Constriction for sets of probabilities Supplementary material

Appendix A. Dubins-deFinetti conditioning

In this section, we show how if we are willing to depart from the classical Kolmogorovian paradigm of probabilities, then we have additional opportunities for constriction. In particular, as we shall see, Bayes' rule can induce constriction if we allow probabilities to be merely finitely additive.

Suppose that we adopt Dubins-deFinetti conditioning (DdFC) framework; an in-depth exposition of DdFC can be found in [10, 14, 35]. Notice also that in [9], the authors compare generalized Bayes', geometric and Dempster's rules for belief functions in DdFC framework. For the sake of the present work, the two main differences with respect to the Kolmogorovian framework is that probabilities need not be countably additive, and that conditioning does not happen on sigma-fields, but rather on events, members of a partition of the state space. The following statements are true; we will provide illustrations for the first one, and the second one is shown similarly.

1. If probability measures are finitely but not countably additive, then constriction can take place for all the elements of a countable partition;
2. Recall that a probability P is completely additive if the measurable union of a set of P-null events is P-null. If probability measures are countably but not completely additive, then constriction can take place for all the elements of an uncountable partition.

Notice also that if probability measures are completely additive, then they must be discrete.

Definition 26 *We say that probability measure P is conglomerable in partition $\mathcal{E}$ when for every event A such that $P(A \mid E)$ is defined for all $E \in \mathcal{E}$, and for all constants k_1, k_2, if $k_1 \leq P(A \mid E) \leq k_2$ for all $E \in \mathcal{E}$, then $k_1 \leq P(A) \leq k_2$.*

Definition 26 asserts that for each event A, if all the conditional probabilities over a partition $\mathcal{E}$ are bounded by two quantities, k_1 and k_2, then the unconditional probability for that event is likewise bounded by these two quantities [25]. De Finetti [10] shows the non-conglomerability of finitely additive probability measures (FAPMs) in denumerable partitions.

Assume that instead of requiring $\mathcal{P}$ to be a set of countably additive probabilities, we allow it to be a set of FAPMs. Then, weak and strict constriction can happen by Bayes updating $\mathcal{P}$ thanks to the non-conglomerability property of FAPMs. The two illustrations that we present in this section build on the example in [14, page 92], which we state here for motivating their construction.

Example 3 ***(Dubins)*** *Let $\Omega = \{A, B\} \times \{N = 1, 2, \dots\}$. Stipulate that*

- *$P(A) = P(B) = 1/2$,*
- *$P(N = n \mid A) = 2^{-n}$, for $n \in \{1, 2, \dots\}$, a countably additive conditional probability,*
- *$P(N = n \mid B) = 0$, for $n \in \{1, 2, \dots\}$, a strongly finitely additive conditional probability.* [8]

Then, $P(N = n) = 2^{-(n+1)} > 0$, for $n \in \{1, 2, \dots\}$, and (marginally) P is merely finitely additive over the subalgebra generated by the partition $\mathcal{E}_N = \{\{N = 1\}, \{N = 2\}, \dots\}$. P displays non-conglomerability for the event A in the partition $\mathcal{E}_N = \{\{N = 1\}, \{N = 2\}, \dots\}$ as $P(A) = 1/2$ and $P(A \mid N = n) = 1$, for $n \in \{1, 2, \dots\}$.

Illustration 1 (weak constriction)

Use Example 3 as follows. Consider a set $\mathcal{P}$ of probabilities on $\Omega = \{A, B\} \times \{N = 1, 2, \dots\}$ such that $\mathcal{P} = \{P_\alpha, 0 < \alpha \leq 1\}$, where

- $P_\alpha(A) = \alpha$,
- $P_\alpha(N = n \mid A) = 2^{-n}$, for $n \in \{1, 2, \dots\}$, a countably additive conditional probability,
- $P_\alpha(N = n \mid B) = 0$, for $n \in \{1, 2, \dots\}$, a strongly finitely additive conditional probability.

Note that neither $P_\alpha(N = n \mid A)$ nor $P_\alpha(N = n \mid B)$ depend upon α. With respect to $\mathcal{P}$, we have $0 < P_\alpha(A) \leq 1$. For each $0 < \alpha < 1$, we have non-conglomerability of P_α for the event A in the partition $\mathcal{E}_N$ as $P_\alpha(A \mid N = n) = 1$, for $n \in \{1, 2, \dots\}$. Observe then that $1 = P_1(A) = P_1(A \mid N = n)$, for $n \in \{1, 2, \dots\}$. Thus, Bayes-updating using the information $\{N = n\}$ from the partition $\mathcal{E}_N$ weakly-constricts $\mathcal{P} = \{P_\alpha, 0 < \alpha \leq 1\}$.

Illustration 2 (strict constriction)

Modify Example 3 as follows. Consider a set $\mathcal{P}$ of probabilities on $\Omega = \{A, B\} \times \{N = 1, 2, \dots\}$ such that $\mathcal{P} = \{P_\alpha, 0 < \alpha < 1\}$, where

- $P_\alpha(A) = \alpha$, so that with respect to $\mathcal{P}$, we have $0 < P_\alpha(A) < 1$,
- $P_\alpha(N = n \mid A) = (1 - \alpha)2^{-n}$, for $n \in \{1, 2, \dots\}$,
- $P_\alpha(N = n \mid B) = \alpha 2^{-n}$, for $n \in \{1, 2, \dots\}$.

[8] Recall that a FAPM is strongly finitely additive if it admits countable partitions by null sets [1].

Note that each of these two conditional probabilities is a merely finitely additive probability distribution over N that depends on α. In addition, observe that $P(N = n) = 2\alpha(1 - \alpha)2^{-n} > 0$, which (for each α) also is a merely finitely additive probability distribution over N.

By a Bayes' updating, for each $0 < \alpha < 1$ and each $n \in \{1, 2, \ldots\}$, $P_\alpha(A \mid N = n) = 1/2$. That is, for each $\alpha \neq 1/2$ with $0 < \alpha < 1$, there is non-conglomerability of P_α for the event A in the partition $\mathcal{E}_N$. Whereas, $1/2 = P_{1/2}(A) = P_{1/2}(A \mid N = n)$. Thus, Bayes-updating using the information $\{N = n\}$ from the partition $\mathcal{E}_N$ strictly-constricts $\mathcal{P} = \{P_\alpha, 0 < \alpha < 1\}$.

Remark 27 *These two illustrations help to explain why Propositions 4 and 5 are restricted to countably additive probabilities.*

Appendix B. Proofs

Proof [Proof of Proposition 4] Let $\underline{P}$ be a probability function that satisfies $\underline{P}(A) = \min_{P \in \mathcal{P}} P(A)$. By Lemma 3, if $\underline{P}(\mathbf{X}^{A+}_{\underline{P}}) > 0$, then $\underline{P}(\mathbf{X}^{A-}_{\underline{P}}) > 0$ and (B, E) does not strictly uniformly constrict A. That is, for each $P_{1i}(A) \in \mathcal{P}(A), \underline{P}(A) \leq P_{1i}(A)$, so (B, E) does not strictly uniformly constrict A when $\underline{P}(A) = P_{1i}(A)$. Hence, if (B, E) weakly uniformly constricts A, we have that $\underline{P}(\mathbf{X}^{A+}_{\underline{P}}) = 0$, and then $\underline{P}(\{x \in \mathbf{X} : \underline{P}^B(A \mid X = x) = \underline{P}(A)\}) = 1$. Let now $\overline{P}$ be a probability function that satisfies $\overline{P}(A) = \max_{P \in \mathcal{P}} P(A)$. By the same reasoning, $\overline{P}(\{x \in \mathbf{X} : \overline{P}^B(A \mid X = x) = \overline{P}(A)\}) = 1$, and then (B, E) does not weakly uniformly constrict A either. ■

Proof [Proof of Proposition 5] Assume (for a reductio proof) that on a set of X values with $\mathcal{P}$-measure 1 (i.e. with P-probability 1, for all $P \in \mathcal{P}$), for each x_i there exist $P_{1i}(A)$ and $P_{2i}(A)$ in $\mathcal{P}(A)$ such that for each $P(A \mid x_i) \in \mathcal{P}(A \mid x_i)$, either $P_{1i}(A) < P(A \mid x_i) \leq P_{2i}(A)$ or $P_{1i}(A) \leq P(A \mid x_i) < P_{2i}(A)$. Since X is a simple random variable, define $P_1(A) := \min_i P_{1i}(A)$ and $P_2(A) := \max_i P_{2i}(A)$. Given Proposition 4, assume $\mathcal{P}(A)$ is not a closed set. Without loss of generality, assume it is open below (the reasoning is parallel if $\mathcal{P}(A)$ is open above). So, $\underline{P}(A) < P_1(A)$. Then, there exists $P_0 \in \mathcal{P}$ with $\underline{P}(A) < P_0(A) < P_1(A)$. Since for each i, $P_{1i}(A) < P_0(A \mid x_i)$ or $P_{1i}(A) \leq P_0(A \mid x_i)$, we also have that for each i, $P_0(A) < P_1(A) \leq P_0(A \mid x_i)$. But then $P_0(\mathbf{X}^{A+}_{P_0}) = 1$, which is a contradiction according to Lemma 3. ■

Proof [Proof of Theorem 7] Immediate from Definition 1 and Theorem 6. ■

Proof [Proof of Theorem 8] We first show that the lower probability (LP) $\inf_{P \in \text{Conv}(\mathcal{P})} P(\cdot)$ of the convex hull of $\mathcal{P}$ and the LP $\inf_{P \in \text{ex}[\text{Conv}(\mathcal{P})]} P(\cdot)$ of the extrema of the convex hull of $\mathcal{P}$ coincide. To see this, pick any $A \in \mathcal{F}$. Since $\text{ex}[\text{Conv}(\mathcal{P})] \subseteq \text{Conv}(\mathcal{P})$, we have that

$$\inf_{P \in \text{Conv}(\mathcal{P})} P(A) \leq \inf_{P \in \text{ex}[\text{Conv}(\mathcal{P})]} P(A). \tag{9}$$

Then, let $\emptyset \neq \text{ex}[\text{Conv}(\mathcal{P})] = \{P^{ex}_j\}_{j \in \mathcal{J}}$. For all $P \in \text{Conv}(\mathcal{P})$ and all $A \in \mathcal{F}$, we have that

$$\begin{aligned} P(A) = \sum_{j \in \mathcal{J}} \alpha_j P^{ex}_j(A) &\geq \sum_{j \in \mathcal{J}} \alpha_j \underline{P}^{ex}(A) \\ &= \underline{P}^{ex}(A) := \inf_{P \in \text{ex}[\text{Conv}(\mathcal{P})]} P(A), \end{aligned}$$

where $\{\alpha_j\}_{j \in \mathcal{J}}$ is a collection of positive reals such that $\sum_{j \in \mathcal{J}} \alpha_j = 1$, which implies that

$$\inf_{P \in \text{Conv}(\mathcal{P})} P(A) \geq \inf_{P \in \text{ex}[\text{Conv}(\mathcal{P})]} P(A). \tag{10}$$

By combining together (9) and (10) we obtained the desired equality.

Now, if $P^\star \in \text{ex}[\text{Conv}(\mathcal{P})]$, then there might be a collection $\{\tilde{A}\} \subseteq \mathcal{F}$ for which $P^\star(\tilde{A}) = \underline{P}(\tilde{A})$ or $P^\star(\tilde{A}) = \overline{P}(\tilde{A})$, so the constriction is weak for the elements of the collection, while $P^\star(A) > \underline{P}(\tilde{A})$ and $P^\star(A) < \overline{P}(\tilde{A})$, for all $A \in \mathcal{F} \setminus \{\tilde{A}\}$. If instead $P^\star = \sum_{j \in \mathcal{J}} \alpha_j P^{ex}_j$, $\alpha_j > 0$ for all j, then $P^\star(A) \in (\underline{P}(A), \overline{P}(A))$, for all $A \in \mathcal{F}$, so we have $\boldsymbol{\jmath} \multimapinv A$, for all $A \in \mathcal{F}$. ■

Proof [Proof of Theorem 9] If $\mathcal{P}(A)$ is closed in the Euclidean topology and $P^\star(A) \in \partial_{\mathcal{B}([0,1])}\mathcal{P}(A)$, then $P^\star(A) = \underline{P}(A)$ or $P^\star(A) = \overline{P}(A)$, so the constriction is weak. If instead $P^\star(A) \in \text{int}_{\mathcal{B}([0,1])}\mathcal{P}(A)$, then $P^\star(A) \in (\underline{P}(A), \overline{P}(A))$, so we have $\boldsymbol{\jmath} \multimapinv A$. ■

Proof [Proof of Lemma 20] This proof draws on that of [19, Lemma 5.1]. Fix any $A \in \mathcal{F}$. Because $\mathcal{P}$ is closed, there exists $P_{(A)} \in \mathcal{P}$ such that $P_{(A)}(A) = \underline{P}(A)$. Notice that subscript (A) reminds us that this probability measure can vary with the choice of A. Then, we have that

$$\begin{aligned} \underline{P}(A) = P_{(A)}(A) &= \sum_{E \in \mathcal{E}} P_{(A)}(A \mid E) P_{(A)}(E) \\ &\geq \sum_{E \in \mathcal{E}} \underline{P}^\times(A \mid E) P_{(A)}(E) \\ &\geq \sum_{E \in \mathcal{E}} \inf_{E \in \mathcal{E}} \underline{P}^\times(A \mid E) P_{(A)}(E) \\ &= \inf_{E \in \mathcal{E}} \underline{P}^\times(A \mid E) \sum_{E \in \mathcal{E}} P_{(A)}(E) \\ &= \inf_{E \in \mathcal{E}} \underline{P}^\times(A \mid E). \end{aligned}$$

The same argument applies for the upper probability of A, that is, if we pick $P'_{(A)} \in \mathcal{P}$ such that $P'_{(A)}(A) = \overline{P}(A)$,

$P'_{(A)}$ possibly different from $P_{(A)}$, then

$$\overline{P}(A) \leq \sum_{E \in \mathcal{E}} \overline{P}^{\times}(A \mid E) P'_{(A)}(E) \leq \sup_{E \in \mathcal{E}} \overline{P}^{\times}(A \mid E).$$

■

Proof [Proof of Theorem 21] Immediate from Lemma 20. ■

Proof [Proof of Theorem 22] This proof comes from that of [36, Theorem 2.3]. Fix an event $A \in \mathcal{F}$ of interest, and let $\times \in \{B, G\}$. Pick any $P_{E_1 \cdots E_{t-k}} \in \mathcal{P}_{\star E_1 \cdots E_{t-k}}(A) \cap \Sigma^-(A, E_{t-k+1} \cap \cdots \cap E_t)$. Then, we have that $P_{E_1 \cdots E_{t-k}}(A) = \underline{P}_{E_1 \cdots E_{t-k}}(A)$ because $P_{E_1 \cdots E_{t-k}} \in \mathcal{P}_{\star E_1 \cdots E_{t-k}}(A)$ and $P_{E_1 \cdots E_{t-k}}(A \cap E_{t-k+1} \cap \cdots \cap E_t) < P_{E_1 \cdots E_{t-k}}(A) P_{E_1 \cdots E_{t-k}}(E_{t-k+1} \cap \cdots \cap E_t)$ because $P_{E_1 \cdots E_{t-k}} \in \Sigma^-(A, E_{t-k+1} \cap \cdots \cap E_t)$. Then,

$$\begin{aligned}
\underline{P}_{E_1 \cdots E_{t-k}}(A) &= P_{E_1 \cdots E_{t-k}}(A) \\
&> \frac{P_{E_1 \cdots E_{t-k}}(A \cap E_{t-k+1} \cap \cdots \cap E_t)}{P_{E_1 \cdots E_{t-k}}(E_{t-k+1} \cap \cdots \cap E_t)} \\
&= P_{E_1 \cdots E_{t-k}}(A \mid E_{t-k+1} \cap \cdots \cap E_t) \\
&\geq \underline{P}^{\times}_{E_1 \cdots E_{t-k}}(A \mid E_{t-k+1} \cap \cdots \cap E_t).
\end{aligned}$$

A similar argument gives us that $\overline{P}_{E_1 \cdots E_{t-k}}(A) < \overline{P}^{\times}_{E_1 \cdots E_{t-k}}(A \mid E_{t-k+1} \cap \cdots \cap E_t)$. So $E_{t-k+1} \cap \cdots \cap E_t$ dilates A regardless of which updating rule $\times \in \{B, G\}$ the agent endorses. In turn, forgetting $E_{t-k+1} \cap \cdots \cap E_t$ constricts A, in symbols $(\text{IF}_\times, E_{t-k+1} \cap \cdots \cap E_t) \looparrowleft A$. ■

Proof [Proof of Corollary 23] Analogous to the proof of Theorem 22. ■

Proof [Proof of Theorem 25] Fix any $A \in \mathcal{F}$. Recall that by (8), we have that $m^I(A \mid E) = \sum_{X \in \mathcal{F}} \mathfrak{f}(A, X) m(X)$, for all $A \in \mathcal{F}$. Also, by property (c) of mass function m associated to LP $\underline{P}$ (see Definition 5), we have that $\underline{P}(A) = \sum_{B \subseteq A} m(B)$. So,

$$\underline{P}^I(A \mid E) = \sum_{B \subseteq A} m^I(B \mid E) = \sum_{B \subseteq A} \sum_{X \in \mathcal{F}} \mathfrak{f}(B, X) m(X).$$

Only if Suppose $(I, E) \looparrowleft A$. Then, by definition of constriction, $\underline{P}^I(A \mid E) > \underline{P}(A)$ and $\overline{P}^I(A \mid E) < \overline{P}(A)$. This happens if and only if

$$\begin{aligned}
&\sum_{B \subseteq A} \sum_{X \in \mathcal{F}} \mathfrak{f}(B, X) m(X) > \sum_{B \subseteq A} m(B) \\
\iff &\sum_{B \subseteq A} \left[\sum_{X \in \mathcal{F}} \mathfrak{f}(B, X) m(X) - m(B) \right] > 0
\end{aligned}$$

and

$$\begin{aligned}
&\sum_{B \subseteq A^c} \sum_{X \in \mathcal{F}} \mathfrak{f}(B, X) m(X) > \sum_{B \subseteq A^c} m(B) \\
\iff &\sum_{B \subseteq A^c} \left[\sum_{X \in \mathcal{F}} \mathfrak{f}(B, X) m(X) - m(B) \right] > 0.
\end{aligned}$$

This latter is true because $\overline{P}^I(A \mid E) < \overline{P}(A) \iff 1 - \underline{P}^I(A^c \mid E) < 1 - \underline{P}(A^c) \iff \underline{P}^I(A^c \mid E) > \underline{P}(A^c)$.

If Assume that

$$\sum_{B \subseteq A} \left[\sum_{X \in \mathcal{F}} \mathfrak{f}(B, X) m(X) - m(B) \right] > 0 \quad \text{and}$$

$$\sum_{B \subseteq A^c} \left[\sum_{X \in \mathcal{F}} \mathfrak{f}(B, X) m(X) - m(B) \right] > 0.$$

Then, we have

$$\underline{P}^I(A \mid E) = \sum_{B \subseteq A} \sum_{X \in \mathcal{F}} \mathfrak{f}(B, X) m(X) > \sum_{B \subseteq A} m(B) = \underline{P}(A)$$

and

$$\begin{aligned}
\underline{P}^I(A^c \mid E) = \sum_{B \subseteq A^c} \sum_{X \in \mathcal{F}} \mathfrak{f}(B, X) m(X) &> \sum_{B \subseteq A^c} m(B) \\
&= \underline{P}(A^c),
\end{aligned}$$

which implies $\overline{P}^I(A \mid E) < \overline{P}(A)$. This concludes the proof. ■