Some Bounds for Conditional Lower Previsions

Renato Pelessoni
University of Trieste
renato.pelessoni@econ.units.it

Paolo Vicig
University of Trieste
paolo.vicig@econ.units.it

Abstract

In this paper we consider some bounds for lower previsions that are either coherent or centered convex. As for coherent conditional previsions, we adopt a structure-free version of Williams' coherence, which we compare with Williams' original version and with other coherence concepts. We then focus on bounds concerning the classical product and Bayes' rules. After discussing some implications of product rule bounds, we generalise a well-known lower bound, which is a (weak) version for coherent lower probabilities of Bayes' theorem, to the case of (centered) convex previsions. We obtain a family of bounds and show that one of them is undominated in all cases.

Keywords. Conditional lower previsions, product rule, Bayes' theorem, Williams' coherence, centered convex previsions.

1 Introduction

Quite recently, P.M. Williams' 1975 seminal paper *Notes on conditional previsions* was published in a slightly revised version [21], preceded by an introductory paper discussing basic aspects and historical motivations for his work [14]. This fact confirms that Williams' ideas on coherence still play a very important role in the theory of conditional imprecise previsions.

One of the aims of this paper is to show that Williams' coherence, while being more general than other coherence concepts that have been developed, may be quite simple to work with in several problems. Precisely, we shall use a variant of Williams' original coherence which does not impose any structural constraint on the set of conditional (bounded) random variables forming the domain of the lower prevision $\underline{P}$ and which is a generalisation of Walley's coherence for unconditional (bounded) random variables (or gambles) [16].

After recalling some preliminary notions in Section 2, we discuss this variant in Section 3, comparing it firstly with Williams' original version and then with other generalisations of Walley's unconditional coherence, either potential or proposed in [16]. When being equivalent to the notion of coherence mainly adopted by Walley in [16], as is the case in the sequel of the paper, Williams' coherence may be conveniently used to prove certain results, which therefore hold in Walley's approach too.

We shall use Williams' coherence to study some bounds for conditional lower previsions. Actually we prove that several results hold also for previsions that are (centered) convex, i.e. satisfy a consistency notion (introduced in [10]) which is more general than Williams' coherence.

We focus on generalisations of product rule and Bayes' rule bounds together with other bounds which we termed sign rules. A motivation for investigating all these bounds is that they may give us some guidance for extending coherent or convex lower previsions. This is particularly relevant when conditioning, given that many rules or standard procedures for inferences or anyway for getting unconditional coherent evaluations do not apply in a conditional framework (for instance, convex combinations of coherent conditional lower previsions are not necessarily coherent).

In Section 4 we discuss some inequalities (product and sign rules), which are essentially known, exploring some of the implications they have for extending $\underline{P}$ under an epistemic irrelevance assumption. It appears here that when the product rule may hold with equality, the lower prevision obtained from this equality is not necessarily the natural extension as in the case of events, but may also coincide with the opposite concept of upper extension. In Section 5 we generalise the well-known lower bound $\underline{P}(A|B) \geq \frac{\underline{P}(A \wedge B)}{\underline{P}(A \wedge B)+\overline{P}(\overline{A} \wedge B)}$ to the case of conditional random variables and of lower previsions that are Williams' coherent or, more generally, centered convex. We derive a family of bounds, proving that one of them, given by equation (11), is the best in all cases.

Section 6 contains some further comments and conclusions.

2 Preliminaries

In the sequel, $\mathcal{D}$ is an *arbitrary* (non-empty) set of *bounded* random variables (also termed gambles [16] or random quantities [21]), or more generally of bounded conditional random variables.

In the conditional case, if $X|B \in \mathcal{D}$, X is a random variable and B a non-impossible event. When $B = \Omega$, we obtain the (unconditional) random variable $X = X|\Omega$.

The supremum $\sup(X|B)$ of $X|B$ may be computed as $\sup_{\omega \Rightarrow B} X(\omega)$ ($\sup_{\omega \in B} X(\omega)$ in the set-theoretic interpretation of events), where all ω belong to a large enough partition (possibility space) $I\!P$. It will be also denoted as $\sup_B X$. Analogously, $\inf(X|B) = \inf_B X$.

We write B for both an event B and its indicator function $|B|$ (de Finetti's convention), appearing from the context which of the two meanings is intended.

A *lower prevision* $\underline{P}$ on $\mathcal{D}$ is a map $\underline{P} : \mathcal{D} \rightarrow \mathbb{R}$. An upper prevision $\overline{P}$ may be defined through the equality $\overline{P}(-X) = -\underline{P}(X)$, which always lets us refer to either lower or upper previsions only. A *precise* prevision P is the special case $\overline{P}(X) = \underline{P}(X) = P(X)$.

The consistency notions we shall consider for $\underline{P}$ are those of *coherence* or *(centered) convexity*. More specifically, when $\mathcal{D}$ is made of unconditional random variables, $\underline{P}$ is said to be *coherent* when satisfying the definition in [16], sec. 2.5.4 (a):

Definition 1 *$\underline{P} : \mathcal{D} \rightarrow \mathbb{R}$ is a* coherent lower prevision *on $\mathcal{D}$ iff, for all $n \in \mathbb{N}^+$, $\forall\ X_0, X_1, \ldots, X_n \in \mathcal{D}$, $\forall\ s_0, s_1, \ldots, s_n$ real and non-negative, defining $\underline{G} = \sum_{i=1}^{n} s_i(X_i - \underline{P}(X_i)) - s_0(X_0 - \underline{P}(X_0))$, $\sup \underline{G} \geq 0$.*

This definition has a well-known behavioural interpretation: $\underline{P}(X)$ is an agent's supremum buying price for X, and $\underline{G}$ is the agent's *gain* resulting from her/his buying $s_i X_i$, for $i = 1, \ldots, n$, and selling $s_0 X_0$. We shall use this terminology too, saying that the agent *bets* on $X_0, \ldots, X_n$ with *stakes* $s_0, \ldots, s_n$ respectively.

In a conditional environment, we adopt the following generalisation of Definition 1 to define a *coherent* $\underline{P}(\cdot|\cdot)$:

Definition 2 *$\underline{P} : \mathcal{D} \rightarrow \mathbb{R}$ is a* coherent conditional lower prevision *on $\mathcal{D}$ iff, for all $n \in \mathbb{N}^+$, $\forall X_0|B_0, \ldots, X_n|B_n \in \mathcal{D}$, $\forall\ s_0, s_1, \ldots, s_n$ real and non-negative, defining $B = \bigvee_{i=0}^{n} B_i$ and $\underline{G} = \sum_{i=1}^{n} s_i B_i(X_i - \underline{P}(X_i|B_i)) - s_0 B_0(X_0 - \underline{P}(X_0|B_0))$, $\sup(\underline{G}|B) \geq 0$.*

Here the gain is $\underline{G}|B$, a conditional random variable itself. Conditioning on B has the meaning of considering only those values for $\underline{G}$ when at least one of $B_0, \ldots, B_n$ is true. It is easy to realise that we would get an equivalent definition (adopted in [18]) by replacing $\underline{G}|B$ with $\underline{G}|S$, where the *support* S is defined as $S = \vee\{B_i : s_i \neq 0, i = 0, \ldots, n\}$.

Throughout the paper, Definition 2 will be referred to as Williams' coherence, or *W-coherence* or simply coherence, but as we will explain in Section 3, it is actually a structure-free version of the original Williams' coherence.

A weaker notion than W-coherence is that of lower prevision that *avoids uniform loss* [16, 18]. It may be obtained from Definition 2 by ruling out the bet on $X_0|B_0$ and modifying B and $\underline{G}$ accordingly. In the unconditional environment it is termed condition of *avoiding sure loss* and is defined in [16], Sec. 2.4.4 a).

The consistency notion of *centered convexity* [10, 11] is weaker than coherence, but sufficiently stronger than the conditions of avoiding sure or uniform loss to allow for interesting properties and applications (for instance, in risk measurement [10]). In fact, several of the results in the next sections apply to centered convex previsions too.

Formally, the definition of *convex lower prevision* is obtained from Definition 1 and Definition 2 by introducing just the extra *convexity constraint* $\sum_{i=1}^{n} s_i = s_0$ (> 0) and eventually by further imposing (this is not restrictive) that $s_0 = 1$ [9, 10]. Again, we could condition $\underline{G}$ on its support S rather than on B, getting an equivalent definition of convex conditional lower prevision. This is done in [10, 11]. *Centered* convexity requires in addition that ($0 \in \mathcal{D}$ and) $\underline{P}(0) = 0$ in the unconditional case, and further that $\forall X|B \in \mathcal{D}$, $0|B \in \mathcal{D}$ and $\underline{P}(0|B) = 0$ in the conditional case.

Centering is quite a natural requirement: non-centered convex previsions have rather weak consistency properties, but special instances of them may be found in the risk literature (cf. [10]).

Proposition 1 *If $\underline{P} : \mathcal{D} \rightarrow \mathbb{R}$ is centered convex, then necessarily [10]:*

P1) $\inf X \leq \underline{P}(X) \leq \sup X$ *(internality);*

P2) $Y \leq X \Rightarrow \underline{P}(Y) \leq \underline{P}(X)$, $\forall X, Y \in \mathcal{D}$ *(monotonicity);*

P3) $\underline{P}(\lambda X + (1-\lambda)Y) \geq \lambda \underline{P}(X) + (1-\lambda)\underline{P}(Y)$, $\forall X, Y \in \mathcal{D}$, $\forall \lambda \in [0, 1]$.

These properties obviously hold for coherent lower previsions too, while P1) might fail for non-centered convex previsions.

Let $\underline{P}$ be a lower prevision defined on an arbitrary set $\mathcal{D}$. Any consistency condition satisfied by $\underline{P}$ should guarantee that there exists an extension of $\underline{P}$ on any $\mathcal{D}' \supset \mathcal{D}$ which satisfies the same consistency condition. If such an extension is not unique, its vaguest or least-committal one, if existing, has a special importance. This peculiar extension is the *natural extension* $\underline{E}$ in the case of coherent or, when conditioning, W-coherent previsions [14, 16, 21], the *convex* natural extension $\underline{E}_c$ for centered convex (unconditional or conditional) previsions [9, 10]. The natural or convex natural extensions always exist for these consistency notions, not necessarily with other ones, like Walley-coherence in [16], Section 7.1.4 (b), or non-centered convexity. Hence, the consistency notions we shall be working with always allow for extensions of the same kind on any superset: we shall often use this fact in the proofs of the results, without always mentioning explicitly that we are performing an extension.

When working with conditional random variables, like $\underline{G}|B$, we shall employ the equality

$$f(X_1, \ldots, X_n)|B = f(X_1|B, \ldots, X_n|B) \quad (1)$$

where f is any real function [3].

3 Two or Three Things on Williams' Coherence

3.1 About Williams' definition

Williams' original definition ([21], Definition 1) differs formally from our definition of W-coherence. One reason is that it refers to upper rather than lower previsions, but this is unimportant, since using the conjugacy relation $\overline{P}(-X) = -\underline{P}(X)$ our condition $\sup \underline{G}|B \geq 0$ corresponds exactly to his inequality in (A^*) of [21]. The true difference is that his notion is not completely structure-free, as it asks that for every $X|B$ in $\mathcal{D}$, $\underline{P}(X|B)$ is assigned for any X in a linear space $\mathcal{X}_B$. It follows for instance that Williams' definition does not formally generalise Walley's coherence for unconditional previsions (our Definition 1), which is structure-free: when $B = \Omega$ for all $X|B \in \mathcal{D}$, the set of all X is constrained to form a linear space $\mathcal{X}_\Omega$. On the contrary, Definition 2 is in particular a generalisation of Walley's unconditional coherence and appears to be, in general, nimbler. For instance, the bounds in Section 4 involve just a few random variables and no structure is actually needed for proving them. The fundamental link between the two versions of Williams' coherence is ensured by the following *extension theorem.*

Proposition 2 *If $\underline{P} : \mathcal{D} \to \mathbb{R}$ is W-coherent on $\mathcal{D}$ (according to Definition 2), it has a W-coherent extension on* any $\mathcal{D}' \supset \mathcal{D}$.

Although we are not aware of any published proof for this proposition, nevertheless it should be regarded as essentially known. In fact, it can be proven by adapting the proofs concerning the convex natural extension in [10], thus proving that there *always* exists the natural extension of a W-coherent lower prevision on any $\mathcal{D}' \supset \mathcal{D}$. Alternatively, the scheme of de Finetti's extension theorem can be followed, with suitable (but basically minor) modifications. After de-Finetti's path-breaking proof concerning precise (unconditional) previsions in [6], this scheme was employed in several generalisations (see e.g. [1, 4]). Its two-step proof shows in the first step that there exist W-coherent extensions on $\mathcal{D}' = \mathcal{D} \cup \{X|B\}$, $\forall X|B$, while the second step generalises the proof to any $\mathcal{D}'$ using Zorn's lemma or equivalent results. A by-product of the first step is that the set of admissible W-coherent extensions on $X|B$ is proved to be a closed interval. Its lower endpoint is the *natural extension* $\underline{E}(X|B)$, while the upper endpoint is the *upper extension* $\underline{U}(X|B)$ of $\underline{P}$. We shall meet again upper extensions in Section 4.

As an important implication of Proposition 2 in our framework, when $\mathcal{D}$ in Definition 2 does not meet Williams' structure requirements in his definition it is always possible to coherently extend $\underline{P}$ on a set $\mathcal{D}'$ such that these requirements hold, and there the two notions of coherence coincide. It follows that our W-coherent lower previsions have all the properties established for Williams' coherence in [21], including the important *envelope theorem*, stating that $\underline{P}$ is coherent on $\mathcal{D}$ if and only if

$$\underline{P}(X|B) = \inf_{P \in \mathcal{M}} P(X|B), \forall X|B \in \mathcal{D}$$

where $\mathcal{M}$ is the set of the coherent precise previsions $P(\cdot|\cdot)$ dominating $\underline{P}(\cdot|\cdot)$ on $\mathcal{D}$ ($P(X|B) \geq \underline{P}(X|B), \forall X|B \in \mathcal{D}$).

3.2 Alternative concepts of coherence

Another issue concerning Definition 2 of W-coherence is: equivalent formulations of Definition 1 are known, so why not rather generalise them in a conditional environment? An answer is that Definition 2 seems more appropriate for further generalisations. In fact, an equivalent version of coherence in Definition 1 is obtained by restricting the stakes $s_0, \ldots, s_n$ to be integer (this is Walley's Definition 2.5.1 in [16]), and this can be done in a conditional environment too. However, considering integer combinations only is not sufficient when the random numbers are unbounded, even in the unconditional case, as shown in [12].

Another definition, less used [1] but equivalent to Definition 1, is:

Definition 3 *$\underline{P}: \mathcal{D} \rightarrow \mathbb{R}$ is a coherent lower prevision on $\mathcal{D}$ iff, for all $n \in \mathbb{N}^+$, $\forall X_0, X_1, \ldots, X_n \in \mathcal{D}$, $\forall s_1, \ldots, s_n \geq 0$, $\forall \lambda_0 \in \mathbb{R}$ such that $X_0 \geq \sum_{i=1}^n s_i X_i + \lambda_0$, it holds that $\underline{P}(X_0) \geq \sum_{i=1}^n s_i \underline{P}(X_i) + \lambda_0$.*

To the best of our knowledge, no generalisation of this definition to a conditional environment is available in the literature, nor does the problem of generalising it to an equivalent version of W-coherence seem to have a straightforward solution.

A further issue is that a number of different generalisations of coherence (Definition 1 or equivalent) to a conditional framework have been proposed in [16]: how do they relate to W-coherence? We observe some basic facts about it.

a) The generalisations in [16] are not structure-free: the conditioning events have some special features. *When being comparable*, W-coherence in Definition 2 is *equivalent* to the following two of them:
   - the concept of *coherence* defined in Sec. 7.1.4 (b) (referred to as *Walley-coherence* here), with the extra assumption that all partitions $\mathcal{B}_i$ in that definition are finite (this equivalence is stated (without proof) in [16]);
   - the concept of *separate coherence* defined in Sec. 6.2.2, *without* any other extra assumption (this equivalence is proved in the Appendix).

b) In general, W-coherence may be weaker than Walley-coherence (when at least one $\mathcal{B}_i$ is infinite). This fact may lead to the disadvantages discussed in [16], but has the non-negligible advantages over Walley-coherence that the natural extension always exists and that the envelope theorem characterises W-coherence. A weaker notion than Walley-coherence, *weak coherence* defined in Sec. 7.1.4 (a), is sometimes stronger and sometimes weaker than W-coherence. This notion is anyway rather counterintuitive, and Walley-coherence is in fact the major conditional coherence condition in [16].

c) At any rate, properties of W-coherence involving only finitely many distinct conditioning events hold for Walley-coherence too (a W-coherent assessment or possibly one of its W-coherent extensions, cf. Proposition 2, may be referred in this case to a finite set of finite partitions $\mathcal{B}_i$). In particular, the bounds we investigate later on hold in Walley's framework too.

Last but not least, we note that the notion of conditional random variable (and of conditional event) is often left at an informal level in the literature, including [16, 21]. A formal approach to these and other descriptive tools of uncertainty is developed in [3, 4].

Although this issue is seemingly not particularly relevant in many matters, a greater formalisation turns out to be useful with other ones. For an example, consider Lemma 6.2.4 in [16]: this lemma states that, if $BX = BY$ and other coherence conditions hold for a lower prevision $\underline{P}$, then $\underline{P}(X|B) = \underline{P}(Y|B)$. But using (1), $BX|B = (B|B) \cdot (X|B) = X|B$, thus condition $BX = BY$ alone implies $X|B = Y|B$. Consequently $\mu(X|B) = \mu(Y|B)$ whatever the uncertainty measure μ is, not because of coherence (μ could even be incoherent), but merely because we are evaluating the same thing.

4 Product and Sign Rules

The *product rule* is among the basic inferential rules in Bayesian statistics. In its simplest version for probabilities, it requires that $P(A \wedge B) = P(A) \cdot P(B|A)$; in a more general version involving a precise prevision P, events A and B and a random variable X, we have $P(AX|B) = P(A|B) \cdot P(X|A \wedge B)$.

We investigate now some generalisations of this rule, and related properties, for coherent lower previsions.

Proposition 3 *Let $\underline{P}$ be coherent on $\mathcal{D} \supset \{AX|B, A|B, X|A \wedge B\}$. Then, necessarily:*

a) (product rule) *if $\underline{P}(X|A \wedge B) > 0$, then*

$$\underline{P}(AX|B) \geq \underline{P}(A|B) \cdot \underline{P}(X|A \wedge B) \tag{2}$$

b) (product rule) *if $\underline{P}(X|A \wedge B) < 0$, then*

$$\underline{P}(AX|B) \leq \underline{P}(A|B) \cdot \underline{P}(X|A \wedge B) \tag{3}$$

c) $\underline{P}(AX|B) = 0$ *iff* $\underline{P}(A|B) \cdot \underline{P}(X|A \wedge B) = 0$

[1] Definition 3 has a curios story: not mentioned explicitly in Walley's book [16], although following directly from results established there, it appears in [2], but without being related to coherence for imprecise previsions. It was then discussed extensively in [7].

d) (sign rules)

$\underline{P}(AX|B) > 0 \Rightarrow \underline{P}(X|A \wedge B) > 0$;

$\underline{P}(AX|B) < 0 \Rightarrow \underline{P}(X|A \wedge B) < 0$;

Proof. Put $p_1 = \underline{P}(AX|B)$, $p_2 = \underline{P}(A|B)$, $p_3 = \underline{P}(X|A \wedge B)$, and consider a gain $\underline{G}$ in Definition 2 arising from betting on $AX|B$, $A|B$, $X|A \wedge B$: $\underline{G} = s_1B(AX - p_1) + s_2B(A - p_2) + s_3AB(X - p_3) = ((s_1 + s_3)AX + (s_2 - s_3p_3)A - s_1p_1 - s_2p_2)B$. Now choose s_1, s_2, s_3 such that

$$s_1 = -s_3, s_2 = s_3p_3 \tag{4}$$

and $\underline{G}$ specialises into

$$\underline{G} = (p_1 - p_2p_3)s_3B. \tag{5}$$

Proof of a). We have $p_3 > 0$. Choose $s_3 > 0$. Then from (4) $s_1 < 0$, $s_2 > 0$. Since only one of the stakes s_1, s_2, s_3 is negative, we have an admissible bet according to Definition 2. To ensure $\sup \underline{G}|B \geq 0$ it is necessary from (5) that $p_1 - p_2p_3 \geq 0$, which is (2).

Proof of b). Analogous to a), after choosing $s_3 < 0$.

Proof of c). To prove the implication $\underline{P}(X|A \wedge B) \cdot \underline{P}(A|B) = 0 \Rightarrow \underline{P}(AX|B) = 0$, note that when $\underline{P}(X|A \wedge B) \cdot \underline{P}(A|B) = p_2p_3 = 0$ the gain $\underline{G}$ in (5) reduces to $\underline{G} = s_3p_1B$, and $\underline{G}|B = s_3p_1$. To ensure $\sup \underline{G}|B \geq 0$, whatever the sign of s_3 may be, it is necessary that $p_1 = \underline{P}(AX|B) = 0$. The proof of the converse implication is similar.

Proof of d). For the first implication, suppose $\underline{P}(AX|B) > 0$. Then $\underline{P}(X|A \wedge B)$ can be neither negative (since then b) would contradictorily imply $\underline{P}(AX|B) \leq 0$), nor zero (c) would imply $\underline{P}(AX|B) = 0$). Hence $\underline{P}(X|A \wedge B) > 0$. The other implication is proven similarly. ■

4.1 Comments

The sign rules are obtained here from the product rule. A simpler version of the first rule holds for convex previsions too, and may be derived from Lemma 1 (cf. Section 5.1). Sign rules introduce some rough inferential constraints. For instance, let $B = \Omega$. Then knowing or assuming that $\underline{P}(AX) > 0$ implies necessarily $\underline{P}(X|A) > 0$ (no matter what sign $\underline{P}(X)$ has).

The product rule has interesting implications, involving the natural and upper extension. To outline this point, let $B = \Omega$, and suppose that A is *epistemically irrelevant* for X, so $\underline{P}(X|A) = \underline{P}(X)$. If we have assessed $\underline{P}(A)$ and $\underline{P}(X)$, but not $\underline{P}(AX)$, it is tempting to extend $\underline{P}$ on AX putting $\underline{P}(AX) = \underline{P}(A) \cdot \underline{P}(X)$ (multiplicative rule). There are instances when this is possible: if X is an event too, under an additional assumption (logical independence of A and X); in this case $\underline{P}(AX)$ is the natural extension $\underline{E}(AX)$ [13]. These properties do not necessarily hold if we further introduce some constraints on $\underline{P}$. For instance, it was shown in [8] that the multiplicative rule holds only in very special cases if we require $\underline{P}$ to be a necessity measure. However, as long as only events are involved, we can hope to simultaneously apply $\underline{P}(AX) = \underline{P}(A) \cdot \underline{P}(X)$ and obtain the natural extension $\underline{E}(AX) = \underline{P}(AX)$. Proposition 3 informs us that in the realm of random variables the situation is more complex: even assuming that $\underline{P}(AX) = \underline{P}(A) \cdot \underline{P}(X)$ is a coherent extension of $\underline{P}$ on AX, there are instances (cf. a)) when $\underline{P}(AX) = \underline{E}(AX)$, but other conditions (cf. b)) imply that $\underline{P}(AX)$ is just the opposite, i.e. the upper extension of $\underline{P}$.[2] This happens in particular when $\sup X < 0$ (hence $\underline{P}(X) < 0$ by P1) of Proposition 1), or also $X \leq 0$ if $\underline{P}(X) \neq 0$.

Finally, note that some sign constraints arise as a joint consequence of *a*), *b*), *c*), depending on the sign of $\underline{P}(A|B)$: if $\underline{P}(A|B) > 0$ then $\underline{P}(AX|B)$ and $\underline{P}(X|A \wedge B)$ must take the same sign (both positive, both negative, or both null), while if $\underline{P}(A|B) = 0$ then $\underline{P}(AX|B) = 0$, but $\underline{P}(X|A \wedge B)$ is unconstrained.

5 Bayes' Rule Bounds for Centered Convex Previsions

The following inequality, which holds if $\underline{P}(B) > 0$ and its terms are well-defined, is well-known in the theory of coherent imprecise probabilities [15, 16, 17]:

$$\underline{P}(A|B) \geq \frac{\underline{P}(A \wedge B)}{\underline{P}(A \wedge B) + \overline{P}(\overline{A} \wedge B)} \tag{6}$$

Together with an analogous bound, eq. (6) generalises Bayes' theorem for precise probabilities (when $\underline{P} = \overline{P} = P$ it reduces to $P(A|B) \geq P(A \wedge B)/P(B)$. The reverse inequality may be obtained from $\overline{P}(A|B) \leq \overline{P}(A \wedge B)/(\overline{P}(A \wedge B) + \underline{P}(\overline{A} \wedge B))$). In fact, an immediate, inferential way of interpreting (6) is to suppose that an unconditional coherent $\underline{P}$ is assigned: then (6) gives a lower bound for extending $\underline{P}$ on $A|B$, hence also a lower bound for its natural extension $\underline{E}(A|B)$. It is well-known [15] that when $\underline{P}$ is defined on an algebra $\mathcal{A}$ and is 2-monotone there (i.e. $\underline{P}(A \vee B) \geq \underline{P}(A) + \underline{P}(B) - \underline{P}(A \wedge B), \forall A, B \in \mathcal{A}$), the bound in (6) is precisely equal to $\underline{E}(A|B)$, which under these assumptions may be written in terms of Choquet integrals. Inequality (6) was also studied in various other papers, including [17], where it is also compared with Dempster's rule of conditioning, and [19].

[2] Upper extensions received little attention in [16], while they were investigated in [20].

Our main purpose in this section is to generalise eq. (6) introducing a more general bound, holding for random variables with corresponding lower previsions that are centered convex. For this, we need a preliminary Lemma, which has also other implications, commented below.

Lemma 1 *Let $\underline{P} : \mathcal{D} \to \mathbb{R}$. Whenever the lower previsions below are defined,*

a) if $\underline{P}$ is convex on $\mathcal{D}$, then for $\lambda \in \mathbb{R}$

$$\underline{P}(B(X-\lambda)) > 0 \Rightarrow \underline{P}(X|B) > \lambda;$$
$$\underline{P}(X|B) > \lambda \Rightarrow \underline{P}(B(X-\lambda)) \geq 0;$$

b) if $\underline{P}$ avoids uniform loss on $\mathcal{D}$, then for $\lambda \in \mathbb{R}$,

$$\underline{P}(B(\lambda - X)) > 0 \Rightarrow \underline{P}(X|B) < \lambda.$$

Proof. To prove a), write the gain $\underline{G}|\Omega = \underline{G}$ for a bet on $B(X-\lambda)$, $X|B$ with stakes $s_1 = s_0$ (note that $s_1 = s_0$ is the convexity condition in this case): $\underline{G} = s_1(B(X-\lambda) - \underline{P}(B(X-\lambda))) - s_1 B(X - \underline{P}(X|B)) = s_1(B(\underline{P}(X|B) - \lambda) - \underline{P}(B(X-\lambda)))$.

To prove the first inequality, put $s_1 = 1$. To ensure $\sup \underline{G} \geq 0$ (note that $\underline{G}$ varies only with B), the following inequality must be false for at least one value of B: $B(\underline{P}(X|B) - \lambda) < \underline{P}(B(X-\lambda))$. If $\underline{P}(B(X-\lambda)) > 0$, then necessarily $\underline{P}(X|B) - \lambda > 0$.

To prove the second inequality, put $s_1 = -1$. To guarantee now that $\sup \underline{G} \geq 0$, the reversed inequality $B(P(X|B) - \lambda) > \underline{P}(B(X-\lambda))$ must be false. If $\underline{P}(X|B) > \lambda$, it is necessary for this that $\underline{P}(B(X-\lambda)) \geq 0$.

To prove b), consider the bet on $B(\lambda - X)$, $X|B$ with gain $\underline{G} = B(\lambda - X) - \underline{P}(B(\lambda - X)) + B(X - \underline{P}(X|B))$, and argue similarly to the preceding cases. ■

Corollary 1 *Under the assumptions of Lemma 1, b), $\overline{P}(B(X-\lambda)) < 0 \Rightarrow \underline{P}(X|B) < \lambda$.*

Proof. Follows from Lemma 1, b) and $\underline{P}(B(\lambda - X)) = -\overline{P}(B(X-\lambda))$. ■

5.1 Comments

Only the first inequality in a) will be actually used to generalise (6), but the three inequalities deserve some comments. The inequalities in a) imply when $\lambda = 0$ a simpler version of the first inequality in Proposition 3 d) (sign rules), but holding under the weaker assumption that $\underline{P}$ is convex. As for the inequality in b), it holds also for centered convex previsions, since these previsions avoid uniform loss [10].

5.2 A Generalised Lower Bound

We obtain now a generalisation of the lower bound (6).

Proposition 4 *Let $\underline{P}$ be an unconditional centered convex lower prevision on $\mathcal{D} \supset \{B, B(X - \sup(X|B)), B(X - \inf(X|B))\}$ and $\underline{P}(B) > 0$. If $\underline{P}(B(X - \inf(X|B))) - \underline{P}(B(X - \sup(X|B))) \neq 0$, any (centered) convex extension of $\underline{P}$ on $X|B$ is such that, $\forall h \leq \inf(X|B)$, $\forall k \geq \sup(X|B)$,*

$$\underline{P}(X|B) \geq \phi(h,k) = \frac{k\underline{P}(B(X-h)) - h\underline{P}(B(X-k))}{\underline{P}(B(X-h)) - \underline{P}(B(X-k))} \quad (7)$$

Proof. We preliminarily observe that the denominator in (7) is positive. This follows from the assumptions and internality and monotonicity of $\underline{P}$ (Proposition 1, P1) and P2)), which imply: $B(X-h) \geq 0 \Rightarrow \underline{P}(B(X-h)) \geq 0$, $B(X-k) \leq 0 \Rightarrow \underline{P}(B(X-k)) \leq 0$, and then $0 < \underline{P}(B(X - \inf(X|B))) - \underline{P}(B(X - \sup(X|B))) \leq \underline{P}(B(X-h)) - \underline{P}(B(X-k))$.

To start now the proof, note that for any $\lambda \in \mathbb{R}$, $B(X - ((1-\lambda)h + \lambda k)) = (1-\lambda)B(X-h) + \lambda B(X-k)$. From this equality, we get for any $\lambda \in [0,1]$ (use P3) of Proposition 1) $\underline{P}(B(X - ((1-\lambda)h + \lambda k))) = \underline{P}((1-\lambda)B(X-h) + \lambda B(X-k)) \geq (1-\lambda)\underline{P}(B(X-h)) + \lambda\underline{P}(B(X-k)) = \underline{P}(B(X-h)) - \lambda(\underline{P}(B(X-h)) - \underline{P}(B(X-k)))$. Defining $\overline{\lambda} = \frac{\underline{P}(B(X-h))}{\underline{P}(B(X-h)) - \underline{P}(B(X-k))}$, $\overline{\lambda} \in [0,1]$. We can therefore replace λ with $\overline{\lambda}$ in the above derivation, getting $\underline{P}(B(X - ((1-\overline{\lambda})h + \overline{\lambda}k)) \geq \underline{P}(B(X-h)) - \overline{\lambda}[\underline{P}(B(X-h)) - \underline{P}(B(X-k))] = 0$.

If $\underline{P}(B(X - ((1-\overline{\lambda})h + \overline{\lambda}k))) > 0$, use Lemma 1, a) to obtain $\underline{P}(X|B) > (1-\overline{\lambda})h + \overline{\lambda}k = \phi(h,k)$.

If $\underline{P}(B(X - ((1-\overline{\lambda})h + \overline{\lambda}k))) = 0$, then $\underline{P}(X|B) = (1-\overline{\lambda})h + \overline{\lambda}k = \phi(h,k)$. We apply here Proposition 9 in [10], which generalises to convex lower previsions a result known for coherent lower previsions [16], ensuring that $r = \underline{P}(X|B)$ is the unique solution of $\underline{P}(B(X-r)) = 0$, if $\underline{P}$ is convex and $\underline{P}(B) > 0$. ■

Notation. When unambiguous we write $S_B = \sup(X|B)$, $I_B = \inf(X|B)$.

Remark 1 *When $\underline{P}$ is coherent, the assumptions in Proposition 4 ensuring that the denominators are non-zero simplify as follows: it is sufficient to ask that*

i) $X|B$ is non-constant;

ii) $\underline{P}(B) > 0$.

In fact, i) and ii) imply $\underline{P}(B(X - I_B)) - \underline{P}(B(X - S_B)) > 0$. To see this, consider a bet on B, $B(X - S_B)$, $B(X - I_B)$ with stakes $S_B - I_B$, 1, -1 respectively. Then $\underline{G} = (S_B - I_B)(B - \underline{P}(B)) + B(X -$

$S_B) - \underline{P}(B(X - S_B)) - B(X - I_B) + \underline{P}(B(X - I_B)) = \underline{P}(B(X - I_B)) - \underline{P}(B(X - S_B)) - (S_B - I_B)\underline{P}(B) = \sup \underline{G}$. *Thus* $\sup \underline{G} \geq 0$ *iff* $\underline{P}(B(X - I_B)) - \underline{P}(B(X - S_B)) \geq (S_B - I_B)\underline{P}(B) > 0$.

As a further remark, note that $\underline{P}(B) = 0$ *(*$\underline{P}$ *coherent) implies* $\underline{P}(B(X - I_B)) - \underline{P}(B(X - S_B)) = 0$*, by Proposition 3 c).*

The lower bound (7) is as a matter of fact a family of lower bounds, indexed on h and k. The immediate question is therefore: which h, k should be chosen? It is not clear a priori that there should be a unique couple (h, k) preferable in all cases, but the following proposition solves the problem in favour of the remarkable couple $h = \inf(X|B)$, $k = \sup(X|B)$.

Proposition 5 *Under the assumptions of Proposition 4,* $\phi(I_B, S_B) \geq \phi(h, k)$, $\forall h \leq I_B$, $\forall k \geq S_B$.

Proof. The proof is made up of two steps. In the first step we prove that for any fixed $h \leq I_B$, $\phi(h, k) \leq \phi(h, S_B)$; in the second that $\phi(h, S_B) \leq \phi(I_B, S_B)$.

To shorten notation, we define $f(r) = \underline{P}(B(X-r))$, so that for instance $f(h) = \underline{P}(B(X - h))$ and $\phi(h, k) = \frac{kf(h)-hf(k)}{f(h)-f(k)}$.

First step. Fix h and define $\delta = \delta(k) = k - h$. We have $\delta \geq S_B - I_B > 0$ (the last inequality is implied by the assumption $\underline{P}(B(X-I_B)) - \underline{P}(B(X-S_B)) \neq 0$ in Proposition 4, which rules out the trivial case that $X|B$ is constant).

We write now $\phi(h, k)$ as a function $u(\delta)$ of δ: $u(\delta) = \frac{(h+\delta)f(h)-hf(h+\delta)}{f(h)-f(h+\delta)}$, or also

$$u(\delta) = \phi(h, h + \delta) = h + \delta \frac{f(h)}{f(h) - f(h + \delta)}. \quad (8)$$

We now consider the function of δ in (8), $\delta/[f(h) - f(h + \delta)]$, proving that:

$$\delta_1 > \delta_2 \ (> 0) \Rightarrow \frac{\delta_1}{f(h)-f(h+\delta_1)} \leq \frac{\delta_2}{f(h)-f(h+\delta_2)}. \quad (9)$$

To prove (9), we first verify that $f(r)$ is concave on $\mathbb{R}$. In fact, for $\lambda \in [0, 1]$ and using also P3) of Proposition 1, $f(\lambda r_1 + (1 - \lambda)r_2) = \underline{P}(B(X - \lambda r_1 - (1 - \lambda)r_2)) = \underline{P}(\lambda B(X - r_1) + (1 - \lambda)B(X - r_2)) \geq \lambda \underline{P}(B(X - r_1)) + (1 - \lambda)\underline{P}(B(X - r_2)) = \lambda f(r_1) + (1 - \lambda)f(r_2)$.

For a standard property of concave real functions, $F(\delta) = \frac{f(h+\delta)-f(h)}{\delta}$ is monotone non-increasing for $\delta \in \mathbb{R}$, hence in particular for $\delta \in I = [S_B - I_B, +\infty[$. Interval I is the domain of δ in our case; here $\delta > 0$ and (cf. the beginning of the proof of Proposition 4) $f(h+\delta) - f(h)$ is negative, thus $F(\delta) < 0$, $\forall \delta \in I$. Recalling this, we easily get (9) from $\delta_1 > \delta_2 \Rightarrow F(\delta_1) \leq F(\delta_2)$.

Using (9), and recalling that $f(h) \geq 0$, $u(\delta)$ is maximised, for a given h, by minimising δ, putting hence $\delta = S_B - h$. This is equivalent to choosing $k = S_B$ in $\phi(h, k)$. Thus $\phi(h, k) \leq \phi(h, S_B)$, $\forall k \geq S_B$.

Second step. Define $\delta = \delta(h) = h - S_B < 0$ and write $\phi(h, S_B)$ as a function $v(\delta)$ of δ:

$$v(\delta) = \phi(S_B + \delta, S_B) = S_B - \delta \frac{f(S_B)}{f(S_B + \delta) - f(S_B)}.$$

We prove now that

$$\delta_1 < \delta_2 \ (< 0) \Rightarrow v(\delta_1) \leq v(\delta_2). \quad (10)$$

For this, we can follow a scheme similar to the proof of the first step (alternatively, a longer proof essentially exploiting the definition of convex prevision is possible). As before, the function $F(\delta) = \frac{f(S_B+\delta)-f(S_B)}{\delta}$ is monotone non-increasing, and negative for $\delta \in]-\infty, I_B - S_B]$. From this and recalling that $f(S_B) \leq 0$, (10) follows straightforwardly.

We conclude that $\phi(h, k) \leq \phi(h, S_B) \leq \phi(I_B, S_B)$, $\forall h \leq I_B$, $\forall k \geq S_B$, where the first inequality follows from step 1, whilst the second is a consequence of step 2. ■

The most notable consequence of Proposition 5 is that we get the following lower bound for $\underline{P}(X|B)$:

$$\underline{P}(X|B) \geq \frac{S_B\underline{P}(B(X-I_B))-I_B\underline{P}(B(X-S_B))}{\underline{P}(B(X-I_B))-\underline{P}(B(X-S_B))}. \quad (11)$$

When X is an event, $X = A$, (11) reduces to

$$\underline{P}(A|B) \geq \frac{\underline{P}(A \wedge B)}{\underline{P}(A \wedge B) - \underline{P}(B(A - 1))}$$

and then to (6), with simple manipulations ($B(A - 1) = -B\overline{A}$).

Thus the lower bound in (11) generalises (6) to random variables and to lower previsions that are centered convex (in particular, W-coherent).

An upper bound for $\overline{P}(X|B)$ can be derived from (11):

Corollary 2 *In the assumptions of Proposition 4 and whenever the relevant previsions are defined* [3]

$$\overline{P}(X|B) \leq \frac{I_B\underline{P}(B(S_B-X))-S_B\underline{P}(B(I_B-X))}{\underline{P}(B(S_B-X))-\underline{P}(B(I_B-X))} \quad (12)$$

Proof. Write (11) for $-X|B$:

$$\underline{P}(-X|B) \geq \frac{-I_B\underline{P}(B(S_B-X))+S_B\underline{P}(B(I_B-X))}{\underline{P}(B(S_B-X))-\underline{P}(B(I_B-X))}.$$

[3] When X is an event A, (12) reduces to $\overline{P}(A|B) \leq \frac{\overline{P}(A\wedge B)}{\overline{P}(A\wedge B)+\underline{P}(\overline{A}\wedge B)}$. We already met this bound in the paragraph following eq. (6).

Eq. (12) follows, reversing signs in the above inequality and since $-\underline{P}(-X|B) = \overline{P}(X|B)$. ■

An issue which remains to be investigated is under what conditions the bound in (11) is sharp, i.e. it is actually equal to the natural extension $\underline{E}(X|B)$ if $\underline{P}$ is coherent, or to the convex natural extension $\underline{E}_c(X|B)$, when $\underline{P}$ is centered convex. The following example illustrates the case of coherence.

Example *Given the partition* $I\!\!P = \{e_1, e_2, e_3, e_4\}$, *define* X *such that* $X(e_1) = 1$, $X(e_2) = -1$, $X(e_3) = 0$, $X(e_4) = 2$. *Given the precise probabilities* P_1, P_2, *having the following values on* $I\!\!P$: $P_1(e_1) = 0.2$; $P_1(e_2) = 0.3$; $P_1(e_3) = 0.2$; $P_1(e_4) = 0.3$; $P_2(e_1) = 0.5$; $P_2(e_2) = 0.1$; $P_2(e_3) = 0$; $P_2(e_4) = 0.4$, *and calling* $\mathcal{A}(I\!\!P)$ *the powerset of* $I\!\!P$, *each of* P_1, P_2 *has a unique coherent extension to a precise prevision on* $U = \mathcal{A}(I\!\!P) \cup \{X\} \cup \{B(X - r) : r \in \mathbb{R}\}$, *where* B *is a given event in* $\mathcal{A}(I\!\!P)$. *A coherent lower prevision* $\underline{P}$ *may be defined on any subset* $\mathcal{D}$ *of* U *as* $\underline{P}(Y) = \min\{P_1(Y), P_2(Y)\}$, $\forall Y \in \mathcal{D}$ *(lower envelope theorem). We choose* $\mathcal{D} = \mathcal{A}(I\!\!P) \cup \{B(X - \inf(X|B)), B(X - \sup(X|B))\}$. *Thus in particular* $\underline{P}(e_1) = 0.2$, $\underline{P}(e_1 \vee e_2 \vee e_3) = 0.6$, *etc. Note that the restriction of* $\underline{P}$ *on* $\mathcal{A}(I\!\!P)$ *is a lower probability which is not 2-monotone (for instance* $\underline{P}(e_1 \vee e_3 \vee e_4) = 0.7 < \underline{P}(e_1 \vee e_3) + \underline{P}(e_3 \vee e_4) - \underline{P}(e_3) = 0.8$). *We have* 10 *non-trivial different choices for the conditioning event* B *in* $\mathcal{A}(I\!\!P)$. *It may be verified that the bound is sharp in all of these but one.*

a) *For instance, let* $B = e_1 \vee e_2 \vee e_3$. *This is one of the* 9 *choices for* B *giving a sharp bound (11). In fact,* $S_B = 1$, $I_B = -1$. *Since* $P_1(B(X-r)) = P_1(BX) - rP_1(B) = -0.1 - 0.7r$ *and* $P_2(B(X - r)) = 0.4 - 0.6r$, *we obtain* $\underline{P}(B(X - r)) = \min\{-0.1 - 0.7r, 0.4 - 0.6r\} = -0.1 - 0.7r$ *iff* $r \geq -5$. *Then the bound (11) is* $\phi(I_B, S_B) = \phi(-1, 1) = -\frac{1}{7} = \underline{E}(X|B)$, *because* $P_1(X|B) = \frac{P_1(BX)}{P_1(B)} = -\frac{1}{7}$. *Note that* $\phi(-1, 1)$ *is not the only sharp bound in the* $\phi(h, k)$ *family:* $\phi(h, k) = -\frac{1}{7}$ *for* $h \in [-5, -1]$, $k \geq 1$.

b) *Let now* $B = e_1 \vee e_4$. *This choice corresponds to the unique non-exact bound (11). In fact, now* $P_1(B(X - r)) = 0.8 - 0.5r$, $P_2(B(X - r)) = 1.3 - 0.9r$, $\underline{P}(B(X-r)) = 1.3 - 0.9r$ *iff* $r \geq \frac{5}{4}$ *and the bound is* $\phi(I_B, S_B) = \phi(1, 2) = \frac{11}{8}$. *To see that the bound cannot be reached, note that [15] if it were sharp, there would be a precise prevision* P, *in the set* $\mathcal{M}_{\mathcal{D}}(\underline{P})$ *of precise previsions dominating* $\underline{P}$ *on* $\mathcal{D}$, *such that its extension on* $X|B$ *ensures that* $P(X|B) = \frac{P(BX)}{P(B)} = \frac{P(e_1)+2P(e_4)}{P(e_1)+P(e_4)} = \frac{11}{8}$, *which means that*

$$P(e_4) = \frac{3}{5}P(e_1) \tag{13}$$

It is then easy to check that no such P *may be found in* $\mathcal{M}_{\mathcal{D}}(\underline{P})$: *just verify that there is no real solution for the system of linear inequalities formed by (13), the dominance constraints* $P \geq \underline{P}$ *on* $\mathcal{D}$, *and the non-negativity and normalisation constraints for* P *on* $I\!\!P$.

c) *Let us introduce partition* $I\!\!P' = \{\omega_1, \omega_{2,3}, \omega_4\}$ *which is a coarsening of* $I\!\!P$: $\omega_1 = e_1$, $\omega_{2,3} = e_2 \vee e_3$, $\omega_4 = e_4$. *We do not modify the uncertainty evaluations of b), defining* P_1, P_2 *on* $I\!\!P'$ *as the restrictions of the previously defined* P_1, P_2 *respectively (thus* $P_1(\omega_1) = 0.2$, $P_1(\omega_{2,3}) = 0.5$, $P_1(\omega_4) = 0.3$, $P_2(\omega_1) = 0.5$, $P_2(\omega_{2,3}) = 0.1$, $P_2(\omega_4) = 0.4$) *and* $\underline{P}$ *as their lower envelope on* $\mathcal{D}' = \mathcal{A}(I\!\!P') \cup \{B(X - \sup(X|B)), B(X - \inf(X|B))\}$, $B \in \mathcal{A}(I\!\!P')$. *Here* $B = \omega_1 \vee \omega_4$, *and* $X(\omega_1) = 1$, $X(\omega_4) = 2$, *while* $X(\omega_{2,3})$ *may take any value, it does not influence the following computations. Note that now* $\underline{P}$ *is* 2*-monotone on* $\mathcal{A}(I\!\!P')$, *since* $I\!\!P'$ *is a three-atom partition [15]. Obviously, the bound (11) is again* $\frac{11}{8}$ *as in b), since, when passing from b) to c), we essentially only grouped together* e_2 *and* e_3, *which are irrelevant in the computation of* $\phi(1, 2)$. *However, there is now a prevision* P *in* $\mathcal{M}_{\mathcal{D}'}(\underline{P})$ *which reaches the bound, i.e. such that* $P(X|B) = \frac{11}{8}$: *its values on* $I\!\!P'$ *are* $P(\omega_1) = 0.5$, $P(\omega_{2,3}) = 0.2$, $P(\omega_4) = 0.3$.

6 Conclusions

In a first part of the paper (Section 3) we related W-coherence with Williams' original definition, and also with other notions of coherence in a conditional framework, especially Walley-coherence. Our main purpose here was to show that W-coherence can be profitably employed to obtain results which hold for Walley-coherence too (Section 4). A more extended comparison between these two coherence concepts is beyond the aims of the present paper, but is an undoubtedly interesting question. It requires analysing further issues, like the role of the conglomerative property or the interpretation of Walley's updating principle.

In the sequel of the paper, we have discussed some implications of product rule bounds and generalised a Bayes' theorem bound to either W-coherent or centered convex lower previsions. Although we did not consider them here, other similar bounds or simple generalisations may be found (for instance, for upper previsions), with analogous properties. A less immediate question is that of investigating further the relationships of these bounds with important concepts in the theory of imprecise previsions: epistemic irrelevance and natural and upper extension for (more general) product rule bounds, 2-monotonicity and possibly Choquet integration for the bound (11). Concerning the latter issue, a generalisation to lower previsions of 2-monotonicity with related results was recently proposed in [5]. There remain anyway two features in our approach which, while ensuring generality, make it difficult to apply pre-existing results to sufficiently general situations, for instance in the problem of establishing when the bound (11) is sharp. One feature is that we are working in a structure-free environment, while 2-monotonicity is customarily referred to algebras of events [15] or (linear) lattices of random variables [5]. With respect to this feature, our example is still rather peculiar: there is a partition $I\!P$ there such that $\mathcal{A}(I\!P) \subset \mathcal{D}$, but this inclusion is generally not required. A second issue is that we consider also the centered convexity condition, and relationships of 2-monotonicity (for previsions) with convexity are still largely to be explored.

Appendix. W-coherence and separate coherence

Let $I\!P$ be an *arbitrary* (finite or not) partition of non-impossible events. We recall the definition of *separate coherence* in [16]:[4]

Definition 4 *The conditional lower previsions $\underline{P}_B(X|B)$, defined for any $B \in I\!P$ and $X \in \mathcal{H}(B)$, where $\mathcal{H}(B)$ is an arbitrary set of gambles containing B, are* separately coherent *iff, for every $B \in I\!P$,*

i) $\underline{P}_B(B|B) = 1$

ii) $\forall s_0, \ldots, s_n \geq 0$, $\forall X_0, \ldots, X_n \in \mathcal{H}(B)$, *defining* $\underline{G} = \sum_{i=1}^n (X_i - \underline{P}(X_i|B)) - s_0(X_0 - \underline{P}(X_0|B))$, *it holds that* $\sup \underline{G} \geq 0$.

When defined on the same domain, separate coherence and W-coherence are equivalent, as we now prove. Define for this the conditional lower prevision $\underline{P}$ such that $\underline{P}(X|B) = \underline{P}_B(X|B)$, $\forall B \in I\!P$, $\forall X \in \mathcal{H}(B)$ ($\underline{P}$ is the collection of all $\underline{P}_B$).

Proposition 6 *The lower previsions $\underline{P}_B$ $(B \in I\!P)$ in Definition 4 are separately coherent iff $\underline{P}$ is W-coherent on $\mathcal{D} = \cup_{B \in I\!P} \mathcal{D}_B$, where $\mathcal{D}_B = \{X|B : X \in \mathcal{H}(B)\}$.*

[4] The integer stakes in [16] may be equivalently replaced by real non-negative ones, as we do here.

Proof. We prove first that W-coherence implies separate coherence. If $\underline{P}$ is W-coherent, i) necessarily holds. As for ii), it follows from $\sup \underline{G} = \max\{\sup_B \underline{G}, \sup_{\overline{B}} \underline{G}\} \geq \sup_B \underline{G} = \sup \underline{G}|B = \sup(B\underline{G}|B) \geq 0$, the last equality holding by (1), the inequality by W-coherence.

To prove the converse implication, suppose that separate coherence holds. Betting on B, $X_0, \ldots, X_n$, it follows then $\sup(s(B - \underline{P}(B|B)) + \sum_{i=1}^n s_i(X_i - \underline{P}(X_i|B)) - s_0(X_0 - \underline{P}(X_0|B))) = \sup(s(B-1) + \underline{G}) = \max(\sup_B(s(B-1) + \underline{G}), \sup_{\overline{B}}(s(B-1) + \underline{G})) \geq 0$. The last inequality implies $\sup_B(s(B-1) + \underline{G}) \geq 0$, if we choose $s > \max(\sup_{\overline{B}} \underline{G}, 0)$, since then $\sup_{\overline{B}}(s(B-1) + \underline{G}) = -s + \sup_{\overline{B}} \underline{G} < 0$. Using also (1), $\sup_B(s(B-1) + \underline{G}) = \sup(\underline{G}|B) = \sup(B\underline{G}|B) = \sup(\sum_{i=1}^n s_i B(X_i - \underline{P}(X_i|B)) - s_0 B(X_0 - \underline{P}(X_0|B))|B) \geq 0$, which means, given the arbitrariness of n, $X_0, \ldots, X_n$ and $s_0, \ldots, s_n \geq 0$, that $\underline{P}$ is W-coherent on $\mathcal{D}_B$. It is then a simple exercise to prove that W-coherence of $\underline{P}$ on each $\mathcal{D}_B$ implies W-coherence of $\underline{P}$ on $\mathcal{D}$, because of the special structure of $\mathcal{D}$. ■

Acknowledgements

We are grateful to the referees for their constructive suggestions.

References

[1] T.E. Armstrong, W.D. Sudderth. Locally coherent rates of exchange. *Annals of Statistics*, 17:1394-1408, 1989.

[2] P. Artzner, F. Delbaen, S. Eber and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9:203-228, 1999.

[3] L. Crisma. Events and conditional events: a subjectivistic approach. In *The notion of event in probabilistic epistemology*, Pubbl. n. 2, Dip. Mat. Appl. 'B. de Finetti', Ediz. Lint, Trieste, 43–89, 1996.

[4] L. Crisma. *Introduzione alla teoria delle probabilità coerenti.* Ediz. EUT, Trieste, 2006.

[5] G. de Cooman, M. C. M. Troffaes and E. Miranda. n-Monotone lower previsions. *Journal of Intelligent & Fuzzy Systems*, 16: 253–263, 2005.

[6] B. de Finetti. Sull'impostazione assiomatica del calcolo delle probabilità. *Annali triestini dell'Università di Trieste*, XIX 2, 29–81, 1949.

[7] S. Maaß. Continuous linear representations of coherent lower previsions. In *Proc. ISIPTA'03*, Lugano (CH), 372–382, 2003.

[8] E. Miranda and G. de Cooman. Epistemic independence in numerical possibility theory. *International Journal of Approximate Reasoning*, 32:23–42, 2003.

[9] R. Pelessoni and P. Vicig. Convex imprecise previsions. *Reliable Computing*, 9(6):465–485, 2003.

[10] R. Pelessoni and P. Vicig. Uncertainty modelling and conditioning with convex imprecise previsions. *International Journal of Approximate Reasoning*, 39(2–3):297–319, 2005.

[11] R. Pelessoni and P. Vicig. Envelope theorems and dilation with convex conditional previsions. In *Proc. ISIPTA'05*, Pittsburgh, PA, 266–275, 2005.

[12] M. C. M. Troffaes and G. de Cooman. Lower previsions for unbounded random variables. In *Soft Methods in Probability, Statistics and Data Analysis*, Advances in Soft Computing, 146–155, Physica-Verlag, New York, 2002.

[13] P. Vicig. Epistemic independence for imprecise probabilities. *International Journal of Approximate Reasoning*, 24:235–250, 2000.

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

[15] P. Walley. Coherent lower (and upper) probabilities. *Research Report*, University of Warwick, Coventry, 1981.

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

[17] P. Walley. Measures of uncertainty in expert systems. *Artificial Intelligence*, 83:1–58, 1996.

[18] 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(1):119-151, 2004.

[19] L.A. Wasserman and J.B. Kadane. Bayes' theorem for Choquet capacities. *Annals of Statistics*, 18:1328-1339, 1990.

[20] K. Weichselberger. *Elementare Grundbegriffe einer allgemeineren Wahrscheinlichkeitsrechnung I – Intervallwahrscheinlichkeit als umfassendes Konzept.* Physika, Heidelberg, 2001.

[21] P. M. Williams. Notes on conditional previsions. *International Journal of Approximate Reasoning*, 44:366–383, 2007 (revised version of: Notes on conditional previsions. *Research Report*, School of Math. and Phys. Science, University of Sussex, 1975).