Updating Probability Intervals with Uncertain Inputs

Karim Tabia
CRIL, Université d'Artois & CNRS, France
tabia@cril.fr

Abstract

Probability intervals provide an intuitive, powerful and unifying setting for encoding and reasoning with imprecise beliefs. This paper addresses the problem of updating uncertain information specified in the form of probability intervals with new uncertain inputs also expressed as probability intervals. We place ourselves in the framework of Jeffrey's rule of conditioning and propose extensions of this conditioning for the interval-based setting. More precisely, we first extend Jeffrey's rule to credal sets then propose extensions of Jeffrey's rule to three common conditioning rules for probability intervals (robust, Dempster and geometric conditionings). While the first extension is based on conditioning the extreme points of the credal sets induced by the probability intervals, the other methods directly revise the interval bounds of the distributions to be updated. Finally, the paper discusses related issues and relates the proposed methods with respect to the state-of-the-art.

1 Introduction

Probability intervals are compact, intuitive and convenient means for representing imprecise probabilities. The theory of imprecise probabilities [Levi, 1980; Walley, 1991] is a unifying uncertainty theory particularly suited for encoding and reasoning with imprecise or ill-known information. It is used to reason with multiple expert information [Nau, 2002], perform sensitivity analysis [Bock *et al.*, 2014], make decisions with partial information [Antonucci *et al.*, 2007], etc. Imprecise probabilities are often associated with a robust Bayesian interpretation [Berger *et al.*, 1994] assuming that the probability measure representing the actual beliefs exists and it is unique but it is unknown due to lack of knowledge. Imprecise probabilities are specified in the form of sets of probability measures, credal sets [Levi, 1980; Walley, 1991] or using other representations such as interval-based probabilities [de Campos *et al.*, 1994], probabilistic logic programs [Lukasiewicz, 2001], belief functions, etc. As for updating probability intervals, it should be noted that there are several conditionings depending on the associated interpretation of probability intervals [Moral and De Campos, 1991]. In this paper, we limit ourselves to the three most common types of conditioning that are the robust Bayesian conditioning [Moral and De Campos, 1991], Dempster conditioning [Shafer, 1976] and the geometric one [Suppes and Zanotti, 1977].

Given an initial belief set, one may learn new information which can be in the form of a hard evidence or in the form of uncertain or soft evidence (e.g. unreliable observation). The focus of the paper is updating[1] a set of probability intervals with new information also expressed by probability intervals. In the classical point-wise probabilistic setting, Jeffrey's rule [Jeffrey, 1965] generalizes the standard probabilistic conditioning to the case where the inputs are uncertain. This conditioning rule has been studied in many uncertainty settings (for instance, see [Dubois and Prade, 1997] for the possibilistic setting and [Smets, 1993; Tang and Zheng, 2006; Ma *et al.*, 2011] for Dempster-Shafer theory).

We place ourselves in the framework of Jeffrey's rule where the new information is pervaded with uncertainty and this latter bears on a partition of the set of possible worlds. We therefore take for granted the principles underlying Jeffrey's rule and we do not focus on the foundations and justifications of such principles. The focus is on extending Jeffrey's rule for updating probability intervals with uncertain inputs. The main contributions of the paper are :

1. Extending Jeffrey's rule of conditioning to update a prior credal set with a new uncertain and imprecise input also in the form of a credal set ;
2. Casting Jeffrey's rule principles (success and probability kinematics) in the probability intervals setting and show that the proposed rules capture such principles ;
3. Proposing Jeffrey's rule counterparts for two alternative conditionings (Dempster and geometric) of probability intervals.

[1] In this paper, we use interchangeably the terms *update, revision* and *conditioning* even if some authors associate different meanings to these belief change processes. See [Dubois and Prade, 1994] for a survey of belief revision and updating rules.

2 Preliminaries

2.1 Probability Intervals

Let us in the following denote by Ω={ω_1,..,ω_n} the set of elementary possible states and denote by $\omega_i \in \Omega$ a given state. Sets of states ϕ, $\psi \subseteq \Omega$ are called events. An interval-based probability distribution (IPD for short) is defined as follows:

Definition 1 (Interval-based probability distribution)**.** *Let Ω be the set of possible states. An interval-based probability distribution P is a function that maps every state $\omega_i \in \Omega$ to a closed interval $P(\omega_i)$=$[l_i, u_i] \subseteq [0, 1]$.*

In an IPD P, every state $\omega_i \in \Omega$ is associated with a probability interval $P(\omega_i)$=$[l_i, u_i]$ where l_i (resp. u_i) denotes the lower (resp. upper) bound of the probability of ω_i. In the following, $\overline{P}(\omega_i)$ (resp. $\underline{P}(\omega_i)$) denotes the upper (resp. lower) bound of $P(\omega_i)$.

In order to be reachable and not empty, the bounds should satisfy the following constraints:

$$\sum_{\omega_i \in \Omega} l_i \leq 1 \leq \sum_{\omega_i \in \Omega} u_i$$

$$\forall \omega_i \in \Omega, l_i + \sum_{\omega_{j \neq i} \in \Omega} u_j \geq 1 \; and \; u_i + \sum_{\omega_{j \neq i} \in \Omega} l_j \leq 1$$

We call a model or a member of IPD P any point-wise (or single-valued) probability distribution p s.t. $\forall \omega_i \in \Omega$, $p(\omega_i) \in P(\omega_i)$=$[l_i,u_i]$. Accordingly, the semantics associated with an IPD is the set of its models, namely all probability measures that comply with the probability intervals. A credal set K on Ω compactly encoded by an IPD P denotes the closed convex set of point-wise probability distributions p that are models of P. Namely,

$$K = \{p \mid l_i \leq p(\omega_i) \leq u_i, \; \forall \omega_i \in \Omega\} \quad (1)$$

A commonly used way to encode a credal set K is the vertex-based representation. This is done by specifying a finite set of standard probability distributions representing the extreme points of K. An extreme point (also called vertex) p of a credal set K is a probability distribution such that it is impossible to find two different probability distributions $p_1 \in K$ and $p_2 \in K$ such that p=α*p_1+(1-α)*p_2 with $\alpha \in]0, 1[$. Enumerating the set of extreme points $ext(P)$ underlying an IPD P requires solving a linear program [de Campos *et al.*, 1994]. An IPD P can be completely represented by the extreme points of the credal set K underlying P.

Reasoning with interval probabilities can either be done at the semantic level using the credal set induced by probability intervals or directly manipulating intervals and using interval arithmetics. Reasoning using a credal set K underlying an IPD P comes down to exploring all the models of that credal set[2]. For instance, marginalizing a credal set $K(X, Y)$ over two subsets of variables X and Y is done as follows:

$$K(X) = \{\sum_{Y} p(X, Y) : p \in K(X, Y)\} \quad (2)$$

[2]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].

For credal sets, the most common form of conditioning on an event $\phi \subseteq \Omega$ is defined as follows :

$$K(\omega_i|\phi) = \{p(\omega_i|\phi) : p \in K \text{ and } p(\phi) > 0\} \quad (3)$$

This is often referred to as cautious or robust Bayesian conditioning (see [Moral and De Campos, 1991] for other forms of conditioning probability intervals).

For 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 extreme points [Walley, 1991; de Campos *et al.*, 1994]. For instance, marginalizing a credal set $K(X, Y)$ over two subsets of variables X and Y is done as follows :

$$K(X) = CH(\{p(X) : p \in ext(K(X, Y)\}), \quad (4)$$

where CH stands for the convex hull operator. Conditioning of Equation 3 comes down to :

$$K(\omega_i|\phi) = CH(\{p(\omega_i|\phi) : p \in ext(K) \text{ and } p(\phi) > 0\}). \quad (5)$$

In the following, we briefly recall Jeffrey's rule in the standard probabilistic setting.

2.2 Jeffrey's Rule of Conditioning

Jeffrey's rule [Jeffrey, 1965] extends the classical probabilistic conditioning to the case where the new information 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={λ_1,..,λ_n} (namely, λ is a partition of Ω). In this setting, the new input is in 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:

Success principle (P1).

$$\forall \lambda_i \in \lambda, p'(\lambda_i) = \alpha_i \quad (6)$$

After the update operation, the posterior probability of each event λ_i must be equal to α_i as required in the new inputs. The uncertain inputs are seen as constraints or an effect once the new information is fully accepted.

Probability kinematics principle (P2).

$$\forall \lambda_i \in \lambda, \forall \phi \subseteq \Omega, p(\phi|\lambda_i) = p'(\phi|\lambda_i). \quad (7)$$

This principle aims to ensure a kind of minimal change by ensuring that the posterior distribution p' should not change the conditional probability degrees of any event ϕ given the uncertain events λ_i. Jeffrey's rule assumes that in spite of the disagreement about the events λ_i in the prior distribution p and the posterior one p', the conditional probability of any event $\phi \subseteq \Omega$ given any uncertain event λ_i should remain the same in the original and the revised distributions.

Given a probability distribution p encoding the initial beliefs and new inputs in the form (λ_i, α_i) for i=1..n, the updated probability degree of any event $\phi \subseteq \Omega$ is obtained as follows:

$$p'(\phi) = \sum_{\lambda_i} \alpha_i * \frac{p(\phi \cap \lambda_i)}{p(\lambda_i)}. \quad (8)$$

The posterior distribution p' obtained using Jeffrey's rule always exists and it is unique [Chan and Darwiche, 2005]. Note that in Jeffrey's rule, the events λ_i should be somewhat possible in the prior distribution (namely, $\forall \lambda_i \in \lambda$, $p(\lambda_i)$>0).

Candidate	Estimate
C_1	[0.24, 0.30]
C_2	[0.22, 0.26]
C_3	[0.14, 0.17]
C_4	[0.30, 0,36]

(a)

Wing	Estimate
$Left$	[0.49, 0.53]
$Right$	[0.47, 0.51]

(b)

Table 1: Example of probability intervals encoding initial information (a) and new uncertain inputs (b).

3 Motivating Example

To illustrate and motivate the interest in expressing and updating uncertain information in the form of probability intervals, let us take the example of the polls generally carried out in the elections and suppose that they are about presidential elections that are going to take place in a given country. Suppose for simplicity that there are only four candidates (denoted C_1, C_2, C_3 and C_4) and that in this country the political polarization is *Left wing* and *Right wing*. Assume now that a first poll of voter preferences for the four candidates on a *small sample* yielded the following results in Table 1 (a) (the estimates are in the form of probability intervals to account for the margin of error).

Assume now that a second poll on a *much larger sample* and where the question is no longer *Which candidate are you going to vote for?* but *Do you prefer a left wing candidate or a right wing one?* To simplify, we assume that candidates C_1 and C_2 are classified in the left wing and the other two in the right wing. Also, the estimates (see Table 1 (b)) are given in the form of probability intervals to take into account the margin of error.

It fully makes sense in this case to update the initial distribution (results of the first poll) with new information (results of the second poll) as this latter is more confident since it is carried out on a much larger sample. It is important to note that the initial information to update is an interval-based probability distribution P and the new input uncertain is also an interval-based probability distribution P_{new} over a partition of the set of candidates. This update task is fully in the spirit of Jeffrey's rule (the new inputs are uncertain information bearing on a partition of Ω and the aim is to give priority to the new inputs). In this example, the task is not a simple belief merging but aims at giving priority to new inputs as in belief revision tasks. This is exactly in line with Jeffrey's rule in the standard probabilistic setting. But this latter handles only point-wise probability distributions and not interval-based ones. There is to the best of our knowledge no counterparts to Jeffrey's rule in the interval-based setting.

4 Updating Credal Sets with Uncertain Inputs

Before addressing updating probability intervals with uncertain inputs, let us first focus on updating credal sets with uncertain inputs using Jeffrey's rule. The idea is to use Jeffrey's rule to update every distribution p from the prior credal set K with every distribution p_{new} from the new input credal set K_{new}. In order to apply Jeffrey's rule on any prior distribution p from K, we assume that $\forall p \in K$, $\forall \lambda_i \in \lambda$, $p(\lambda_i) > 0$.

This section answers two fundamental questions: i) does updating a credal set with another credal result in a credal set and ii) can this update be done by manipulating only the vertices of the credal sets as in the case of conditioning with hard evidence. The following are the first main results of this paper (proofs are provided as supplementary material):

Proposition 1. *Let K be the closed convex set to update. Let p_{new} be a single probability distribution over the partition λ and encoding the new information. Let K' be such that*

$$K' = \{p' : p' = p \oplus p_{new}; p \in K\}.$$

Then K' is a convex credal set.

In Proposition 1, $p \oplus p_{new}$ denotes conditioning a point-wise probability distribution p with new inputs encoded in the form of a probability distribution p_{new} over a partition of Ω using Jeffrey's rule. This result states that using Jeffrey's rule to update each member of a convex credal set K with the same input distribution p_{new} results in a convex credal set K'. Hence, we generalize the result that conditioning a credal set with a hard evidence gives a convex credal set.

Proposition 2. *Let K be the closed convex set to update and let $p \in K$ be a single probability distribution. Let K_{new} be a credal set over the partition λ and encoding the new information at hand. Let K' be such that*

$$K' = \{p' : p' = p \oplus p_{new}; p_{new} \in K_{new}\}.$$

Then K' is a convex credal set.

Proposition 2 states that updating using Jeffrey's rule a single probability distribution p over Ω with each member of credal set K_{new} over a partition λ of Ω results in a convex credal set. Let us now generalize to the case where both K and K_{new} are arbitrary convex sets.

Proposition 3. *Let K be the closed convex set to update and K_{new} be a closed convex set over the partition λ on Ω encoding the new information. Let*

$$K' = \{p' : p' = p \oplus p_{new};\ p \in K \text{ and } p_{new} \in K_{new}\}.$$

Then K' is a convex credal set.

Proposition 3 states that using Jeffrey's rule to update each member of a prior credal set K with each member of credal set K_{new} results in a convex credal set K'.

The following proposition states that the extreme points of the updated credal set K' correspond to the updated set of extreme points of K with the new input using Jeffrey's rule.

Proposition 4. *Let K be the closed convex set to update and let $ext(K)$ be its set of extreme points. Let p_{new} be a single probability distribution over the partition λ and encoding the new information. Let K' be the updated credal set following Proposition 1. Then*

$$ext(K') = \{p' : p' = p \oplus p_{new}; p \in ext(K)\}$$

is the set of extreme points of K'.

Lemma 1. *Let K be the closed convex set to update and let $ext(K)$ be its set of extreme points. Let p_{new} be a single probability distribution over the partition λ and encoding the* *new information. Let K' be the updated credal set following Proposition 1. Then*

$$K' = CH(\{p' : p' = p \oplus p_{new}; p \in ext(K)\}).$$

In fact, given that from Proposition 1 K' is convex (namely, it contains every convex combination of its extreme points) and the fact that $CH(ext(K'))$ is by definition equal to all the convex combinations of the finite set $ext(K')$ computed following Proposition 4 then K' can be recovered using only such extreme points.

Proposition 5 and Lemma 2 generalize conditioning credal sets based on extreme points to the case where the new input is also a credal set.

Proposition 5. *Let K be the closed convex set to update and K_{new} be a closed convex set over the partition λ encoding the new information. Then*

$$ext(K') \subseteq \{p' : p' = p \oplus p_{new}; p \in ext(K), p_{new} \in ext(K_{new})\}$$

This means that an extreme point of K' is necessarily a combination using Jeffrey's rule of an extreme point of K with an extreme point of K_{new} as illustrated in Example 1. But it is not true that every combination of $p \oplus p_{new}$ such that $p \in ext(K)$ and $p_{new} \in ext(K_{new})$ is an extreme point of K'. This is due to the union operation of convex sets. Indeed, in the case where the union of convex sets is convex, the extreme points of the resulting convex set (here $ext(K')$) is included in the union of sets of extreme points of the starting sets.

Lemma 2. *Let K be the closed convex set to update and let $ext(K)$ be its set of extreme points. Let p_{new} be a single probability distribution over the partition λ over Ω and encoding the new information. Let K' be the updated credal set following Proposition 3. Then*

$$K' = CH(\{p' : p' = p \oplus p_{new}; p_{new} \in ext(K_{new}), p \in ext(K)\})$$

Up to now, we showed that updating a credal set K with another uncertain input in the form of a credal set K_{new} over a partition of Ω results in a credal set K' and amounts to updating the extreme points of K with those of K_{new} using Jeffrey's rule. Let us now use such findings to update prior probability intervals P with new uncertain inputs also in the form of probability intervals P_{new}.

5 Updating Probability Intervals with Uncertain Inputs

Let us first place Jeffrey's conditioning rule principles in the context of probability intervals then provide update methods complying with such principles.

Let P be the IPD encoding the prior beliefs. Let P_{new} be the IPD encoding the new information in hand. P_{new} is of the form $((\lambda_1, [l_1, u_1]),..,(\lambda_n, [l_n, u_n]))$. This notation means that $P_{new}(\lambda_i)=[l_i, u_i]$. As in Jeffrey's rule, the set $\lambda=\{\lambda_1,..,\lambda_n\}$ is a partition of Ω and we assume that the events λ_i are somewhat possible in the prior beliefs P (namely, we assume that $\forall\lambda_i\in\lambda$, $\underline{P}(\lambda_i)>0$). The properties to satisfy stated in terms of interval-based probabilities are:

- **Success principle (IP1)** : It ensures that in the updated IPD P', the new information is fully accepted. Namely,

  $$\forall\lambda_i\in\lambda, P'(\lambda_i)=P_{new}(\lambda_i).$$

  This principle enforces posterior intervals to be equal to the input intervals $P_{new}(\lambda_i)$.
- **Probability Kinematics principle (IP2)** : The objective here is to ensure that the intervals of P are modified as little as possible to accept the new information while respecting the probability kinematics principle in order to avoid affecting non relevant information.

  $$\forall\lambda_i \in \lambda, \forall\phi \subseteq \Omega, P(\phi|\lambda_i) = P'(\phi|\lambda_i)$$

Obviously, in the case of IPDs where lower bounds coincide with upper bounds (namely, if the initial belief set consists in a point-wise distribution over Ω and the new inputs are a point-wise distribution over the partition λ, the principles **IP1** and **IP2** collapse to **P1** and **P2** of Jeffrey's rule of Section 2.2. The success principle **IP1** may be questionable but it may be a desired property in some applications such as in [Skulj, 2006]. In order to keep in Jeffrey's rule line, we just rephrase **P1** in the context of probability intervals.

Now given an IPD P encoding the current information and new information P_{new}, there are basically two possible ways to update P with P_{new}: a credal-set based method applying at the semantic level and an interval-based one manipulating directly the bounds of the IPDs. In this paper, we rely on the former for robust conditioning while the latter is used for Dempster and geometric conditionings.

5.1 Updating Probability Intervals via Robust Conditioning on the Underlying Credal Sets

Our objective here is to update a belief set consisting of a prior IPD P with new inputs P_{new} also provided in the form of an IPD over a partition λ of Ω. The belief update here consists in updating the credal set K underlying P (containing all the models of P) by the credal set K_{new} underlying P_{new} using Jeffrey's rule. Indeed, one direct way to update a set of probability measures is to apply Jeffrey's rule on each member of the set where the new input is a single probability measure over the partition λ of Ω.

Following Lemma 2, we can define the update operation using only the extreme points of K and K_{new}, namely update each $p \in ext(K)$ with each $p_{new} \in ext(K_{new})$ using Jeffrey's rule then recover the credal set K' using the *convex hull* operator. Once K' computed, the IPD P' is computed as lower and upper bounds from K'. Hence, the credal-based update method is defined as follows:

Definition 2. *Let P be IPD to update and P_{new} be the new input IPD over a partition λ of Ω. Let K' be the updated credal set computed according to Lemma 2 on K and K_{new} underlying respectively P and P_{new}. P' is an IPD on Ω such that $\forall\omega_i\in\Omega$,*

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

Example 1. *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 2, we have* *the marginal distribution of A (namely, $P(A)$), the one of B (namely, $P(B)$) and the conditional distribution of A given B (namely, $P(A|B)$).*

$A\ B$	$P(AB)$
$a_1\ b_1$	[.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 2: Example of an initial IPD P and the underlying marginal and conditional distributions.

Assume now that we have new uncertain inputs given in probability distribution $P_{new}(B)$ as follows:

B	$P_{new}(B)$
b_1	[.7, .8]
b_2	[.2, .3]

In order to update P to absorb P_{new} using Definition 2, we update K with K_{new} using Lemma 2. Note that K has two extreme points p_1=$(.70, .05, .1, .15)$ and p_2=$(.50, .25, .1, .15)$ and K_{new} has also two extreme points, namely p_{new_1}=$(.7, .3)$ and p_{new_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 in P' of Table 3.

$A\ B$	$P'(AB)$
$a_1\ b_1$	[.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 3: Updated beliefs of the distribution given in Table 2.

Table 2 and 3 show that the new input beliefs $P_{new}(B)$ are fully accepted in the posterior IPD P' (see the marginal distribution $P'(B)$ computed from the updated distribution $P'(AB)$). Moreover, one can also check that $P(A|B)$=$P'(A|B)$ $\forall a_i{\in}D_A$, $\forall b_i{\in}D_B$.

It is obvious that in case the IPDs P and P_{new} underlying credal sets K and K_{new} each consisting of only one single probability distribution then updating P with P_{new} following Definition 2 is equivalent to updating using Jeffrey's rule.

Lemma 3. *If $\forall\omega{\in}\Omega$, $\underline{P}(\omega)$=$\overline{P}(\omega)$ and $\forall\lambda_i{\in}\lambda$, $\underline{P}_{new}(\lambda_i)$=$\overline{P}_{new}(\lambda_i)$ then the posterior distribution P' computed following Definition 2 amounts to using Jeffrey's rule to update the unique point-wise distribution $p{\in}K$ underlying P with $p_{new}{\in}K_{new}$ underlying P_{new}.*

In the general case, Proposition 6 states that updating using Definition 2 ensures that principles **IP1** and **IP2** are satisfied.

Proposition 6. *Let P be IPD to update. Let the new information be the IPD P_{new} over the partition λ of Ω. Let P' be the posterior IPD computed from P and P_{new} following Definition 2. Then P' satisfies* **IP1** *and* **IP2**.

5.2 Alternatives to Robust Conditioning

Robust conditioning of probability intervals based on updating at the credal set level suffers from some drawbacks. For instance, it manipulates extreme points of credal sets underlying the IPDs while the number of such extreme points for an IPD with n states can be up to $n!$ [Wallner, 2007]. More over, it requires that the events λ_i must be somewhat possible in the prior beliefs (namely, $\forall\lambda_i{\in}\lambda$, $\underline{P}(\lambda_i){>}0$) which may be seen as a strong constraint. The alternative then is to update directly the bounds of the intervals of the IPD to accommodate the input P_{new} and use conditionings that do not require such strong assumptions. For instance, Dempster conditioning [Shafer, 1976] directly manipulates the bounds of the prior IPD P to condition on a new evidence $\psi{\subseteq}\Omega$ and it is defined as long as $\overline{P}(\psi){>}0$. It is defined as follows:

$$\overline{P}(\phi||\psi) = \frac{\overline{P}(\phi \cap \psi)}{\overline{P}(\psi)} \tag{9}$$

The lower bound can be easily obtained by duality (namely, $\underline{P}(\phi||\psi)$=1-$\overline{P}(\phi^C||\psi)$=1-$\frac{\overline{P}(\phi^C \cap \psi)}{\overline{P}(\psi)}$. Here ϕ^C denotes the complement of ϕ in Ω. Proposition 7 extends Jeffrey's rule to probability intervals using Dempster conditioning.

Proposition 7. *Let P be the IPD to update. Let the new information be the IPD P_{new} on a partition λ of Ω. Let P' be the posterior IPD computed from P and P_{new} as follows: $\forall\phi{\subseteq}\Omega$,*

$$P'(\phi) = [1-\sum_{\lambda_i} \overline{P}(\phi^C||\lambda_i){*}\overline{P}_{new}(\lambda_i), \sum_{\lambda_i} \overline{P}(\phi||\lambda_i){*}\overline{P}_{new}(\lambda_i)]$$

Then P' satisfies principles **IP1**-**IP2**.

Another common definition of conditioning of probability intervals is the one of geometric conditioning [Suppes and Zanotti, 1977] defined as follows:

$$\underline{P}(\phi|_G\psi) = \frac{\underline{P}(\phi \cap \psi)}{\underline{P}(\psi)} \tag{10}$$

The upper bound is obtained by duality $\underline{P}(\phi|_G\psi)$=$1 - \overline{P}(\phi^C|_G\psi)$=1-$\frac{\overline{P}(\phi^C \cap \psi)}{\overline{P}(\psi)}$.

Proposition 8. *Let P be the IPD to update. Let the new input be the IPD P_{new} on a partition λ of Ω. Let P' be the posterior IPD computed from P and P_{new} as follows: $\forall\phi{\subseteq}\Omega$,*

$$P'(\phi) = [\sum_{\lambda_i} \underline{P}(\phi|_G\lambda_i){*}\underline{P}_{new}(\lambda_i), 1-\sum_{\lambda_i} \underline{P}(\phi^C|_G\lambda_i){*}\underline{P}_{new}(\lambda_i)]$$

Then P' satisfies principles **IP1**-**IP2**.

It is easy to check that in case where $\forall\omega{\in}\Omega$, $\underline{P}(\omega)$=$\overline{P}(\omega)$ and $\forall\lambda_i{\in}\lambda$, $\underline{P}_{new}(\lambda_i)$=$\overline{P}_{new}(\lambda_i)$ then updating following Propositions 7 and 8 collapses to updating using Jeffrey's rule in the standard setting (see Equation 8).

6 Related Work and Discussions

Jeffrey's rule generalizes the standard probabilistic conditioning to the case where the inputs are uncertain. This conditioning rule has been studied in many uncertainty settings such as possibility theory [Dubois and Prade, 1997] and Dempster-Shafer theory [Smets, 1993; Ma *et al.*, 2011]. 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. 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. This work is mostly dealing with revising probabilistic logic programs and the proposed extension coincides with Jeffrey's rule and Bayesian conditioning only when the updated probabilistic logic program induces a single probability distribution. In [Marchetti and Antonucci, 2018], the authors propose a family of generalized adjustment operators tailored to update convex sets of probability measures with inconsistent piece of evidence. In case where the imprecise knowledge is compactly encoded by means of belief graphical models called credal networks [Cozman, 2000], there is only one work [da Rocha *et al.*, 2008] dealing with updating a credal network with soft evidence. Note that the new information here is not a set of probability intervals but consists in likelihood ratios that can be cast into a single point-wise probability distribution. The authors in [Ma and Liu, 2011] deal with belief merging and focus on the influence of the strengths of inputs. The unique operator complying with the proposed postulates was proved to be a merging operator.

Updating sets of probability measures is not a new topic [Grove and Halpern, 1998; Levi, 1980; Walley, 1991; Tabia, 2018]. However, most of this work update sets of probability measures with hard evidence while the focus of the current work is updating sets of probability measures with new inputs both expressed by means of probability intervals. There are essentially two types of generalizations of Jeffrey's rule in the literature. The first one concerns rather generalizations or counterparts of this conditioning rule in some uncertainty theories such as belief functions [Smets, 1993] or possibility theory [Dubois and Prade, 1997] (note that such settings do not generalize the one of probability intervals and credal sets.). The second type of generalizations concerns rather the relaxation of the requirements of Jeffrey's rule. For example, in [Smets, 1993] it is no longer necessary that the new inputs be a partition of the set of possible states. Moreover, in [Smets, 1993; Wagner, 1992] the authors deal with the generalization of the conditionings (including Dempster and the geometric ones) of belief functions where both the prior information and new inputs are encoded by means of basic probability assignments over 2^{Ω} (the powerset of Ω).

Regarding iterative revision, it is well-known that Jeffrey's rule is not commutative (since the new inputs are fully accepted, then updating first with (λ_i,$[l_i, u_i]$) then with (λ_i,$[l'_i, u'_i]$) will be different from updating first with (λ_i,$[l'_i, u'_i]$) followed by updating with ((λ_i,$[l_i, u_i]$)).

Among the questions still open, there is that of the uniqueness of the proposed extensions (the paper only answers the question of the existence of extensions) as well as their optimality in terms of minimizing the belief change. As mentioned earlier, the objective of probability kinematics principle is to ensure some kind of minimal change. In the standard probabilistic setting, Jeffrey's rule was shown to provide the optimal solution in the sense of a distance measure between a prior distribution p and the posterior one p' [Chan and Darwiche, 2005]. One can directly extend such a distance measure to credal sets as follows :

Proposition 9. *Let K and K' be two credal sets over the same states space Ω. Let*

$$D(K, K') = \max(\max_{p \in ext(K)} \Delta(p, K'), \max_{p' \in ext(K')} \Delta(p', K)),$$

where

$$\begin{cases} \Delta(p, K') = \min_{p' \in ext(K')} d(p, p') \\ d(p, p') = ln \max_{\omega \in \Omega} \frac{p'(\omega)}{p(\omega)} - ln \min_{\omega \in \Omega} \frac{p'(\omega)}{p(\omega)}. \end{cases}$$

Then D is a distance measure.

The distance measure D combines Hausdorff distance (measuring distances between sets) with a distance for point-wise probability distributions specifically designed to bound belief change yielded by Jeffrey's rule in the standard probabilistic stetting [Chan and Darwiche, 2005]. Then there remains to show that updating a prior credal set K with new inputs in the form of credal set K_{new} following Lemma 2 guarantees that the posterior credal set K' is optimal in the sense of distance $D(K, K')$ of Proposition 9. These two questions of optimality and uniqueness of the solution are part of the avenues for future work.

Another important open question concerns certain strong requirements of Jeffrey's conditioning in particular that the new information relates only to a partition of the set of possible states and the one requiring that the events $\lambda_i \in \lambda$ be somewhat possible in the prior beliefs. We have proposed to use Dempster conditioning for this purpose but there are other tracks and other conditionings that address this problem such as the conditioning based on *adjustment* [Marchetti and Antonucci, 2018]. This latter is supposed to absorb even new inconsistent inputs (for instance in case of updating the probabilities of an event λ_i which could be impossible in the prior belief set in the case where $\underline{P}(\lambda_i)$=0). In addition to these open questions, other avenues of future work will concern computational complexity and study of extensions of Jeffrey's rule with Pearl method of virtual evidence [Pearl, 1988] in the context of imprecise belief graphical models [Cozman, 2000]. Another avenue will be extending Jeffrey's rule to the case where the probabilistic beliefs are compactly encoded by probabilistic constraints as in conditional logic programs.

Acknowledgements

This research has received support from the ANR CROQUIS project, grant ANR-21-CE23-0004 of the French research funding agency ANR.

References

[Antonucci *et al.*, 2007] Alessandro Antonucci, Alberto Piatti, and Marco Zaffalon. Credal networks for operational risk measurement and management. In *Knowledge-Based Intelligent Information and Engineering Systems*, volume 4693 of *LNCS*, pages 604–611. 2007.

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

[Berger *et al.*, 1994] JamesO. Berger, Elías Moreno, Luis-Raul Pericchi, M.Jesús Bayarri, JoséM. Bernardo, JuanA. Cano, Julián De la Horra, Jacinto Martín, David Ríos-Insúa, Bruno Betrò, A. Dasgupta, Paul Gustafson, and et all Wasserman. An overview of robust bayesian analysis. *Test*, 3(1):5–124, 1994.

[Bock *et al.*, 2014] Jasper De Bock, Cassio Polpo de Campos, and Alessandro Antonucci. Global sensitivity analysis for MAP inference in graphical models. In *Annual Conference on Neural Information Processing Systems, Dec. 8-13 2014, Montreal, Canada*, 2014.

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

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

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

[de Campos *et al.*, 1994] Luis M. de Campos, Juan F. Huete, and Serafin Moral. Probability intervals: A tool for uncertain reasoning. *Int. J. of Uncertainty, Fuzziness and Knowledge-Based Systems*, 02(02):167–196, 1994.

[Dubois and Prade, 1994] Didier Dubois and Henri Prade. A survey of belief revision and updating rules in various uncertainty models. *Int. J. of Intelligent Systems*, 9, 1994.

[Dubois and Prade, 1997] Didier Dubois and Henri Prade. 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, 1997.

[Grove and Halpern, 1998] Adam J. Grove and Joseph Y. Halpern. Updating sets of probabilities. In *UAI '98: Proceedings of UAI98, Madison, Wisconsin, USA, July 24-26, 1998*, pages 173–182, 1998.

[Jeffrey, 1965] R.C. Jeffrey. *The Logic of Decision*. Mc. Graw Hill, New York, 1965.

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

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

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

[Ma *et al.*, 2011] Jianbing Ma, Weiru Liu, Didier Dubois, and Henri Prade. Bridging jeffrey's rule, AGM revision and dempster conditioning in the theory of evidence. *Int. Jour. on Art. Intel. Tools*, 20(4):691–720, 2011.

[Marchetti and Antonucci, 2018] Sabina Marchetti and Alessandro Antonucci. Imaginary kinematics. In *Proceedings of UAI 2018, Monterey, California, USA, August 6-10, 2018*, pages 104–113, 2018.

[Moral and De Campos, 1991] Serafín Moral and Luis M. De Campos. Updating uncertain information. In *Uncertainty in Knowledge Bases*, pages 58–67, 1991.

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

[Pearl, 1988] Judea Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann Publishers Inc., San Francisco, USA, 1988.

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

[Skulj, 2006] Damjan Skulj. Jeffrey's conditioning rule in neighbourhood models. *International Journal of Approximate Reasoning*, 42(3):192 – 211, 2006.

[Smets, 1993] Philippe Smets. Jeffrey's rule of conditioning generalized to belief functions. UAI'93, page 500–505, San Francisco, CA, USA, 1993. Morgan Kaufmann.

[Suppes and Zanotti, 1977] Patrick Suppes and Mario Zanotti. On using random relations to generate upper and lower probabilities. *Synthese*, 36(4):427–440, 1977.

[Tabia, 2018] Karim Tabia. Analysis of jeffrey's rule of conditioning in an imprecise probabilistic setting. In *The Thirty-First International Flairs Conference*, 2018.

[Tang and Zheng, 2006] Yongchuan Tang and Jiacheng Zheng. Generalized jeffrey's rule of conditioning and evidence combining rule for a priori probabilistic knowledge in conditional evidence theory. *Information Sciences*, 176(11):1590 – 1606, 2006.

[Wagner, 1992] Carl G. Wagner. Generalizing jeffrey conditionalization. In *Uncertainty in Artificial Intelligence*, pages 331–335. Morgan Kaufmann, 1992.

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

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

[Yue and Liu, 2008] Anbu Yue and Weiru Liu. Revising imprecise probabilistic beliefs in the framework of probabilistic logic programming. In *Proceedings of AAAI 2008, Chicago, USA, July 13-17, 2008*, pages 590–596, 2008.