Z. Wahrscheinlichkeitstheorie verw. Gebiete
66, 205-226 (1984)

Wahrscheinlichkeitstheorie
und verwandte Gebiete

The Extent of Non-Conglomerability of Finitely Additive Probabilities

Mark J. Schervish, Teddy Seidenfeld*, and Joseph B. Kadane

Statistics Department, Carnegie-Mellon University, Pittsburgh, PA 15213, USA

Summary. An arbitrary finitely additive probability can be decomposed uniquely into a convex combination of a countably additive probability and a purely finitely additive (PFA) one. The coefficient of the PFA probability is an upper bound on the extent to which conglomerability may fail in a finitely additive probability with that decomposition. If the probability is defined on a σ-field, the bound is sharp. Hence, non-conglomerability (or equivalently non-disintegrability) characterizes finitely as opposed to countably additive probability. Nonetheless, there exists a PFA probability which is simultaneously conglomerable over an arbitrary finite set of partitions.

Neither conglomerability nor non-conglomerability in a given partition is closed under convex combinations. But the convex combination of PFA ultrafilter probabilities, each of which *cannot* be made conglomerable in a common margin, is singular with respect to any finitely additive probability that is conglomerable in that margin.

1. Introduction

Kolmogorov's [11] classic treatment of the theory of probability from a frequency view-point justifies a finitely additive probability. Nonetheless he assumes a countably additive probability "for expedience", and nearly all modern writers in probability have followed him. Similarly deFinetti [7] derives a finitely additive subjective probability from axioms of coherence, although many Bayesians regard only countably additive probabilities as "proper" (Lindley [12, 13]).

Whether countable additivity is a convenient regularity condition whose assumption does not change essential results is, then, a reasonable matter to

* T. Seidenfeld is presently Associate Professor of Philosophy, Washington University, St. Louis, Missouri

explore. In this paper we examine one aspect of finite additivity, namely non-conglomerability, first noted by deFinetti [6, p. 99] some fifty years ago. Consider a partition $\pi=\{h_i: i<\alpha\}$ of pairwise exclusive and mutually exhaustive sets. If α is an infinite cardinal, it may be that for some event E

$$k_1 \leqq P(E \mid h) \leqq k_2 \quad \text{for all } h \in \pi, \tag{1.1}$$

and yet,

$$P(E) < k_1 \quad \text{or} \quad P(E) > k_2. \tag{1.2}$$

In this case we say P is not conglomerable with respect to π.

An example of non-conglomerability (attributed to P. Lévy, see [6, 5.30]) is as follows. Consider the denumerable set of points $\{\langle i, j\rangle: i, j$ positive integers$\}$. Let P be a finitely (and not countably) additive probability defined for a field $\mathscr{F}$ that includes all finite and complements of finite sets of points, subject to these two constraints:

(i) $P(\langle i, j\rangle)=0$ for all singletons,
(ii) $P(\langle i, j\rangle \mid A)=0$ if A is infinite.

Consider the event $E=\{\langle i, j\rangle: i<j\}$, i.e. E is the region of the first quadrant (in $\mathscr{F}$) above the line $i=j$. Now, note that

$$P(E \mid i=k<\omega_0)=1$$

but

$$P(E \mid j=k<\omega_0)=0.$$

Thus, depending upon the partition (by first or by second coordinate), conglomerability fixes the probability of E as 1 or 0. Hence, conglomerability fails for at least one of these partitions.

In a recent paper Dubins [8] reports that conglomerability (for random variables) in a partition π is equivalent to "disintegrability" in π. As a consequence P is conglomerable in π for all events E just in case:

$$P(E)=\int_{h \in \pi} P(E \mid h)\, dP(h), \tag{1.3}$$

i.e. iff P is the "average" of conditional probabilities $P(E \mid h)$, for $h \in \pi$. We remind the readers of the definition of integral we use with finitely additive measures. Following [9]

$$\int f(h_i)\, dP(h_i)=\sup \int g(h_i)\, dP(h_i),$$

where the supremum is taken over simple $g \leqq f$. Now g is simple iff it has the form $g=\sum_{j \in D} a_j I_{h_j}$ where D is finite. Thence

$$\int g(h_i)\, dP(h_i)=\sum_{j \in D} a_j P(h_j).$$

The principal questions addressed in this paper are these: For an arbitrary finitely additive probability P (that is not countably additive), is there a denumerable partition where conglomerability fails for some event? That is, does there exist an event E and partition $\pi=\{h_i: i=1,\dots\}$ such that, for some $\eta>0$,

$$P(E)-P(E|h_i)>\eta? \tag{1.4}$$

Second, what is the l.u.b. of η such that (1.4) holds?

In order to answer these questions, we begin in Sect. 2 with a unique decomposition of finitely additive probabilities into a convex combination given in [16]:

$$P=\alpha P_C+\beta P_D: \quad \alpha+\beta=1; \quad \alpha,\beta\geqq 0, \tag{1.5}$$

where P_C is countably additive and P_D is purely finitely additive (PFA) (c.f. [9, p. 163]). We then prove that β is an upper bound for all failures of conglomerability, i.e. the left-hand side of (1.4) is bounded above by β, in all denumerable partitions.

In Sect. 3 we find that if $\beta\neq 0$, if the range of P is not limited to finitely many distinct values and if P is defined on a σ-field of events, then the upper bound on the failure of conglomerability, β, must be approached. That is, we show that for $\eta<\beta$ there is an event E and partition $\pi=\{h_i(i<\omega_0)\}$ such that (1.4) holds. Next we consider the case in which P assumes only finitely many values. We assume that P is defined on the power set, that all conditional probabilities $P(\cdot|\cdot)$ are specified consistently and satisfy a certain "principle of conditional coherence". Then again we show in Theorem 3.3 that the upper bound on the failure of conglomerability, β, must be approached.

In Sect. 4, we show first, that there exist finitely additive (and not countably additive) probabilities that, when specified merely unconditionally, can be extended to conditional probabilities simultaneously conglomerable in any finite set of margins chosen antecedently. Thus, for this class of finitely additive distributions, the question of conglomerability in a particular margin is not determined by the unconditional distributions. Our investigation (in Sect. 4) into what may occur with respect to conglomerability in a particular margin shows also that the convex combination of two finitely additive distributions, each conglomerable in a common margin, may fail to be conglomerable in that margin. Similarly, the convex combination of two distributions that each fail to be conglomerable in a common margin may, nonetheless, be conglomerable in that margin. Hence, neither conglomerability nor non-conglomerability in a margin is closed under simple convex combinations.

Lastly, in Sect. 5, we consider a question of the connection between non-conglomerability and strong non-approximability by conglomerable distributions. It follows quickly that all countably additive probabilities are *singular* with respect to all PFA ones, and that all continuous distributions are singular with respect to ultrafilter ones. A distribution *cannot be made conglomerable* in a margin if, when specified merely unconditionally, it cannot be extended to include consistent conditional probabilities that are conglomerable in that margin. We show that the convex combination of PFA ultrafilter probabilities, each of which cannot be made conglomerable in a common margin, is singular with respect to all finitely additive probabilities that are conglomerable in that margin.

Since conditioning arguments are so common and so important in all aspects of probability theory, and its statistical applications, our results lead us to be curious about the extent to which standard countably-additive probability results carry over to the finitely additive case. To the extent that they do not, the assumption of countable additivity as a regularity condition is not innocuous.

2. Upper Bound on Failures of Conglomerability

In this section we review some important theorems about finitely additive probabilities. These theorems, in turn, lead to an upper bound on failures of conglomerability. The probabilities discussed in this section need only be defined over a field $\mathscr{F}$ of subsets of some space Ω. In later sections, we will require that probabilities be defined over a σ-field.

Definition 1.1. A probability P is *purely finitely additive* (PFA) if the only non-negative countably additive set function Q which satisfies $P \geqq Q \geqq 0$ is $Q \equiv 0$.

Yosida and Hewitt [16] prove a theorem from which the following follows trivially.

Theorem 2.1. *For every finitely additive probability P defined on a field of events $\mathscr{F}$, there exist P_C, a countably additive probability, P_D, a PFA probability, $\alpha \geqq 0$ and $\beta \geqq 0$ such that $P = \alpha P_C + \beta P_D$ and $\alpha + \beta = 1$. The numbers α and β are unique, and, if $\alpha \neq 0$, P_C is unique. Similarly, if $\beta \neq 0$, P_D is unique.*

In view of Theorem 2.1, for any finitely additive probability P, we will denote by $\beta(P)$ the coefficient of the PFA probability P_D. A different characterization of PFA probabilities than is given in Definition 1.1 proves to be more useful in proving the theorems of this paper. First we need the following definition.

Definition 2.2. A probability P is *strongly finitely additive* (SFA) if there exists a partition $\pi = \{h_1, h_2, \ldots\}$ such that $P(h_i) = 0$ for every i.

Theorem 2.2. *A probability P defined on a field of events $\mathscr{F}$ is PFA if and only if for every $\varepsilon > 0$, there exists a partition $\pi = \{h_1, h_2, \ldots\}$ such that $\sum_{i=1}^{\infty} P(h_i) < \varepsilon$.*

Proof. "if" Let Q be a countably additive non-negative set function satisfying $P \geqq Q \geqq 0$ with $Q(\Omega) = a$. For every $\varepsilon > 0$ there exists a partition as in the statement of the theorem such that $\varepsilon > \sum_{i=1}^{\infty} P(h_i) \geqq \sum_{i=1}^{\infty} Q(h_i) = a$. Hence $a < \varepsilon$ for every $\varepsilon > 0$, so $a = 0$ and $Q \equiv 0$ implying P is PFA.

"only if" Let P be PFA. Lemma 1 of [1] shows that there exist countably many SFA probabilities

$$P_i,\ i = 1, 2, \ldots \quad \text{such that } P = \sum_{i=1}^{\infty} \alpha_i P_i, \quad \text{and} \quad \sum_{i=1}^{\infty} \alpha_i = 1.$$

Let $\varepsilon>0$ be given and let K be large enough so that $\sum_{i=1}^{K} \alpha_i > 1-\varepsilon$. For each P_i, $i=1, \ldots, K$ there exists a partition π_i as described in Definition 2.2. Let π be the countable partition consisting of events each of which is the intersection of K events, one from each of the partitions π_i. Set $\pi=\{h_1, h_2, \ldots\}$. It follows that $P_i(h_j)=0$ for all j and for $i=1, \ldots, K$. Hence

$$\sum_{j=1}^{\infty} P(h_j) = \sum_{j=1}^{\infty} \sum_{i=K+1}^{\infty} \alpha_i P_i(h_j) = \sum_{i=K+1}^{\infty} \alpha_i \sum_{j=1}^{\infty} P_i(h_j) \leqq \sum_{i=K+1}^{\infty} \alpha_i < \varepsilon,$$

and the lemma is proven.

The following corollary to Theorems 2.1 and 2.2 is needed to derive the upper bound on failures of conglomerability. The proof is trivial and is omitted.

Corollary 2.1. *If P is a probability defined on a field of events and $\pi=\{h_1, h_2, \ldots\}$ is a countable partition, then* $\sum_{i=1}^{\infty} P(h_i) \geqq 1-\beta(P)$.

Theorem 2.3. *Let P be a probability defined on a field of events and $\pi=\{h_1, h_2, \ldots\}$ be a countable partition such that $P(\cdot|h_i)$ is defined for every i. If E is an event such that $P(E)-P(E|h_i) \geqq b$ for all i [or such that $P(E|h_i)-P(E) \geqq b$ for all i], then $b \leqq \beta(P)$.*

Proof. Since E may be replaced by E^c, we need only prove that $P(E)-P(E|h_i) \geqq b$ implies $b \leqq \beta(P)$. If $P(E)-P(E|h_i) \geqq b$ for all i, then

$$P(E|h_i) \leqq P(E)-b, \quad \text{for all } i. \tag{2.1}$$

Multiply both sides of (2.1) by $P(h_i) \geqq 0$ and sum over i.

$$\begin{aligned}\sum_{i=1}^{\infty} P(E|h_i)\, P(h_i) &= \sum_{i=1}^{\infty} P(E \cap h_i) \\ &\leqq [P(E)-b] \sum_{i=1}^{\infty} P(h_i) \leqq P(E)-b.\end{aligned} \tag{2.2}$$

By Corollary 2.1, $P(E^c)+\sum_{i=1}^{\infty} P(E \cap h_i) \geqq 1-\beta(P)$ since $\{E^c, E \cap h_1, E \cap h_2, \ldots\}$ is a partition. Adding $P(E^c)$ to the extremes of (2.2) yields $1-\beta(P) \leqq 1-b$, hence $b \leqq \beta(P)$ and the proof is complete.

Theorem 2.3 states that failures of conglomerability of a finitely additive probability P cannot exceed $\beta(P)$. It follows trivially that failures of disintegrability cannot exceed $\beta(P)$ either.

3. The Extent to Which Conglomerability Fails

The results of this section pertain mostly to probabilities defined on a σ-field of events. The first lemma gives a construction which is useful in the sequel.

Lemma 3.1. *If P is a probability defined on a σ-field $\mathcal{F}$ whose range is an infinite set, then for every $\varepsilon>0$, there exists a set A such that $P(A)\leqq\varepsilon$ and $A=\bigcup_{i=1}^{\infty} A_i$ with $A_i\in\mathcal{F}$, $A_i\cap A_j=\phi$ for $i\neq j$ and $P(A_i)>0$ for all i.*

Proof. Since the range of P is infinite there must exist an event $A\in\mathcal{F}$ such that $P(A)>0$ and $P(A^c)>0$. Either A or A^c (or both) has subsets with infinitely many distinct values. Let it be A^c. Set $B_1=A^c$ and $C_1=A$. Partition B_1 into $B_1=C_2\cup B_2$ where $P(C_2)>0$, $P(B_2)>0$ and B_2 contains subsets with infinitely many distinct values. Partition B_2 into $B_2=C_3\cup B_3$, etc.

The procedure above produces a sequence of disjoint events $\{C_i: i=1,\dots,\infty\}$ with $P(C_n)>0$ for all n. Let $m>2/\varepsilon$ be an integer. Define $E_i(k)=C_{(i-1)m+k}$ for $k=1,\dots,m$, $i=1,\dots,\infty$. Since $P\left(\bigcup_{n=1}^{\infty} C_n\right)=\sum_{k=1}^{m} P\left\{\bigcup_{i=1}^{\infty} E_i(k)\right\}$, there must exist a k_0 such that $P\left\{\bigcup_{i=1}^{\infty} E_i(k_0)\right\}-\sum_{i=1}^{\infty} P\{E_i(k_0)\}<\varepsilon/2$. Let n be large enough so that $\sum_{i=n}^{\infty} P\{E_i(k_0)\}<\varepsilon/2$. Set $A=\bigcup_{i=n}^{\infty} E_i(k_0)$ and $A_i=E_{i-1+n}(k_0)$ for $i=1,\dots,\infty$, and the lemma is proven.

Theorem 3.1. *Let P be a finitely additive probability satisfying:*

(i) *P is defined on a σ-field*

(ii) $\beta(P)>0$

(iii) *P assumes more than finitely many values.*

Then for every $\varepsilon>0$, there is an event E and a partition $\pi=\{h_i, i=1,\dots\}$ such that $P(E)-P(E|h_i)>\beta(P)-\varepsilon$ for all i.

Proof. Let $\varepsilon>0$ be given. Consider the decomposition given by Theorem 2.1, $P=\alpha P_C+\beta P_D$. Choose $n>2/\varepsilon$ large enough so that $\delta\equiv 1/2n<\varepsilon/(4\beta)$. Now use Lemma 3.1 to find an event $A=\bigcup_{i=1}^{\infty} A_i$ with $P(A)<\beta\delta$, $A_i\cap A_j=\emptyset$ for $i\neq j$, and $P(A_i)>0$ $\forall i$. Let $\gamma=\sum_{i=1}^{\infty} P(A_i)$. Since P_D is PFA, we can write $A^C=\bigcup_{i=1}^{\infty} C_i$ with $\sum_{i=1}^{\infty} P_D(C_i)<\delta\gamma$, by Theorem 2.2. Let k be large enough so that $\sum_{i=k+1}^{\infty} P_C(C_i)<\delta\gamma$. Let $A_1^1=\bigcup_{i=1}^{k} C_i$, $A_i^1=A_{i-1}$ for $i=2,3,\dots$, and let $\{A_i^2: i=1,\dots\}$ be the sequence of events C_{k+i}, $i=1,\dots$, arranged in order of decreasing P. It is clear that $\Omega=\bigcup_{j=1}^{2}\bigcup_{i=1}^{\infty} A_i^j$, and the A_i^j are disjoint. Set $E=\bigcup_{i=1}^{\infty} A_i^2$. Then $P(E)\geqq\beta P_D(E)=\beta\left(P_D(A^C)-\sum_{i=1}^{k} P_D(C_i)\right)$. Now $P_D(A)\leqq P(A)/\beta<\delta$, so $P_D(A^C)>1-\delta$. We know $\sum_{i=1}^{k} P_D(C_i)<\delta$. So $P(E)\geqq\beta(1-2\delta)>\beta-\varepsilon/2$. Notice that

$$\sum_{i=1}^{\infty} P(A_i^1)\geqq\sum_{i=2}^{\infty} P(A_i^1)=\gamma$$

and

$$\sum_{i=1}^{\infty} P(A_i^2) = \alpha \sum_{i=1}^{\infty} P_C(A_i^2) + \beta \sum_{i=1}^{\infty} P_D(A_i^2) < \alpha\delta\gamma + \beta\delta\gamma = \delta\gamma$$

so $\sum_{i=1}^{\infty} P(A_i^1) \Big/ \sum_{i=1}^{\infty} P(A_i^2) \geqq \gamma/(\delta\gamma) = 1/\delta = 2n$, i.e.

$$\sum_{i=1}^{\infty} P(A_i^1) \geqq 2n \sum_{i=1}^{\infty} P(A_i^2). \tag{3.1}$$

Let k_1 be the smallest integer such that $\sum_{i=1}^{k_1} P(A_i^1) \geqq n P(A_1^2)$. Clearly $k_1 < \infty$ by (3.1). Also let m_1 be the smallest integer greater than zero such that

$$\sum_{i=k_1+1}^{\infty} P(A_i^1) \geqq 2n \sum_{i=m_1+1}^{\infty} P(A_i^2), \tag{3.2}$$

which is finite since the sum on the right hand side of (3.1) converges and $P(A_i^1) > 0$ for all i.

Define $F_1 = \bigcup_{i=1}^{k_1} A_i^1$, $G_1 = \bigcup_{i=1}^{m_1} A_i^2$, and $h_1 = F_1 \cup G_1$. We claim that $P(F_1) \geqq n P(G_1)$. To see this, reason as follows: If $P(F_1) = \sum_{i=1}^{k_1} P(A_i^1) \leqq 2n P(A_1^2)$ then by subtracting from (3.1) we have $\sum_{i=k_1+1}^{\infty} P(A_i^1) \geqq 2n \sum_{i=2}^{\infty} P(A_i^2)$ which implies $m_1 = 1$, $G_1 = A_1^2$ and $P(F_1) \geqq n P(G_1)$.

If $P(F_1) > 2n P(A_1^2)$ then $P(F_1) > 2n P(A_i^2)$ for all $i \geqq 1$ since the A_i^2 are arranged by decreasing P value. Let m be the smallest integer $\geqq 1$ such that $P(F_1) \leqq 2n \sum_{i=1}^{m} P(A_i^2)$. If $m = \infty$, then $P(F_1) > 2n P(G_1) \geqq n P(G_1)$ trivially. If $m < \infty$, then by subtraction from (3.1) we get $\sum_{i=k_1+1}^{\infty} P(A_i^1) \geqq 2n \sum_{i=m+1}^{\infty} P(A_i^2)$, hence $m \geqq m_1$. So

$$2n P(G_1) \leqq 2n \sum_{i=1}^{m} P(A_i^2) = 2n P(A_m^2) + 2n \sum_{i=1}^{m-1} P(A_i^2) < 2P(F_1),$$

hence $n P(G_1) \leqq P(F_1)$ in this case also. It now follows that

$$P(E \mid h_1) = P(G_1)/[P(F_1) + P(G_1)] \leqq P(G_1)/P(F_1) \leqq n^{-1} \leqq \varepsilon/2.$$

So $P(E) - P(E \mid h_1) \geqq \beta - \varepsilon/2 - \varepsilon/2 = \beta - \varepsilon$.

Now use the fact that (3.2) is just like (3.1) but for the sequences $\{A_i^1 : i = k_1 + 1, \ldots\}$ and $\{A_i^2 : i = m_1 + 1, \ldots\}$. So repeat the above process finding $k_2 > k_1$ and $m_2 > m_1$. Set $F_2 = \bigcup_{i=k_1+1}^{k_2} A_i^1$, $G_2 = \bigcup_{i=n_1+1}^{m_2} A_i^2$, $h_2 = F_2 \cup G_2$. It follows that (3.2) and (3.3) hold with subscripts 1 replaced by 2, etc. This generates the necessary partition $\pi = \{h_i : i = 1, \ldots\}$. □

Theorem 3.1 states that $\beta(P)$ is the least upper bound on failures of conglomerability for a probability P defined on a σ-field $\mathscr{F}$ and assuming infinitely many values. It is trivial to see that $\beta(P)$ is also the least upper bound on failures of disintegrability under the same conditions. Next we state a result (Theorem 3.2, Corollary 3.1) that applies to all finitely additive probabilities. Then we state and prove Theorem 3.3, which shows that $\beta(P)$ is the least upper bound on failures of conglomerability for probabilities taking only finitely many values, as well.

Definition 3.1. An *ultrafilter* $\mathscr{U}$ is a collection of subsets of Ω satisfying

(i) if $A\in\mathscr{U}$ and $A\subset B$, then $B\in\mathscr{U}$

(ii) if $A, B\in\mathscr{U}$ then $A\cap B\in\mathscr{U}$

(iii) for every A, either $A\in\mathscr{U}$ or $A^C\in\mathscr{U}$, but not both.

Definition 3.2. A probability, P, is called an *ultrafilter probability* with atoms in an ultrafilter $\mathscr{U}$ if $\forall E\in\mathscr{U}$, $P(E)=1$. It is trivial to see that for every ultrafilter $\mathscr{U}$ there is a unique ultrafilter probability with atoms in $\mathscr{U}$.

Definition 3.3. A probability P is said to be *non-atomic* if for every event E and every $\varepsilon>0$, there exist a finite number of disjoint subsets $E_1,\ldots,E_n$ of E, such that $E=\bigcup_{i=1}^{n} E_i$ and $P(E_i)\leqq\varepsilon$, $i=1,\ldots,n$. The following result is proven in [15].

Theorem 3.2. *For every finitely additive probability P defined on a field $\mathscr{F}$, there exists a non-atomic probability P_0 and at most countably many ultrafilter probabilities $P_i (1\leqq i\leqq N\leqq\infty)$, each with atoms in an ultrafilter $\mathscr{U}_i$, such that $P=\sum_{i=0}^{N}\gamma_i P_i$ where $\sum_{i=0}^{N}\gamma_i=1$. The ultrafilters $\mathscr{U}_i$ are distinct, i.e. $\forall i<j$, $\exists E\in\mathscr{U}_i \ni E\notin\mathscr{U}_j$. Suppose the γ's are ordered so that $\gamma_1\geqq\gamma_2\geqq\ldots$. The sequence $(\gamma_i;\ 0\leqq i\leqq\infty)$ is uniquely determined. If $\gamma_0\neq 0$, P_0 is uniquely determined. Suppose $I=\{i\,|\,\gamma_i=\gamma>0\}$. I is of course a finite set. Then $\mathscr{U}_i$ is unique up to possible permutation within the set I.*

We call $\mathscr{U}_i$ the *constituent* ultrafilters of P. The following corollary to Theorem 3.2 is trivial and its proof is omitted.

Corollary 3.1. *For every finitely additive probability P defined on a field $\mathscr{F}$, there exists a continuous probability P_0, at most countably many PFA ultrafilter probabilities P_i $(1\leqq i\leqq N\leqq\infty)$ and at most countably many ultrafilter probabilities R_j $(1\leqq j\leqq M\leqq\infty)$ each countably additive such that*

$$P=\gamma_0 P_0+\sum_{i=1}^{N}\gamma_i P_i+\sum_{j=1}^{M}\delta_j R_j,$$

where $\sum_{i=1}^{N}\gamma_i+\sum_{j=1}^{M}\delta_j=1$, $\gamma_i\geqq 0$ for all $i\geqq 0$ and $\delta_j\geqq 0$ for all $j\geqq 1$. Uniqueness for γ's, δ's, $\mathscr{U}_i$'s and $\mathscr{U}_j$'s is similar to that of Theorem 3.2.

In contrast to the case considered earlier in this section, if P assumes only finitely many values, then there will not exist a partition $\pi=\{h_i\colon i=1,\ldots,\infty\}$

with $P(h_i)>0$ for all i. Hence, in order to find a failure of conglomerability, we must assume that $P(\cdot\,|\,h)$ is defined for events h with $P(h)=0$. We will assume further, the following.

Principle of Conditional Coherence. For all pairs of events A, B such that $A\cap B\neq\emptyset$, $Q(\cdot)=P(\cdot\,|\,B)$ is a finitely additive probability and $Q(\cdot\,|\,A)=P(\cdot\,|\,A\cap B)$.

The principle of conditional coherence applies trivially to events with positive probability. We assume it also applies to events of zero probability. It is a simple consequence of this principle that if h_1 and h_2 are disjoint and $P(h_1\,|\,h_1\cup h_2)\geqq 1/2$, then $P(h_1\,|\,h_1\cup h_2\cup A)\geqq P(h_2\,|\,h_1\cup h_2\cup A)$ for all sets A. Dubins (1975, sec. 3) proves that for any probability P conditional probabilities can be defined in such a way that conditional coherence holds.

We are now ready to state the final result of this section.

Theorem 3.3. *If P is a finitely additive probability, defined on the power set of a space Ω, with $\beta(P)=\beta>0$ which assumes only finitely many values and satisfies the principle of conditional coherence, then for every $\varepsilon>0$, there exists an event E and a partition $\pi=\{h_1, h_2, \ldots\}$ with $P(E)-P(E\,|\,h_i)>\beta(P)-\varepsilon$ for all i.*

The proof of Theorem 3.3 rests on the following lemma.

Lemma 3.2. *Let P be a PFA ultrafilter probability which satisfies the conditional coherence principle and has atoms in $\mathscr{U}$. Let $h_{-1}\notin\mathscr{U}$. For every $\varepsilon>0$, there exists a partition $\pi^*=\{h_{-1},\ h_0^*, h_1^*, \ldots\}$ and an event E, such that $P(E)=1$ and $P(E|h_i^*)<\varepsilon$ for all $i\geqq 0$ and $E\cap h_{-1}=\phi$.*

An outline of the proof of this lemma follows, but full details are given in the appendix. The idea is to start with a partition $\{h_{-1}, h_0, h_1, \ldots\}$ with $P(h_i)=0$ for all i and begin grouping together successive h_i (starting with h_1). The h_i's are grouped into finite collections in order that one of two results occurs. The first is that an event $E\in\mathscr{U}$ can be formed by taking the union of a few of the h_i from each collection and the partition π^* formed by the unions of the h_i's in each collection satisfies the conclusion of the Lemma. The other result that might occur is that one finite collection of h_i's may emerge with special properties. If this occurs, the special collection is removed from the partition and the remaining h_i's are grouped as before. This process continues until either the first result occurs or an infinite sequence of special collections is generated. In this last case an event E and a partition π^* can be formed in the same manner as above, either from the special collections or from the set of h_i's which are not in the special collections.

Now the proof of Theorem 3.3 can be given. Since P assumes only finitely many values, the decomposition of Theorem 3.2 has $\gamma_0=0$ and $N<\infty$. That is $P=\sum_{i=1}^{N}\gamma_i P_i$ with P_i having atoms in ultrafilter $\mathscr{U}_i$ and the $\mathscr{U}_i$ are all distinct. It follows from Corollary 3.1 that if $\beta>0$, then there are some P_i, say $P_1, \ldots, P_k$ for convenience, with P_i PFA for $i=1,\ldots,k$ and $\beta=\sum_{i=1}^{k}\gamma_i$. Since $\mathscr{U}_1$ and $\mathscr{U}_2$ are distinct, there exists an event $A\in\mathscr{U}_1$ with $A^C\in\mathscr{U}_2$. Either $A\in\mathscr{U}_3$ or $A^C\in\mathscr{U}_3$. For convenience suppose $A \in \mathscr{U}_3$. Find $B \in U_3$ such that $B^C \in \mathscr{U}_1$. Set $A' = A \cap B$, $A^* = A \cap B^C$. Then $A^* \in \mathscr{U}_1$, $A^C \in \mathscr{U}_2$, $A' \in \mathscr{U}_3$ and the three events form a partition. This process can be repeated to find a finite partition $\{A_1, \ldots, A_k\}$ such that $A_i \in \mathscr{U}_i$, $i = 1, \ldots, k$ and $A_i \notin \mathscr{U}_j$ for $i \neq j$. Define, for $i = 1, \ldots, k$ $h^i_{-1} = \bigcup_{j \neq i} A_j$. It follows that $P_i(h^i_{-1}) = 0$ for $i = 1, \ldots, k$. Define k new probabilities P_i^* by $P_i^*(\cdot \mid h) = P(\cdot \mid h)$ for all $h \notin \mathscr{U}_i$ and $P_i^*(\cdot \mid h) = P_i$ for $h \in \mathscr{U}_i$. It follows that $P_i^* = P_i$ for $i = 1, \ldots, k$ but P_i^* may differ from P_i in the way conditional probabilities are assigned. Since P satisfies the conditional coherence principle, so do all of the P_i^*. Apply Lemma 3.2 to the pair (P_i^*, h^i_{-1}) for each i to obtain events $E_1, \ldots, E_k$ and partitions $\{h^i_{-1}, h^i_0, \ldots\}$ $i = 1, \ldots, k$ with $P_i^*(E_i) = 1$, $P_i^*(E_i \mid h^i_j) < \varepsilon$ for each i, j and $E_i \cap h^i_{-1} = \emptyset$. Since $A_i = \bigcup_{j=0}^{\infty} h^i_j$ for $i = 1, \ldots, k$, it follows that $\pi = \{h^i_j : i = 1, \ldots, k;\ j = 0, \ldots, \infty\}$ forms a partition. Since $E_i \cap h^i_{-1} = \emptyset$, $E_i \cap h^j_s = \emptyset$ for all $s \geqq 0$ and $j \neq i$. So

$$P_j^*(E_i \mid h^j_s) = 0, \qquad \text{for } s \geqq 0,\ j \neq i. \tag{3.3}$$

Write $E = \bigcup_{i=1}^{k} E_i$ so that $P(E) = \sum_{i=1}^{k} P(E_i) = \sum_{i=1}^{k} \sum_{j=1}^{N} \gamma_j P_j(E_i) \geqq \sum_{i=1}^{k} \gamma_i = \beta$. Write $\pi = \{h_1, h_2, \ldots\}$ where each $h_i = h^s_j$ for some s and $j \geqq 0$. Then $P(E \mid h_i) = P(E \mid h^s_j)$ for some s and j. Hence $P(E \mid h_i) = \sum_{t=1}^{k} P(E_t \mid h^s_j) = \sum_{t=1}^{k} P_s^*(E_t \mid h^s_j) = P_s^*(E_s \mid h^s_j)$ by (3.3). Since $P_s^*(E_s \mid h^s_j) < \varepsilon$, $P(E \mid h_i) < \varepsilon$ for each i. Together with the fact that $P(E) \geqq \beta$, it follows that $P(E) - P(E \mid h_i) \geqq \beta - \varepsilon$ for all i. □

4. Conglomerability in Particular Margins When P Assumes Only Finitely Many Values

In Sect. 3, Theorem 3.3 we showed for each $\varepsilon > 0$ the existence, under mild regularity conditions on P (definition on a σ-field, satisfying conditional coherence), of a partition π and an event E such that the inequality (1.4) is satisfied by nearly as much as is allowed by Corollary 2.1, namely $\beta(P) - \varepsilon$. In this section, we take up a related topic, namely whether conditional probabilities can be defined so that, for specified partitions, conglomerability holds. That is, when P is given unconditionally and a *finite* collection of partitions specified, can conditional probabilities be defined for P (subject to conditional coherence) so that conglomerability is satisfied in each partition in the collection? In other words, is P *simultaneously conglomerable* in each partition in the collection of partitions? Sufficient conditions are given for this in Theorem 4.1, and Corollary 4.1 establishes the existence of finitely additive (and not countably additive) probabilities meeting these conditions. Then examples are given to show that the convex combination of two finitely additive probabilities, each $\left\langle \begin{matrix}\text{not conglomerable}\\ \text{conglomerable}\end{matrix} \right\rangle$ in a particular partition π, may $\left\langle \begin{matrix}\text{be}\\ \text{fail to be}\end{matrix} \right\rangle$ conglomerable in π. Hence neither finitely additive probabilities conglomerable in a partition π, nor those failing to be conglomerable in π, constitute convex sets.

Definition 4.1. A partition $\pi=\{h_i: i<\delta\leqq\omega_0\}$ has the *minimal order property* for ultrafilter $\mathscr{U}$ if there exists an $A\in\mathscr{U}$ such that

$$|\{i<\delta: |h_i\cap A|>1\}|\leqq 1.$$

Informally, a partition π has the minimal order property for $\mathscr{U}$ if there is a set A in $\mathscr{U}$ for which there is at most one member of π whose intersection with A has more than one element.

Theorem 4.1. *Let $\Pi=\{\pi_1, \pi_2, \ldots, \pi_m\}$ be a finite set of partitions. Suppose P takes only finitely many values. Then P is simultaneously conglomerable in all $\pi_i\in\Pi$ if each of its constitutent ultrafilters satisfies the minimal order property in each $\pi_i\in\Pi$.*

Proof. First we state and prove a helpful lemma.

Lemma 4.1. *Suppose P is an ultrafilter probability with atoms in $\mathscr{U}$, and suppose each $\pi_j\in\Pi$ satisfies the minimal order property for $\mathscr{U}$. Then P is simultaneously conglomerable in all $\pi_j\in\Pi$.*

Proof of Lemma 4.1. Choose mA_j's $(1\leqq j\leqq m)$ such that $A_j\in\mathscr{U}$ and

$$|\{i<\delta_j: |A_j\cap h_i^j|>1\}|\leqq 1 \qquad \text{for } h_i^j\in\pi_j \ (1\leqq j\leqq m).$$

Let h_*^j be that member of the partition π_j, unique if it exists, such that $|A_j\cap h_*^j|>1$. Since P has atoms in $\mathscr{U}$, either $P(h_*^j)=0$ or $P(h_*^j)=1$. Let

$$A_j^*=\begin{cases} A_j-h_*^j & \text{if } P(h_*^j)=0 \\ A_j & \text{if not} \end{cases}.$$

Now $A_j^*\in\mathscr{U}$.

Let $A=\bigcap_{j=1}^{m} A_j^*$. Clearly $A\in\mathscr{U}$, and, for each $\pi_j\in\Pi$,

$$|\{i<\delta_j: |A\cap h_i^j|>1\}|\leqq 1.$$

Next, construct conditional probability functions $P(\cdot\,|\,h_i^j)$ as follows: Each h_i^j belongs to exactly one of the following sets:

(i) $h_i^j\in\Pi_j^1$ if $|A\cap h_i^j|=1$. In this case, let $P(\cdot\,|\,h_i^j)$ be the discrete, countably additive probability with all its mass concentrated on that element $(A\cap h_i^j)$.

(ii) $h_i^j\in\Pi_j^2$ if $|A\cap h_i^j|>1$, and, by construction above, $P(h_i^j)=1$. In this case $P(\cdot\,|\,h_i^j)$ is determined by P and the multiplication theorem,

$$P(\cdot\,|\,h_i^j)=\frac{P(\cdot\cap h_i^j)}{P(h_i^j)}=B(\cdot\cap h_i^j)=P.$$

(iii) $h_i^j\in\Pi_j^3$ if $A\cap h_i^j=\emptyset$. Then choose one finitely additive probability defined on the power set $\mathscr{P}(\Omega-A)$, with associated conditional probability functions $P(\cdot\,|\,Y)$ for $Y\in\mathscr{P}(\Omega-A)$. Then $P(\cdot\,|\,h_i^j)$ is given by this conditional probability.

Finally, for every $E\in\mathscr{F}$ and every $\pi_j\in\Pi$, if there is an $h_*^j\in\Pi_j^2$, then

$$P(E)=P(E\,|\,A)=P(E\,|\,h_*^j)=\int_{h_i^j\in\pi_j} P(E\,|\,h_i^j)\,dP(h_i^j)$$

if not, then

$$P(E)=P(E\mid A)=\int_{h_i^j\in\Pi_j^1} P(E\mid h_i^j)\,dP(h_i^j)=\int_{h_i^j\in\pi_j} P(E\mid h_i^j)\,dP(h_i^j). \quad \square$$

Resuming the proof of Theorem 4.1, we remind the reader that by Theorem 3.2, since P takes only finitely many values, P can be written uniquely as

$$P=\sum_{\ell=1}^{r}\gamma_\ell P_\ell$$

where each P_ℓ is an ultrafilter probability with atoms in ultrafilter $\mathscr{U}_\ell$, which by hypothesis of the theorem satisfies the minimal order property. Applying the lemma, we can take P_ℓ to be conglomerable in all $\pi_j\in\Pi$, and seek to define $P(\cdot\mid h_i^j)$, $h_i^j\in\pi_j$ $(1\leqq j\leqq m)$ so that P is conglomerable in all $\pi_j\in\Pi$.

Note that for each j, there are at most finitely many $h_i^j\in\Pi_j$ such that $P(h_i^j)>0$. For these, $P(\cdot\mid h_i^j)$ is determined by P and the multiplication theorem. We assume $P(\{h_i^j\colon P(h_i^j)=0\})>0$, since otherwise P is already conglomerable in that π_j.

Since, by hypothesis, each $\mathscr{U}_\ell$ $(1\leqq\ell\leqq r)$ satisfies the minimal order property in each $\pi_j\in\Pi$, we may choose $A_\ell=\bigcap_{j=1}^{m}A^*_{\ell j}$, as in Lemma 4.1 where $A_\ell\in\mathscr{U}_\ell$ and for each π_j $(1\leqq j\leqq m)$

$$|\{i<\delta\colon |A_\ell\cap h_i|>1\}|\leqq 1, \qquad h_i\in\pi_j.$$

Since we are considering only those $h\in\pi_j\ni P(h)=0$, without loss of generality we may assume that $|A_\ell\cap h|\leqq 1$. For each such h, identify which of the (at most r) A_ℓ satisfy $|A_\ell\cap h|=1$. Let these be $A_{\ell 1},\ldots,A_{\ell q}$ $(q\leqq r)$, with associated $\mathscr{U}_{\ell 1},\ldots,\mathscr{U}_{\ell q}$, and coefficients $\gamma_{\ell 1},\ldots,\gamma_{\ell q}$. Let

$$P(\cdot\mid h)=\frac{\gamma_{\ell 1}P_{\ell 1}(\cdot\mid h)+\ldots+\gamma_{\ell q}P_{\ell q}(\cdot\mid h)}{\sum_{k=1}^{q}\gamma_{\ell k}}$$

where $P_{\ell k}(\cdot\mid h)$ is that discrete, countably additive conditional probability function, provided by Lemma 4.1, for $P_{\ell k}$.

Note. For all h such that $P(h)=0$ and $(A_\ell\cap h)=\emptyset$ $(1\leqq\ell\leqq r)$, then $P(\cdot\mid h)$ is given by any one finitely additive probability function defined on

$$\mathscr{P}\left(\Omega-\bigcup_{\ell=1}^{r}A_\ell\right).$$

Finally, by straightforward arithmetic, $P(E)=\int_{h\in\pi_j}P(E\mid h)\,dP(h)$ $\forall\,\pi_j\in\Pi$, $\forall\,E\in\mathscr{F}$. That is, for any $E\in\mathscr{F}$, E belongs to $q\leqq r$ of the $\mathscr{U}_\ell$ and for each P_ℓ, $(1\leqq\ell\leqq r)$, $\pi_j\in\Pi$,

$$P_\ell(E)=P_\ell(E\mid A_\ell)=\int_{h\in\pi_j}P_\ell(E\mid h)\,dP(h),$$

where $P_\ell(\cdot\mid h)$ is defined by Lemma 4.1. $\square$ Theorem 4.1.

[M.A.]

Corollary 4.1. *Let $\Pi = \{\pi_1, \ldots, \pi_m\}$ be a finite set of partitions. There exists a finitely additive (but not countably additive) probability which is simultaneously conglomerable over each partition in Π.*

Proof of Corollary. Let $|\Omega| = \omega_0$ and let $\mathscr{U}$ be an ultrafilter of subsets of Ω. By Theorem 9.6 of [5], *every* π satisfies the minimal order property for ultrafilter $\mathscr{U}$ if $\mathscr{U}$ is a Ramsey ultrafilter. By Theorem 9.13, *ibid.*, there are $2^{2^{\omega}}$ distinct Ramsey ultrafilters on a cardinal ω. $\square$

We remind the reader that Theorem 3.3 and Corollary 4.1 do not conflict. The latter establishes the existence of a probability P which is simultaneously conglomerable in an arbitrary *finite* collection of denumerable partitions. The former result shows that, subject to conditional coherence, the collection cannot be extended to include all denumerable partitions unless P is countably additive.

Example 4.1. The following example (to be found in [6, p. 205] and reported in [8, p. 92]) illustrates the failure of conglomerability in partition π of the convex combination of two finitely additive probabilities, each of which is (separately) conglomerable in π. Let the field be the set of (all) subsets of $X = \{a_{ij}: i = 1, 2; j = 1, 2, \ldots\}$. Let $P_C(a_{ij}) = \begin{cases} \dfrac{1}{2^j} & \text{if } i = 1 \\ 0 & \text{otherwise} \end{cases}$; hence P_C is countably additive and lives on the set $X_1 = \{a_{ij}: i = 1\}$.

Let $P_D(a_{ij}) = 0\ \forall i, j$ and let $P_D(X_1) = 0$. Finally, let $P = \alpha P_C + \beta P_D$ ($\alpha, \beta > 0$; $\alpha + \beta = 1$). Fix $\pi = \{h_j: h_j = \{a_{1j}, a_{2j}\}\}$.

Then P is *not* conglomerable in π as $P(X_1) = \alpha$ whereas $P(X_1 \mid h_j) = 1$, $\forall j$. Clearly P_C is conglomerable in π, as in P_D. (Just set $P_D(a_{2j} \mid h_j) = 1$, $\forall j$.) A similar example is reported in [2, Example 3.1].

Example 4.2. This example shows that non-conglomerability in a margin is *not* closed under simple convex combinations.

Let P_1 be P as above. That is, $P_{C_1}(a_{ij}) = \begin{cases} \dfrac{1}{2^j} & \text{if } i = 1 \\ 0 & \text{if } i = 2 \end{cases}$, so that $P_{C_1}(X_1) = 1$. Let $P_{D_1}(a_{ij}) = 0\ \forall ij$ and let $P_{D_1}(X_1) = 0$. Choose $\alpha, \beta > 0$; $\alpha + \beta = 1$ and fix $P_1 = \alpha P_{C_1} + \beta P_{D_1}$. Hence (as shown above), P_1 is not conglomerable in $\pi = \{h_j: h_j = \{a_{1j}, a_{2j}\}\}$.

Next, let P_2 be the mirror image of P_1 in X_1 and $X_2 (= X_1^c)$. That is, define $a'_{ij} = \begin{cases} a_{2j} & \text{if } i = 1 \\ a_{1j} & \text{if } i = 2 \end{cases}$ and define $X' = \{a'_{ij}: a_{ij} \in X\}$. Let $P_2(X) = P_1(X')$. Specifically $P_{C_2}(a_{ij}) = \left\{\begin{matrix} 0 & \text{if } i = 1 \\ \dfrac{1}{2^j} & \text{if } i = 2 \end{matrix}\right\}$, $P_{D_2}(X_1) = 0$, and $P_2 = \alpha P_{C_2} + \beta P_{D_2}$. Hence, P_2 likewise is not conglomerable in π. Finally, set $P^* = \frac{1}{2}P_1 + \frac{1}{2}P_2$. Then P^* *is* conglomerable in π. For example

$$P^*(a_{ij}) = \left\{\begin{matrix} \frac{1}{2}P_1(a_{ij}) & \text{if } i = 1 \\ \frac{1}{2}P_2(a_{ij}) & \text{if } i = 2 \end{matrix}\right\} = \alpha \cdot \frac{1}{2^{j+1}},$$

so that, for each $j P^*(a_{ij}|h_j)=\frac{1}{2}$ $(i=1,2)$. Moreover, $P^*(X_1)=P^*(X_2)=\frac{1}{2}$. Thus

$$P^*(X_1)=\int_{h\in\pi} P^*(X_1|h)\,dP^*(h)=\tfrac{1}{2}.$$

Last, we note that the collection of nearly disintegrable measures is convex [2].

5. Approximation by Probabilities Conglomerable in a Margin

A question left open by the previous work is whether, even though finitely additive probabilities non-congomerable in a particular margin π exist, they can be approximated well by distributions that are conglomerable in that margin. The result of this section is that in one sense, at least, they cannot.

Definition 5.1. Two finitely additive probabilities P_1 and P_2 are *singular* $P_1 \perp P_2$ if for every $\varepsilon>0$, $\exists E\in\mathscr{F}$ such that $|P_1(E)-P_2(E)|>1-\varepsilon$.

Thus P_1 and P_2 are singular if there are sets E on which they differ by nearly as much as possible.

Definition 5.2. A finitely additive probability P defined only unconditionally, *cannot be made conglomerable in* $\pi=\{h_1,\ldots\}$ if, for every set of conditional probabilities consistent with P, $\{P(\cdot|h), h\in\pi\}$ there is an $E\in\mathscr{F}$ such that, for some $\varepsilon>0$,

$$|P(E)-\int_{h\in\pi} P(E|h)\,dP(h)|>\varepsilon.$$

By saying that $P(\cdot|h)$ is consistent with P, we mean that P satisfies the principle of conditional coherence (see Sect. 3) when extended to include $P(\cdot|h)$.

Section 4 shows that if P takes only finitely many values and cannot be made conglomerable in π, then at least one of its constituents lacks the minimal order property.

Examples of distributions P that cannot be made conglomerable in particular partitions π are given in [8, Sect. 2] and [14]. Now we can state

Theorem 5.1. *Let P be a (finite or countable) convex combination of PFA ultrafilter probabilities and suppose that each cannot be made conglomerable in some partition π. Then P is singular with respect to any probability P' that is conglomerable in π.*

We begin the proof by reminding the reader of several elementary facts about singularity and absolute continuity, which we state for completeness.

Lemma 5.1. [16] *Every PFA probability P_D is singular with respect to every countably additive probability P_C.*

Lemma 5.2. [4] *Every continuous probability P_S is singular with reapect to every probability P_A that can be written as a (finite or infinite) convex combination of ultrafilter probabilities.*

Lemma 5.3. *Let* $P=\sum_{i=1}^{\infty}\gamma_i P_i$, $\sum_{i=1}^{\infty}\gamma_i=1$, *and* $Q\perp P_i$ *for all i. Then* $Q\perp P$.

The proof is simple and is omitted.

Lemma 5.4. *Suppose that P is an ultrafilter probability with atoms in* $\mathcal{U}$. *Suppose that* P^* *is conglomerable in* π, *is not singular with respect to P, and satisfies*

$$P^*=\int_{h\in\pi} P^*(\cdot\mid h)\,dP(h).$$

Then P is conglomerable in π.

Proof. Decompose P^* as follows:

$$P^*=\alpha^* P_C^*+(1-\alpha^*)\sum_{\ell=0}\gamma_\ell^* P_\ell^*,$$

where P_C^* is countably additive, P_0^* is PFA and nonatomic, and for $\ell\geqq 1$, P_ℓ^* is PFA and an ultrafilter probability with atoms in $\mathcal{U}_\ell^*$. ($\mathcal{U}_\ell^*\neq\mathcal{U}_{\ell'}^*$, if $\ell\neq\ell'$). Also $\sum_{\ell=0}\gamma_\ell^*=1$. Since, by assumption, P and P^* are not singular, there is an $\varepsilon>0$ such that

$$\sup_E |P(E)-P^*(E)|\leqq 1-\varepsilon. \tag{5.1}$$

Using Lemmas 5.1, 5.2, and 5.3, we have $P\not\perp\sum_{\ell=1}\gamma_\ell^* P_\ell^*$. Again, using Lemma 5.3 and the fact that ultrafilter probabilities are singular or identical, we have $P=P_\ell^*$ for some $\ell\geqq 1$, where $(1-\alpha^*)\gamma_\ell^*\geqq\varepsilon$. Without loss of generality we take this $\ell=1$. Next we use the decomposition Theorems 2.1 and 3.2 on each conditional probability in the set $\{P^*(\cdot\mid h_i): h_i\in\pi\}$. Thus, for each $h_i\in\pi$, we write

$$P^*(\cdot\mid h_i)=\alpha_i^* P_{ic}^*(\cdot\mid h_i)+(1-\alpha_i^*)\sum_{\ell=0}\gamma_{i\ell}^* P_{i\ell}^*(\cdot\mid h_i),$$

where $P_{ic}^*(\cdot\mid h_i)$ is countably additive, $P_{i0}^*(\cdot\mid h_i)$ is PFA and nonatomic, $P_{i\ell}^*(\cdot\mid h_i)$ ($\ell\geqq 1$) is a PFA ultrafilter probability with atoms in $\mathcal{U}_{i\ell}^*$, and where $\sum_{\ell=0}\gamma_{i\ell}^*=1$, $\gamma_{i\ell}^*\geqq\gamma_{i,\ell+1}^*$ ($\ell\geqq 1$), for all i.

Claim. For each $h_i\in\pi$ one may choose one conditional probability, say $P_{i\ell}^*(\cdot\mid h_i)$ such that

$$P_1^*=\int_{h_i\in\pi} P_{i\ell_i}^*(\cdot\mid h_i)\,dP(h_i).$$

We show the claim indirectly. First, choose λ so that $\varepsilon>\lambda>0$. Then partition π into (at most) $\frac{1}{\varepsilon-\lambda}+1$ disjoint sets $R_j\left(j=0,\ldots,\frac{1}{\varepsilon-\lambda}\right)$ where $R_j=\{h_i: P^*(\cdot\mid h_i)$ has exactly j PFA ultrafilter probabilities with coefficients greater than $\varepsilon-\lambda\}$. Observe that $P_1^*(R_j)=0$ for all but one value of j, say $P_1^*(R_{j'})=1$.

Suppose, first, $j'\neq 0$. Consider the j' selections $\{P_{i\ell}^*(\cdot\mid h_i): h_i\in R_{j'}\}$ $\ell=1,\ldots,j'$. Assume the *claim* is false. Then $P_1^*(\cdot)\neq\int_{h_i\in R_{j'}} P_{i\ell}^*(\cdot\mid h_i)\,dP(h_i)$ for $\ell=1,\ldots,j'$. But since each of these j' distributions, $\int_{h_i \in R_{j'}} P^*_{i\ell}(\cdot \mid h_i)\, dP(h_i)$ $(\ell = 1, \ldots, j')$, is a PFA ultrafilter probability with atoms in one ultrafilter, there exist j' sets $E_1, \ldots, E_{j'}$, each with $P_1^*(E_\ell) = 1$ $(\ell = 1, \ldots, j')$ and where

$$\int_{h_i \in R_{j'}} P^*_{i\ell}(E_\ell \mid h_i)\, dP(h_i) = 0 \qquad \text{for } \ell = 1, \ldots, j'.$$

Let $E' = \bigcap_{\ell=1}^{j'} E_\ell$. Then $P_1^*(E') = 1$ yet

$$\int_{h_i \in R_{j'}} P^*_{i\ell}(E' \mid h_i)\, dP(h_i) = 0 \qquad \text{for } \ell = 1, \ldots, j'.$$

Moreover, for each (unrestricted) selection of conditional, ultrafilter probabilities $\{P^*_{i\ell_i}(\cdot \mid h_i) \colon h_i \in R_{j'} \text{ and } 1 \leqq \ell_i \leqq j'\}$ (of which there are 2^{ω_0} selections if $j' > 1$),

$$\int_{h_i \in R_{j'}} P^*_{i\ell_i}(E' \mid h_i)\, dP(h_i) = 0.$$

[This follows since, if there is a selection where

$$\int_{h_i \in R_{j'}} P^*_{i\ell_i}(E' \mid h_i)\, dP(h_i) = 1 \qquad (1 \leqq \ell_i \leqq j'),$$

this selection is one of the j' distributions on a subset $R'_{j'} \subseteq R_{j'}$, where $P_1^*(R'_{j'}) = 1$. That is, for one of the j' distributions (for one value of ℓ)

$$\int_{h_i \in R'_{j'}} P^*_{i\ell_i}(\cdot \mid h_i)\, dP(h_i) = \int_{h_i \in R'_{j'}} P^*_{i\ell}(\cdot \mid h_i)\, dP(h_i),$$

where $R'_{j'} \subseteq R_{j'}$ and $P_1^*(R'_{j'}) = 1$. But, for each of the j' distributions and for each $R'_{j'} \subseteq R_{j'}$ with $P_1^*(R'_{j'}) = 1$, $\int_{h_i \in R'_{j'}} P^*_{i\ell}(E \mid h_i)\, dP(h_i) = 0$.]

Let $\pi_0 = \{h_i \colon h_i \in R_{j'} \text{ and } P^*_{i\ell}(E' \mid h_i) = 0 \text{ for } \ell = 1, \ldots, j'.\}$ Then $P_1^*(\pi_0) = 1$. If $j' = 0$, let $\pi_0 = R_0$.

Thus, there exists an E, $P_1^*(E) = 1$, yet for every $h_i \in \pi_0$ $P^*_{i\ell}(E \mid h_i) = 0$ for all ℓ such that $(1 - \alpha^*)\gamma^*_{i\ell} > \varepsilon - \lambda$.

For each $h_i \in \pi_0$ partition h_i into the biggest integer $J \leqq \dfrac{2}{\varepsilon - \lambda}$ disjoint sets h_{ik} $(k = 1, \ldots, J)$, such that $P^*(E \cap h_{ik} \mid h_i) \leqq \varepsilon - \lambda$. Consider the J disjoint sets $E_k^\bullet = \bigcup_{h_i \in \pi_0} (E \cap h_{ik})$ $(k = 1, \ldots, J)$. $P_1^*(E_k^\bullet) = 0$ for all but one value of k. Without loss of generality, let $P_1^*(E_1^\bullet) = 1$.

However,

$$\begin{aligned} P^*(E_1^\bullet) &= \int_{h_i \in \pi} P^*(E_1^\bullet \mid h_i)\, dP(h_i) \\ &= \int_{h_i \in \pi_0} P^*(E_1^\bullet \mid h_i)\, dP(h_i) = \int_{h_i \in \pi_0} P^*(E \cap h_{i1} \mid h_i)\, dP(h_i) \leqq \varepsilon - \lambda. \end{aligned}$$

Thus, $(P_1^*(E_1^\bullet) - P^*(E_1^\bullet)) > 1 - \varepsilon$ in contradiction with the initial assumptions. This established the *claim*.

Now, since $P_1^* = P$, P is conglomerable in π using the set of conditional probabilities $\{P_{i\ell_i}^*(\cdot \mid h_i) : h_i \in \pi\}$ provided by the claim. □

Lemma 5.5. *Suppose P' is conglomerable in π, i.e.*

$$P' = \int_{h\in\pi} P'(\cdot \mid h)\, dP'(h),$$

and let

$$P^+ = \int_{h\in\pi} P'(\cdot \mid h)\, dP(h).$$

Suppose $\sup_E |P(E) - P'(E)| \leqq 1 - \varepsilon$ *for some* $\varepsilon > 0$.

Then $\sup_E |P(E) - P^+(E)| \leqq 1 - \varepsilon$.

Proof. If not, there is an $E' \ni P(E') = 1$ and a set

$$S' = \{h_i \in \pi : P'(E' \mid h_i) \leqq \varepsilon - \delta\} \qquad (\delta > 0),$$

where $P(S') = 1$. But then

$$\begin{aligned} P'(E' \cap S') &= \int_{h\in S'} P'(E' \cap S' \mid h)\, dP'(h) + \int_{h\in S'^c} P'(E' \cap S' \mid h)\, dP'(h) \\ &= \int_{h\in S'} P'(E' \cap S' \mid h)\, dP'(h) + 0 \leqq \varepsilon - \delta. \end{aligned}$$

But $P(E' \cap S') = 1$, so $P(E' \cap S') - P'(E' \cap S') > 1 - \varepsilon$, a contradiction. □

Lemma 5.6. *Let P be a PFA ultrafilter probability with atoms in $\mathscr{U}$, and suppose P cannot be made conglomerable in π. Then P is singular with respect to every P' that is conglomerable in π.*

Proof. Suppose the contrary, that is, that P' is conglomerable in π and not singular with respect to P. Then Lemma 5.5 applies, and shows that P^+ is not singular with respect to P. Let $h' \in \pi$. Then $P'(h' \mid h) = 0$ if $h \neq h'$ and 1 if $h = h'$.

Hence

$$P^+(h') = P(h') \qquad \text{for all } h' \in \pi.$$

Thus taking $P^+(\cdot \mid h) = P'(\cdot \mid h)$, P^+ is conglomerable in π, and satisfies

$$P^+ = \int_{h\in\pi} P^+(\cdot \mid h)\, dP(h).$$

Consequently P^+ satisfies the requirements for P^* in Lemma 5.4, so P is conglomerable in π, which contradicts the assumption. □

The proof of Theorem 5.1 is now immediate from Lemmas 5.6 and 5.3. □

Last we note that Lemma 5.6 may fail for P which are PFA but nonatomic. For instance, let P_D of Example 4.1 be nonatomic in addition to the stated conditions. Then P retains its non-conglomerability in π. However, as shown in [8, Sect. 2, p. 95], since π is simple the finitely additive probabilities which are conglomerable in π are norm-dense. That is, P is approximable by a sequence of finitely additive probabilities, each conglomerable in π.

In response to a question we posed, both a refere and W. Sudderth point out that a consequence of (1.2) in [14] is the existence of a PFA, continuous probability P' and a partition π in which P' cannot be made conglomerable, but where P' is *not* singular with respect to a P^* which is conglomerable in π yet where

$$\sup_X |P'(X) - P^*(X)| \geqq 0.5.$$

7. Conclusion

The combination of Theorem 2.3, Theorem 3.1, Theorem 3.3 and the regularity condition that P is defined on a σ-field, shows that for every $\varepsilon > 0$, there exist partitions π and events E such that P fails to be conglomerable in E with respect to π by as much as possible, that is, by as much as P fails to be countably additive, minus ε. Should we be concerned about the regularity condition? A simple application of the Hahn-Banach Theorem (see [3]) shows that every finitely additive probability defined on a field can be extended (perhaps in many ways) to a σ-field, and in fact, to the power set. Consequently our result says that for a finitely additive probability defined on a field, every extension to a σ-field must fail conglomerability by no more than our bound. The only way to escape our conclusion, then, is to refuse to extend the finitely additive probability to a σ-field. Furthermore, there exist finitely additive probabilities defined on fields which are not σ-fields, but for which the conclusion to Theorem 3.1 is true. For such probabilities, maximal failure of conglomerability is inescapable.

Failure of conglomerability, then, rather than being an aberration is typical of finitely additive probabilities that are not countably additive. We believe that this has important statistical consequences which we discuss in [10].

Acknowledgement. We thank W. Sudderth for his many helpful comments. Also, we are grateful to W. Fleissner for suggesting we consult Comfort and Negrepontis's book on ultrafilters. In addition we appreciate the constructive advice offered by a referee.

References

1. Armstrong, T.E., Prikry, K.: κ-Finiteness and κ-Additivity of Measures on Sets and Left Invariant Measures on Discrete Groups. Proc. Amer. Math. Soc. **80**, 105–112 (1980)
2. Armstrong, T.E., Sudderth, W.: Nearly Strategic Measures. Pacific J of Math. Forthcoming (University of Minnesota School of Statistics Technical Report No. 334), (1981)
3. Ash, R.B.: Real Analysis and Probability. New York: Academic Press 1972
4. Bochner, S., Phillips, R.S.: Additive Set Functions and Vector Lattices. Ann. Math. **42**, 316–324 (1941)
5. Comfort, W.W., Negrepontis, S.: The Theory of Ultrafilters. Berlin-Heidelberg-New York: Springer 1974
6. DeFinetti, B.: Probability, Induction and Statistics. New York: Wiley 1972
7. DeFinetti, B.: The Theory of Probability. (2 volumes) New York: Wiley 1974
8. Dubins, L.E.: Finitely Additive Conditional Probabilities, Conglomerability and Disintegrations. The Ann. Probability **3**, 89–99 (1975)

9. Dunford, N., Schwartz, J.T.: Linear Operations. Part I: General Theory. New York: Interscience 1958
10. Kadane, J.B., Schervish, M.J., Seidenfeld, T.: Statistical implications of finitely additive probability. In: Bayesian Inference and Decision Techniques with Applications: Essays in Honor of Bruno de Finetti, Goel, P.K., Zellner, A. (eds.) (to appear)
11. Kolmogorov, A.N.: Foundations of the Theory of Probability. New York: Chelsea 1956
12. Lindley, D.: Bayesian Statistics, A Review. Philadelphia: SIAM, 1971
13. Lindley, D.: Comment on 'Strong Inconsistency from Uniform Priors'. J. Amer. Statist. Assoc. 71, 120–121 (1976)
14. Prikry, K., Sudderth, W.: Singularity with Respect to Strategic Measures. Unpublished manuscript (1980)
15. Sobczyk, A., Hammer, P.: A Decomposition of Additive Set Functions. Duke J. Math. 11, 839–846 (1944)
16. Yosida, K., Hewitt, E.: Finitely Additive Measures. Trans. Amer. Math. Soc. **72**, 46–66 (1952)

Received August 15, 1981; 2n revised form August 15, 1983

Appendix. Proof of Lemma 3.2.

The proof of Lemma 3.2, as outlined in Sect. 3, requires two other lemmas which are stated and proven first. All of the results of this appendix apply to a PFA ultrafilter P (satisfying the principle of conditional coherence), and whose atoms are in an ultrafilter $\mathscr{U}$ of subsets of a space Ω.

Lemma A.1. Let $\pi=\{h_{-1}, h_0, h_1, \dots\}$ be a partition such that $P(h_i)=0$ for all i. Let G_k be the union of finitely many h_i, say $G_k=\bigcup_{j=1}^{n_k} a_j^k$ with each $a_j^k=h_i$ for some i, for $k=1,\dots,\infty$. Suppose the G_k are disjoint, and define $G_0=\left(\bigcup_{k=1}^{\infty} G_k\right)^c\setminus h_{-1}$. Assume $G_0\notin\mathscr{U}$. Suppose that $P(a_j^k|G_k)$ is maximized over j at exactly q values of j for each $k\geqq 1$. Then $\mathscr{A}=\{h_{-1}, G_0, G_1, \dots\}$ is a partition, and there exists an event E such that $P(E)=1$, $E\cap h_{-1}=\emptyset$, and $P(E|G_i)\leqq 1/q$ for all i.

Proof. It is clear that $\mathscr{A}$ is a partition. Let X equal j times the indicator of whether a_j^k occurs. Conditional on G_k, X is a random variable taking one of the values $1,\dots,n_k$ for $k\geqq 1$. Let $m_k(i)$ be the conditional i/q quantile of X given G_k for $k=1,2,\dots$, $i=1,\dots,q$. Define $m_k(0)=1$ and $m_k(q+1)=n_k$. Set

$$C_i=\bigcup_{k=1}^{\infty}\bigcup_{j=m_k(i)+1}^{m_k(i+1)-1} a_j^k,\ i=0,\dots,q \quad\text{and}\quad D_i=\bigcup_{k=1}^{\infty} a_{m_k(i)}^k,\quad i=0,\dots,q+1.$$

Each event C_i is defined so that $P(C_i|G_k)<1/q$. Each D_i is the union of one a_j^k for each k. So for each k, $P(D_i|G_k)=P(a_j^k|G_k)$ for some j. Since $P(a_j^k|G_k)$ is maximized at exactly q values of j, $P(D_i|G_k)\leqq 1/q$ for all i, k. It is easy to see that $\Omega=h_{-1}\cup G_0\cup\left(\bigcup_{i=0}^{q+1} D_i\right)\cup\left(\bigcup_{i=0}^{q} C_i\right)$. Since $h_{-1}\notin\mathscr{U}$ and $G_0\in\mathscr{U}$, it must be that either one of the C_i or one of the D_i is in $\mathscr{U}$. Set one such event equal to E and note that $E\cap h_{-1}=\emptyset$ to finish the proof.

Lemma A.2. *Let $\pi=\{h_{-1}, h_0, h_1, \ldots\}$ be a partition with $P(h_i)=0$ for all i. Let G_k be the union of finitely many h_i, say $G_k=\bigcup_{j=1}^{n_k} a_j^k$ with each $a_j^k=h_i$ for some i, for $k=1,\ldots,\infty$. Suppose the G_k are disjoint, and define $G_0=\left(\bigcup_{k=1}^{\infty} G_k\right)^c \backslash h_{-1}$. Assume $G_0 \notin \mathscr{U}$. Suppose that $P(a_j^k|G_k)$ is maximized at S_k distinct values of j for each $k\geqq 1$, with $1\leqq S_k\leqq n$. Suppose in addition that $P(a_i^k|a_i^k\cup a_j^t)<\frac{1}{2}$ whenever $t>k$ and j maximizes $P[a_j^t|G_t]$ (or $>\frac{1}{2}$ whenever $t>k$ and i maximizes $P[a_i^k|G_k]$). Then for every $\varepsilon>0$, there exist a partition $\pi^*=\{h_{-1}, h_0^*, h_1^*, \ldots\}$ and an event E such that $P(E)=1$, $E\cap h_{-1}=\emptyset$, and $P(E|h_i^*)<\varepsilon$ for all i.*

Proof. It is clear that $\mathscr{A}=\{h_{-1}, G_0, G_1, \ldots\}$ is a partition. Let $q>1/\varepsilon$ and construct events $C_0, \ldots, C_q$ and $D_0, \ldots, D_{q+1}$ exactly as in the proof of Lemma A.1. Each C_i has $P(C_i|G_k)<1/q$ for all k. If one of the $C_i\in\mathscr{U}$, the proof is complete. Assume then that $D_x\in\mathscr{U}$ for some x. Write $D_x=\bigcup_{k=1}^{\infty} b_k$, where $b_k=a_{m_k}^k(x)\subseteq G_k$ in the notation of Lemma A.1. Define $H_i=\bigcup_{j=0}^{\infty} b_{(q+1)j+i}$ for $i=1,\ldots,q+1$. The H_i are disjoint and their union is D_x, so exactly one of them is in $\mathscr{U}$, say $H_s\in\mathscr{U}$.

Consider first the case in which $P(a_i^k|a_i^k\cup a_j^t)<1/2$ whenever $t>k$ and j maximizes $P(a_j^t|G_t)$. Define $h_{j+1}^*=\bigcup_{i=0}^{q} G_{(q+1)j+s+i}$ for $j=0,1,\ldots$. Set $E=H_s$ and notice that $E\cap h_{j+1}^*=b_{(q+1)j+s}$. For each pair (i,j), let $C_{j,i}=a_m^{(q+1)j+s+i}$ for any m which maximizes $P(a_m^{(q+1)j+s+i}|G_{(q+1)j+s+i})$. It follows that

$$P(b_{(q+1)j+s}|b_{(q+1)j+s}\cup C_{j,i})<1/2 \qquad \text{for } i=1,\ldots,q \text{ and all } j.$$

Since the $C_{j,i}$ are disjoint and each $C_{j,i}\subseteq h_{j+1}^*$, it follows from conditional coherence that $P(E|h_{j+1}^*)<1/q<\varepsilon$ for $j=0,1,\ldots$. Let $h_0^*=\bigcup_{i=0}^{s-1} G_i$ and notice $E\cap h_{-1}=\phi$ to finish the proof.

Next let $P(a_i^k|a_i^k\cup a_j^t)>1/2$ whenever $t>k$ and i maximizes $P(a_i^k|G_k)$. Define $h_j^*=\bigcup_{i=0}^{q} G_{(q+1)j+s+i}$ for $j=1,2,\ldots$. Set $E=H_s\backslash b_s$ so that $P(E)=1$ also. Notice that $E\cap h_j^*=b_{(q+1)j+s}$. For each pair (i,j), let $C_{j,i}=a_m^{(q+1)j+s-i}$ for any m which maximizes $P(a_m^{(q+1)j+s-i}|G_{(q+1)j+s-i})$. It follows that

$$P(b_{(q+1)j+s}|b_{(q+1)j+s}\cup C_{j,i})<1/2 \qquad \text{for } i=1,\ldots,q \text{ and all } j.$$

Since $C_{j,i}$ are disjoint and each $C_{j,i}\subseteq h_{j+1}^*$, it follows from conditional coherence that $P(E|h_j^*)<1/q<\varepsilon$ for $j=1,2,\ldots$. Let $h_0^*=\bigcup_{i=0}^{s} G_i$ and notice $E\cap h_{-1}=\phi$ to finish the proof.

Proof of Lemma 3.2. Since P is a PFA ultrafilter probability, there exists a partition $\pi=\{h_{-1}, h_0, h_1, \ldots\}$ such that $P(h_i)=0$ for all i, with h_{-1} any event as specified in the statement of the lemma, i.e., $h_{-1}\notin\mathscr{U}$. Let $q>1/\varepsilon$. Construct a sequence of partitions $\{\pi_n^1\}_{n=1}^{\infty}$ as follows. First set $k=0$, $k^*=q+1$, and $n=1$. These will be indices to make the notation simpler. Define $H_1(1)=\bigcup_{i=1}^{q} h_i$, $G_0(1)=h_0$, and $\pi_n^1=\{h_{-1}, G_0(1), \dots, G_k(1), H_n(1), h_{k^*}, h_{k^*+1}, \dots\}$. To form partition π_{n+1}^1, consider two cases.

Case 1) $\max_i P[h_j|H_n(1)]$ occurs for at most $q-1$ values of i. In this case set $H_{n+1}(1)=H_n(1)\cup h_{k^*}$, $\pi_{n+1}^1=\{h_{-1}, G_0(1), \dots, G_k(1), H_{n+1}(1), h_{k^*+1}, \dots\}$, $n=n+1$, and $k^*=k^*+1$.

Case 2) $\max_i P[h_i|H_n(1)]$ occurs for exactly q values of i. In this case set $G_{k+1}(1)=H_n(1)$, $H_{n+1}(1)=\bigcup_{i=k^*}^{k^*+q-1} h_i$, $\pi_{n+1}^1=\{G_0(1), \dots, G_{k+1}(1), H_{n+1}(1), h_{k^*+q}, \dots\}$, $n=n+1$, $k^*=k^*+q$, and $k=k+1$. It follows from conditional coherence that at each step, either case 1 or case 2 will apply.

Continue the above process, generating the desired sequence $\{\pi_n^1\}_{n=1}^{\infty}$ with the properties that if $G_i(1)\in\pi_n^1$ then $G_i(1)\in\pi_m^1$ for all $m>n$, that $G_i(1)$ is the union of finitely many h_i, and that $P[h_j|G_i(1)]$ is maximized at exactly q values of j for all $i\geqq 1$. The sequence also has the property that if there are only finitely many $G_i(1)$, then there exists n^* such that for all $m\geqq n^*$, $P[h_i|H_m(1)]$ is maximized for at most $q-1$ values of i, say

$$J_m=\{i_m(1), \dots, i_m(t_m)\}$$

with $t_m<q$. If $i_m(1)<j<i_m(s)\,\&\,j\notin J_m$, then $P[h_j|h_j\cup h_{i_m}(s)\,i_m(s)]<1/2$ for all s. There are now three cases to consider.

Case I) There are infinitely many $G_i(1)$. Set $G_i=G_i(1)$, $i=0, 1, \dots$, and apply Lemma A.1 to complete the proof.

Case II) There are only finitely many $G_i(1)$ and there are infinitely many distinct sets J_m. For each $m>n^*$ $H_{m+1}(1)=H_m(1)\cup h_{k^*}$; hence, by conditional coherence either $J_m\subseteq J_{m+1}$ or $J_m\cap J_{m+1}=\phi$ and $J_{m+1}=\{k^*\}$. Note also that if $J_m\cap J_n=\phi$ with $n>m$, then for all $i\in J_m$, $j\in J_n$, $i<j$. It follows, then, that there are infinitely many disjoint J_m with $m>n^*$. Rename them $L_1, L_2, \dots$ with $L_i=\{j_i(1), \dots, j_i(s_i)\}$, $j_i(s)<j_i(t)$ for $s<t$, $s_i<q$ for all i, and $j_i(s)<j_k(t)$ for all s, t if $i<k$. Then $P[h_i|h_i\cup h_{j_k(s)}]<1/2$ for all s and $j_1(1)\leqq i<j_k(1)$. Define $G_k=\bigcup_{i=j_k(1)}^{j_{k+1}(1)-1} h_i$ for $k=1, 2, \dots$. With $G_0=\bigcup_{i=0}^{j_1(1)-1} h_i\notin\mathscr{U}$, we can apply Lemma A.2 to finish the proof.

Case III) There are finitely many $G_i(1)$, say n_1, and there are only finitely many distinct sets J_m. It follows that the last such J_m, call it J^1 has the property that for all h_i such that $i\notin J^1$, $i\geqq 1$, and $h_i\nsubseteq\bigcup_{j=1}^{n_1} G_j(1)$; $P[h_i|h_i\cup h_j]<1/2$ for all $j\in J^1$. Let $P_1(\cdot)$ be $P(\cdot)$ restricted to $S\setminus\bigcup_{i\in J^1} h_i$ and let $\pi_1=\pi\setminus\{h_i\colon i\in J^1\}=\{h_{-1}, h_0, h_{i_1}, h_{i_2}, \dots\}$. Perform the construction (as above) of partitions $\{\pi_n^2\}_{n=1}^{\infty}$ beginning with P_1 and π_1. That is $G_0(2)=h_0$, $k=0$, $k^*=q+1$, $n=1$, $H_1(2)=\bigcup_{j=1}^{q} h_{i_j}$, $\pi_1^2=\{h_{-1}, G_0(2), \dots, G_k(2), H_n(2), h_{i_{k^*}}, \dots\}$ etc. Stop this procedure if ever case I or case II prevails for some sequence $\{\pi_n^k\}_{n=1}^{\infty}$. Set $h_0^1 = h_0 \cup \bigcup_{i=1}^{k=1} \bigcup_{j\in J^i} h_j$, and apply Lemma A.1 for case I or A.2 for case II. If neither case I nor case II prevails, then a sequence of sets $J^1, J^2, \ldots$ has been generated, with the property that $P[h_i | h_i \cup h_j] < 1/2$ if $i \in J^s$, $j \in J^t$ for $t < s$. Because the procedure which generates $\{\pi_n^{k+1}\}_{n=1}^{\infty}$ is identical to that which produces $\{\pi_n^k\}_{n=1}^{\infty}$ up to the step at which the latter encounters the first $i \in J^k$, at which time all $G_j(k)$ have already been formed, it follows that $G_i(k) = G_i(k+1)$ for $i = 0, \ldots, n_k$ and there may be additionally $G_{n_k+1}(k+1), \ldots, G_{n_{k+1}}(k+1)$. If

$$\bigcup_{k=1}^{\infty} \bigcup_{i=0}^{n_k} G_i(k) = A \in \mathscr{U}, \quad \text{order the } G_i(k) \text{ and apply Lemma A.1.}$$

to finish the proof. If $A^c \in \mathscr{U}$, then note that A^c consists of h_{-1}, all h_i with $i \in \bigcup_{j=1}^{\infty} J^j = B$, and all other $h_i \nsubseteq G_j(k)$ for any j or k at all. These last h_i have the property that $P(h_i | h_i \cup h_j) < 1/2$ for all $j \in B$. Number all such h_i as $h_{v_1}, \ldots, h_{v_N}$ with $N = \infty$ possible. Define $G_k = (\bigcup_{i \in J^k} h_i) \cup h_{v_k}$ for $k = 1, 2, \ldots$. Lemma A.2 now applies and the proof is complete.