 1938.* Oxford University Press, 1956. Translated by J.H. Woodger.

[85] B. Vantaggi. Statistical matching of multiple sources: A look through coherence. *International Journal of Approximate Reasoning*, 49:701–711, 2008.

[86] J. Vejnarová. A note on the interval-valued marginal problem and its maximum entropy solution. *Kybernetika*, 34(1):19–26, 1998.

[87] P. Vicig, M. Zaffalon, and F. Cozman. Notes on 'Notes on conditional previsions'. *International Journal of Approximate Reasoning*, 44(3):358–365, 2007.

[88] N. N. Vorob'ev. Consistent families of measures and their extensions. *Theory of Probability and its Applications*, VII(2):147–163, 1962.

[89] P. Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.

[90] P. Walley. Towards a unified theory of imprecise probability. *International Journal of Approximate Reasoning*, 24:125–148, 2000.

[91] P. Walley, R. Pelessoni, and P. Vicig. Direct algorithms for checking consistency and making inferences from conditional probability assessments. *Journal of Statistical Planning and Inference*, 126:119–151, 2004.

[92] Y. J. Wang and E. H. Ip. Conditionally specified continuous distributions. *Biometrika*, 95(3):735–746, 2008.

[93] Y. J. Wang and K.-L. Kuo. Compatibility of discrete conditional distributions with structural zeros. *Journal of Multivariate Analysis*, 101(1):191–199, 2010.

[94] P. M. Williams. Notes on conditional previsions. Technical report, School of Mathematical and Physical Science, University of Sussex, UK, 1975. Reprinted in [95].

[95] P. M. Williams. Notes on conditional previsions. *International Journal of Approximate Reasoning*, 44:366–383, 2007. Revised journal version of [94].

[96] M. Zaffalon and E. Miranda. Desirability foundations of robust rational decision making. *Synthese*. Accepted for publication.

[97] M. Zaffalon and E. Miranda. Conservative inference rule for uncertain reasoning under incompleteness. *Journal of Artificial Intelligence Research*, 34:757–821, 2009.

[98] M. Zaffalon and E. Miranda. Probability and time. *Artificial Intelligence*, 198(1):1–51, 2013.

[99] M. Zaffalon and E. Miranda. Axiomatising incomplete preferences through sets of desirable gambles. *Journal of Artificial Intelligence Research*, 60:1057–1126, 2017.

[100] M. Zahn. Solving of constraint satisfaction problems using join trees. In *Proceedings of International Conference on Expert Systems for Development*, pages 159–164, 1994.