Non-Conglomerability for Finite-Valued, Finitely Additive Probability

Teddy Seidenfeld
Carnegie Mellon University, teddy@stat.cmu.edu

Mark J. Schervish
Carnegie Mellon University, mark@stat.cmu.edu

Joseph B. Kadane
Carnegie Mellon University, kadane@stat.cmu.edu

Follow this and additional works at: http://repository.cmu.edu/statistics

Published In
NON-CONGLOMERABILITY FOR FINITE-VALUED, FINITELY ADDITIVE PROBABILITY

By T. SEIDENFELD
M.J. SCHERVISH
and
J.B. KADANE
Carnegie Mellon University, Pittsburgh

SUMMARY. We consider how an unconditional, finite-valued, finitely additive probability $P$ on a countable set may localize its non-conglomerability (non-disintegrability). Non-conglomerability, a characteristic of merely finitely additive probability, occurs when the unconditional probability of an event $P(E)$ lies outside the closed interval of conditional probability values,

$$\min_{h \in \pi} P(E|H), \quad \sup_{h \in \pi} P(E|H),$$

$P(E|H)$, taken from a countable partition $\pi = \{h_j : j = 1, \ldots\}$. The problem we address is how to identify events and partitions where a finite-valued, finitely additive probability fails to satisfy conglomerability. We focus on the extreme case of 2-valued finitely additive probabilities that are not countably additive. These are, equivalently, non-principal ultrafilters. Evidently, the challenge we face is that a countable partition, at most one of its elements has positive probability under $P$. Thus, we must find ways of regulating the coherent conditional probabilities, given null events, that cohere with the unconditional probability $P$. Our analysis of $P$ proceeds by the use of combinatorial properties of the associated non-principal ultrafilter $U_P$. We show that when ultrafilter $U_P$ is not minimal in the Rudin-Keisler partial order of $\beta(\omega,\omega)$, we may locate a partition in which $P$ fails to satisfy the conglomerability principle by examining (at most) countably many partitions. This result is then applied to finitely additive probabilities that assume only finitely many values. By contrast, if ultrafilter $U_P$ in Rudin-Keisler minimal, then $P$ is simultaneously conglomerable in each finite collection of partitions, though not simultaneously conglomerable in all partitions.

1. Introduction to Finitely Additive [f.a.] Probability

Let $\mathcal{F}$ be a $\sigma$-field of sets, of subsets of $\Omega$. Kolmogorov’s (1936) axiomatization of probability requires that $\forall(A,B) \in \mathcal{F}$:

AMS (1991) subject classifications. 03A05, 04A20.

Key words and phrases: finitely-additive probability, non-conglomerability, non-principal ultrafilter, Ramsey ultrafilter, Rudin-Keisler partial order, ultrafilter
(1) $0 \leq P(A) \leq 1$.
(2) $P(\Omega) = 1$.
(3) If $A \cap B = \phi$, then $P(A) + P(B) = P(A \cup B)$.

A probability satisfying axioms (1)-(3) is said to be finitely additive [f.a.].

Last, consider a fourth axiom, $\sigma$-additivity (taken by Kolmogorov as an "expeditious")

(4) If $(A_i \cap A_j) = \phi$ whenever $i \neq j$, then $P(\bigcup_i A_i) = \Sigma_i P(A_i)$.

What is distinctive about a f.a. probability that is not $\sigma$-additive, a merely finitely additive probability? The following illustrates a hallmark of merely finitely additive probability.

**Example 1** (deFinetti, 1930 and attributed to Levy by Cantelli, 1935). Consider a f.a. probability $P$ on the set of all pairs $<s,t>$, for $s$ and $t$ positive integers, with the following two restrictions:

$P(<s,t>) = 0$, that is, $P$ is 0 on finite sets

and

$P(<s,t>|B) = 0$ if $B$ is an infinite set.

Define the events: $E = \{<s,t>: s > t\}$; $S_m = \{<s,t>: s = m\}$ ($m = 1, \ldots$) and $T_n = \{<s,t>: t = n\}$ ($n = 1, \ldots$).

Diagram for deFinetti - Levy Example

Event $E$ corresponds to pairs $<s,t>$

below the main diagonal

Event $E'$
Then
\[ P(E \mid S_m) = 0 \text{ for } m = 1, \ldots \]
yet
\[ P(E \mid T_n) = 1 \text{ for } n = 1, \ldots \]

Let \( \pi_s \) be the partition by vertical strips: \( \pi_s = \{S_m : m = 1, \ldots \} \) and let \( \pi_t \) be the partition by horizontal strips: \( \pi_t = \{T_n : n = 1, \ldots \} \). Thus, we see that the following principle (evidently valid for each countably additive probability) is invalid for each f.a. probability that satisfies the two constraints of Example 1, as such a probability must violate the principle in at least one of the two partitions \( \pi_s \) and \( \pi_t \). Let \( \pi = \{h_n : n = 1, \ldots \} \) be an exhaustive partition.

**Principle of (\( \omega^- \)) conglomerability for events** (deFinetti, 1972, p.99)
\[ \forall (A \in \mathcal{F}) \text{ if } c_1 \leq P(A \mid h_n) \leq c_2 \text{ (n = 1, \ldots)}, \text{ then } c_1 \leq P(A) \leq c_2. \]

Dubins (1975) strengthens the conglomerability principle to apply to all bounded random variables, rather than applying it solely to the indicator functions for events. He establishes that \( P \) is conglomerable (in the stronger sense) in a partition \( \pi \) iff \( P \) is disintegrable in \( \pi \), which concept we review next.

For a f.a. probability \( P \), partition \( \pi = \{h_j : j = 1, \ldots \} \), and bounded random variable \( X \), let \( E_p[X] \) and \( E_p[X \mid h] \) denote the \( P \)-expectation of \( X \) and conditional \( P \)-expectation of \( X \) given \( h \), respectively. (We discuss coherence of conditional probability, below.)

**Definition.** \( P \) is disintegrable in the partition \( \pi \) provided that, for each bounded random variable \( X \),
\[ E_p[X] = \int_{h \in \pi} E_p[X \mid h] dP(h). \]

In this paper, whenever we show that a f.a. probability \( P \) is conglomerable in a partition \( \pi \), we do so for the strong (Dubins') sense of the principle and when we show that \( P \) fails to satisfy conglomerability in \( \pi \), we give the failure with respect to an event, i.e., then we show that (deFinetti's) weak conglomerability principle fails.

Providing that preference is according to subjective expected utility, when conglomerability fails in a partition \( \pi \), then such basic decision theoretic principles as simple dominance (or “admissibility”) fail in \( \pi \) as well. Therefore, it is important to understand not only whether but where, i.e., in which partitions does a particular merely finitely additive probability fail to satisfy the principle of conglomerability.

In fact, (weak) conglomerability characterizes countable additivity. That is, non-conglomerability of merely finitely additive probability is a hallmark, as the following result reports.
Theorem 1 (Schervish et al., 1984). Each merely f.a. probability fails conglomorability for some event in some denumerable partition.

This result quantifies over all denumerable partitions and over all events. These are large sets, e.g., of cardinality of the continuum when the underlying set is countable. (Of course, if the underlying set is finite, there is no issue to discuss.) To locate where non-conglomerability occurs, according either to our proof of Theorem 1, or Zane’s (1988) simplified proof, depends for some $P$ upon how the conditional probabilities $P(\bullet \mid \bullet)$ are defined, separate from the unconditional probability, $P(\bullet)$.

For an illustration of this issue, reconsider Example 1. There are two constraints on $P$ that lead to non-conglomerability in (at least) one of the two partitions, $\pi_1$ and $\pi_2$, for the event $E$. The first is a constraint on the unconditional probability $P(\bullet)$; that $P$ is 0 on finite sets. The second is a constraint on the conditional probabilities $P(\bullet \mid \bullet)$: that $P(<s,t>|B) = 0$ if $B$ is an infinite set. The first constraint insures that $P$ is a purely finitely additive probability on $\omega$. However, if $Q = \{Q: Q$ is a finitely additive probability satisfying the first constraint$\}$ then only for a proper subset of $Q$ does the unconditional probability $Q$ specify even the two families of conditional probabilities, given $\pi_1$ and given $\pi_2$. Only for some $Q \in Q$ does the unconditional probability entail the two sets of conditional probabilities used in Example 1.

In this paper, we investigate the following issue:

**Thematic question.** Given a merely f.a. probability $P$, is it determined where conglomerability fails based solely on the unconditional probability values $P(\bullet)$? That is, given $P$ as an unconditional finitely additive probability, can event $A$ and partition $\pi$ be found where $P$ is not conglomerable in $\pi$ for event $A$, i.e., where $P$ is not $\pi$-conglomerable?

By contrast with the situation in Example 1, a positive answer to the question is available (Schervish et al., 1984, p 210) whenever the range of $P$ is an infinite set, i.e., whenever $P(\bullet)$ assumes infinitely many values. The following (see Dubins, 1975) illustrates what happens.

**Example 2.** Let $E_1$ be the event of flipping a “fair” coin until heads shows. Let $E_2$ be the event of picking a positive integer “at random,” according to some (purely) f.a. probability that assigns each integer 0 probability. (There are very many such purely f.a. probabilities, indeed!) Assume $P(E_1) = P(E_2) = 0.5$; for example, which of $E_1$ or $E_2$ occurs may be determined by an extraneous flip of a “fair” coin. Let $x_1$ be the random variable of the number of flips in case $E_1$ obtains, 0 otherwise, and let $x_2$ be the integer chosen at random in case $E_2$ obtains, 0 otherwise. Then $P(\{x_2 = n\}) = 0$, for $n = 1, \ldots$ and $P(\{x_1 = n\}) = 2^{-(n+1)}$. Let $k$ denote the random variable of the positive integer that (with $P$-probability 1) results. So $P(k = n) = P(\{x_1 = n\}) = 2^{-n-1}$. Thus, $P(E_1 | k) = 1 (k = 1, 2, \ldots)$, and conglomerability fails in the partition $\pi = \{\{h_k : k = 1, \ldots\}$.

Note that here each conditioning event, each partition
element $h_k$ has positive (unconditional) probability: $P(h_k) = 2^{-(k+1)} > 0$. Here the conditional probabilities $P(\bullet \mid k)$ are fixed by the unconditional probabilities and non-conglomerability in $\pi$ is fixed by the unconditional probability $P$. $P$ cannot be made conglomerable in $\pi$.

When the (unconditional) probability $P$ assumes infinitely many distinct values, the unconditional probabilities for events identify an event $E$ and countable partition, $\pi = \{h_j : j = 1, \ldots\}$, where each $h_j$ has positive $P$-probability, and where $P(E \mid h_j) > P(E) + \epsilon$, for $\epsilon > 0$ and $j = 1, \ldots$. That is, when the range of $P$ is an infinite set, there exists a partition where $P$ cannot be made conglomerable.

The remaining case, thus, is where a merely f.a. probability $P$ assumes only finitely many values. A special sub-case that we consider in the next section is for a two-valued merely f.a. (unconditional) probability $P$: $\forall \in \mathcal{F} P(A) = 0$ or $P(A) = 1$. This is the difficult case because then the conditional probabilities $P(\bullet \mid \bullet)$ are determined by the unconditional probability $P(\bullet)$ only up to conditioning sets of measure 1. With a two-valued probability, these form only a sparse collection. Given a partition $\pi = \{h_i : i = 1, \ldots\}$, either one or none of its elements has positive $P$-probability. To compensate, we use the following principle in dealing with conditional probabilities $P(\bullet \mid \bullet)$, especially for conditional probability given events of (unconditional) probability 0.

**Principle of conditional coherence:** For all pairs of events, $A$ and $B$ such that $A \cap B \neq \emptyset$, $Q(\bullet) = P(\bullet \mid B)$ is a finitely additive probability with $Q(B) = 1$, and $Q(\bullet \mid A) = P(\bullet \mid A \cap B)$.

When $P(A \cap B) > 0$, the principle applies, trivially. The principle of conditional coherence helps to formalize de Finetti’s (1972) concern with conditional probability given an event, rather than given a field, in that $P(\bullet \mid B)$ does not depend upon how we partition the contrary to event $B$.

Dubins (1975, §3, Corollary 1) reproves an important result of P. Krauss (1968): for each finitely additive (unconditional) probability $P$, a full set of conditional probabilities may be define that satisfy the principle of conditional coherence. Moreover, Dubins (1975, Theorem 5) establishes that when $P$ is disintegrable in a partition $\pi$, then there is a full set of coherent conditional probabilities extending the set of conditional probabilities $\{P(\bullet \mid h) : h \in \pi\}$. Throughout this paper we adopt the principle of conditional coherence, relying on Dubins’ Theorem 5 to show that whenever $P$ is conglomerable in $\pi$, then the set of conditional probabilities $\{P(\bullet \mid h) : h \in \pi\}$ that make it conglomerable are also coherent conditional probabilities.

**Remark.** Though the merely f.a. probability $P$ is two-valued, given $B$ with $P(B) = 0$, a coherent conditional probability $P(\bullet \mid B)$ may have an infinite range and may be countably additive.

In section 2 we establish conditions when a two-valued merely f.a. probability $P$ on $\omega$ cannot be made conglomerable in a specific partition, based on combinatoric properties of its associated non-principal ultrafilter $Up$. We show that
for each non-Ramsey ultrafilter $U_P$, either $P$ cannot be made conglomerable in a specific partition (based on $U_P$'s combinatorics), or else a countable sequence of partitions suffice to locate a partition where $P$ does not satisfy the (weak) principle of conglomerability. We refer the reader to Comfort and Negrepontis (1974), especially chapters 9 and 16 for basic facts about ultrafilters and the Rudin-Keisler ordering of $\beta(\omega) \setminus \omega$.

2. **Ultrafilters on $\omega$**

We investigate our thematic question first for two-valued, merely f.a. unconditional probabilities by considering combinatorial properties of their associated non-principal ultrafilters.

**Definitions.** An ultrafilter $U$ (on $\omega$) is a non-empty family of non-empty subsets (of $\omega$), such that if $A, B \in U$, then $(A \cap B) \in U$; if $A \in U$ and $C \supset A$, then $C \in U$; and $\forall A (\subset \omega) A \in U$ or $A^c \in U$.

A principal ultrafilter is one generated by an element $i \in \omega$, i.e., the set of all subsets of $\omega$ that contain $i$.

A non-principal ultrafilter contains no finite subsets.

Evidently, each 2-valued, unconditional f.a. probability $P$ on the powerset of $\omega$, with $P(\omega) = 0$ ($n = 1, \ldots$), corresponds uniquely to a non-principal ultrafilter $U_P$ on $\omega$, as determined solely by its unconditional probability values.

**Fact (ZFC).** There are $2^{2^\omega}$ non-principal ultrafilters on $\omega$. (See Comfort-Negrepontis, 1974, p. 146.) Hereafter, we focus on non-principal ultrafilters in our discussion of ultrafilters on $\omega$.

2.1 **Some elementary combinatorics.** Consider a function $f : \omega \to \omega$. The function $f$ induces a denumerable (finite or countable) partition $\pi_f = \{h_n : n = 1, \ldots \}$ of $\omega$ by considering the inverse images, $f^{-1}(n)$.

**Definition.** Call an ultrafilter $U$ selective in a partition $\pi = \{h_1, h_2, \ldots \}$ if there is an $A \in U$ so that:

$$\exists h \in \pi \text{ with either } A \subseteq h \text{ or } |A \cap h_n| \leq 1 \ (n = 1, \ldots).$$

**Definition.** When an ultrafilter $U_P$ is selective in a partition $\pi$, we call the following (one version of) its natural conditional probability given $\pi$:

In the first case, when $P(h) = 1$ then $P(\bullet | h) = P(\bullet)$ and for $h_n \neq h$, let $P(\bullet | h_n)$ be an arbitrary f.a. probability defined on $h_n$.

In the second case, let $P(\bullet | h_n)$ be concentrated, with probability 1, at the singleton $(A \cap h_n) = \{a_n\}$, if it exists. Otherwise, let $P(\{a_n\} | h_n) = 1$ for an arbitrary $a_n \in h_n$. Thus $P(\bullet | h_n)$ is a 0-1 principal ultrafilter probability.

If $U_P$ is selective in partition $\pi$ then there are many different versions of its natural conditional probability, depending upon which $A \in U_P$ is chosen. We
note that each two versions differ only on a set of elements of \( \pi \) that lie outside the ultrafilter \( U_P \), hence; each two versions differ for a set of conditioning events of \( P \)-measure 0.

**Lemma 1.** When \( U_P \) is selective in a partition \( \pi \), it is conglomerable there using (any version of) its natural conditional probability for \( P \) given \( \pi \).

**Proof.** In the first case, it is immediate that \( P \) is conglomerable in \( \pi \) from Dubins' equivalence with disintegrability and the obvious equality, \( E_P [X \mid h] = E_P [X] \).

In the second case, let \( X \) be a bounded random variable with \( E_P [X] = c \). Then, as \( P \) is an ultrafilter probability, for each \( \epsilon > 0 \), \( P(\{\omega \in \Omega : |X(\omega)| \leq c + \epsilon\}) = 1 \). Each natural conditional probability satisfies \( E_P [X \mid h_n] = E_P [X(a_n) \mid h_n] = X(a_n) \) (for \( n = 1, \ldots \)). However, \( \cup \{a_n\} = A \in U_P \). Then \( P(\cup h_n : |E_P [X|h_n]| \leq c + \epsilon) = 1 \) and \( P \) is disintegrable in \( \pi \); hence, \( P \) is conglomerable in \( \pi \) using (any version of) its natural conditional probability.

**Definition.** A non-principal ultrafilter \( U \) on \( \omega \) is said to be a Ramsey ultrafilter if it satisfies Ramsey's partition theorem for a set in \( U \). That is, for each integer \( n \) and binary partition \( \{h_0, h_1\} \) of \( [\omega]^n \), there is an \( A \in U \) such that either \([A]^n \subseteq h_0\) or \([A]^n \subseteq h_1\).

**Theorem 2.** (attributed to Kunen). A non-principal ultrafilter on \( \omega \) is Ramsey if and only if it is selective in each partition \( \pi \). (See Comfort-Negrepontis, 1974, p. 212.)

**Remark.** In (ZFC + CH) there are \( 2^{2^\omega} \) Ramsey ultrafilters on \( \omega \). (Comfort-Negrepontis, 1974, p. 220)

**Theorem 3.** If \( U_P \) is a Ramsey ultrafilter, \( P \) can be made conglomerable (simultaneously) in any finite number of partitions, \( \pi_i (i = 1, \ldots, k) \), by using natural conditional probabilities.

**Proof.** Given \( \pi_i = \{h_i^n : n = 1, \ldots \}(i = 1, \ldots, k) \), let \( A_i \in U_P \) satisfy either \( \exists \bar{x} \in \pi_i \) with \( A_i \subseteq \bar{x} \) or \( |A_i \cap h_i^n| \leq 1 \) (\( n = 1, \ldots \)). Let \( \cap_i A_i = A \in U_P \). Then, corresponding to the respective case, given \( \pi_i \) for \( i = 1, \ldots, k \), either \( A \subseteq \bar{x} \) or \( |A \cap h_i^n| \leq 1 \) (\( n = 1, \ldots \)). Using \( A \), consider a (version of the) natural conditional probability for \( P \) given \( \pi_i (i = 1, \ldots, k) \). By Lemma 1, \( P \) is conglomerable (simultaneously) in each of the \( \pi_i \).

Hence, when Ramsey ultrafilters exist, they provide a negative answer to our thematic question. That is, given any specific partition \( \pi \), each Ramsey ultrafilter probability has its associated (natural) conditional probability that makes it conglomerable in \( \pi \). Next, we show that for non-Ramsey ultrafilters, their combinatorial properties pinpoint some of their non-conglomerability.

**Definition.** An ultrafilter \( U \) is weakly selective in a partition \( \pi \) if there exists an \( A \in U \) such that: either \( \exists x \in \pi \) with \( A \subseteq x \) or \( |A \cap h_n| < \omega \) (\( n = 1, \ldots \)).
**Remark.** An ultrafilter $U$ is a $P$-point *iff* it is weakly selective in each $\pi$. (Blass, 1973) Trivially, each Ramsey ultrafilter is a $P$-point.

**Corollary 1.** If $U_P$ is a $P$-point, but not Ramsey, then $P$ fails to be conglomerable in each partition where it lacks the Ramsey (selective) property.

**Proof.** The claim is immediate from the following lemma.

**Lemma 2.** If $U_P$ is weakly selective but not selective in partition $\pi$ then $U_P$ is not conglomerable in $\pi$, nor is it even approximately conglomerable in $\pi$ (and the extent of non-conglomerability is 1).

**Proof.** (based on an argument by Dubins (1975, pp. 92-93). Let $\pi$ be a partition where $U_P$ is weakly selective but not selective. Thus, no element $(h_n)$ of $\pi$ belongs to $U_P$ and no graph of a function $g : \omega \to \omega$ belongs to $U_P$ either. But, as $U_P$ is weakly selective in $\pi$, there is a function $f : \omega \to \omega$ whose graph in $\pi$ bounds $U_P$ from above. That is, let $S_n = \{x : x \in h_n$ and $x \leq f(n)\}$ and define $S = \cup S_n$. Then $S \in U_P$.

Let $Q$ be a finitely additive probability that is disintegrable in $\pi$ so that, for event $A$, $Q(A) = \int_{h \in \pi} Q(A \mid h) dQ(h)$. Then we see that $P$ and $Q$ are singular, as follows. Fix $\epsilon > 0$ and let $m$ satisfy $m\epsilon \geq 1$. Consider the finite partition of $S$ according to the $1/m$ quantiles of the conditional distributions $Q(S \mid h)(h \in \pi)$. That is, given $h \in \pi$ and integer $k$ $(1 \leq k \leq m)$ the $k/m$-th quantile point is the least element of $h$, $x^k$, such that $Q(\{x \leq x^k\} \mid h) \geq k/m$. By the reasoning above, for each $k$, the graph of the $k/m$-th quantile points does not belong to $U_P$. Hence, one of the (at most) $m$ regions strictly between these (at most) $m$-graphs, call it $R_k$, belongs to $U_P$. But $Q(R_k) \leq \epsilon$ yet $P(R_k) = 1$. Hence, $P$ is not at all approximable by f.a. probabilities $Q$ that are conglomerable (in Dubins' sense) in $\pi$. That is, failure of conglomerability in $\pi$ is maximal. For each coherent conditional probability $P(\bullet \mid \bullet)$, for each $\epsilon > 0$, there exists $A$ with $P(A) = 1$, but $P(A \mid h) < \epsilon Vh \in \pi$.

Thus, for non-Ramsey $P$-points, we can identify a partition where conglomerability of events fails maximally.

3. **Non-Ramsey Ultrafilters and Non-conglomerability**

Next we explore consequences of Lemma 2 for the Rudin-Keisler partial order of ultrafilters. The Rudin-Keisler partial order of ultrafilters, $\preceq$, is defined as follows:

Let $f : \omega \to \omega$ and ultrafilter $U$ be given. Define ultrafilter $V = f(U)$ by, $X \in V$ if $f^{-1}(X) \in U$. Then say that $V \preceq U$ if there is some $g : \omega \to \omega$ with $g(U) = V$.

The Rudin-Keisler partial order $\preceq$ is reflexive and transitive. Denote by $U \approx V$ the equivalence relation $(V \preceq U$ and $U \preceq V)$ and denote by $V < U$ the
strict partial order \((V \preceq U \text{ and } V \not\preceq U)\). It is well known that \(U \approx V\) iff there is a function \(g(U) = V\) where \(g\) is 1-1 on a set in \(U\). (See Comfort-Negrepontis, 1974, p.209.) That is, \(U \approx V\) obtains iff there is a mapping \(g(U) = V\) where \(U\) is selective in \(\pi_0\) but no element of \(\pi_0\) belongs to \(U\). Two other familiar results about the Rudin-Keisler partial order on \(\beta(\omega)\) \(\setminus\omega\) are that Ramsey ultrafilters are minimal, and if \(U\) is a \(P\)-point and \(V \preceq U\), then \(V\) is a \(P\)-point too. Thus, Lemma 2 provides us with the following:

**Corollary 2.** If \(V \prec U_P\) and \(V\) is a non-Ramsey \(P\)-point then, by mapping \(U_P\) to \(V\) and locating where \(V\) is not-selective, we fix a partition \(\pi\) based on the combinatorial properties of \(U_P\) (and \(V\)) in which \(P\) is non-conglomerable. \(\square\)

Recall, too, that in Example 1 the non-conglomerability for event \(E\) is localized to one of two orthogonal partitions, i.e., partitions whose elements meet each other in singleton sets, at most. Since an ultrafilter is weakly selective in (at least) one of each pair of orthogonal partitions, we can generalize this feature of Example 1 using Lemma 2 as follows:

**Corollary 3.** Let \(V \prec U\) and \(W \prec U\) with \(f(U) = V\) and \(g(U) = W\). If \(\pi_U\) and \(\pi_W\) are orthogonal partitions, then \(U\) is non-conglomerable in (at least) one of these two partitions. \(\square\)

Unfortunately, we do not know whether, for each non-Ramsey ultrafilter \(U\) there exist \(V\) and \(W\) satisfying the hypothesis of Corollary 3. Next, we show that when \(U_P\) is not a Ramsey ultrafilter and we use conditionally coherent versions of \(P\)'s natural conditional probabilities whenever \(U_P\) is selective in a partition, then the combinatorial properties of \(U_P\) locate partitions and events where \(P\) does not satisfy conglomerability.

**Theorem 4.** If \(U_P\) is not Rudin-Keisler minimal then, either there exists a partition where \(U_P\) is weakly selective but not selective (where \(P\) is maximally not conglomerable), or else a coherent set of \(P\)'s natural conditional probabilities, coming from a countable sequence of partitions where \(U_P\) is selective, lead to a failure of conglomerability also to the maximum possible extent.

**Proof.** By assumption, there exists \(V \prec U_P\). Let \(\pi_0\) be a partition

\[
\pi_0 = \{h^0_j : h^0_j = \{f^{-1}(j), j = 1, \ldots\}\}
\]

induced by the mapping \(f(U_P) = V\). Since \(V \prec U_P\) and \(V\) is non-principal, \(U_P\) is not selective in \(\pi_0\). Consider a (canonical) 1-1 map \(m\) between \(\omega\) and \(\omega \times \omega\) where \(m(h^0_j) = \{(i, j) : i = 1, \ldots\}\) for \(j = 1, \ldots\), and let \(\Delta_0\) be the diagonal of \(\pi_0\) under \(m\). Then, \(m^{-1}(\Delta_0) \not\prec U_P\). If \(P\) is conglomerable in \(\pi_0\) then (Lemma 2) \(U_P\) is not weakly selective in \(\pi_0\) either. Hence, we may assume that (under \(m^{-1}\)) the set \(\gamma_0\) of points above \(\Delta_0\) in \(\pi_0\) belong to \(U_P\). (Hereafter we suppress \(m\) in our discussion of partitions and, where notationally convenient, we identify unit sets with their members, as in the last sentence of this paragraph.) Let \(\pi_1\)
be the partition orthogonal to $\pi_0$, i.e., $\pi_1 = \{b_{i,j}^1 : h_{i,j}^1 \in \{i^j\} \}$ where some elements of $\pi_0$ and $\pi_1$ may be disjoint. Since $U_P$ is not selective in $\pi_0$, for each $h_i \in \pi_1$, $h_i \notin U_P$. But $U_P$ is weakly selective in $\pi_1$, since $Y_0 \cup U_P$ and $|Y_0 \cap h_i^1| < \omega (n-1, \ldots)$. Thus, if $U_P$ is not selective in $\pi_1$ then $P$ is not conglomerable in $\pi_1$. Hence, we assume there exists a set $A_1 \subseteq U_P$ such that $|A_1 \cap h_i^1| \leq 1 (n-1, \ldots)$. If we consider the mapping $f^1 : \omega \to \omega$ associated with $\pi_1$, then $f^1(U_P)$ is the ultrafilter $U_1$ and since $f^1$ is 1-1 on $A_1 \subseteq U_P$, $U_P \approx U_1$. Based on $A_1$, choose a (version of the) natural conditional probability for $P$, given $\pi_1$, with set $B_1 = \{b_{i,n} \in h_i^1(n = 1, \ldots)\}, B_1 \supseteq A_1$ with $P(b_{i,n} \mid h_i^1) = 1$.

Let $h_i^1$ be the least element of $\pi_0$ that meets $B_1$ at some element $b_{i,k}$, and define set $C_1 = B_1 - \{b_{i,k}\}$. Obviously, $C_1 \subseteq U_P$, though this depends upon the version of the natural conditional probability used. Call $h_i^1$ that element of $\pi_1$ which contains $b_{i,k}$. Make a partially ordered tree $T_i$ from $h_i^1$ by rooting it in $b_{i,k}$ at level 0, i.e. setting level 0 equal to the unit set $\{b_{i,k}\}$ and making level 1 the unit set $\{h_i^1 \setminus \{b_{i,k}\}\}$. This partial order coincides in an obvious way with the qualitative order from the natural conditional probability: lower levels have (much) higher probability. Specifically, $P(h_i^1 \setminus \{b_{i,k}\} \mid h_i^1) = 0$.

Iterate this procedure to create partitions $\pi_i (i = 2, \ldots)$ of the sets $C_{i-1}$ (where $\lim_i C_i = \emptyset$) and where each $\pi_i$ is orthogonal to $\pi_0$, i.e., where each element of the partition $\pi_i$ meets each element of the partition $\pi_0$ in at most one element of $\omega$. (See the figure below.) Thus, the set $Y_i$ (the points above the diagonal $\Delta_i$ of $\pi_0 \times \pi_i$) belongs to $U_P$; hence, $U_P$ is weakly selective in $\pi_i$. Then, $P$ is not conglomerable in $\pi_i$ unless $U_P$ is selective in that partition. Then, if $U_P$ is selective in $\pi_i$, the function $f^i$ (which is associated with $\pi_i$) yields the ultrafilter $U_i = f^i(U_P)$ and $U_P \approx U_i$.

Assuming that $P$ is selective in $\pi_i$, we arrive at the sets $A_i$ and $B_i$, the element $b_{i,k}$, and the set $C_i$, just as in the case $i = 1$, above. In this way we produce another R-K equivalent ultrafilter $U_i$, i.e., $U_P \approx U_i$. Also, we create the tree $T_i$, as described below. If this process continues, i.e., if $U_P$ is selective in each $\pi_i (i = 1, \ldots)$ then the infinite set of trees, $\{T_i : i = 1, \ldots\}$, which partition $\omega$, identify a partition $\pi^* = \{h_i^* : i = 1, \ldots\}$ in which the (chosen versions of the) natural conditional probabilities associated with the $\pi_i$s fail deFinetti’s conglomerability principle. Moreover, the extent of the failure is maximal, i.e. there is an event $E$ with $P(E) = 1$ and $P(\{h_i^* \mid E\} = 0 (i = 1, \ldots)$. Next, we give the details of the partial order for the tree $T_i$. We define $\pi_i, A_i, B_i, b_{i,k}, T_i$, and $C_i$, inductively, as follows: For $C_{i-1} \subseteq U_P$, let

$$\pi_i = \{h_i^j : h_i^j \in \{i^j\}\} \setminus h_i^1 : n = 1, \ldots\}, j = 1, \ldots\}.$$ 

(Note: $\pi_i$ is a partition of $C_{i-1}$, not of the full set $\omega$.) Thus, each partition element $h_i \in \pi_i$ is orthogonal to $\pi_0$ and is the graph of a (partial) function. If $k > j$, then $h_i^j$ lies above $h_i^j$ in $\pi_0$. Also, the elements $h_i^j$ (that is, functions) in partition $\pi_i$ grow more rapidly than do those in $\pi_j$. Since $U_P$ is not selective
nor even weakly selective in $\pi_0$, $U_P$ is weakly selective in each $\pi_i$. As argued in
the base case ($i = 1$), $\Delta_i \not\subseteq U_P$, $Y_i \subseteq U_P$, and for each $h^i \in \pi_i$, $h^i \not\subseteq U_P$. Thus, $P$
is conglomerate in $\pi_i$ if and only if $U_P$ is selective there.

Let $A_i$ be a set in $U_P$ meeting the condition that $|A_i \cap h^i_n| \leq 1$ ($n = 1, \ldots$).
Since there exists a coherent version of the natural conditional probability for
$P$, given $\pi_i$, there is a set $B_i = \{b_{in} : b_{in} \in h^i_n (n = 1, \ldots)\}$, $B_i \supseteq A_i$ with
$P(b_{in} | h^i_n) = 1$. Let $h^i_{n,j}$ be the least element of $\pi_0$ that meets $B_i$ at some
element $b_{ik}$ and let $C_i = B_i - \{b_{ik}\}$. Evidently, $C_i \in U_P$. Denote by $h^i_{n,j}$ that
element of $\pi_i$ which contains $b_{ik}$, and for $j < i$, denote by $h^i_{n,j}$ that element of
$\pi_j$ containing $b_{ik}$.
Make a tree $T_i$ of height $i$ by rooting it in $b_{ik}$ at level 0 and making level 1 the $i$-element set $\{h^1_j = \{b_{ik}\}, h^2_j = \{b_{ik}\}, \ldots, h_i^j = \{b_{ik}\}\}$. Finite additivity assures that $P(\cup \text{ level } 1 \mid \cup \text{ level } 0) = 0$. Level 2 of $T_i$ is formed by adjoining to each $b \in h^2_j - \{b_{ik}\} (j = 2, \ldots, i)$ the $(j - 1)$-many sets $\{h^j - b\}, h^{j-2} - b, \ldots, h^1 - b\}$, where $b \in h^n \in \pi^n(n = 1, \ldots, j - 1)$. Again, finite additivity assures that $P(h^j - b \cup h^{j-2} - b \cup \ldots \cup h^1 - b) = 0$. Continue this way to extend the branches of $T_i$ by adjoining to each $b \in \text{ level } m (m < i)$ the $m$-many sets $\{h^i - b\}$ for the $m$-many partition elements $h$ that contain $b (1 \leq v \leq m - 1)$. Finite additivity assures that $P(h^i - b \cup \ldots \cup h^1 - b) = 0$. The tree $T_i$ has branches ending at each level and the branches terminate in sets of the form $h^i - b$, for $b \in h^i \in \pi_i$.

Now, either this inductive procedure terminates after finitely many steps in a partition $\pi_k$ where $U_P$ is weakly selective but not selective, or else it leads to an infinite forest of trees $\{T_i : i = 1, \ldots\}$. In the latter case the trees partition the space, $\omega$, because: (1) Elements of a tree are disjoint subsets of $\omega$. (2) $(\cup T_i) \cap (\cup T_j) = \emptyset$ whenever $i \neq j$. And (3) for each $b \in \omega$, $b$ belongs only to a non-empty finite sequence of partition elements $\{h^j : b \in h^j, h^{j-2} \in \pi_j : j = 1, \ldots, k\}$ where $P(b \mid h^j) = 1$, for $i < k$, and either $b$ is the root of tree $T_k$ or $P(b \mid h^k) = 0$.

Note that the union of sets in a tree does not belong to $U_P$ since $\omega - C_{i+1} \subseteq \cup T_i$ but $C_{i+1} \subseteq \cup T_i$. Moreover, the set of all tree-roots $R = \{b_{ik} : b_{ik} \text{ root of } T_i, i = 1, \ldots\}$ does not belong to $U_P$ either. This is so because the $b_{ik}$ are selected from decreasing sets $C_i$ in order to have $h^i_1$ meet $\pi_0$ in its least element. Thus, either all but a finite number of the $b_{ik}$ belong to one partition element of $\pi_0$, or else $|R \cap h^n | < \omega (n = 1, \ldots)$. Since $U_P$ is not weakly selective in $\pi_0, R \notin U_P$.

Last, consider the binary partition of $\omega \setminus R$ formed by taking the union of the sets in the odd levels of all trees, $L_O$, and the union of the sets in the even levels (excluding $R$, the set of roots) $L_E$. Exactly one of these two countable sets belongs to $U_P$, since $R$ does not. Without loss of generality, assume that the union of sets from the odd levels is a set in $U_P$. Observe, next, that each $b$ (an element of a set at level $2i$) has adjoined to it at level $2i + 1$ the $v$-many sets $\{h^i - b, \ldots, h^1 - b\}$ for the $v$-many partition elements $h$ that contain $b (1 \leq v \leq 2i(i = 1, \ldots)$.

Consider the denumerable set, $\pi^i = \{h^i_j : j = 1, \ldots\}$ where each $h^i_j$ contains finitely many subsets of $\omega$, one of which is $\{b\}$ for some $b \in L_E$, that is, $b$ is an element of a set from level $2i$ for some $i$ (or from $R$), and the other sets in $h^i_j$ are the finitely many disjoint sets (disjoint "events") that are adjoined to $\{b\}$ at level $2i + 1$ (or at level 1). Evidently, $(\cup_{A \in h^i_j} A) \cap (\cup_{B \in h^i_j} B) = \emptyset$ whenever $i \neq j$. Let $g : \omega \to \omega$ be any function that is constant on each element of $h^i_j$ for every $j$, such that $g^{-1}(\{i\}) \neq g^{-1}(\{j\})$ when $i \neq j$, and call $V_g$ the ultrafilter defined by $g(U_P)$. 

487

FINITELY ADDITIVE PROBABILITY
Since each \( h^j_i \) is a finite collection of disjoint subsets of \( \omega \), trivially, \( V_p \) is weakly selective in the partition \( \pi^+ = \{ h^j_i : h^j_i = \{ i : g(m) = i, m \in B \in h^j_i \}, j = 1, \ldots \} \). Either \( V_p \) is not selective in \( \pi^+ \) and \( P \) cannot be made conglomerate there, by the reasoning of Corollary 2, or else there is a set \( A^+ \in V_p \) such that \( |A^+ \cap h^j_i| \leq 1 (j = 1, \ldots) \) and we may assume that \( g^{-1}(A^+) \) is a subset of sets belonging to \( L_O \). Thus for each set \( h^j_i - \{ b \} \) that meets \( g^{-1}(A^+) \), from an odd level \( 2i + 1 \) (or from level 1) there is a unique \( b \), an element of an element of level \( 2i \) (or an element of \( R \)) where \( p(h^j_i - \{ b \}) = 0, (i = 1, \ldots) \). Hence, using these natural conditional probabilities for \( P \) violates conglomerability in the partition, \( \pi^+ = \{ h^j_i : h^j_i - \{ b \} \in g^{-1}(A^+) \}, j = 1, \ldots \). That is, there is a set \( g^{-1}(A^+) \) belonging to \( U_P \) where, for each \( h^j_i \in \pi^+, P(g^{-1}(A^+) \mid h^j_i) = 0 \) but \( P(g^{-1}(A^+)) = 1 \). When \( \cup L_E \in U_P \), just reverse the roles of even and odd levels in the trees.

Next, we show that the combinatorial properties of \( U_P \) serve to localize the non-conglomerability in \( P \) even when the (coherent) conditional probabilities \( P(\bullet \mid \bullet) \) are not so-called “natural.”

**Theorem 5.** If \( U_P \) is not Rudin-Keisler minimal then, either (i) there exists a partition where \( U_P \) is weakly selective but not selective, i.e., a partition \( \pi^+ \) in which \( P \) is singular with respect to each f.a. \( Q \) that is conglomerable in \( \pi^+ \) (by Lemma 2), or else (ii) \( P \)'s conditional probabilities (taken from no more than a sequence of partitions where \( U_P \) is selective) lead to a failure of conglomerability of extent at least \( 1/2 \), i.e.,

\[
\exists (E \in U_P) \exists n \forall (h \in \pi^+) P(E \mid h) \leq .5.
\]

**Proof.** We follow the reasoning of the previous theorem, identifying partitions \( \pi_i (i = 1, \ldots) \) of the nested sets \( C_i \subset U_P \) (where \( C_{i+1} \supset C_i \)) in which \( U_P \) is weakly selective. If \( U_P \) fails to be selective in \( \pi_i \), clause (i) is established. However, unlike the situation in Theorem 4, when \( U_P \) is selective in \( \pi_i \), the conditional probability \( P(\bullet \mid h^j_i) \) need not be the “natural” one. Nonetheless, one of two cases arises.

Case (a): \( \exists (E \in U_P) \) such that for each \( h^j_i \in \pi_i \), \( P(E \mid h^j_i) \leq .5 \). Then conditional probability given \( \pi_i \), is assigned so that \( P \) fails to be conglomerable in \( \pi_0 \) (by at least .5) - clause (i).

Case (b): There exists \( A_i \in U_P \) such that \( |A_i \cap h^j_i| \leq 1 (n = 1, \ldots) \) and the conditional probability, given \( \pi_i \), satisfies \( P(a_{in} \mid h^j_i) > .5 \) whenever \( A_i \cap h^j_i \neq \emptyset \).

We continue the argument with Case (b). Let \( C_0 = \omega \) and denote by \( B_i = \{ a_{in} : P(a_{in} \mid h^j_i) > .5 \} \) so that \( B_i \subseteq U_P \). As before, let \( h^0_{jk} \) be the least element of \( \pi_0 \) that meets \( B_i \) at some element \( b_{ik} \) and let \( C_i = B_i - \{ h^0_{jk} \} \). Evidently, \( C_i \subseteq U_P \). Again, denote by \( h^j_{ik} \) that element of \( \pi_j \) which contains \( b_{ik} \), and for \( j < i \), denote by \( h^j_{jk} \) that element of \( \pi_j \) containing \( b_{ik} \). We construct the
trees $T'_i$ as before; however, now the $T'_i$ may fail to be a partition of $\omega$. That is, let $D'_i = \{h^i : h^i \cap C'_i = \emptyset\}$ and $D'_i \neq \emptyset$ is possible. To accommodate the sets of these $h^i$, each a partition element where conditional probability fails to concentrate above $\frac{1}{2}$ on any element, we form a second array of partially ordered sets, $S'_i(i = 1, \ldots)$, analogous to the $T'_i$. Each $S'_i$ has a base (level 0) rooted in the set $D'_i$ (rather than the singleton root $b_{ik}$ of $T'_i$). The tree structure in $S'_i$ above $D'_i$ is analogous to that above $b_{ik}$ in $T'_i$. There are two sub-cases to consider:

(b.1) $T = \{T'_i\} \in U_P$. Then, by reasoning and notation of the previous theorem, i.e., dividing between odd and even levels of $T$, $\exists(g^{-1}(A^+)) \in U_P)\exists\pi^*v(h^* \in \pi^*) P(g^{-1}(A^+) \mid h^*) \leq \frac{1}{2}$.

(b.2) $S = \{S'_i\} \in U_P$. Whereas the roots $R$ of $T$ (or roots $R'$ of $T'$) do not belong to $U_P$, the set $D = \cup_i D_i$ may belong to $U_P$. However, as $D$ is a collection of (disjoint) partition elements $\{h^i : h^i \in D_i; i = 1, \ldots\}$, each with $h^i$ orthogonal to $\pi_0$, consider the partition $\pi_D$ of $D$ formed by these sets $h^i$. If $D \in U_P$, then $U_P$ is weakly selective in $\pi_D$. So, either

(b.2.1) $U_P$ is not selective in $\pi_D$ and $P$ is not even approximately $\pi_D$-conglomerable or

(b.2.2) $\exists(A \in U_P) \mid A \cap h^i \mid \leq 1$ for each $h^i \in \pi_D$. But, since $\pi_D \cap C'_i = \emptyset$ (i.e., for each $h^i \in \pi_D$, $P(A \mid h^i) \leq \frac{1}{2}$ and the conditional probability $P$, given $\pi_D$, fails the conglomerability principle by an extent $\frac{1}{2}$ (at least).

Thus, without loss of generality in case (b.2), assume that $D \not\in U_P$. That is, the union of the level 0 sets of $S$ does not belong to $U_P$. Then, as the structure of $S$ at higher levels if the same as in $T$, conclude by reasoning (with notation as in Theorem 4) that upon dividing between the odd and even levels of $S$, $\exists(g^{-1}(A^+) \in U_P)\exists\pi^*v(h^* \in \pi^*) P(g^{-1}(A^+) \mid h) \leq \frac{1}{2}$.

4. Non-conglomerability for Finitely Valued Merely f.a. Probabilities

Let $P$ be a merely f.a. probability that assumes only finitely many values. Then (Schervish et al. 1984 p. 213) $P$ may be written as $P = \sum_{i=1}^{k} \gamma_i P_i$, where $\gamma_i > 0$, $\sum_i \gamma_i = 1$, and each $P_i$ is an ultrafilter probability. Since $P$ is merely finitely additive, there exists an integer $k_1$, $0 \leq k_1 < k$ where each $P_i(i \leq k_1)$ is a principal ultrafilter probability, and each $P_i(k_1 + 1 \leq i \leq k)$ is a non-principal ultrafilter probability. Denote each of these ultrafilters by $U_i$.

Theorem 3.3. (Schervish et al., 1984) establishes that $\beta = \gamma_{k_1+1} + \ldots + \gamma_k$ is the least upper bound on the extent of non-conglomerable possible with $P$, over all events and all countable partitions. It is an elementary fact that we may find $k$ disjoint sets $\{A_i : A_i \in U_i(i = 1, \ldots, k)\}$, with $A_i \cap A_j = \emptyset$ for $i \neq j$. Thus, $\gamma_i P_i(\bullet) = P(\bullet \mid A_i)$, so that $P(\bullet) = \sum_i \gamma_i P(\bullet \mid A_i)$. 


Let $M$ be the (possibly empty) set of integers that index the non-Ramsey (non-principal) ultrafilters, and let $\gamma_M = \sum_{i \in M} \gamma_i$. Then we may apply Theorem 4 to obtain the following result.

**Corollary 4.** Let $P$ be as above, together with its decomposition as a mixture of ultrafilter probabilities. Then, (i) either there is a determinate partition where $P$ fails to be even approximately conglomerable (up to the extent $\gamma_M$), or (ii) based on the natural conditional probabilities for $P$, there is a determinate partition where $P$'s extent of non-conglomerability is $\gamma_M$.

**Proof.** Use the fact that the $A_i$s are disjoint to apply Theorem 4 to each $U_i(i \in M)$. Whenever (i): according to the proof of that theorem we encounter a partition of $A_i$ where $U_i$ is weakly selective but not selective, there $P$ fails to be (even approximately) conglomerable to the extent $\gamma_i$. We may concatenate these (disjoint) partitions to form a single partition where $P$ cannot be made conglomerable. If (ii): for a given $U_i(i \in M)$, it is selective in each of the (countably many) partitions used in the proof of Theorem 4, then as previously shown, when $P_i$'s natural conditional probabilities are used, it fails conglomerability in the partition $\pi^*$ of $A_i$, and the extent of nonconglomerability there is the maximum possible value, 1.

Likewise, we may apply Theorem 5 to obtain the following:

**Corollary 5.** Let $P$ be as above, together with its decomposition as a mixture of ultrafilter probabilities. Then, either there is a determinate partition where $P$ fails to be even approximately conglomerable (up to the extent $\gamma_M$), or, based on the conditional probabilities for $P_i$, there is a determinate partition where $P$'s extent of non-conglomerability is $\gamma_M/2$, at least.

**Proof.** Use Theorem 5 with each of the disjoint sets $A_i(i \in M)$.

5. **Conclusion**

We have shown how to locate partitions in which a finite-valued, merely finitely additive probability $P$ displays non-conglomerability. Our approach is to use some combinatorial properties of the associated non-principal ultrafilters for $P$ to regulate all the coherent conditional probabilities for $P$. These combinatorial properties of the associated ultrafilters are given by the unconditional probability $P$. This analysis improves upon our previous result in two ways. It demonstrates where $P$ displays non-conglomerability according to its unconditional probability even when $P$ is only two-valued, and it avoids quantifying over a continuum of partitions and conditional probabilities. Also, we hope we have indicated how some basic set-theoretic combinatorial properties of ultrafilters carry interesting consequences for finitely additive probabilities.
References


T. Seidenfeld, M.J. Schervish, and J.B. Kadane

Department of Statistics

Carnegie Mellon University

Pittsburgh, PA 15213

e-mail : tody@stat.cmu.edu
mark@stat.cmu.edu
kadane@stat.cmu.edu