Proceedings of Machine Learning Research 147:42–50, 2021 ISIPTA 2021

Probability Filters as a Model of Belief; Comparisons to the Framework of Desirable Gambles.

Catrin Campbell-Moore CATRIN.CAMPBELL-MOORE@BRISTOL.AC.UK
Department of Philosophy, University of Bristol, UK

Abstract

We propose a model of uncertain belief. This models coherent beliefs by a filter, F, on the set of probabilities. That is, it is given by a collection of sets of probabilities which are closed under supersets and finite intersections. This can naturally capture your probabilistic judgements. When you think that it is more likely to be sunny than rainy, we have $\{p \mid p(\textsc{Sunny}) > p(\textsc{Rainy})\} \in F$. When you think that a gamble g is desirable, we have $\{p \mid \mathrm{Exp}_p[g] > 0\} \in F$. It naturally extends the model of credal sets; and we will show it captures all the expressive power of the desirable gambles model. It also captures the expressive power of sets of desirable gamble sets (with a mixing axiom, but no Archimedean axiom).

1. Introduction

We propose a new model of belief based on probabilities which is very expressively powerful. This models coherent beliefs by a filter, F, on the set of probabilities. That is, it is given by a collection of sets of probabilities which are closed under supersets and finite intersections.

We can use this to directly capture your opinions. To capture your comparative judgement that it is more likely to be sunny than rainy, we have

$$\{p \mid p(\textsc{Sunny}) > p(\textsc{Rainy})\} \in F.$$

We can also capture judgements regarding gambles by considering probabilistic expectations. To capture that you think one gamble, g, is (strictly) preferable to another gamble, f, we have

$$\{p \mid \mathrm{Exp}_p[g] > \mathrm{Exp}_p[f]\} \in F.$$

To capture that you think that a gamble, g, is desirable (strictly preferable to the status quo) we have

$$\{p \mid \mathrm{Exp}_p[g] > 0\} \in F.$$

This model is expressively powerful. It can, for example, capture all the expressive power of the framework of desirable gambles. This involves some non-Archimedeanicity by permitting the filters to be non-principal, i.e., not generated by a credal set. It can also capture all the expressive power of the model of choice functions, or sets of desirable gamble sets (we include a mixing axiom, but no Archimedean axiom).

Using other terms, this model was proposed and discussed in a joint paper with Jason Konek, (Campbell-Moore and Konek, 2019), using the interpretation of beliefs with probabilistic contents as outlined in Moss (2018). The main results of this paper were stated there without proof.

The paper proceeds as follows. We introduce the model of probability filters in section 2. We observe that it extends the model of credal sets (section 2.2) and show when and how to extend a given set of judgements to a probability filter (section 2.4). In section 3 we show that the model of probability filters captures all the expressive power of framework of desirable gambles and in section 4 we show it captures all the expressive power of the framework of sets of desirable gamble sets or choice functions. Section 5 notes a case where it goes beyond the framework of choice functions. Full investigation of its expressive power remains future work.

2. Probability Filters

Fix Ω as a finite non-empty set. Our results are all within the context of finite sample spaces. Extending these results is future work.

Setup 1 *A* probability function $p : \wp(\Omega) \to \mathbb{R}$ *is characterised by a probability mass function, $p : \Omega \to \mathbb{R}$, with $p(\omega) \geq 0$ and $\sum_{\omega\in\Omega} p(\omega) = 1$.*

p is regular *iff $p(\omega) > 0$ for every $\omega \in \Omega$. We restrict attention to regular probability functions.*

rProbs *is the collection of all regular probability functions on Ω.*

We model your belief by a set of sets of probabilities, $F \subseteq \wp(\mathrm{rProbs})$. This captures your (probabilistic) opinions. If you think that it is less than 0.3-likely to be rainy, then $\{p \mid p(\textsc{Rainy}) < 0.3\} \in F$.

Further reflections on exactly what it means to have $P \in F$ is future work, but we work with the idea that we use it to naturally encode certain (probability-based) judgements.

We define what it is for F to be coherent as being a proper filter. It must be closed under supersets (F3) and finite intersections (F2). It should also be non-trivial: it should be non-empty (F1) and it must not contain the emptyset (F4) otherwise it contains every $P \subseteq \text{rProbs}$. (Axiom F4 is what it is for a filter to be *proper*).

Definition 2 *Say $F \subseteq \wp(\text{rProbs})$ is* coherent *iff it is a proper filter; that is:*

F1. $F \neq \emptyset$.

F2. *If $P, Q \in F$ then $P \cap Q \in F$.*

F3. *If $P \in F$ and $Q \supseteq P$ then $Q \in F$.*

F4. $\emptyset \notin F$.

Axioms F2 and F3 ensure that F is closed under finite probabilistic consequence. That is, if $P_1, \ldots, P_n \in F$ and $Q \supseteq P_1 \cap \ldots \cap P_n$ then also $Q \in F$. This is the characteristic feature of a filter.

2.1. Regularity

We have essentially encoded an axiom in the setup by only considering filters on *regular* probabilities. We have decided to adopt this to match the assumption used in the desirable gambles setting that *weakly* dominating 0 is sufficient to be (strictly) desirable.

If g merely weakly dominates 0 and p is not regular, then we can have $\text{Exp}_p[g] = 0$. Since we will capture the judgement that g is desirable with $\{p \mid \text{Exp}_p[g] > 0\} \in F$, we want to ensure that this is in F for any g weakly dominating 0, which we obtain by restricting to regular probabilities.[1]

I conjecture that all the results will also work in the setting without regularity (with the desirability axioms also modified).[2] The only step which is then more delicate is to apply the separating hyperplane results, as the relevant sets might not be closed.

It would be more explicit (and preferable) to simply have F a collection of sets of possibly non-regular probabilities and then have an axiom:

- For each $\omega \in \Omega$, $\{p \mid p(\omega) > 0\} \in F$

However this makes the notion of natural extension more complex and it is easier to simply impose it as part of the setup.

1. We can prove that this restriction is required: For $\mathbf{I}_{\{\omega\}}$ (the indicator gamble) to be desirable we need $\{p \mid \text{Exp}_p[\mathbf{I}_{\{\omega\}}] > 0\} = \{p \mid p(\omega) > 0\} \in F$. When Ω is finite, that means $\{p \mid p(\omega) > 0 \text{ for all } \omega \in \Omega\} = \text{rProbs} \in F$.
2. An alternative to maintain the usual desirability axioms but drop the regularity assumption is to instead put $g \in D_F$ iff $\{p \mid p \cdot g \geq 0\} \in F$ and $\{p \mid p \cdot g \leq 0\} \notin F$. So long as $\{p \mid p(\omega) = 0\} \notin F$ for any ω, this will obtain axiom D2. Further investigation remains future work.

2.2. Principal Filters and Sets of Probabilities

Special kinds of filters are those which are *principal*; that is, there is some set of probabilities, $\mathbb{P} \subseteq \text{rProbs}$, with

$$P \in F \text{ iff } P \supseteq \mathbb{P} \tag{1}$$

Principal filters are equivalent to the model of belief given by credal sets, where we capture your beliefs with a single set of probabilities. Strictly speaking, since we have restricted to regular probabilities, it is equivalent to (arbitrary) sets of regular probabilities, but by relaxing this assumption the more general equivalence will hold.

The restriction to principal filters, and thus the model of belief of credal sets, is given by strengthening the axiom of finite intersections (axiom F2) to arbitrary intersections:

$\text{F}_{\text{inf}\cap}$. If $P_i \in F$ for each $i \in I$ then $\bigcap_{i \in I} P_i \in F$. (Where I can be infinite.)

We will not adopt this strengthened axiom. Non-principal filters are important to capture the expressive power of the desirable gambles framework as it allows for non-Archimadean behaviour when we consider free filters, where $\bigcap F = \emptyset$.

2.3. Totality / Ultrafilters

Another interesting class of filters are *ultrafilters*. These are the maximal filters, where there is no filter $F' \supset F$. Ultrafilters satisfy the axiom of totality:

F_{tot}. $P \in F$ or $\overline{P} \in F$, where $\overline{P} := \text{rProbs} \setminus P$.

This says that you are maximally opinionated. For each probabilistic property you either endorse it or its complement.

Principal ultrafilters give us precise probabilities. Non-principal ultrafilters roughly give us hyperreal probabilities.[3]

2.4. Natural Extensions

If we obtain some of your probabilistic judgements but not a complete description, in what cases that it be extended coherently?

We define what the resultant extension will be, if possible. This just closes a set under finite intersections and supersets.

Definition 3 *Let $E \subseteq \wp(\text{rProbs})$. $\text{ext}(E)$, is defined by $Q \in \text{ext}(E)$ iff there are some $P_1, \ldots, P_n \in E$ and $Q \supseteq P_1 \cap \ldots \cap P_n$.*

3. This is not a statement of a formal claim, although a formal result is hopefully possible. This would then mean that any probability filter can be given by a set of hyperreal probabilities because any filter is characterised by the set of ultrafilters refining it.

This is coherent whenever $\emptyset \notin \text{ext}(E)$. This is when E satisfies the finite intersection property: whenever $P_1, \dots, P_n \in E$, $P_1 \cap \dots \cap P_n \neq \emptyset$. Such E are called filter subbases, and $\text{ext}(E)$ is the filter generated by it.

Proposition 4 *There is a proper filter $F \supseteq E$ iff $\emptyset \notin \text{ext}(E)$; and the minimal such filter is $\text{ext}(E)$.*

Another useful case is when E is already closed under finite intersections. In this case, $\text{ext}(E)$ just takes supersets. For example, if we have a descending chain $P_1 \supseteq P_2 \supseteq P_3 \supseteq \dots$, then $Q \in \text{ext}(\{P_1, P_2, \dots\})$ iff $Q \supseteq P_i$ for some i.

3. Probability Filters and Desirable Gambles

One of the most prominent models of belief in the imprecise probability literature is to model one's belief by a set of desirable gambles (Walley, 2000). In this section, we will show that the probability filters model contains all the representational power of such sets of desirable gambles.

Setup 5 *$\mathscr{G}$ is the set of all gambles, which are the bounded functions from Ω to $\mathbb{R}$.*[4]

When $f(\omega) \geq g(\omega)$ for all $\omega \in \Omega$, we will say $f \geq g$.

$\mathscr{G}_{\gtrsim 0}$ is the set of gambles which weakly dominate 0. I.e., where $f \geq 0$ and $f(\omega) > 0$ for some $\omega \in \Omega$.[5]

For p probabilistic and g a gamble, we use $p \cdot g$ for $\sum_{\omega \in \Omega} p(\omega) g(\omega)$, which is just probabilistic expectation, also denoted $\text{Exp}_p[g]$.

We will also make use of the positive linear hull of a set: $\text{posi}(B) := \{\sum_{i=1}^n \lambda_i g_i \mid n > 0, \lambda_i > 0, g_i \in B\}$.

Finally, we use $\mathbf{I}_X$ as the indicator gamble for $X \subseteq \Omega$. I.e., the gamble taking value 1 for $\omega \in X$ and value 0 for $\omega \notin X$.

Definition 6 *D is* coherent *if:*

D1. $0 \notin D$

D2. *If $g \in \mathscr{G}_{\gtrsim 0}$, then $g \in D$*

D3. *If $g \in D$ and $\lambda > 0$, then $\lambda g \in D$*

D4. *If $f, g \in D$, then $f + g \in D$*

Given a probability filter, we can extract a set of desirable gambles using:

$$g \in D_F \text{ iff } \{p \mid p \cdot g > 0\} \in F \tag{2}$$

Any coherent probability filter gives a coherent set of desirable gambles.

Theorem 7 *If F is coherent, then D_F is a coherent set of desirable gambles.*

Proof Axiom D1 follows from axiom F4 by observing that $\{p \mid p \cdot 0 > 0\} = \emptyset$.

Axiom D2: If $g \in \mathscr{G}_{\gtrsim 0}$, any $p \in \text{rProbs}$ has $p \cdot g > 0$; so this follows from $\text{rProbs} \in F$ (using our choice to restrict to regular probabilities, and axioms F1 and F3). Note that this relies on our restriction to regular probabilities.

Axiom D3 holds because when $\lambda > 0$, $\{p \mid p \cdot g > 0\} = \{p \mid p \cdot \lambda g > 0\}$.

Axiom D4: If $g \in D_F$ and $f \in D_F$, then $\{p \mid p \cdot g > 0\} \in F$ and $\{p \mid p \cdot f > 0\} \in F$. So $\{p \mid p \cdot g > 0 \text{ and } p \cdot f > 0\} \in F$ by axiom F2. If $p \cdot g > 0$ and $p \cdot f > 0$ then also $p \cdot (g + f) > 0$. So by axiom F3, $\{p \mid p \cdot (g + f) > 0\} \in F$, and thus $g + f \in D_F$. ■

Given a set of desirable gambles, D, we can construct a filter F_D which is the minimal filter evaluating each $g \in D$ as desirable. That is, it is the minimal filter where $\{p \mid p \cdot g > 0\} \in F$ for each $g \in D$. Formally, then, F_D is the filter generated by this collection, that is, it just closes it under finite intersection and supersets.

We are able to prove that in doing this we do not go beyond the constraints of coherence in the desirable gamble model; that is, we do not add evaluations that gambles are desirable for any gambles not already in D.

Theorem 8 *If D is coherent, then*

$$F_D := \text{ext}(\{\{p \mid p \cdot g > 0\} \mid g \in D\}) \tag{3}$$

is coherent and

$$f \in D \text{ iff } \{p \mid p \cdot f > 0\} \in F_D. \tag{4}$$

Thus, for distinct coherent D and D', F_D and $F_{D'}$ are distinct. And $D_{F_D} = D$.

Proof Any $f \in D$ has $\{p \mid p \cdot f > 0\} \in F_D$ by construction. We need to show the converse.

By definition of ext, $\{p \mid p \cdot f > 0\} \in F_D$ iff there are $g_1, \dots, g_n \in D$ with

$$\{p \mid p \cdot f > 0\} \supseteq \{p \mid p \cdot g_1 > 0\} \cap \dots \cap \{p \mid p \cdot g_n > 0\}. \tag{5}$$

The main work of this proof is to show that this entails that $f \in \text{posi}(\{g_1, \dots, g_n\} \cup \mathscr{G}_{\gtrsim 0})$.

Suppose $f \notin \text{posi}(\{g_1, \dots, g_n\} \cup \mathscr{G}_{\gtrsim 0})$. This is a convex cone, so by a separating hyperplane theorem (Klee, 1955), we can find a separating linear functional that can be normalised to obtain a probabilistic p with $p \cdot g \geq 0$ for each $g \in \text{posi}(\{g_1, \dots, g_n\} \cup \mathscr{G}_{\gtrsim 0})$ and $p \cdot f \leq 0$ (see figure 1). Since $0 \notin \text{posi}(\{g_1, \dots, g_n\} \cup \mathscr{G}_{\gtrsim 0})$ by coherence of D, we can in fact ensure $p \cdot g > 0$ for each g in this set. Since each $\mathbf{I}_{\{\omega\}}$ is in this set, each $p(\omega)$ is thus positive, so this gives a regular probabilistic p with $p \cdot g_i > 0$ for $g_1, \dots, g_n$, but $p \cdot f \leq 0$; and thus $\{p \mid p \cdot f > 0\} \not\supseteq \{p \mid p \cdot g_1 > 0\} \cap \dots \cap \{p \mid p \cdot g_n > 0\}$.

4. Since Ω is finite, they are automatically bounded.
5. These are usually simply denoted with $\mathscr{G}_{>0}$, but I keep the $\gtrsim$ to highlight the *weak* dominance component, rather than that $g(\omega) > 0$ for all $\omega \in \Omega$.

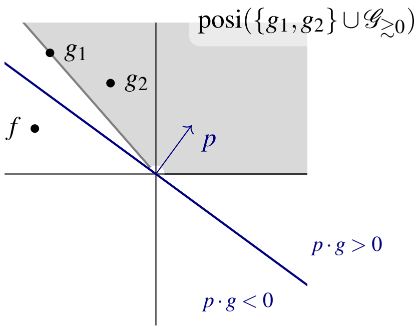

Figure 1: When $f \notin \text{posi}(\{g_1,\dots,g_n\} \cup \mathscr{G}_{\gtrsim 0})$, we can separate them by a hyperplane.

So we have shown that $\{p \mid p \cdot f > 0\} \supseteq \{p \mid p \cdot g_1 > 0\} \cap \dots \cap \{p \mid p \cdot g_n > 0\}$ implies $f \in \text{posi}(\{g_1,\dots,g_n\} \cup \mathscr{G}_{\gtrsim 0})$. Thus, for $f \in F_D$, there are some $g_1,\dots,g_n \in D$ with $f \in \text{posi}(\{g_1,\dots,g_n\} \cup \mathscr{G}_{\gtrsim 0})$; and so $f \in D$ by the coherence of D. So we have shown equation (4).

For the coherence of F_D we just need to use equation (4) and the coherence of D to observe that $\{p \mid p \cdot -1 > 0\} = \emptyset \notin F_D$. ∎

Theorem 7 tells us that we can represent all coherent sets of desirable gambles in the probability filter model, including those that are not representable by credal sets. Consider $\Omega = \{\omega_t, \omega_f\}$ and the coherent set of desirable gambles with $h := \langle \frac{1}{2}, -\frac{1}{2} \rangle \in D$, but also each $g_n \in D$ where $g_n := -h + \langle \frac{1}{n}, \frac{1}{n} \rangle = \langle -\frac{1}{2} + \frac{1}{n}, \frac{1}{2} + \frac{1}{n} \rangle$; as in figure 2.

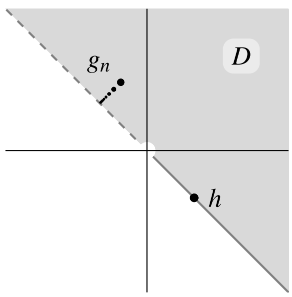

Figure 2: D with $\langle \frac{1}{2}, -\frac{1}{2} \rangle \in D$ and $\langle -\frac{1}{2} + \varepsilon, \frac{1}{2} + \varepsilon \rangle \in D$ for each $\varepsilon > 0$.

This set of desirable gambles cannot be captured in the credal set model, but it can be captured by a probability filter. Since $h \in D$,

$$\left\{ p \,\middle|\, p(\omega_t) > \frac{1}{2} \right\} \in F_D.$$

Since $f_n \in D$,

$$\left\{ p \,\middle|\, p(\omega_t) < \frac{1}{2} + \frac{1}{n} \right\} \in F_D.$$

This can be given a very natural gloss. You think that the probability exceeds $\frac{1}{2}$ but not by any particular amount. This filter is non-principal (indeed it is free: $\bigcap F_D = \emptyset$). This gives it non-Archimadean behaviour. If we were to ask what one thinks the probability of ω_t is, we would have to give a hyperreal value. The model based on probability filters does not have such hyperreals as values, i.e., it is based on standard-valued probabilities, but something like hyperreals are captured by the filter structure.

The probability filters model goes beyond the model of desirable gambles. It can, for example, capture arbitrary credal sets rather than just convex ones.

4. Probability Filters and Choice Functions

4.1. Choice Functions

We now consider a more general model of uncertainty: choice functions (cf. Seidenfeld et al., 2010). We follow De Bock and de Cooman (2018) and consider these under a desirability based characterisation so it easily extends the model of desirable gambles that we have just considered. (The connection to choice is spelled out in De Bock and de Cooman.)

Whereas the desirable gambles framework just considers whether an individual gamble is desirable or not, in this more general framework we also consider whether a set of gambles contains at least one desirable gamble. We model your uncertainty with a set of sets of gambles, $K \subseteq \wp(\mathscr{G})$, where a set of gambles (also called a gamble set), A, is in K if you think that at least one member of A is desirable.

Given a probability filter, F, we can extract a collection of gamble sets, K_F, with $A \in K_F$ if you think that at least one member of A is desirable, using F to determine what your opinions are. However, there are different ways to understand what it is to think that some member of A is desirable. Does there need to be some particular member of the set which you think is desirable, or does it suffice that you think that something is desirable without being able to identify any particular one? In the special case when filters are principal filters, this is equivalent to the question of whether a choice function is extracted from a credal set according to Maximality (Walley/Sen) or E-admissibility (Levi).

To apply the analogue of E-admissibility in this setting, there doesn't need to be a particular gamble that you think is desirable, so long as you think that at least one of them is.[6] This is captured by putting:

$$A \in K_F \text{ iff } \{p \mid \text{there is some } g \in A \text{ with } p \cdot g > 0\} \in F \tag{6}$$

To obtain Maximality rule, we could instead use:

$$A \in K_F^{\text{Max}} \text{ iff there is some } g \in A \text{ with } \{p \mid p \cdot g > 0\} \in F. \tag{7}$$

We will associate a filter with a collection of desirable gamble sets using equation (6). This is similar to Seidenfeld et al. (2010) who axiomatise choice functions associated with credal sets by E-admissibility. It is in contrast to De Bock and de Cooman (2018) who, following van Camp (2018), want their axioms to allow for both.[7]

We will now give analogous results to those offered in section 3. We give axioms for when a set of desirable gamble sets is coherent (including a mixing axiom but no Archimadean axiom) and show that the filter model can capture any coherent set of desirable gamble sets.

4.2. Axioms on Desirable Gamble Sets

Our axioms are based on De Bock and de Cooman (2018), with two important differences. Firstly, we include a so-called mixing axiom (our axiom K7). This is because of our choice to associate a filter with a set of desirable gamble sets using equation (6), generalising E-admissibility from credal sets. Secondly, we do not restrict attention to finite gamble sets.

Note that like De Bock and de Cooman, we do not impose any Archimadeanicity axiom.[8] This is in contrast to Seidenfeld et al. (2010), who are axiomatising choice functions obtained from credal sets. Our more general model of probability filters allows us to drop the Archimadean axiom.

The model of sets of desirable gamble sets (strictly) extends the desirable gambles model of section 3.

Before stating the axioms, there are two notions used in the axioms:

- $\text{posi}(B)$ is the set of positive finite linear combinations of members of B. That is:

$$\text{posi}(B) = \left\{ \sum_{i=1}^{n} \lambda_i g_i \,\middle|\, n > 0, \lambda_i > 0, g_i \in B \right\} \tag{8}$$

  This is the smallest convex cone extending B.

- $\text{clposi}(B)$ is the closure of $\text{posi}(B)$. This is the smallest closed and convex cone extending B.

Definition 9 *$K \subseteq \wp(\mathscr{G})$ is* coherent *if it satisfies*

K1. $\emptyset \notin K$

K2. *If $A \in K$ then $A \setminus \{0\} \in K$.*

K3. *If $g \in \mathscr{G}_{\gtrsim 0}$, then $\{g\} \in K$*

K4. *If $A \in K$ and $B \supseteq A$, then $B \in K$*

K5. *If $A \in K$ and for each $g \in A$, f_g is some gamble where $f_g \geq g$, then $\{f_g \mid g \in A\} \in K$.*

K6. *If $A_1, \ldots, A_n \in K$ and for each sequence $g_1 \in A_1, \ldots, g_n \in A_n$, $f_{\langle g_1,\ldots,g_n\rangle}$ is some member of $\text{posi}(\{g_1, \ldots, g_n\})$, then*

$$\{f_{\langle g_1,\ldots,g_n\rangle} \mid g_1 \in A_1, \ldots, g_n \in A_n\} \in K$$

K7. *If $A \in K$ and $\text{clposi}(B) \supseteq A \supseteq B$ then $B \in K$.*

Our axioms K1 to K6 are very close to the axioms of De Bock and de Cooman (2018), with two differences. Firstly, our axiom K6 is already generalised to apply to finitely many sets, rather than just for a pair A_1 and A_2. Secondly, we have added the axiom, K5, which says that if replace some members of a desirable gamble set by weak dominators, then the set remains desirable. In the setting where gamble sets must be finite these are both derivable from the De Bock and de Cooman axioms, but in the infinite setting their status is less clear. We will further discuss them in section 4.4.

Axiom K7 is the so-called mixing axiom. It restricts to special kinds of choice functions (see De Bock and de Cooman, 2019, §8). Our formulation of this axiom is similar to the formulation of De Bock and de Cooman (2019, Sec. 8) except that it follows Seidenfeld et al. (2010, axiom 2b) in taking the *closure*. This is not present in De Bock and de Cooman as they are restricting attention to finite sets of gambles. Further discussion of the formulation of this axiom is left to section 4.5, and we move to our results.

Our first result shows us that these axioms follow from our probability filter axioms, given the interpretation of equation (6).

Theorem 10 *If F is coherent, then K_F is coherent; where K_F is given by:*

$$A \in K_F \textit{ iff } \{p \mid \textit{there is some } g \in A \textit{ with } p \cdot g > 0\} \in F \tag{6}$$

In the proof, we use the notation

$$[\![A]\!] := \{p \mid \text{there is some } g \in A \text{ with } p \cdot g > 0\} \tag{9}$$

Proof

6. Linking to choice functions, this says you reject an option, o, from a set, O, if you think it is non-optimal, in the sense that $\{p \mid \text{there is some } o' \in O \text{ with } \text{Exp}_p U(o') > \text{Exp}_p U(o)\} \in F$. If the filter is principal, given by credal set $\mathbb{P}$, this is just when every $p \in \mathbb{P}$ has some $o' \in O$ with $\text{Exp}_p U(o') > \text{Exp}_p U(o)$.
7. To allow for both in the probability filter framework, we can hold the probability filter fixed as the model of belief but allow for varying choice procedures.
8. See also De Bock and de Cooman (2019, §9).

Axiom K1 follows from axiom F4 as $[\![\emptyset]\!] = \emptyset \notin F$.

Axiom K2: Note that $p \cdot 0 = 0$. So for every p, if $g \in A$ with $p \cdot g > 0$, then $g \in A \setminus \{0\}$; thus $[\![A]\!] \subseteq [\![A \setminus \{0\}]\!]$; so $[\![A]\!] \in F$ implies $[\![A \setminus \{0\}]\!] \in F$ by axiom F3, as required.

Axiom K3: If $g \in \mathscr{G}_{\gtrsim 0}$, then every $p \in \mathrm{rProbs}$ has $p \cdot g > 0$, so $[\![\{g\}]\!] = \mathrm{rProbs} \in F$ (axioms F1 and F3).

Axiom K4 follows from axiom F3 because $B \supseteq A$ implies $[\![B]\!] \supseteq [\![A]\!]$.

Axiom K5: observe that if $p \cdot g > 0$ then $p \cdot f_g > 0$, so it holds by axiom F3.

Axiom K6: If $A_1, \ldots, A_n \in K_F$, then $[\![A_1]\!], \ldots, [\![A_n]\!] \in F$ so by axiom F2, $[\![A_1]\!] \cap \ldots \cap [\![A_n]\!] \in F$. For any $p \in [\![A_1]\!] \cap \ldots \cap [\![A_n]\!]$, there is some $g_1 \in A_1, \ldots, g_n \in A_n$ with $p \cdot g_i > 0$ for each i, and thus for $f_{\langle g_1, \ldots, g_n \rangle} \in \mathrm{posi}(\{g_1, \ldots, g_n\})$, also $p \cdot f_{\langle g_1, \ldots, g_n \rangle} > 0$. So it follows from axiom F3.

Now consider axiom K7. We consider the closure and positive hull parts separately, observing that they can be combined to obtain axiom K7 (see proposition 14).

Let $\mathrm{posi}(B) \supseteq A$ and $[\![A]\!] \in F$. We need to show that $[\![B]\!] \in F$. By axiom K4, it suffices to show that $[\![A]\!] \subseteq [\![B]\!]$. So we need to show that if $p \cdot f > 0$ for some $f \in A$ then $p \cdot g > 0$ for some $g \in B$. Since we have assumed that $A \subseteq \mathrm{posi}(B)$, if $f \in A$ then there are some $g_1, \ldots, g_n \in B$ and λ_i positive with $f = \lambda_1 g_1 + \ldots + \lambda_n g_n$. If $p \cdot f > 0$, then also $p \cdot g_i > 0$ for some i. And thus we have found some $g \in B$ with $p \cdot g > 0$, as required.

If $f \in A \subseteq \mathrm{cl}(B)$, then there is a sequence, $\langle g_n \rangle$, of members of B with g_n converging to f. If $p \cdot f > 0$, then there is some g_m (in fact a tail) with $p \cdot g_m > 0$. So again $[\![A]\!] \subseteq [\![B]\!]$, as required. ■

4.3. Expressive Power of Probability Filters

The probability filter model captures all the representational power of sets of desirable gamble sets, as given by our axioms, which include the mixing axiom.

We show that for any coherent set of gamble sets, we can find a filter which evaluates exactly those gamble sets as desirable.

Theorem 11 *For any coherent K,*

$$F_K := \mathrm{ext}(\{p \mid \text{there is some } g \in A \text{ with } p \cdot g > 0\} \mid A \in K) \tag{10}$$

is coherent with

$$B \in K \text{ iff } \{p \mid \text{there is some } g \in B \text{ with } p \cdot g > 0\} \in F_K \tag{11}$$

Thus, for distinct coherent K and K', F_K and $F_{K'}$ are distinct. And $K_{F_K} = K$.

The proof extends the proof used for theorem 8; but we will first state a lemma:

Lemma 12 *Suppose K is coherent. If $A_1, \ldots, A_n \in K$ and for each sequence $g_1 \in A_1, \ldots, g_n \in A_n$, we have some $f_{\langle g_1, \ldots, g_n \rangle} \in \mathrm{clposi}(B) \cap \mathrm{posi}(\{g_1, \ldots, g_n\} \cup \mathscr{G}_{\gtrsim 0})$, then $B \in K$.*

Proof If $f_{\langle g_i \rangle} \in \mathrm{posi}(\{g_1, \ldots, g_n\} \cup \mathscr{G}_{\gtrsim 0})$, then unless $f_{\langle g_i \rangle} \in \mathscr{G}_{\gtrsim 0}$, there is some $h_{\langle g_i \rangle} \in \mathrm{posi}(\{g_1, \ldots, g_n\})$ with $f_{\langle g_i \rangle} \geq h_{\langle g_i \rangle}$.

If there is some $f \in \mathrm{clposi}(B)$ with $f \in \mathscr{G}_{\gtrsim 0}$, then $\{f\} \in K$ (axiom K3) so $\mathrm{clposi}(B) \in K$ (axiom K4), thus $B \in K$ (axiom K7).

So suppose we have each $f_{\langle g_i \rangle} \notin \mathscr{G}_{\gtrsim 0}$, and let $h_{\langle g_i \rangle}$ be as above. By axiom K6, $\{h_{\langle g_i \rangle} \mid g_1 \in A_1, \ldots, g_n \in A_n\} \in K$; and then by axiom K5, $\{f_{\langle g_i \rangle} \mid g_1 \in A_1, \ldots, g_n \in A_n\} \in K$. And thus $\mathrm{clposi}(B) \in K$ by axiom K4. So $B \in K$ by axiom K7. ■

Proof of Theorem 11. We continue to use the notation $[\![B]\!] := \{p \mid \text{there is some } g \in A \text{ with } p \cdot g > 0\}$. Any $B \in K$ has $[\![B]\!] \in F_K$ by construction, which is one direction of equation (11). We need to show the other direction.

By definition of ext and F_K, $[\![B]\!] \in F_K$ iff there are $A_1, \ldots, A_n \in K$ with $[\![B]\!] \supseteq [\![A_1]\!] \cap \ldots \cap [\![A_n]\!]$. We will now show that if $[\![B]\!] \supseteq [\![A_1]\!] \cap \ldots \cap [\![A_n]\!]$, then for each sequence $g_1 \in A_1, \ldots, g_n \in A_n$ there is some $f_{\langle g_i \rangle} \in \mathrm{clposi}(B) \cap \mathrm{posi}(\{g_1, \ldots, g_n\} \cup \mathscr{G}_{\gtrsim 0})$, which will let us use Lemma 12 to get that $B \in K$, as required.

Suppose $\mathrm{clposi}(B)$ and $C := \mathrm{posi}(\{g_1, \ldots, g_n\} \cup \mathscr{G}_{\gtrsim 0})$ are disjoint. Since we specified $\mathscr{G}_{\gtrsim 0}$ to be specified by weak dominance, these are both closed cones when 0 is added, and are disjoint except at 0. Thus we can use a separation result for convex cones (Klee, 1955, Theorem 2.7) to find a linear functional separating them (see figure 3).[9] By

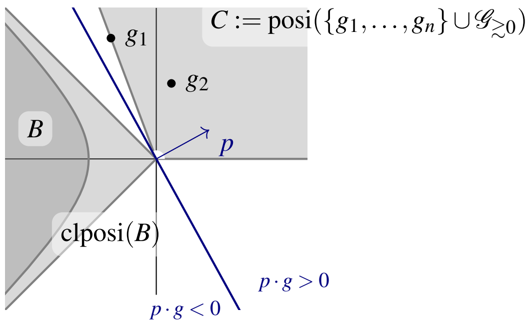

Figure 3: When $\mathrm{clposi}(B)$ and $C := \mathrm{posi}(\{g_1, \ldots, g_n\} \cup \mathscr{G}_{\gtrsim 0})$ are disjoint they can be separated.

normalising this, we have a probabilistic p with $p \cdot g \geq 0$ for each $g \in C$ and $p \cdot f \leq 0$ for each $f \in \mathrm{clposi}(B)$. Since $0 \in \mathrm{clposi}(B)$ and we assumed C and $\mathrm{clposi}(B)$ are disjoint, we have that $0 \notin C$, and thus for every $g \in C$, in fact $p \cdot g > 0$. Since each $\mathbf{I}_{\{\omega\}} \in C$, each $p(\omega)$ is positive, so we have our regular probabilistic p with $p \cdot g_i > 0$ for $g_1, \ldots, g_n$, but

9. Local compactness of C follows from the local compactness of $\mathbb{R}^\Omega$ and the fact that C is closed.

$p \cdot f \leq 0$ for each $f \in B$. Thus $p \in [\![A_1]\!] \cap \ldots \cap [\![A_n]\!]$ but $p \notin [\![B]\!]$, so $[\![B]\!] \not\supseteq [\![A_1]\!] \cap \ldots \cap [\![A_n]\!]$.

So we have shown that if $[\![B]\!] \supseteq [\![A_1]\!] \cap \ldots \cap [\![A_n]\!]$, then for each sequence $g_1 \in A_1, \ldots, g_n \in A_n$ there is some $f_{\langle g_i \rangle} \in \text{clposi}(B)$ with $f \in \text{posi}(\{g_1, \ldots, g_n\} \cup \mathscr{G}_{\gtrsim 0})$. By lemma 12, this implies $B \in K$. So we have shown equation (11).

For the coherence of F_K we just need to use equation (11) and the coherence of K to observe that $[\![\{-1\}]\!] = \emptyset \notin F_K$. ■

4.4. Our Alternative Axioms

We now discuss how our main axioms diverge from those of De Bock and de Cooman (2018).

Axiom K6 We have formulated axiom K6 for finitely many sets instead of the formulation of De Bock and de Cooman (2018), who state it as it applies just to a pair of sets (their K_3):

- If $A_1, A_2 \in K$ and for each $g_1 \in A_1$, $g_2 \in A_2$, $f_{\langle g_1, g_2 \rangle}$ is some member of $\text{posi}(\{g_1, g_2\})$, then $\{f_{\langle g_1, g_2 \rangle} \mid g_1 \in A_1, g_2 \in A_2\} \in K$.

When the gamble sets must be finite, it can be iterated to obtain axiom K6. This is a consequence of their representation result (De Bock and de Cooman, 2018, Theorem 7). But since we are working in the infinite setting, we do not have access to these results and so we have opted to simply include the generalisation directly. It is an open question whether it is derivable from the pair-version in the infinite setting.

One might consider extending axiom K6 to infinitely many sets:

✗ If $A_i \in K$ for each $i \in I$ (possibly infinite) and for each sequence $\langle g_i \rangle_{i \in I}$ with each $g_i \in A_i$, $f_{\langle g_i \rangle}$ is some member of $\text{posi}(\{g_i \mid i \in I\})$, then

$$\{f_{\langle g_i \rangle} \mid \langle g_i \rangle \in \prod_{i \in I} A_i\} \in K$$

This is adopted in De Bock (2020). But it is not valid when one uses non-principal filters and the E-admissibility-style interpretation given by equation (6).

Proof Consider some F as in section 3, with $\{p \mid p(\omega_t) > \frac{1}{2}\} \in F$ and each $\{p \mid p(\omega_t) < \frac{1}{2} + \frac{1}{n}\} \in F$, (see figure 4).

This will have each singleton $\{g_n\} \in K_F$.

It also has $\{-g_n \mid n \in \mathbb{N}\} \in K_F$. To see this, note that for p with $p(\omega_t) > \frac{1}{2}$, when $\frac{1}{n} < p(\omega_t) - \frac{1}{2}$, then $p \cdot -g_n > 0$. Thus $[\![\{-g_n \mid n \in \mathbb{N}\}]\!] \in F$.

This leads to a violation of the infinite extension of axiom K6: Any sequence selecting a member of each of the (infinitely many) singletons $\{g_n\}$ as well as a member of $\{-g_n \mid n \in \mathbb{N}\}$ includes some $-g_n$ but also includes each

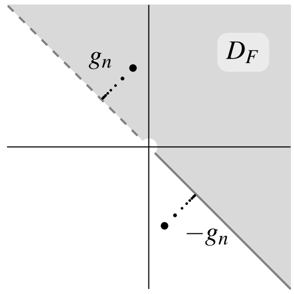

Figure 4: $g_n = \langle -\frac{1}{2} + \frac{1}{n}, \frac{1}{2} + \frac{1}{n} \rangle$.

g_n; so $0 = g_n + -g_n$ is in each relevant posi but $\{0\} \notin K_F$. ■

Taking Dominators - Axiom K5 Our axiom K5 is used in lemma 12 to allow us to replace members of a set by weak dominators and remain desirable. This has a clear intuitive motivation: it can only improve the situation.

When gamble sets are finite, one can use De Bock and de Cooman (2019, Lemma 34) to see that it follows from the other axioms. They show that one can use axioms K3 and K6 to replace a single member of a desirable set with a dominator. In the finite setting, this can be iterated to derive axiom K5, but not when can be infinite.

In fact, it is nonetheless derivable from the other axioms when gamble sets can be infinite but where Ω is finite.

Proposition 13 *When Ω is finite, axiom K5 follows from axioms K3 and K6.*

Proof Let $\Omega = \{\omega_1, \ldots, \omega_n\}$. If $f_g \geq g$ then $f_g \in \text{posi}(\{g, \mathbf{I}_{\omega_1}, \ldots, \mathbf{I}_{\omega_n}\})$. So since each $\{\mathbf{I}_{\omega_i}\} \in K$ by axiom K3, also $\{f_g \mid g \in A\} \in K$ by axiom K6. ■

Since we have assumed that Ω is finite, we could omit this axiom. However, it is not clear that an analogous argument will work when Ω is infinite. It is also not clear that an analogous argument will work when regularity is dropped and we use strict dominance for axiom K3. Since I do not want to unnecessarily restrict the results, I have added this axiom.

We could have opted to include this strength in other ways, which some readers may find more natural. In particular, we could add it by strengthening axiom K6 by instead of requiring $f_{\langle g_1, \ldots, g_n \rangle}$ be in $\text{posi}(\{g_1, \ldots, g_n\})$, allowing it to be in $\text{posi}(\{g_1, \ldots, g_n\}) + \mathscr{G}_{\geq 0}$, or directly $\text{posi}(\{g_1, \ldots, g_n\} \cup \mathscr{G}_{\gtrsim 0})$.

It would also follow if we extend axiom K6 (with just $\text{posi}(\{g_1, \ldots, g_n\})$) to apply to infinitely many sets. But as we showed in section 4.4, this is not valid.

4.5. The Formulation of the Mixing Axiom - Axiom K7

We have adopted the mixing axiom in the form:

K7. If $A \in K$ and $\mathrm{clposi}(B) \supseteq A \supseteq B$ then $B \in K$.

There are various closely related axioms that differ just in which operation they use in place of clposi. For O an operator such as conv, posi, clconv, let M_O be the axiom:

M_O. If $A \in K$ and $O(B) \supseteq A \supseteq B$ then $B \in K$.

Given axiom K4, we could equivalently state any of these as

M_O. If $O(B) \in K$ then $B \in K$.

De Bock and de Cooman (2019) choose to use the axiom $\mathrm{M}_{\mathrm{posi}}$ instead of $\mathrm{M}_{\mathrm{conv}}$. Similarly, we have chosen to use $\mathrm{M}_{\mathrm{clposi}}$ instead of Seidenfeld et al. (2010)'s use of $\mathrm{M}_{\mathrm{clconv}}$. (De Bock and de Cooman can omit the closure component because they restrict to finite sets of gambles.) The choice of posi or conv doesn't make a difference.

Proposition 14 *Given axioms K4 and K6, the following are equivalent:*

1. $\mathrm{M}_{\mathrm{clposi}} =$ *axiom* K7
2. M_{cl} *and* $\mathrm{M}_{\mathrm{posi}}$
3. $\mathrm{M}_{\mathrm{clconv}}$
4. M_{cl} *and* $\mathrm{M}_{\mathrm{conv}}$

Proof Item 2 $\Longrightarrow$ item 1: $\mathrm{clposi}(B) = \mathrm{cl}(\mathrm{posi}(B))$. If $\mathrm{clposi}(B) \in K$, then by M_{cl}, $\mathrm{posi}(B) \in K$ and by $\mathrm{M}_{\mathrm{posi}}$, $B \in K$.

Item 1 $\Longrightarrow$ item 2: If $\mathrm{cl}(B) \in K$, then $\mathrm{clposi}(B) \in K$ by axiom K4, so $B \in K$. Similarly, if $\mathrm{posi}(B) \in K$ then $\mathrm{clposi}(B) \in K$ so $B \in K$.

The argument for item 3 $\Longleftrightarrow$ item 4 is exactly analogous.

For item 4 $\Longleftrightarrow$ item 2, we need to show that $\mathrm{conv}(B) \in K$ iff $\mathrm{posi}(B) \in K$. Since $\mathrm{posi}(B) \supseteq \mathrm{conv}(B)$, one direction follows from axiom K4. For the other direction, assume $\mathrm{posi}(B) \in K$. For each $g \in \mathrm{posi}(B)$ there is some $f_g \in \mathrm{conv}(B)$ where $f_g \in \mathrm{posi}(\{g\})$ (it can simply be normalised). So by axiom K6, $\mathrm{conv}(B) \in K$. ■

It is worth noting that for $\mathrm{scalar}(B) = \{\lambda_g g \mid \lambda_g > 0, g \in B\}$, $\mathrm{M}_{\mathrm{scalar}}$ follows directly from axiom K6.

5. Beyond

An encoding of which gamble sets are desirable, using equation (6), does not suffice to tell us everything about the opinion state, as given by a probability filter. This contrasts to the case of credal sets, where Archimadean (mixing) choice functions are equivalent to sets of probabilities (Seidenfeld et al., 2010). The expressive power of probability filters goes strictly beyond that of sets of desirable gamble sets.

One example of this can be given by considering weak desirability. In the desirable gambles model, we only directly encode whether a gamble is strictly preferable to the status quo, or equivalently, whether one gamble is strictly preferred to another. From this, we cannot always recover judgements of weak desirability, or whether one gamble is weakly preferred to another.

In section 3 we considered a case where you think that the probability exceeds $\frac{1}{2}$ but not by any particular amount. The desirable gambles that this leads to are identical to those where you think the probability does not exceed $\frac{1}{2}$ by any particular amount but leave open whether it is strictly above $\frac{1}{2}$ or equal to $\frac{1}{2}$. The move to sets of desirable gamble sets also does not help distinguish these. It cannot distinguish what Schervish et al. (ms) call semi-preference from indifference.

We could add such expressive power by extending the desirability based accounts to also explicitly include judgements of weak desirability. (See also Quaeghebeur et al., 2015.) For the gamble sets framework this means encoding not only when you judge that the set contains a (strictly) desirable gamble but also whether you judge it contains a weakly desirable gamble. Equivalently we could encode judgements on whether the set of gambles contains only (strictly) desirable gambles.

This would allow the expressive power of the desirability based frameworks to come closer to that of probability filters. However, it will still fall short of the full expressive power of probability filters. Consider again the filter where you think that the probability exceeds $\frac{1}{2}$ but not by any particular amount. There are many distinct refinements of it (though we need the axiom of choice to obtain them). For example, an ultrafilter needs to specify whether you think that the probability is in $\{.51, .501, .5001, .50001, \dots\}$ or not.

I conjecture that any such refinements do not distinguish any matters of whether sets contain some or only contain strictly or weakly desirable gambles.

A more detailed story of exactly what it means to think that the probability is in $\{.51, .501, .5001, \dots\}$ or be represented by F with $\{p \mid p(\omega_t) \in \{.51, .501, .5001, \dots\}\} \in F$ remains future work.

In the desirability-based spirit, one proposal for aiding this interpretation question is to also consider whether a set of gambles contains at least one gamble that's equivalent to the status quo. I conjecture that this allows the full power of probability filters to be captured.

6. Conclusion

We have proposed representing beliefs by probability filters. It is a natural model that directly captures probabilistic judgements; and it is easy to work with.

It directly extends the credal sets framework, and we have shown that it captures the power of the desirable gambles framework, thus providing a nice unifier. Further models such as comparative frameworks are also easily captured in the framework.

In summary, it is a natural and powerful framework for modelling uncertain belief.

Acknowledgments

These ideas were developed in joint work with Jason Konek (Campbell-Moore and Konek, 2019) and I am hugely grateful to him. Thanks also to Arthur van Camp, Kevin Blackwell and Teddy Seidenfeld for discussions, and to the referees for ISIPTA for detailed and helpful comments.

References

Catrin Campbell-Moore and Jason Konek. Believing Probabilistic Contents: On the Expressive Power and Coherence of Sets of Sets of Probabilities. *Analysis*, 80(2): 316–331, 11 2019.

Jasper De Bock. Choice functions based on sets of strict partial orders: an axiomatic characterisation. *arXiv*, 2003.11631v2, 2020.

Jasper De Bock and Gert de Cooman. A desirability-based axiomatisation for coherent choice functions. In S. Destercke, T. Denoeux, M.A. Gil, P. Grzergorzewski, and O. Hryniewicz, editors, *Uncertainty Modelling in Data Science: Proceedings of SMPS*, pages 46–53. Springer, 2018.

Jasper De Bock and Gert de Cooman. Interpreting, axiomatising and representing coherent choice functions in terms of desirability. In *International Symposium on Imprecise Probabilities: Theories and Applications*, pages 125–134. PMLR, 2019.

VL Klee. Separation properties of convex cones. *Proceedings of the American Mathematical Society*, 6(2): 313–318, 1955.

Sarah Moss. *Probabilistic knowledge*. Oxford University Press, 2018.

Erik Quaeghebeur, Gert De Cooman, and Filip Hermans. Accept & reject statement-based uncertainty models. *International Journal of Approximate Reasoning*, 57:69–102, 2015.

Mark J Schervish, Teddy Seidenfeld, Ruobin Gong, Joseph B Kadane, and Rafael B Stern. When no price is righ, ms.

Teddy Seidenfeld, Mark J Schervish, and Joseph B Kadane. Coherent choice functions under uncertainty. *Synthese*, 172(1):157–176, 2010.

Arthur van Camp. *Choice functions as a tool to model uncertainty*. PhD thesis, Ghent University, Faculty of Engineering and Architecture., 2018.

Peter Walley. Towards a unified theory of imprecise probability. *International Journal of Approximate Reasoning*, 24(2-3):125–148, 2000.