The Thirty-First International Florida
Artificial Intelligence Research Society Conference (FLAIRS-31)

Analysis of Jeffrey's Rule of Conditioning in an Imprecise Probabilistic Setting

Karim Tabia
Artois University - Nord de France,
CRIL CNRS UMR 8188, Artois, F-62307 Lens, France
tabia@cril.univ-artois.fr

Abstract

While sets of probability measures and imprecise probabilities in general, are widely accepted as a powerful and unifying framework for handling uncertain and incomplete information, updating such belief sets with new uncertain inputs has not received enough attention. In this paper, we provide an analysis of Jeffrey's rule of conditioning for updating sets of probability measures with new information, possibly uncertain and imprecise, also expressed as sets of probability measures. The paper first provides properties for updating sets of probability measures in the spirit of Jeffrey's rule, then provides and analyses extensions of Jeffrey's rule to three main imprecise probability representations: i) finite sets of probability measures and ii) convex set of probability measures specified by extreme points. The proposed extensions capture the proposed postulates and recover the standard Jeffrey's rule in case where the updated set and the new input are single probability measures.

Introduction

Imprecise probability theory (Levi 1980; Walley 2000) is a unifying uncertainty theory particularly suited for encoding and reasoning with imprecise or ill-known information. This framework is often seen as a probabilistic setting with relaxed parameters and it is typically used to reason with multiple expert information (Nau 2002), perform sensitivity analysis (Bock, de Campos, and Antonucci 2014), decision making with incomplete or scarce information (Antonucci, Piatti, and Zaffalon 2007), etc. Imprecise probabilities are often associated with a robust Bayesian interpretation (Berger et al. 1994) assuming that the probability measure corresponding to the actual beliefs exists and it is unique but it is unknown, that's why it is expressed in an imprecise way using the concept of sets of probability measures, credal sets (Levi 1980; Walley 2000) or using other representations (such as interval-based probabilities (de Campos, Huete, and Moral 1994) and probabilistic logic programs (Lukasiewicz 2001)).

Given a set of initial uncertain beliefs, one may have new information which can be in the form of a hard evidence or in the form of uncertain or soft evidence (e.g. unreliable input) or simply new uncertain information regarding some events[1]. Our focus is on updating a set of probability measures with new information expressed also as a set of probability measures. In the standard probabilistic setting, Jeffrey's rule (Jeffrey 1965) generalizes the standard probabilistic conditioning to the case of uncertain inputs. This conditioning rule has been adapted and studied in many uncertainty settings (for instances, see (Dubois and Prade 1997) for the possibilistic setting, (Ma et al. 2011) for Dempster-Shafer theory). In (Benferhat et al. 2010), it is claimed that this rule can successfully recover most of belief revision rules such as natural and drastic belief revision.

Many works highlighted the necessity of updating probabilistic information with sets of probability measures (Karlsson, Johansson, and Andler 2011; Skulj 2006; Tang and Zheng 2006; Rens, Meyer, and Casini 2016). For example, in (Skulj 2006), the author updates a probability measure to create some *neighborhood* of imprecise probabilities for some events. In (Karlsson, Johansson, and Andler 2011), the authors study combining multiple evidences provided in the form of credal sets. There is to the best of our knowledge no study on an extension of Jeffrey's rule to sets of probability measures. The main contributions of the paper are:

i) We provide natural properties that an extension of Jeffrey's rule to imprecise probabilistic settings should satisfy.
ii) We provide extensions of Jeffrey's rule to sets of probability measures and convex credal sets.
iii) We study the properties of the proposed extensions. Interestingly enough, the proposed extensions satisfy the defined postulates and collapse to the standard Jeffrey's rule in case where the prior belief set and the new input consist only in single probability measures.

Imprecise probabilities: Basic concepts

In the following, Ω=$\{\omega_0, \omega_1, .., \omega_m\}$ denotes the universe of discourse (all possible states of the world) and ω_i denotes a given state (also called interpretation). Sets of interpretations $\phi\subseteq\Omega$, $\psi\subseteq\Omega$ are called events.

Sets of probability measures and credal sets

Let Δ denote the set of all probability measures over Ω. A set of probability measures K is a subset of Δ. K denotes

[1] On the different meanings of hard, soft and uncertain evidence, see (Ma and Liu 2011).

a finite[2] or inifinite set of classical probability measures p over Ω. In order to avoid heavy notations, a set of probability measures will be denoted K, the same notation used for credal sets. In this paper, a credal set represents all the probability measures satisfying some requirements or constraints. More precisely, a credal set is defined as follows:

Definition 1 (Credal set) *A credal set K is a closed convex set of probability distributions.*

Credal sets are generally induced by probabilistic beliefs encoded by means of interval-based probabilities (de Campos, Huete, and Moral 1994) or probabilistic constraints as in conditional logic programs (Lukasiewicz 2001). Intuitively, if K is a convex set of probability measures, then linearly $mixing$[3] any two distributions p_1 and p_2 from K will result in a distribution p belonging to K. Given that a credal set may contain an infinite number of probability measures, there are three main commonly used ways to encode imprecise beliefs. *i) Vertex-based representation* where the uncertainty is encoded by a finite set of standard probability distributions representing extreme points of the convex set K. *ii) Interval-based representation* where every interpretation $\omega \in \Omega$ is associated with upper and lower probabilities. *iii) Constraint-based representation* where the uncertainty is specified by means of constraints as in the comparative probabilities framework (Miranda and Destercke 2013) or in probabilistic logic programs (Lukasiewicz 2001). In this paper, we focus only on the vertex representation.

Vertex-based representation

This representation defines a convex credal set K by a finite number of probability measures called extreme points. Such a credal set is called a finitely generated credal set. Any probability measure of K can be expressed as linear combination of extreme points.

Definition 2 (Extreme point) *An extreme point (also called vertex) p of a credal set K is a probability measure such that it is impossible to find two different probability distributions $p_1 \in K$ and $p_2 \in K$ such that $p=\alpha * p_1 + (1-\alpha) * p_2$ with $\alpha \in]0,1[$.*

This representation is for instance used in the JavaBayes[4] platform for modeling and reasoning with Bayesian and credal networks (Cozman 2000). In the following, $ext(K)$ denotes the set of extreme points of the credal set K.

Definition 3 (Convex hull) *The convex hull of credal set K, denoted $CH(K)$ is the closed set of probability measures whose polytope is characterized by the set of extreme points $ext(K)$.*

Example 1 *In Figure 1, a convex credal set K is depicted using a barycentric representation. Here $\Omega = \{\omega_0, \omega_1, \omega_2\}$ and K is finitely generated by three extreme points $p_1=(.45, 0, .55)$, $p_2=(.7, .1, .2)$ and $p_3=(.1, .6, .3)$.*

[2]For instance, in case we have 10 experts where each expert i having his own beliefs in the form of a probability measure p_i then this set is composed by the 10 probability measures p_1,..,p_{10}.

[3]Mixing here means linearly combining a set of distributions p_1 .. p_k as follows: $p=\sum_{i=1}^{k}(a_i * p_i)$ where $\sum_{i=1}^{k} a_i=1$.

[4]http://www.cs.cmu.edu/~javabayes/Home/

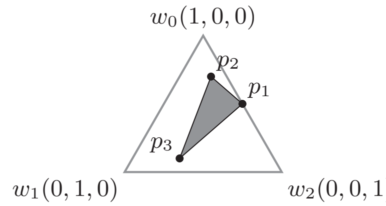

Figure 1: Example of extreme points p_1, p_2 and p_3 using a barycentric representation.

Note that any closed convex set can be encoded by a finite number of extreme points (Levi 1980; Wallner 2007).

Reasoning with credal sets

Reasoning tasks are performed on sets of probability measures or credal sets by exploring all the models of that credal set[5]. For instance, marginalizing a credal set $K(X,Y)$ on two sub-sets of variables X and Y is done as follows:

$$K(X) = \{\sum_{Y} p(X,Y) : p \in K(X,Y)\} \quad (1)$$

Conditioning on an event $\phi \subseteq \Omega$ is defined as follows:

$$K(\omega_i|\phi) = \{p(\omega_i|\phi) : p \in K \text{ and } p(\phi) > 0\} \quad (2)$$

Note that for practical computational reasons, reasoning on K is done on $ext(K)$ which provides an equivalent representation. Indeed, inference on a credal set K is equivalent to inference on its extremes points (de Campos, Huete, and Moral 1994). For instance, for marginalization, given a credal set $K(X,Y)$ on two subsets of variables X and Y. Then,

$$K(X) = CH(\{p(X) : p \in ext(K(X,Y)\}) \quad (3)$$

We assume that K is a finitely generated credal set, namely K is the convex hull of its set of extreme points $ext(K)$. In the following, we propose extensions of Jeffrey's rule to sets of probability measures.

Extending Jeffrey's rule to sets of probability measures

This section analyzes a straightforward extension of Jeffrey's rule to sets of probability measures.

Motivating example

Let us assume we are dealing with learning probabilities from a dataset over five boolean variables X_1, X_2, X_3, X_4 and X_5. Suppose a small dataset D is collected where the values of some variables are missing. Assume also we are interested in deriving an imprecise probability distribution (here an interval-based probability distribution) from this small dataset. In this case, for each configuration $x_1..x_5$ of the five variables $(X_1,..,X_5)$, we will have a lower bound

[5]Alternative approaches consist for instance in selecting the most informative model (in the sense of information entropy for example) of K to draw inferences as it is done in (Lukasiewicz 2001).

l corresponding to the frequency of $x_1..x_5$ in D and an upper bound u corresponding to the proportion of entries of D that are either $x_1..x_5$ or that can be $x_1..x_5$ (for instance, if the value of variable X_5 is missing in a given entry of D then it can be any value of X_5). Let P be the interval-based probability distribution derived from D.

X_1	X_2	X_3	X_4	X_5	$P(X_1X_2X_3X_4X_5)$
0	0	0	0	0	[.001, .04]
0	0	0	0	1	[.02, .1]
...	...	...	...	...	...
1	1	1	1	1	[.005, .035]

Assume now that we have a latest and bigger dataset D' but only on a subset of variables $X_1..X_2$. D' also contains some missing data. Let P' be the interval-based probability distribution computed from D'.

X_1	X_2	$P(X_1X_2)$
0	0	[.025, .059]
0	1	[.2, .35]
...	...	...
1	1	[.5, .7]

It fully makes sense to revise the initial distribution P by P' since this latter is more recent and more representative of the problem as it concerns a large amount of data. It is important to note that the information to update is a set of probablity measues (all probability measures complying with the intervals of P) and the new input is also a set of probability measures. This update task is fully in the spirit of Jeffrey's rule but there is to the best of our knowledge no extension of Jeffrey's rule to sets of probability measures. Of course, the need to revise sets of probabilities by new sets of probabilities can be encountered either when dealing with empirical data (typical situations are dealing with missing data and small datasets) or when dealing with subjective beliefs of agents.

Jeffrey's rule

Jeffrey's rule (Jeffrey 1965) is an extension of the classical probabilistic conditioning to the case where the new observation is uncertain. It allows to update an initial probability distribution p into a posterior one p' given the uncertainty bearing on a set of mutually exclusive and exhaustive events λ_1,..,λ_n. The new input is of the form (λ_i, α_i), i=1..n where α_i denotes the new probability of λ_i. Jeffrey's rule lies on the two following principles:

i) Success principle: After the update operation, the posterior probability of each event λ_i must be equal to α_i, namely $\forall\lambda_i$, $p'(\lambda_i)$=α_i.

ii) Probability kinematics principle: This constraint ensures a kind of minimal change principle. Jeffrey's method assumes that in spite of the disagreement about the events λ_i in the initial distribution p and the new one p', the conditional probability of any event $\phi\subseteq\Omega$ given any uncertain event λ_i remains the same in the original and the revised distributions. Namely,

$$\forall\lambda_i \subseteq \Omega, \forall\phi \subseteq \Omega, p(\phi|\lambda_i) = p'(\phi|\lambda_i). \quad (4)$$

Given a probability measure p encoding the initial beliefs and new inputs the form (λ_i, α_i). The updated probability degree of any event $\phi\subseteq\Omega$, is done as follows:

$$p'(\phi) = \sum_{\lambda_i} \alpha_i * \frac{p(\phi, \lambda_i)}{p(\lambda_i)}. \quad (5)$$

The posterior distribution p' obtained using Jeffrey's rule always exists and it is unique (Chan and Darwiche 2005).

Jeffrey's rule for sets of probability measures

Recall that our focus is not the foundations and justifications of Jeffrey's rule in imprecise probabilities. Interested readers can refer for instance to (Chan and Darwiche 2005; Grove and Halpern 1998; Skulj 2006; Yue and Liu 2008). For the sake of simplicity, the input, the belief set to update is given in the form of a credal set denoted K. The new information is also given in the form of a credal set K_{in} over a partition of Ω. This form for the inputs is general enough to capture sure observations, uncertain observations and imprecise ones. Moreover, we assume that K and K_{in} are not empty sets. Let K' be a the updated set obtained by updating K with K_{in}. Let us now see what an extension of Jeffrey's rule could aim to satisfy in an imprecise probabilistic setting.

(P1) $K'(\lambda_1..\lambda_n)\subseteq K_{in}$

(P2) $\forall\lambda_i\subseteq\Omega$, $\forall\phi\subseteq\Omega$, $K(\phi|\lambda_i)$=$K'(\phi|\lambda_i)$

(P3) $udp(K, K_{in})$=$\bigcup_{p\in K, p_{in}\in K_{in}} udp(p, p_{in})$.

Postulate **P1** corresponds to the success postulate ensuring that the new information should be accepted (the inputs are seen as constraints to be satisfied). Of course, the success postulate may be questionable in some contexts, but it may be a desired property in some applications such as in (Skulj 2006). In order to stay in Jeffrey's rule spirit, we just rephrase this postulate in the context of sets of probabilities. The converse inclusion $K_{in}\subseteq K'(\lambda_1..\lambda_n)$ is strong as there may exist λ_i and $p_{in}\in K_{in}$ such that $p_{in}(\lambda_i)$>0 while $\forall p\in K$, $p(\lambda_i)$=0 preventing the application of Jeffrey's rule on an a priori impossible event as in the standard case.

P2 is the statement of kinematics principle adapted to the case of sets of probability measures. This postulate aims to ensure that K' and K preserve the conditional credal sets on the events λ_1,..,λ_n.

P3 extends the one proposed in (Grove and Halpern 1998) in order to capture the fact that updating a set of probability measures by another set of measures should take into account every measure in the initial set and every measure in the new input. This makes sense within a robust Bayesian interpretation of sets of probability measures.

Lemma 1 *If* $|K|$=$|K_{in}|$=*1 then postulates* **P1** *and* **P2** *recover with the success and probability kinematics principles of Jeffrey's rule respectively.*

Obviously, if the credal sets K and K_{in} are singletons (namely, each one composed of only one probability measure), then **P1** will recover the success principle (the input K_{in} is fully accepted as in Jeffrey's rule) while **P2** will recover the probability kinematics principle. Consequently, the only solution satifying these properties is the one obtained using Jeffrey's rule and it always exists (Chan and Darwiche 2005).

Conditioning sets of probabiliy measures with uncertain inputs

One direct way to extend Jeffrey's rule to finite sets of probability measures is to update every member of the belief set K by every member of the new input K_{in} as follows:

Definition 4 *Let K be a set of probability measures representing the current beliefs over the universe of discourse Ω. Let the new information be K_{in}.*

$$K' = \{p' : p' = \textit{Jeffrey}(p, p_{in}), p \in K : \forall\lambda_i, p(\lambda_i) > 0, p_{in} \in K_{in}\} \tag{6}$$

where Jeffrey(p, p_{in}) is the update according to Jeffrey's rule given in Equation 5 of the probability measure p with the new input p_{in}=$(p_{in}(\lambda_1), .., p_{in}(\lambda_n))$.

Updating using Definition 4 in straightforward in case where the belief sets K and K_{in} consist of finite sets of probability measures. It is clear that if both K and K_{in} contain only one probability measure then Definition 4 comes down to Jeffrey's rule in the standard probabilistic setting.

Example 2 *Let us assume that Ω=$\{a_1b_1, a_1b_2, a_2b_1, a_2b_2\}$ and that the current beliefs about a given problem over two binary variables A and B is a set composed of three probability distributions p_1, p_2 and p_3. Suppose that we receive*

A	B	$p_1(AB)$	$p_2(AB)$	$p_3(AB)$
a_1	b_1	.6	.65	.7
a_2	b_1	.15	.1	.1
a_1	b_2	.1	.1	.1
a_2	b_2	.15	.15	.1

Table 1: Example of a belief set K characterized by three extreme points p_1, p_2 and p_3.

new information (for example new data) saying that the probability $p_{in}(b_1)$=.9 and $p_{in}(b_2)$=.1. Applying Jeffrey's rule to each probability measure p_1, p_2 and p_3 will give three updated distributions p'_1, p'_2 and p'_3.

A	B	$p'_1(AB)$	$p'_2(AB)$	$p'_3(AB)$
a_1	b_1	.72	.78	.79
a_2	b_1	.18	.12	.11
a_1	b_2	.04	.04	.05
a_2	b_2	.06	.06	.05

Table 2: The posterior set K' obtained from K of Table 1.

Proposition 1 *Let K be a finite set probability measures over Ω. Let the new information be K_{in} which is a set on an exhaustive and mutually exclusive set of events λ_1,..λ_n. Let K' be the results of updating K' with K_{in} using Definition 4. Then K' satisfies postulates $P1$, $P2$ and $P3$.*

Proof 1 (Sketch)

- *For* **P1***, to show that $K'(\lambda_1..\lambda_n)$$\subseteq$$K_{in}$, let $p'$$\in$$K'$ and show that $\forall\lambda_i$, $\exists p_{in}$$\in$$K_{in}$ s.t. $p'(\lambda_i)$=$p_{in}(\lambda_i)$. If $p'$$\in$$K'$ then $\exists p$$\in$$K$ and $\exists p_{in}$$\in$$K_{in}$ such that p'=Jeffrey(p, p_{in}). Since p' is obtained by updating p with p_{in} with Jeffrey's rule, then $\forall\lambda_i$, $p'(\lambda_i)$=$p_{in}(\lambda_i)$.*
- *For* **P2** *and* **P3***, the proof is also straightforward for finite sets of probability measures since by Definition 4 the update is done using Jeffrey's rule applied individually on each member of K and on each member of K_{in}.*

In practice, the credal set K to update may be finite or infinite (in case of convex sets). In the following, we extend Jeffrey's rule to closed convex credal sets.

Conditioning credal sets with uncertain inputs

In this section, the belief set to update is a closed convex set K specified by its extreme points $ext(K)$ and the new input K_{in} is also a closed convex set specified by its extreme points $ext(K_{in})$. One direct way to extend Jeffrey's rule is to update only extreme points of K with the ones of K_{in}, namely update each $p$$\in$$ext(K)$ with each $p_{in}$$\in$$ext(K_{in})$ using Jeffrey's rule.

Definition 5 *Let K be the closed convex set to update. Let the new information be K_{in} which is a closed convex set on set of exhaustive and mutually exclusive events λ_1,..,λ_n.*

$$K' = CH(\{p' : p' = \textit{jeffrey}(p, p_{in}); p \in ext(K) \textit{ and } p_{in} \in ext(K_{in}), \forall\lambda_i, p(\lambda_i) > 0\}), \tag{7}$$

Given that it is impossible to update every $p$$\in$$K$, update of Definition 5 proceeds by updating only the set of extreme points of K by the set of extreme points of K_{in} then recovers a convex set using the *convex hull* operator.

Example 3 (Example 1 continued) *Let us reuse the credal set K of Example 1 where Ω=$\{\omega_0, \omega_1, \omega_2\}$ and K is finitely generated by three extreme points p_1=$(.45, 0, .55)$, p_2=$(.7, .1, .2)$ and p_3=$(.1, .6, .3)$. Assume now that new information K_{in} regarding two events λ_1=$\{\omega_1, \omega_2\}$ and λ_2=$\{\omega_3\}$ has become available. Assume also that $ext(K_{in})$ consists of two extreme points $\{(.7, .3); (.6, .4)\}$.*

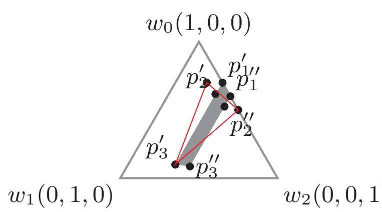

Figure 2: Credal set K' obtained by updating K of Figure 1 with K_{in} whose extreme points are $\{(.7, .3); (.6, .4)\}$.

As shown in Figure 2, the number of extreme points of K' is at most to $|ext(K)|$*$|ext(K_{in})|$.

Proposition 2 *Let K be the closed convex set to update and K_{in} be the new input. Let K' be the updated set computed according to Definition 5 then K' satisfies postulates* **P1**-**P3**.

Proof 2 (Proof sketch)
For **P1***, in order to show that $K'(\lambda_1..\lambda_n)$$\subseteq$$K_{in}$, let $p'$$\in$$K'$ and show that $\exists p_{in}$$\in$$K_{in}$ s.t. $p'(\lambda_1..\lambda_n)$=$p_{in}(\lambda_1..\lambda_n)$. Since $p'$$\in$$K'$ then $p'$$\in$$CH(\{p' : p'$=Jeffrey$(p, p_{in}); p \in ext(K), p_{in} \in ext(K_{in})\})$. It is clear that in case where p'=Jeffrey(p, p_{in}) with $p$$\in$$ext(K)$ and $p_{in}$$\in$$ext(K_{in})$*

then $p'(\lambda_1..\lambda_n)\in K_{in}$ since p' is obtained by updating p with p_{in} using Jeffrey's rule. Now, for any $p'\in K'$ that is not an extreme point of K', p' can be expressed using the extreme points $ext(K')$ as a convex combination of extreme points of K': $p'(\omega)=\sum_{i,j}\alpha_{i,j} * p'_{i,j}(\omega)$ where $p'_{i,j}\in ext(K')$ obtained by updating the extreme point $p_i\in ext(K)$ with the extreme point $p_{in\ j}\in ext(K_{in})$ using Jeffrey's rule, namely $p'_{i,j}(\omega)=\frac{p_i(\omega)*p_{in\ j}(\lambda)}{p_i(\lambda)}$. Hence, $p'(\lambda)=\sum_{\omega\in\lambda}(\sum_{i,j}\alpha_{i,j} * p'_{i,j}(\omega))$ with $\sum_{i,j}\alpha_{i,j}$=1. The proof is consists in starting with expressing $p'(\lambda)$ as a convex combination of extreme points of $ext(K')$ and ending up with expressing $p'(\lambda)$ as a convex combination of $ext(K_{in})$.

For **P2**, *it is enough to see that if $p'\in ext(K')$ then necessarily $\exists p\in ext(K)$ and $\exists p_{in}\in ext(K_{in})$ such that p'=Jeffrey(p, p_{in}). Hence, $\forall\psi\subseteq\Omega$, $\forall\lambda_i\subseteq\Omega$, $p(\phi|\lambda_i)=p'(\phi|\lambda_i)$. Since K and K_{in} are convex sets, then $\forall p'\in K'$, $\exists p\in K$ and $\exists p_{in}\in K_{in}$ such that p'=Jeffrey(p, p_{in}). Hence, $\forall\psi\subseteq\Omega$, $\forall\lambda_i\subseteq\Omega$, $p(\phi|\lambda_i)=p'(\phi|\lambda_i)$.*

For **P3**, *the idea of the proof is based on the convexity of K' obtained by combining two convex sets K and K_{in}. Indeed, K' is obtained by the convex hull operator on a kind of cartesian product of elements of $ext(K)$ and $ext(K_{in})$. Let $K'_{p_{in}}$ be the credal set obtained by updating K with only one point $p_{in}\in K_{in}$. Then $\forall p\in K$, $\exists p'_{in}K'_{p_{in}}$ s.t. p'_{in}=Jeffrey(p, p_{in}). Now, by updating K by every member of K_{in} and taking the convex hull of the obtained points, it holds that $\forall p\in K$, $\forall p_{in}\in K_{in}$, $\exists p'\in K'$ s.t. p'=Jeffrey(p, p_{in}).* □

In the following, we study Jeffrey's rule extension in another widely used representation of imprecise probabilities, namely interval-based probability distributions.

Updating interval-based probability measures

Let us see now how to uptate interval-based probability distributions as the ones of the motivating example. Let P be an interval-based probability distribution (IPD for short) encoding the initial beliefs where each interpretation $\omega\in\Omega$ is associated with a sub-interval of [0, 1]. Given an IPD P encoding the current knowledge and new information P_{in}, there are basically two possible ways to update P with P_{in}:

- *A credal-based method*: This consists in updating the credal set K underlying P (denoted $K(P)$ and containing all the models of P) by the credal set K_{in} underlying P_{in} (denoted $K_{in}(P_{in})$) using Definition 5. Once K' computed, the IPD P' can be computed from K'.
- *An interval-based method*: The main drawback of updating at the credal level is that it manipulates extremes points of IPDs while the number of such extreme points for an IPD with m interpretations can be up to $m!$ (Wallner 2007). The alternative then is to manipulate directly the intervals of the IPD to accommodate the input P_{in}. This method will be addressed in future works.

The credal-based update method is defined as follows:

Definition 6 *Let P be IPD to update and P_{in} be the new input IPD on set of exhaustive and mutually exclusive events λ_1,..,λ_n. Let K' be the updated credal set computed according to Definition 5 on $K(P)$ and $K_{in}(P_{in})$. P' is an IPD on Ω such that $\forall\omega_i\in\Omega$,*

$$P'(\omega_i) = [inf_{p'\in K'}(p'(\omega_i)), sup_{p'\in K'}(p'(\omega_i))]. \quad (8)$$

Example 4 *Let us assume in this example that the current beliefs about a given problem over two binary variables A and B are given by the IPD $P(AB)$. In Table 3, we have the marginal distribution of A (namely, $P(A)$), the one of B (namely, $P(B)$) and the conditional distribution of B given A (namely, $P(B|A)$). Let us now assume that we have new uncertain inputs given in probability distribution $P_{in}(B)$ such that $P_{in}(B{=}b_1)$=[.7, .8] and $P_{in}(B{=}b_2)$=[.2, .3]. In order to update P to accommodate P_{in} using Definition 6, we update $K(P)$ with $K_{in}(P_{in})$ using Definition 5. Note that $K(P)$) has two extreme points p_1=(.70, .05, .1, .15) and p_2=(.50, .25, .1, .15) and $K_{in}(P_{in})$ has also two extreme points, namely p_{in_1}=(.7, .3) and p_{in_2}=(.8, .2). p_1 will be updated into p'_1=(.65, .05, .12, .18) and p''_1=(.75, .05, .08, .12) and p_2 will be updated into p'_2=(.47, .23, .12, .18) and p''_2=(.53, .27, .08, .12). Hence $K'=CH(\{p'_1, p''_1, p'_2, p''_2\})$.*

The updated distribution is given by P' of Table 4. Table 3 and 4 show that the input beliefs encoded by $P'(B)$ are fully accepted (see the marginal distribution $P'(B)$ computed from the updated distribution $P'(AB)$).

A	B	$P(AB)$
a_1	b_1	[.50, .70]
a_2	b_1	[.05, .25]
a_1	b_2	[.10, .10]
a_2	b_2	[.15, .15]

A	$P(A)$
a_1	[.60, .80]
a_2	[.20, .40]

B	$P(B)$
b_1	[.75, .75]
b_2	[.25, .25]

A	B	$P(A\|B)$
a_1	b_1	[.67, .93]
a_2	b_1	[.07, .33]
a_1	b_2	[.40, .40]
a_2	b_2	[.60, .60]

Table 3: Example of an initial IPD P and the underlying marginal and conditional distributions.

A	B	$P'(AB)$
a_1	b_1	[.47, .75]
a_2	b_1	[.05, .27]
a_1	b_2	[.08, .12]
a_2	b_2	[.12, .18]

A	$P'(A)$
a_1	[.59, .83]
a_2	[.17, .41]

B	$P'(B)$
b_1	[.7, .8]
b_2	[.2, .3]

A	B	$P'(A\|B)$
a_1	b_1	[.67, .93]
a_2	b_1	[.07, .33]
a_1	b_2	[.40, .40]
a_2	b_2	[.60, .60]

Table 4: Updated beliefs of the distribution given in Table 3.

Proposition 3 states that this updating ensures that the postulates **P1-P3** are satisfied.

Proposition 3 *Let P be IPD to update. Let the new information be the IPD P_{in} on set of exhaustive and mutually exclusive events λ_1,..,λ_n. Let K' be the updated credal set computed according to Definition 5 on $K(P)$ and $K_{in}(P_{in})$. Let P' the posterior IPD computed from P and P_{in} following Definition 6. Then P' satisfies* **P1-P3**.

Proof 3 (Sketch) *The proof directly follows from the fact that $K(P)$ is an equivalent representation of models of P and the fact that updating using Definition 5 satisfies* **P1-P3**.

Related works and concluding remarks

This paper proposed extensions of Jeffrey's rule of conditioning to the case where the information is encoded in an imprecise probabilistic setting. More precisely, the paper rephrases the two postulates of Jeffrey's rule and added another one to enforce the update operation to take into account every member of the initial set of probability measures and every member of the new input set. The paper extends Jeffrey's rule to i) sets of probability measures and ii) convex credal sets in a vertex-based representation. These extensions are shown to satisfy the proposed postulates and collapse to standard Jeffrey's rule when the initial set and new information are singleton distributions.

Updating sets of probability measures is not a new topic (Grove and Halpern 1998)(Levi 1980)(Walley 2000). However, all these works update sets of probability measures with hard evidence or observations while the focus of the work is updating sets of probability measures with new inputs expressed by means of a set of probability measures. The existing extensions of Jeffrey's rule are limited to special imprecise probabilistic information such as the extensions proposed in (Ma et al. 2011) for Dempster-Shafer theory or the possibilistic extension of Jeffrey's rule proposed in (Dubois and Prade 1997). In (Skulj 2006), the author use Jeffrey's rule to update a single probability distribution in order to obtain the desired neighborhood of events of interest expressed only in terms of interval probabilities. In (Yue and Liu 2008), the authors dealt with updating imprecise knowledge in the framework of probabilistic logic programming. In case where the imprecise knowledge is compactly encoded by means of belief graphical models called credal networks, there is only one work (J. C. F. da Rocha and de Campos 2008) dealing with updating with soft evidence.

References

Antonucci, A.; Piatti, A.; and Zaffalon, M. 2007. Credal networks for operational risk measurement and management. In Apolloni, B.; Howlett, R.; and Jain, L., eds., *Knowledge-Based Intelligent Information and Engineering Systems*, volume 4693 of *Lecture Notes in Computer Science*. 604–611.

Benferhat, S.; Dubois, D.; Prade, H.; and Williams, M.-A. 2010. A framework for iterated belief revision using possibilistic counterparts to jeffrey's rule. *Fundam. Inform.* 99(2):147–168.

Berger, J.; Moreno, E.; Pericchi, L.; Bayarri, M.; Bernardo, J.; Cano, J.; De la Horra, J.; Martn, J.; Ros-Insa, D.; Betr, B.; Dasgupta, A.; Gustafson, P.; Wasserman, L.; Kadane, J.; Srinivasan, C.; Lavine, M.; OHagan, A.; Polasek, W.; Robert, C.; Goutis, C.; Ruggeri, F.; Salinetti, G.; and Sivaganesan, S. 1994. An overview of robust bayesian analysis. *Test* 3(1):5–124.

Bock, J. D.; de Campos, C. P.; and Antonucci, A. 2014. Global sensitivity analysis for MAP inference in graphical models. In *Advances in Neural Information Processing Systems 27: Annual Conference on Neural Information Processing Systems 2014, December 8-13 2014, Montreal, Quebec, Canada*, 2690–2698.

Chan, H., and Darwiche, A. 2005. On the revision of probabilistic beliefs using uncertain evidence. *Artif. Intell.* 163(1):67–90.

Cozman, F. G. 2000. Credal networks. *Artificial Intelligence* 120(2):199 – 233.

de Campos, L. M.; Huete, J. F.; and Moral, S. 1994. Probability intervals: A tool for uncertain reasoning. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 02(02):167–196.

Dubois, D., and Prade, H. 1997. A synthetic view of belief revision with uncertain inputs in the framework of possibility theory. *Int. J. of Approximate Reasoning* 17(2–3):295–324.

Grove, A. J., and Halpern, J. Y. 1998. Updating sets of probabilities. In *UAI '98: Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence, University of Wisconsin Business School, Madison, Wisconsin, USA, July 24-26, 1998*, 173–182.

J. C. F. da Rocha, A. M. G., and de Campos, C. P. 2008. Dealing with soft evidence in credal networks. In *Proceedings of the Conferencia Latinoamericana de Informatica*.

Jeffrey, R. 1965. The logic of decision.

Karlsson, A.; Johansson, R.; and Andler, S. F. 2011. Characterization and empirical evaluation of bayesian and credal combination operators. *J. Adv. Inf. Fusion* 6(2):150–166.

Levi, I. 1980. *The enterprise of knowledge : an essay on knowledge, credal probability, and chance / Isaac Levi*. MIT Press Cambridge, Mass.

Lukasiewicz, T. 2001. Probabilistic logic programming with conditional constraints. *ACM Trans. Comput. Logic* 2(3):289–339.

Ma, J., and Liu, W. 2011. A framework for managing uncertain inputs: An axiomization of rewarding. *Int. J. Approx. Reasoning* 52(7):917–934.

Ma, J.; Liu, W.; Dubois, D.; and Prade, H. 2011. Bridging jeffrey's rule, AGM revision and dempster conditioning in the theory of evidence. *International Journal on Artificial Intelligence Tools* 20(4):691–720.

Miranda, E., and Destercke, S. 2013. Extreme points of the credal sets generated by elementary comparative probabilities. In van der Gaag, L., ed., *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, volume 7958 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg. 424–435.

Nau, R. F. 2002. The aggregation of imprecise probabilities. *Journal of Statistical Planning and Inference* 105(1):265–282.

Rens, G.; Meyer, T. A.; and Casini, G. 2016. On revision of partially specified convex probabilistic belief bases. In *ECAI 2016 - 22nd European Conference on Artificial Intelligence, 29 August-2 September 2016, The Hague, The Netherlands*, 921–929.

Skulj, D. 2006. Jeffreys conditioning rule in neighbourhood models. *International Journal of Approximate Reasoning* 42(3):192 – 211.

Tang, Y., and Zheng, J. 2006. Generalized jeffreys rule of conditioning and evidence combining rule for a priori probabilistic knowledge in conditional evidence theory. *Information Sciences* 176(11):1590 – 1606.

Walley, P. 2000. Towards a unified theory of imprecise probability. *International Journal of Approximate Reasoning* 24(23):125 – 148.

Wallner, A. 2007. Extreme points of coherent probabilities in finite spaces. *Int. J. Approx. Reasoning* 44(3):339–357.

Yue, A., and Liu, W. 2008. Revising imprecise probabilistic beliefs in the framework of probabilistic logic programming. In *Proceedings AAAI 2008, Chicago, Illinois, USA, July 13-17, 2008*, 590–596.