Full conglomerability, continuity and marginal extension

Enrique Miranda[1] and Marco Zaffalon[2]

Abstract We investigate fully conglomerable coherent lower previsions in the sense of Walley, and some particular cases of interest: envelopes of fully conglomerable linear previsions, envelopes of countably additive linear previsions and fully disintegrable linear previsions. We study the connections with continuity and countable super-additivity, and show that full conglomerability can be characterised in terms of a supremum of marginal extension models.

Key words: Coherent lower previsions, conglomerability, disintegrability, marginal extension.

1 Introduction

Conglomerability of a probability P was first discussed by Bruno de Finetti in [4]. If we consider a partition $\mathcal{B}$ of the possibility space Ω such that $P(B) > 0$ for every $B \in \mathcal{B}$, conglomerability means that

$$(\forall A \subseteq \Omega) \inf_{B \in \mathcal{B}} P(A|B) \leq P(A) \leq \sup_{B \in \mathcal{B}} P(A|B). \tag{1}$$

A related (but stronger) notion was later studied by Dubins, with the name *disintegrability* [3]. Other studies in the precise case were made in [1, 2, 9, 10].

Imposing as well as checking conglomerability can be technically difficult. Partly for this reason, there are different schools of thought about the previous question: those who reject that conglomerability should be a rationality requirement—among them looms the figure of de Finetti himself; and those

University of Oviedo. C-Calvo Sotelo, s/n 33007 Oviedo (Spain) Tel: (+34) 985102955 Fax: (+34) 985103354 mirandaenrique@uniovi.es · IDSIA (Lugano, Switzerland) zaffalon@idsia.ch

who think it should be imposed, often in the light of the paradoxical situations that the lack of conglomerability may lead to. Among the latter stands Peter Walley, who has proposed a behavioural theory of *imprecise* probabilities, where the core modelling unit is a closed convex set of finitely additive probabilities [11]. This theory is essentially Peter Williams' earlier theory of imprecise probability [12] with an additional axiom of conglomerability for sets of probabilities, which coincides with Eq. (1) in the special case of precise probability (and with disintegrability if we require that the conditional model is also precise). The notion of conglomerability is nonetheless not univocally defined in the literature; for this reason, in Section 3 we try to sort out the situation by examining and comparing the different proposals in some detail.

In previous papers we have provided a behavioural support for conglomerability [13] and we have showed that it may be a difficult condition to work with in practice [7, 8]. Here we investigate whether at least the notion of *full* conglomerability (that is, conglomerability with respect to every partition) admits a simple treatment. To this end, we make a thorough mathematical study of the properties of full conglomerability and its relations to other notions: continuity (in various forms), countable super-additivity, and marginal extension. Due to limitations of space, the proofs of the results as well as some relevant counterexamples have been omitted.

2 Preliminary notions

Let us introduce the basic elements of the theory of coherent lower previsions. We refer to [11] for more details. Consider a possibility space Ω. A *gamble* is a bounded map $f : \Omega \to \mathbb{R}$. One instance of gambles are the *indicator* gambles of sets $B \subseteq \Omega$, which we shall denote by I_B or B. We denote by $\mathcal{L}(\Omega)$ the space of all gambles on Ω.

A *linear prevision* on $\mathcal{L}(\Omega)$ is a linear operator satisfying $P(f) \geq \inf f$ for all $f \in \mathcal{L}(\Omega)$. It is the expectation operator with respect to a finitely additive probability. When its restriction to events is countably additive, meaning that $P(\cup_n B_n) = \sum_n P(B_n)$ for any countable family $(B_n)_n$ of pairwise disjoint events, we say that P is a *countably additive linear prevision*.

A *coherent lower prevision* $\underline{P}$ on $\mathcal{L}(\Omega)$ is the lower envelope of a closed and convex set of linear previsions. The conjugate upper envelope $\overline{P}$ is called a *coherent upper prevision*, and it holds that $\overline{P}(f) = -\underline{P}(-f)$ for all f. We let $\mathcal{M}(\underline{P}) := \{P \text{ linear prevision} : (\forall f)\ P(f) \geq \underline{P}(f)\}$ and call it the *credal set* associated with $\underline{P}$. More generally, we say that a map $\underline{P} : \mathcal{L}(\Omega) \to \mathbb{R}$ *avoids sure loss* when it is dominated by some coherent lower prevision. The smallest such prevision is called its *natural extension*, and it coincides with the lower envelope of the non-empty set $\mathcal{M}(\underline{P})$.

A coherent lower prevision is in a one-to-one correspondence with its associated set of *strictly desirable gambles* $\underline{\mathcal{R}} := \{f : \underline{P}(f) > 0 \text{ or } f \gneq 0\}$, in the sense that $\underline{P}(f) = \sup\{\mu : f - \mu \in \underline{\mathcal{R}}\}$ for all $f \in \mathcal{L}(\Omega)$; the closure $\overline{\mathcal{R}}$ of the set of strictly desirable gambles in the topology of uniform convergence is called the set of *almost-desirable gambles*, and it satisfies $\overline{\mathcal{R}} = \{f : \underline{P}(f) \geq 0\}$.

The notion of coherence can also be extended to the conditional case. Let $\mathcal{B}$ be a partition of Ω. A *separately coherent* conditional lower prevision is a map $\underline{P}(\cdot|\mathcal{B}) := \sum_{B\in\mathcal{B}} I_B \underline{P}(\cdot|B)$, and where for every $B \in \mathcal{B}$ the functional $\underline{P}(\cdot|B) : \mathcal{L}(\Omega) \to \mathbb{R}$ is a coherent lower prevision satisfying $\underline{P}(B|B) = 1$.

Given a coherent lower prevision $\underline{P}$ and a separately coherent conditional lower prevision $\underline{P}(\cdot|\mathcal{B})$, they are (jointly) *coherent* when $\underline{P}(G(f|B)) = 0$ for all $f \in \mathcal{L}(\Omega), B \in \mathcal{B}$ and $\underline{P}(G(f|\mathcal{B})) \geq 0$ for all $f \in \mathcal{L}(\Omega)$, where $G(f|B) := B(f - \underline{P}(f|B))$ and $G(f|\mathcal{B}) := \sum_B G(f|B) = f - \underline{P}(f|\mathcal{B})$.

This notion is based on what Walley called the *conglomerative principle*, which means that if a gamble f satisfies that $I_B f$ is desirable for any $B \in \mathcal{B}$, then f should also be desirable. This is the main point of controversy between Walley's and de Finetti's approaches. The latter only requires that a finite sum of desirable gambles is again desirable, and this yields a different notion of conditional coherence, usually referred to as *Williams coherence* [12].

The notion of natural extension can also be considered in the conditional case. Given a coherent lower prevision $\underline{P}$ and a partition $\mathcal{B}$ of Ω, its *conditional natural extension* $\underline{P}(\cdot|\mathcal{B})$ is given by

$$\underline{P}(f|B) := \begin{cases} \inf_B f & \text{if } \underline{P}(B) = 0, \\ \sup\{\mu : \underline{P}(B(f-\mu)) \geq 0\} & \text{otherwise} \end{cases} \tag{2}$$

for any $f \in \mathcal{L}(\Omega)$. It always holds that $\underline{P}(G(f|B)) = 0$ for all $f \in \mathcal{L}(\Omega), B \in \mathcal{B}$, so $\underline{P}, \underline{P}(\cdot|\mathcal{B})$ are coherent if and only if $\underline{P}(G(f|\mathcal{B})) \geq 0$ for all $f \in \mathcal{L}(\Omega)$.

3 Different notions of conglomerability in the literature

As we mentioned in the Introduction, conglomerability was first introduced by de Finetti in [4] in terms of Eq. (1). The conditional probability $P(A|B)$ in that equation is derived from the unconditional one by Bayes' rule, so that $P(A|B) = P(A \cap B)/P(B)$, whenever $P(B) \neq 0$. However, de Finetti argued [5, Ch. 5] that it also makes sense to consider the conditional probability $P(A|B)$ when the event B has probability 0 but is not deemed impossible. In that case, he suggested to define a *full conditional measure* as that considered in [3, Sect. 3].

There exists a connection between full conditional measures and the theory of coherent previsions: if we represent a full conditional measure on $\mathcal{P}(\Omega) \times (\mathcal{P}(\Omega) \setminus \emptyset)$ as a family of conditional and unconditional assessments $\{P(\cdot|B) : B \subseteq \Omega\}$, then these conditional previsions satisfy the notion of Williams coherence [12, Prop. 6]. On the other hand, as Schervisch, Seidenfeld and Kadane have established in [9, 10], if the linear prevision that results from restricting a full conditional measure to $\mathcal{P}(\Omega)$ is not countably additive, then there is some partition $\mathcal{B}$ of Ω where Eq. (1) is violated. In other words, under this approach the only fully conglomerable models are the countably additive ones.

On the other hand, Walley [11, Sect. 6.8.1] calls a coherent lower prevision $\underline{P}$ on $\mathcal{L}(\Omega)$ $\mathcal{B}$*-conglomerable* if for any gamble f such that $\underline{P}(Bf) \geq 0$ for all $B \in \mathcal{B}$ with $\underline{P}(B) > 0$, it holds that $\underline{P}(\sum_{\underline{P}(B)>0} Bf) \geq 0$. This is equivalent to the existence of a conditional lower prevision $\underline{P}(\cdot|\mathcal{B})$ such that $\underline{P}, \underline{P}(\cdot|\mathcal{B})$ are jointly coherent, and also to the coherence of $\underline{P}$ with its conditional natural extension. Thus, conglomerability means that the coherent lower prevision $\underline{P}$ can be updated in a coherent way to a conditional lower prevision $\underline{P}(\cdot|\mathcal{B})$. The notion can be applied in particular to linear previsions. However, in that case we may also require that the linear prevision can be updated into a *linear* model; this gives rise to a stronger notion, called $\mathcal{B}$*-disintegrability*. From [11, Thm. 6.5.7], the $\mathcal{B}$-disintegrability of a linear prevision is equivalent to the existence of a conditional linear prevision $P(\cdot|\mathcal{B})$ such that $P = P(P(\cdot|\mathcal{B}))$.

We say that $\underline{P}$ is *fully conglomerable* when it is $\mathcal{B}$-conglomerable for every partition $\mathcal{B}$ of Ω. In a similar manner, we say that a linear prevision P is *fully disintegrable* when for every partition $\mathcal{B}$ there is some conditional linear prevision $P(\cdot|\mathcal{B})$ such that $P = P(P(\cdot|\mathcal{B}))$.

If a lower prevision $\underline{P}$ is fully conglomerable, then we can define a family of conditional lower previsions $\mathcal{H} := \{\underline{P}(\cdot|\mathcal{B}) : \mathcal{B} \text{ partition of } \Omega\}$ with the property that $\underline{P}, \underline{P}(\cdot|\mathcal{B})$ are coherent for every partition $\mathcal{B}$. It can be checked that these conditional lower previsions are also coherent with each other, in the sense that they can all be induced by a common fully conglomerable set of desirable gambles. This means that when we consider the family of all partitions, coherence becomes equivalent to the notion of *conglomerable coherence* studied in much detail in [7]. In the same manner that the natural extension of a lower prevision is the smallest dominating coherent lower prevision, we shall call the *fully conglomerable natural extension* the smallest fully conglomerable coherent lower prevision that dominates $\underline{P}$, in case it exists.

We see then that the two approaches are different, basically because of the manner the problem of conditioning on sets of (lower) probability zero is dealt with. In de Finetti's case, it is advocated to use full conditional measures, while in Walley's case these sets are not taken into account (in the lower prevision approach we are considering here; a more informative model would be that of sets of desirable gambles). In this sense, Walley's condition is close to what Armstrong called *positive conglomerability* in [1]. The different approach means for instance that a linear prevision whose restriction to events is $\{0, 1\}$-valued will always be fully conglomerable for Walley, while it may not be so for de Finetti. Another key difference is in the rejection by de Finetti of the conglomerative principle, that makes the conditional models subject to a different consistency condition (Williams coherence for de Finetti, and the stronger version of Walley in his case).

4 Full conglomerability in the precise case

In the precise case, we consider three properties for a linear prevision P:

M1. P is countably additive.
M2. P is fully disintegrable.
M3. P is fully conglomerable.

By [11, Thm. 6.9.1], condition M1 implies M2; on the other hand, it follows from its definition that a fully disintegrable linear prevision is in particular fully conglomerable. With respect to the converse implication, we shall consider two cases: linear previsions whose restrictions to events have a finite range (called *molecular* in [2]) and those whose restrictions to events have infinite range (called *non-molecular* in [2]).

Proposition 1. *Let P be a linear prevision on $\mathcal{L}(\Omega)$.*

1. *If P is molecular, then for every partition $\mathcal{B}$ of $\Omega, |\{B \in \mathcal{B} : P(B) > 0\}| < +\infty$, and as a consequence, P is fully conglomerable.*
2. *If P is non-molecular, then it is countably additive if and only if it is fully conglomerable. In that case, $P(\{\omega \in \Omega : P(\omega) > 0\}) = 1$.*

In [9, Thm. 3.3] it is proven that any full conditional measure whose associated unconditional probability is molecular and not countably additive is not fully disintegrable. In other words, countable additivity and full disintegrability are equivalent in the molecular case provided we enter the framework of full conditional measures.

Next we study the connection with continuity. We consider the following continuity conditions:

C1. $(f_n)_{n\in\mathbb{N}} \to f \Rightarrow (\underline{P}(f_n))_{n\in\mathbb{N}} \to \underline{P}(f)$.
C2. $(f_n)_{n\in\mathbb{N}} \downarrow f \Rightarrow (\underline{P}(f_n))_{n\in\mathbb{N}} \downarrow \underline{P}(f)$.
C3. $(f_n)_{n\in\mathbb{N}} \downarrow 0 \Rightarrow (\underline{P}(f_n))_{n\in\mathbb{N}} \downarrow 0$.
C4. $(f_n)_{n\in\mathbb{N}} \uparrow f \Rightarrow (\underline{P}(f_n))_{n\in\mathbb{N}} \uparrow \underline{P}(f)$.

It is not difficult to show the following:

Proposition 2. *For any linear prevision P, M1 $\Leftrightarrow$ C2 $\Leftrightarrow$ C3 $\Leftrightarrow$ C4.*

We deduce from this that condition C1 is sufficient for P to be countably additive. However, it is not necessary. On the other hand, any of the conditions C2–C4 is sufficient for P to be fully disintegrable, and as a consequence also fully conglomerable.

The only open problem at this stage would be the equivalence between M2 and M1. A counterexample would require the definition of a family of conditional linear previsions $\{P(\cdot|\mathcal{B}) : \mathcal{B}$ partition of $\Omega\}$ and an unconditional linear prevision P such that $P = P(P(\cdot|\mathcal{B}))$ for every $\mathcal{B}$ (so P is fully disintegrable) while there exists a finite sub-family of $\{P(\cdot|\mathcal{B}) : \mathcal{B}$ partition of $\Omega\}$ which violates Williams coherence (so that we cannot make a representation in terms of full conditional measures, because if we could, then P would be countably additive by [9]). Such an example seems unlikely, in our opinion.

5 Full conglomerability in the imprecise case

In the imprecise case, we consider three properties of a coherent lower prevision $\underline{P}$:

M4. $\underline{P}$ is the lower envelope of a family of countably additive linear previsions.
M5. $\underline{P}$ is the lower envelope of a family of fully conglomerable linear previsions.
M6. $\underline{P}$ is fully conglomerable.

Analogous conditions to M4, M5 (in terms of upper envelopes) can be established for a coherent upper prevision $\overline{P}$. It is immediate to see that

$$\text{M1} \Rightarrow \text{M3} \Rightarrow \text{M5} \Rightarrow \text{M6 and M1} \Rightarrow \text{M4} \Rightarrow \text{M5} \Rightarrow \text{M6}.$$

However, the remaining implications do not hold: on the one hand, a linear prevision may be fully conglomerable without being countably additive; moreover, there are fully conglomerable coherent lower previsions that are not dominated by any fully conglomerable (and as consequence by any countably additive) linear prevision [11, Ex. 6.9.6].

With respect to M4, Krätschmer established in [6, Sect. 5] that a 2-alternating upper probability on $\mathcal{P}(\Omega)$ is the upper envelope of a family of countably additive probabilities if and only if $\overline{P}(A) = \sup\{\overline{P}(B) : A \supseteq B \text{ finite}\}$ for every $A \subseteq \Omega$. However, we have shown that the above condition does not characterise M4 in general. Nevertheless, we can give a necessary and sufficient condition in the particular case where $\Omega = \mathbb{N}$:

Proposition 3. *Let $\overline{P}$ be a coherent upper prevision on $\mathcal{L}(\mathbb{N})$. Then $\overline{P}$ satisfies M4 $\Leftrightarrow (\forall n \in \mathbb{N})\ \overline{P} = \sup \mathcal{M}_n \Leftrightarrow (\forall f \geq 0)\ \overline{P}(f) = \lim_n \overline{P}(fI_{\{1,\dots,n\}}) \Leftrightarrow (\forall f \geq 0)\ \overline{P}(f) = \sup\{\overline{P}(g) : g \leq f, \mathtt{supp}(g) \text{ finite}\}$, where $\mathcal{M}_n := \{P \leq \overline{P} : \lim_m P(\{1,\dots,m\}) \geq 1 - \frac{1}{n}\}$ and $(\forall g)\ \mathtt{supp}(g) = \{n : g(n) \neq 0\}$.*

Next, we study the connection with the continuity properties C1–C4. On the one hand, we deduce from the precise case that none of them is necessary for $\underline{P}$ to belong to M5, M6. On the other hand, we have that:

Proposition 4. *C1 $\Rightarrow$ C4 $\Rightarrow$ M4 $\Rightarrow$ C2, M5 $\Rightarrow$ M6 and C2 $\Rightarrow$ C3. Moreover, no additional implication other than the ones that immediately follow from these holds.*

Next we investigate the connection with the following condition:

M7. $(\forall (f_n)_n \subseteq \mathcal{L}(\Omega)$ such that $\sum_n f_n \in \mathcal{L}(\Omega))\ \underline{P}\left(\sum_n f_n\right) \geq \sum_n \underline{P}(f_n)$.

The reason for our investigation is that both countable super-additivity and conglomerability are quite related to the closedness of a set of desirable gambles under countable sums. Specifically, we have proven the following:

Proposition 5. *Let $\underline{P}$ be a coherent lower prevision and let $\underline{\mathcal{R}}, \overline{\mathcal{R}}$ denote its associated sets of strictly desirable and almost desirable gambles, respectively. Then each of the following statements implies the next:*

1. $\underline{P}$ *satisfies M7.*
2. $(\forall (f_n)_n \subseteq \underline{\mathcal{R}} : \sum_n f_n \in \mathcal{L}(\Omega))\ \ \sum_n f_n \in \underline{\mathcal{R}}$.
3. $(\forall (f_n)_n \subseteq \underline{\mathcal{R}} : \sum_n f_n \in \mathcal{L}(\Omega))\ \ \sum_n f_n \in \overline{\underline{\mathcal{R}}}$.
4. $\underline{P}$ *satisfies C3.*

The connection between M7 and the other conditions is given by

$$\text{C2} \Rightarrow \text{M7} \Rightarrow \text{C3 and M7} \Rightarrow \text{M6},$$

together with those derived from Prop. 4. We deduce that if P is linear,

$$\text{C1} \Rightarrow \text{M1} \Leftrightarrow \text{C2} \Leftrightarrow \text{M7} \Leftrightarrow \text{C3} \Leftrightarrow \text{C4} \Rightarrow \text{M2} \Rightarrow \text{M3}.$$

The only open problem left at this stage is whether M7 and C2 are equivalent.

6 Full conglomerability and marginal extension

From [11, Thm. 6.8.2], given a coherent lower prevision $\underline{P}$ and a partition $\mathcal{B}$ of Ω, it holds that $\underline{P}$ is $\mathcal{B}$-conglomerable if and only if $\underline{P} \geq \underline{P}(\underline{P}(\cdot|\mathcal{B}))$, where $\underline{P}(\cdot|\mathcal{B})$ is the conditional natural extension of $\underline{P}$, given by Eq. (2). Thus, $\underline{P}$ is fully conglomerable if and only if $\underline{P} \geq \sup_{\mathcal{B} \text{ partition}} \underline{P}(\underline{P}(\cdot|\mathcal{B})) := \underline{Q}$.

The concatenation $\underline{P}(\underline{P}(\cdot|\mathcal{B}))$ of a marginal and a conditional lower prevision is called a *marginal extension model* [11, Sect. 6.7]; this is an extension of the product rule to the imprecise case. The condition above tells us then that fully conglomerable lower previsions are always the supremum of a family of marginal extension models. Our next proposition summarizes the relationship between $\underline{P}$ and the functional $\underline{Q}$ it determines:

Proposition 6. *Let $\underline{P}$ be a coherent lower prevision and $\underline{F}$ its fully conglomerable natural extension (if it exists), and define $\underline{Q}$ as above.*

1. $\underline{P} \leq \underline{Q} \leq \underline{F}$.
2. $\underline{P}$ *is fully conglomerable* $\Leftrightarrow \underline{P} = \underline{Q}$.
3. $\underline{Q}$ *does not avoid sure loss in general, and* $\mathcal{M}(\underline{Q}) \neq \emptyset \not\Rightarrow \underline{P}$ *satisfies M6.*

Thus, the full conglomerability of $\underline{P}$ implies the coherence of $\underline{Q}$. Although it is an open problem whether the converse holds in general, it is easy to see that when P is linear, then $\underline{Q} \geq P$ is coherent if and only if $\underline{Q} = P$ (it cannot be that $\underline{Q}(f) > P(f)$ and still be that $\underline{Q}$ is coherent), so in the precise case we have the equivalence.

7 Conclusions

Our results show that countably additive models and their envelopes seem to be easier to use in practice than fully conglomerable ones; although the connection with continuity in the precise case is well known, as it follows almost immediately from existing results from probability theory, in the imprecise case we have given a necessary and a sufficient condition, as well as a characterisation in terms of the natural extension from gambles with a finite range. In our view, this indicates that envelopes of countably additive linear previsions may be more interesting in practice, and they could be a tool to guarantee the property of full conglomerability.

The definition of joint coherence of a conditional and an unconditional lower prevision has led us to define the functional $\underline{Q}$ as a supremum of marginal extensions. A deeper study of this functional is one of the main open problems for future work; in particular, we would like to determine whether the existence of the fully conglomerable natural extension is equivalent (and not only sufficient) to $\underline{Q}$ avoiding sure loss, and whether the coherence of $\underline{Q}$ is sufficient (and not only necessary) for its equality with the fully conglomerable natural extension of $\underline{P}$.

More generally, it would be interesting to make a deeper comparison between our results and the ones established by Seidenfeld et al. for the precise case by means of full conditional measures.

Acknowledgements The research reported in this paper has been supported by project TIN2014-59543-P.

References

1. Armstrong T (1990) Conglomerability of probability measures on Boolean algebras. J. Math. Anal. Appl. 150:335–358
2. Armstrong T, Prikry K (1982) The semi-metric on a Boolean algebra induced by a finitely additive probability measure. Pac. J. Math. 99:249–264
3. Dubins L (1975) Finitely additive conditional probabilities, conglomerability and disintegrations. Ann. Prob. 3:88–99
4. de Finetti B (1930) Sulla proprietà conglomerativa delle probabilità subordinate. Rend. Real. Inst. Lomb. 63:414–418
5. de Finetti B. (1972) Probability, Induction and Statistics. Wiley, London
6. Krätschmer V (2003) When fuzzy measures are upper envelopes of probability measures. Fuz. Sets Syst. 138:455–468
7. Miranda E, Zaffalon M (2013) Conglomerable coherence. Int. J. App. Reas. 54:1322–1350
8. Miranda E, Zaffalon M, de Cooman G (2012) Conglomerable natural extension. Int. J. App. Reas. 53:1200–1227
9. Schervisch M, Seidenfeld T, Kadane, J (1984) The extent of nonconglomerability of finitely additive probabilities. Zeit. Wahr. Verw. Geb. 66:205–226
10. Seidenfeld T, Schervisch M, Kadane J (1998) Non-conglomerability for finite-valued finitely additive probability. Sank. 60:476–491
11. Walley P (1991) Statistical Reasoning with Imprecise Probabilities. Chapman and Hall, London
12. Williams P (2007) Notes on conditional previsions. Int. J. App. Reas. 44:366–383
13. Zaffalon M, Miranda E (2013) Probability and time. Art. Int. 198:1–51