A COUPLE OF REMARKS ON THE CONVERGENCE OF $\sigma$-FIELDS ON PROBABILITY SPACES

MATIJA VIDMAR

ABSTRACT. The following modes of convergence of sub-$\sigma$-fields on a given probability space have been studied in the literature: weak convergence, strong convergence, convergence with respect to the Hausdorff metric, almost-sure convergence, set-theoretic convergence, monotone convergence. It is noted that all preserve independence in the limit, and all are invariant under passage to an equivalent probability measure. Partial results for the case of operator-norm convergence obtain.

1. INTRODUCTION

Fix a (not necessarily complete) probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and set $\mathcal{N} := \mathbb{P}^{-1}([0]) = \{F \in \mathcal{F} : \mathbb{P}(F) = 0\}$. For a sub-$\sigma$-field $\mathcal{A}$ of $\mathcal{F}$, $\overline{\mathcal{A}} := \mathcal{A} \vee \sigma(\mathcal{N}) = \sigma(\mathcal{A} \cup \mathcal{N})$, the $\mathbb{P}$-completion of $\mathcal{A}$. A $\sigma$-subfield means a $\mathbb{P}$-complete sub-$\sigma$-field of $\mathcal{F}$, i.e. a sub-$\sigma$-field of $\mathcal{F}$ that is equal to its $\mathbb{P}$-completion. The collection of all $\sigma$-subfields is denoted $\mathfrak{F}$. For an $\mathcal{F}/\mathcal{B}([-\infty, \infty])$-measurable map $f$ satisfying $\int f^+ d\mathbb{P} \land \int f^- d\mathbb{P} < \infty$, $P_f := \mathbb{E}^\mathbb{P}[f]$; if further $\mathcal{A} \in \mathfrak{F}$, then $P_A f := \mathbb{E}^\mathbb{P}[f|\mathcal{A}]$, the conditional expectation of $f$ w.r.t. $\mathcal{A}$ under $\mathbb{P}$.

Recall now that for a sequence $(\mathcal{B}_n)_{n \in \mathbb{N}} \subset \mathfrak{F}$, classical martingale theory gives, for any $f \in L^1(\mathbb{P})$, the convergence $P_{\mathcal{B}_n} f \to P_{\mathcal{B}_0} f$ in $L^1(\mathbb{P})$ and $\mathbb{P}$-a.s. as $n \to \infty$, provided one has

(MC) Monotone convergence. $\mathcal{B}_n \subset \mathcal{B}_{n+1}$ for all $n \in \mathbb{N}$ and $\mathcal{B}_0 = \bigvee_{n \in \mathbb{N}} \mathcal{B}_n$, or $\mathcal{B}_n \supset \mathcal{B}_{n+1}$ for all $n \in \mathbb{N}$ and $\mathcal{B}_0 = \bigwedge_{n \in \mathbb{N}} \mathcal{B}_n$ [7, Theorem 6.23].

Generalizing/complementing this monotone convergence, the following ways of making precise the concept of convergence of a sequence of $\sigma$-subfields $(\mathcal{B}_n)_{n \in \mathbb{N}}$ to a $\sigma$-subfield $\mathcal{B}_0$ (under $\mathbb{P}$), have been proposed and studied in the literature (among others; all the convergences are as $n \to \infty$):

(WC) Weak convergence. $\mathcal{B}_n$ converges weakly to $\mathcal{B}_0$ if $P_{\mathcal{B}_n} 1_A \to P_{\mathcal{B}_0} 1_A = 1_A$ in $\mathbb{P}$-probability for every $A \in \mathcal{B}_0$ [11, 2].

(SC) Strong convergence. $\mathcal{B}_n$ converges strongly to $\mathcal{B}_0$ if $P_{\mathcal{B}_n} 1_A \to P_{\mathcal{B}_0} 1_A$ in $\mathbb{P}$-probability for every $A \in \mathcal{F}$ [9, 15, 3, 2, 13, Problem IV.3.2] [19, Section 2] [6, Section VIII.2].

(HC) Hausdorff convergence. $\mathcal{B}_n$ converges to $\mathcal{B}_0$ w.r.t. the Hausdorff metric if $D(\mathcal{B}_n, \mathcal{B}_0) \to 0$, where for $\mathcal{A} \in \mathfrak{F}$ and $\mathcal{B} \in \mathfrak{F}$, $D(\mathcal{A}, \mathcal{B}) := \rho(\mathcal{A}, \mathcal{B}) + \rho(\mathcal{B}, \mathcal{A})$ with $\rho(\mathcal{A}, \mathcal{B}) := \sup_{A \in \mathcal{A}} \inf_{B \in \mathcal{B}} \mathbb{P}(A \Delta B)$ [4, 17, 12, 14, 20, 10, 6, Section VIII.2].

2010 Mathematics Subject Classification. Primary: 60A05, 46B28; Secondary: 28A05, 28A20.

Key words and phrases. Convergence of $\sigma$-fields; negligible sets of probability measures; independence.
A COUPLE OF REMARKS ON THE CONVERGENCE OF $\sigma$-FIELDS ON PROBABILITY SPACES

(STC) Set-theoretic convergence. $B_n$ converges to $B_0$ in the set-theoretic sense if
\[
\liminf_{n \to \infty} B_n := \bigvee_{n \geq 1} \bigcap_{k \geq n} B_k = B_0 = \bigcap_{n \geq 1} \bigvee_{k \geq n} B_k =: \limsup_{n \to \infty} B_n \tag{5,16 Problem II.6}.
\]

(ASC) Almost-sure convergence. $B_n$ converges to $B_0$ in the almost-sure sense if $P$-a.s. $P_{B_n} f \to P_{B_0} f$ for any $f \in L^1(P) \tag{1}$.

And, for $\{p, q\} \subset [1, \infty]$, $q \leq p$:

(ONC$_p^q$) Operator-norm convergence. $B_n$ converges to $B_0$ in the operator-norm sense if $P_{B_n} \to P_{B_0}$ in the operator norm $\| \cdot \|_{L^p \to L^q}$ when viewed as mappings between the (real) normed spaces $(L^p(P), \| \cdot \|_{L^p(P)})$ and $(L^q(P), \| \cdot \|_{L^q(P)}) \tag{17}$.

Beyond the obvious relevance of these convergence types to the issue of continuity of conditional expectations w.r.t. the conditioning $\sigma$-field, we note applications in statistics $\tag{8,18}$, studying closeness and convergence of information $\tag{20}$ $\tag{6 Section VIII.2}$, to the theory of noises $\tag{19}$ (see also the references therein).

As for our contribution, we shall demonstrate two desirable properties shared by (essentially) all these modes of convergence. First, they all preserve independence in the limit — and this claim generalizes to conditional independence, save for $\tag{WC}$ — see Section $\tag{3}$ (in particular Remark $\tag{3.2}$ for the precise meaning of preservation of independence in the limit). Second, excepting (perhaps) only $\tag{ONC_p^q}$ when $p = q$, all are invariant under passage to an equivalent probability measure (the latter is trivial for $\tag{MC}$ and $\tag{STC}$ but not obvious for the others) — see Section $\tag{4}$ Given that $\sigma$-subfields are often interpreted as bodies of information, and hence convergence of these as a convergence of information, it is certainly note-worthy that all these types of convergence do in fact depend on the underlying probability measure $P$ only via $\mathcal{N} = P^{-1} (\{0\})$ (which, short of dispensing with the probability measure altogether, is surely the best we can hope for). Likewise, independence is a fundamental probabilistic property — its preservation in the limit of $\sigma$-fields deserves to be made explicit. We will indeed see in relation to this, that the simultaneous consideration of the various convergence types enunciated above allows for a great economy of argument.

We will also show en passant that $\tag{ONC_p^q}$ with $q < p$ is equivalent to $\tag{HC}$, while the case $p = q = 1$ or $p = q = \infty$ is vacuous (apart from the trivial case of $(B_n)_{n \in \mathbb{N}}$ ultimately constant), but the case $p = q \in (1, \infty)$ is not $\tag{Section 2}$.

Finally, in terms of what has been noted in the literature in connexion to this thus far:

- $\tag{19}$ Corollary 3.6 gives, assuming $L^2(P)$ is separable, preservation of pairwise independence in the strong limit. (Were the join operation $\bigvee$ (sequentially) continuous under $\tag{SC}$ then preservation of independence would be an essentially immediate corollary. But it is not, see e.g. the example of $\tag{19}$ Section 1.2.)

- $\tag{6}$ Theorem VIII.2.23 gives invariance of $\tag{HC}$ under passage to a “uniformly absolutely continuous” (see $\tag{6}$ Definition VIII.2.22) probability measure.

- $\tag{6}$ Theorem VIII.2.40 gives invariance of $\tag{SC}$ under passage to an equivalent probability measure, assuming $L^1(P)$ is separable.
(Note that for \( r \in [1, \infty) \), the separability of \( L^r(P) \) is equivalent to \( \mathcal{F} = \sigma(\mathcal{A})^p = \sigma(\mathcal{A} \vee \mathcal{N}) \) for some denumerable \( \mathcal{A} \subset \mathcal{F} \).)

2. Preliminaries

We gather some relevant results scattered in the literature and make some observations. Let \( (\mathcal{B}_n)_{n \in \mathbb{N}_0} \) be a sequence of \( \sigma \)-subfields.

(1) Hausdorff metric. Thanks to the insistence on the \( P \)-completeness of the \( \sigma \)-subfields, \( D \) is a metric on \( \mathfrak{H} \) [4] Theorem 1 & Corollary 1. By taking \( \vee \) instead of \( + \) in its definition (see [HC]), one obtains an equivalent metric \( \delta \leq 1 \wedge D \), which is the restriction to \( \mathfrak{H} \times \mathfrak{H} \) of the usual Hausdorff distance on closed subsets of \( \mathcal{F} \), associated to the pseudometric \( \mathcal{F} \times \mathcal{F} \ni (A, B) \to P(A \Delta B) \in [0, 1] \).

(2) Implications and non-implications between the various convergence types.

As already observed in the Introduction, [MC] \( \Rightarrow \) [ASC] (and, of course, [MC] \( \Rightarrow \) [STC]).

\([SC] \Rightarrow [WC] \) trivially, but not the other way around — not even when [WC] is to the largest \( \sigma \)-subfield \( \mathcal{B}_0 \) for which it holds, cf. [4], second bullet point, below — [9] Example 3.1.

\([HC] \Rightarrow [SC] \) [4] Theorem 4 and [STC] \( \Rightarrow \) [SC] [5], though neither conversely [2] Example 4.4; [ASC] \( \Rightarrow \) [SC] trivially, again the converse fails [15] Example 3.5.

Of [STC] [HC] [ASC] no one implies another: [STC] \( \Rightarrow \) [ASC] [1]; [ASC] \( \Rightarrow \) [STC] [2] Example 4.1; [STC] \( \Rightarrow \) [HC] [2] Example 4.2; [HC] \( \Rightarrow \) [STC] [2] Example 4.3; [HC] \( \Rightarrow \) [ASC] [4] penultimate paragraph; [MC] (hence [ASC] [STC] \( \Rightarrow \) [HC] [2] Example 4.2.

\([ONC_\rho^q] \Rightarrow \) [HC] \( \Rightarrow \) [ONC_\rho^q]

Proof. For \( \sigma \)-subfields \( \mathcal{A} \) and \( \mathcal{B} \), and for \( A \in \mathcal{A} \), \( \inf_{B \in \mathcal{B}} P(A \Delta B) = \inf_{B \in \mathcal{B}} P[1_A - 1_B] \leq P[1_A - 1_{\{P_B \geq 1/2\}}] \). Now if \( q < \infty \), we obtain \( \inf_{B \in \mathcal{B}} P(A \Delta B) \leq 2^q P[1_A - P_B 1_A]^q = 2^q P[1_A - P_B 1_A]^q = 2^q P[1_A - P_B 1_A] \), where we have used \( |1_A - 1_{\{P_B \geq 1/2\}}| \leq 2^q |1_A - P_B 1_A| \). When \( q = \infty \), we have simply \( \inf_{B \in \mathcal{B}} P(A \Delta B) \leq 2 P[1_A - P_B 1_A] \leq 2 P[1_A - P_B 1_A] \leq \|P_A - P_B\| L^\infty \to L^\infty \).

Finally, \([HC] \Rightarrow \) [ONC_\rho^q] assuming that \( p > q \).

Proof. Recall the metric \( \delta \) from [1] and \( \rho \) from [HC] We quote the following two results from the literature:

(a) Let \( a \in (0, \infty) \), \( r \in [1, \infty) \), \( H \subset L^r(P) \) and define \( \delta_{H,r}(a) := \sup\{ f 1_{\{|f| > a\}} \| L^r(P) : f \in H \} \). Then, for \( \sigma \)-subfields \( \mathcal{A} \) and \( \mathcal{B} \) satisfying \( \mathcal{A} \subset \mathcal{B} \), one has the inequality

\[
\sup\{ P_B f - P_A f \| L^r(P) : f \in H \} \leq C_r a[\delta(\mathcal{A}, \mathcal{B})(1 - \delta(\mathcal{A}, \mathcal{B}))]^{1/r} + 2 \delta_{H,r}(a),
\]

where \( C_r = 2 \cdot 2^{1/r} \) if \( r < 2 \) and \( C_r = 2 \) if \( r \geq 2 \), [17] Theorem 4, items (i) & (ii)

(b) \( \rho(\mathcal{A} \vee \mathcal{B}, \mathcal{B}) \leq 4 \rho(\mathcal{A}, \mathcal{B}) \) for \( \sigma \)-subfields \( \mathcal{A} \) and \( \mathcal{B} \). [10] Corollary 4]
Set \( H := \{ f \in L^p(P) : \|f\|_{L^p(P)} \leq 1 \} \). For any \( f \in H \), by the triangle inequality,
\[
\|P_{\mathcal{B}_n} f - P_{\mathcal{B}_0} f\|_{L^q(P)} \leq \|P_{\mathcal{B}_n} f - P_{\mathcal{B}_0} f\|_{L^q(P)} + \|P_{\mathcal{B}_n \vee \mathcal{B}_0} f - P_{\mathcal{B}_0} f\|_{L^q(P)}.
\]

Then, by (a) above, for any \( a \in (0, \infty) \),
\[
\sup\{\|P_{\mathcal{B}_n} f - P_{\mathcal{B}_0} f\|_{L^q(P)} : f \in H\} \\
\leq 2 \cdot 2^{1/q} \{\delta(\mathcal{B}_n, \mathcal{B}_0 \vee \mathcal{B}_n)(1 - \delta(\mathcal{B}_n, \mathcal{B}_0 \vee \mathcal{B}_n))\}^{1/q} + 2\delta_{H,q}(a) + \\
2 \cdot 2^{1/q} \{\delta(\mathcal{B}_n \vee \mathcal{B}_0, \mathcal{B}_0)(1 - \delta(\mathcal{B}_n \vee \mathcal{B}_0, \mathcal{B}_0))\}^{1/q} + 2\delta_{H,q}(a)
\]
(where the notation \( \delta_{H,q}(a) \) is that of (a) above, which by (b) is (note that for \( \sigma \)-subfields \( \mathcal{A} \supset \mathcal{B} \), \( \delta(\mathcal{A}, \mathcal{B}) = \delta(\mathcal{B}, \mathcal{A}) = \rho(\mathcal{A}, \mathcal{B}) \))
\[
\leq 4 \cdot 8^{1/q} \delta(\mathcal{B}_n, \mathcal{B}_0)^{1/q} + 4\delta_{H,q}(a).
\]

Since \( \delta_{H,q}(a) \leq a^{-(\frac{p}{q} - 1)} \) when \( p < \infty \) and \( \delta_{H,q}(a) \leq 1_{[0,1]}(a) \) when \( p = \infty \), it follows that
\[
\limsup_n \sup\{\|P_{\mathcal{B}_n} f - P_{\mathcal{B}_0} f\|_{L^q(P)} : f \in H\} \leq 4\delta_{H,q}(a) \to 0 \text{ as } a \to \infty.
\]

Hence
\[
\lim_{n \to \infty} \sup\{\|P_{\mathcal{B}_n} f - P_{\mathcal{B}_0} f\|_{L^q(P)} : f \in H\} = 0,
\]
which is the desired operator norm convergence. \( \square \)

(The last two implications, in the case \( q = 1, \ p = \infty \) can be found e.g. in [1, Theorem VIII.2.21].)

Up to trivial corollaries, this exhausts the mutual implications and non-implications of the convergence types \( \{\text{ONC}_p\} \) for \( p > q \), \( \{\text{MC}\}, \{\text{HC}\}, \{\text{STC}\}, \{\text{ASC}\}, \{\text{SC}\}, \{\text{WC}\} \)
(Figure 1).

(3) Uniqueness of limits. Excepting \( \{\text{WC}\} \) the limits are unique. Indeed, the results of e.g. [9] imply uniqueness of the limit in the case of \( \{\text{SC}\} \).

(4) Weak convergence.
- For any \( p \in [1, \infty) \), \( \mathcal{B}_n \to \mathcal{B}_0 \) weakly iff \( P_{\mathcal{B}_n} f \to f \) in \( L^p(P) \) for every \( f \in L^p(P|_{\mathcal{B}_0}) \) (i.e. for every \( f \in L^p(P) \) that is \( \mathcal{B}_0 \)-measurable).

Proof. By linearity, for sure \( P_{\mathcal{B}_n} f \to P_{\mathcal{B}_0} f \) in \( P \)-probability for any bounded simple \( \mathcal{B}_0 \)-measurable \( f \). Now let \( f \in L^p(P|_{\mathcal{B}_0}); \delta > 0 \). The simple functions being dense in \( L^p(P) \),

\[
\begin{align*}
\]
there exists a simple $\mathcal{B}_0$-measurable $f^\delta$ for which $P|f - f^\delta|^p < \delta$. Then it follows from the decomposition

$$P_{B_n} f - P_{B_0} f = P_{B_n} (f - f^\delta) + P_{B_n} f^\delta - P_{B_0} f^\delta + P_{B_0} (f^\delta - f),$$

from the elementary estimate $|x + y|^p \leq 2^{p-1}(|x|^p + |y|^p)$ for $\{x, y\} \subseteq \mathbb{R}$, from conditional Jensen’s inequality, finally from the fact that boundedness implies uniform integrability (hence coupled with convergence in $P$-probability, $L^1(P)$ convergence), that

$$\limsup_{n \to \infty} P|P_{B_n} f - P_{B_0} f|^p \leq C \delta,$$

for some constant $C \in (0, \infty)$ depending only on $p$. Let $\delta \downarrow 0$. □

- Then, according to [2] Lemmas 1.1 and 1.3, $B_n \to B_0$ weakly as $n \to \infty$ iff $B_0 \subset B_p$, where

$$B_p := \{A \in \mathcal{F} : \lim_{n \to \infty} \inf_{B \in B_n} P(A \Delta B) = 0\}.$$ 

The $\sigma$-subfield $B_p$ coincides with the $P$-liminf$_{n \to \infty} B_n$ of [3], see [3] Theorem 3.2.

- The join (sup) operation $\vee$ is sequentially continuous under weak convergence [11 Proposition 2.3] (but the meet (inf) $\wedge$ is not [11 Proposition 2.1]). It means that for sequences $(A_n)_{n \in \mathbb{N}}$ and $(B_n)_{n \in \mathbb{N}}$ in $\mathcal{F}$, if $A_n \to A_0$ and $B_n \to B_0$ weakly, then also $A_n \vee B_n \to A_0 \vee B_0$ weakly (but in general this fails if $\wedge$ replaces $\vee$).

5 Strong convergence

With an analogous justification to the one in [11], for any $p \in [1, \infty)$, $B_n \to B_0$ strongly iff $P_{B_n} f \to P_{B_0} f$ in $L^p(P)$ for every $f \in L^p(P)$. Thus strong convergence is nothing but the strong operator convergence of the conditional expectation operators in the spaces $L^p(P)$, $p \in [1, \infty)$ (but not in $L^\infty(P)$; the latter fails even for monotone increasing sequences, see [3] final paragraph).

6 Operator convergence $L^p \to L^p$. Convergence in the operator norm $\| \cdot \|_{L^p \to L^p}$, $p \in [1, \infty]$, appears elusive.

- For one, $\textbf{[ONC]}_p$ of not-ultimately-constant sequences of $\sigma$-subfields fails always when $p = q = 1$ or when $p = q = \infty$. Indeed, if $A$ and $B$ are two distinct $\sigma$-subfields, then $\|P_A - P_B\|_{L^1 \to L^1} \geq 1$ and $\|P_A - P_B\|_{L^\infty \to L^\infty} \geq 1/2$. To see this, let $B \in \mathcal{B} \setminus A$. Set first $f = 1_B - P_A 1_B$, so that $0 \neq f \in L^1(P)$. Then P-a.s., $P_A f = 0$, while $f$ (hence $P_B f$) is $\geq 0$ on $B$ and $\leq 0$ on $\Omega \setminus B$. It follows that $P|P_B f - P_A f| = P|P_B f| = P|(P_B f) 1_B - (P_B 1_B) f 1_B| = P f 1_B - P f 1_B = P|f|$, viz. $\|P_A - P_B\|_{L^1 \to L^1} \geq 1$. Set now $f = 1_B$, noting that $\|f\|_{L^\infty(P)} = 1$. Then $P(B \Delta \{P_A 1_B > 1/2\}) > 0$ so that either $P(B \setminus \{P_A 1_B > 1/2\}) > 0$ or $P(\{P_A 1_B > 1/2\} \setminus B) > 0$, which coupled with the P-a.s. equality $P_B f = 1_B$, yields $\|P_A - P_B\|_{L^\infty \to L^\infty} \geq 1/2$.

- If for infinitely many $n \in \mathbb{N}$, $B_n \subsetneq B_0$ or $B_0 \subsetneq B_n$ — or if $B_n \subsetneq B_m$ for arbitrarily large $n \in \mathbb{N}$ and $m \in \mathbb{N}$ — then again $P_{B_n}$ does not converge to $P_{B_0}$ in the $\| \cdot \|_{L^p \to L^p}$ norm. For if $A$ is a $\sigma$-subfield that is strictly contained in the $\sigma$-subfield $B$, one can take $B \in \mathcal{B} \setminus A$, set $f = 1_B - P_A 1_B$, which is then not P-a.s. equal to 0, and finds that P-a.s. $P_B f - P_A f = f$. In particular, one sees that $\textbf{[ONC]}_p$ simply precludes $\textbf{[MC]}_p$ of not-ultimately-constant sequences of $\sigma$-subfields altogether.
• If, for infinitely many \( n \in \mathbb{N} \), \( \mathcal{B}_n \) contains a non-\( \mathbb{P} \)-trivial event independent of \( \mathcal{B}_0 \) or vice versa — or if this obtains with \( \mathcal{B}_m \) in place of \( \mathcal{B}_0 \) for arbitrarily large \( m \) and \( n \) — then \( \mathbb{P}_{\mathcal{B}_n} \) does not converge to \( \mathbb{P}_{\mathcal{B}_0} \) in the \( \| \cdot \|_{L^p \to L^p} \) norm. For if \( A \) belongs to a \( \sigma \)-subfield \( \mathcal{A} \), is independent of a \( \sigma \)-subfield \( \mathcal{B} \), and has \( \mathbb{P}(A) \in (0,1) \), then taking \( f = \mathbb{P}(A)^{-1} \mathbbm{1}_A - (1 - \mathbb{P}(A))^{-1} \mathbbm{1}_{\Omega \setminus A} \), one has \( \mathbb{P}(A) \text{-a.s.} \) \( \mathbb{P}_{\mathcal{A}} f - \mathbb{P}_{\mathcal{B}} f = f - \mathbb{P} f = f \), the latter not being \( \mathbb{P} \)-a.s. equal to 0.

• Still, for \( p \in (1, \infty) \), this convergence type is not vacuous with respect to not-ultimately-constant sequences of \( \sigma \)-subfields:

**Example 2.1.** \(^{1}\) Fix \( p \in (1, \infty) \). Let for \( r \in [0,1] \), \( \text{Ber}(r) \) denote the Bernoulli law (on the space \( \{0,1\} \)) with success parameter \( r \): \( \text{Ber}(r)(\{0\}) = 1 - r \) and \( \text{Ber}(r)(\{1\}) = r \). Let \( \Omega = \{0,1\}^{\mathbb{N}_0} \); let \( X_i \), \( i \in \mathbb{N}_0 \), be the canonical projections; \( \mathcal{F} \) the \( \sigma \)-field generated by them; \( \mathbb{P} = \text{Ber}(1/2) \times \times_{n \in \mathbb{N}} \text{Ber}(1/2^n) \). For \( n \in \mathbb{N}_0 \), set \( Y_n := X_0 \vee X_n \) and then let \( \mathcal{B}_n = \sigma(Y_n) \), the \( \sigma \)-complete \( \sigma \)-subfield generated by \( Y_n \). Note that by the tower property of conditional expectations and the independence of the \( (X_m)_{m \in \mathbb{N}_0} \), for \( n \in \mathbb{N} \), \( \sup\{\mathbb{P}|\mathcal{F}_{\mathcal{B}_n} f - \mathbb{P}_{\mathcal{B}_0} f|^p : f \in L^p(\mathbb{P}), \|f\|_{L^p(\mathbb{P})} \leq 1\} = \sup\{\mathbb{Q}^{1/2^n}(|Q_{\mathcal{A}_2}^c f - Q_{\mathcal{A}_1}^c| f|^p : f \in L^p(\mathbb{Q}^{1/2^n}), \|f\|_{L^p(\mathbb{Q}^{1/2^n})} \leq 1\}, \) where \( Q^c := \text{Ber}(1/2) \times \text{Ber}(\epsilon) \) for \( \epsilon \in [0,1] \), and where, with \( Z_1 : \{0,1\} \times \{0,1\} \rightarrow \{0,1\} \) and \( Z_2 : \{0,1\} \times \{0,1\} \rightarrow \{0,1\} \) being the projections onto the first and second coordinate respectively, \( \mathcal{A}_1 := \sigma(Z_1) \) and \( \mathcal{A}_2 := \sigma(Z_1 \vee Z_2) \). Now fix \( \epsilon \in [0,1] \). For \( f \in L^1(\mathbb{Q}^c) \), we have

\[
Q_{\mathcal{A}_2}^c f = 1_{\{(0,0)\}} f(0,0) + 1_{\{(0,1),(1,0),(1,1)\}} \frac{f(0,1) \epsilon + f(1,0)(1 - \epsilon) + f(1,1) \epsilon}{(1 + \epsilon)},
\]
whilst

\[
Q_{\mathcal{A}_1}^c f = 1_{\{(0,0),(0,1)\}} \left[ f(0,0)(1 - \epsilon) + f(0,1) \epsilon \right] + 1_{\{(1,0),(1,1)\}} \left[ f(1,0)(1 - \epsilon) + f(1,1) \epsilon \right].
\]

In view of the equality \( Q^c Q_{\mathcal{A}_2}^c f - Q_{\mathcal{A}_1}^c f|^p = \sum_{\omega \in \{0,1\} \times \{0,1\}} Q^c (Q_{\mathcal{A}_2}^c f - Q_{\mathcal{A}_1}^c f|^p 1_{\{\omega\}}) \), and using the elementary estimate \( (x + y)^r \leq 2^{-r-1}(x^r + y^r) \) for \( \{x, y\} \subset [0, \infty) \) and \( r \in [1, \infty) \), it is now straightforward to see that \( Q^c [Q_{\mathcal{A}_2}^c f - Q_{\mathcal{A}_1}^c f|^p] \leq C_p \epsilon^{(p-1)\Lambda} \text{Lip}[f]^p \) for some \( C_p \in (0, \infty) \) depending only on \( p \). It follows that \( \mathbb{P}_{\mathcal{B}_n} \to \mathbb{P}_{\mathcal{B}_0} \) in the \( \| \cdot \|_{L^p \to L^p} \) operator norm.

\[
\Box
\]

3. Preservation of independence in the limit

**Proposition 3.1.** Let \( (\mathcal{C}_n)_{n \in \mathbb{N}_0} \) be a sequence in \( \mathfrak{S} \) with \( \mathcal{C}_n \to \mathcal{C}_0 \) strongly as \( n \to \infty \); \( \mathcal{I} \) an arbitrary index set; finally \( (\mathcal{B}_n^i)_{(n,i) \in \mathbb{N}_0 \times \mathcal{I}} \) a collection of \( \sigma \)-subfields of \( \mathcal{F} \) with

\( \mathcal{B}_n^i \to \mathcal{B}_0^i \) weakly as \( n \to \infty \) for each \( i \in \mathcal{I} \)

and with

the family \( (\mathcal{B}_n^i)_{i \in \mathcal{I}} \) conditionally independent given \( \mathcal{C}_n \) for each \( n \in \mathbb{N} \).

Then the family \( (\mathcal{B}_0^i)_{i \in \mathcal{I}} \) is conditionally independent given \( \mathcal{C}_0 \).

**Remark 3.2.** One says conditional independence is preserved in the limit, under a convergence type (iC), when the statement of Proposition 3.1 prevails for any choice of the \( \mathcal{C}_n \)

---

\(^{1}\) Due to J. Warren, private communication.
and the $B_i$'s, and with the words “strongly” and “weakly” replaced by “for the convergence type (iC)” therein. This specializes to (unconditional) independence when $C_n$ is $\mathbb{P}$-trivial for every $n \in \mathbb{N}$. By (2) it follows that (unconditional) independence is preserved in the limit under any of the convergence types $\text{(ONC)}^\circ, \text{(MC)}, \text{(ASC)}, \text{(STC)}, \text{(HC)}, \text{(SC)}, \text{(WC)}$. For conditional independence, we must except $\text{(WC)}$; cf. Example 3.3 below.

**Proof.** We may assume $\mathcal{I}$ is finite; then, thanks to (4), last bullet point, via mathematical induction and properties of conditional independence, we reduce to the case $\mathcal{I} = \{0, 1\}$; finally, on account of (4), second bullet point, there is no loss of generality in taking $B_0^i = B_0^i$ for each $i \in \mathcal{I}$. Now take $A_0^0 \in B_0^0$ and $A_0^1 \in B_0^1$ arbitrary. Then for each $i \in \{0, 1\}$ we find a sequence $(A_n^i)_{n \in \mathbb{N}}$ with $A_n^i \in B_n^i$ for each $n \in \mathbb{N}$ and with $\lim_{n \to \infty} \mathbb{P}(A_n^1 \Delta A_n^0) = 0$. Now, $P_{C_n}(A_0^0 \cap A_1^1) \to P_{C_n}(A_0^0 \cap A_1^0)$, $P_{C_n}(A_0^0) \to P_{C_0}(A_0^0)$ and $P_{C_n}(A_1^0) \to P_{C_0}(A_1^0)$ in $\mathbb{P}$-probability. Also, for each $n \in \mathbb{N}$,

- thanks to independence, $\mathbb{P}$-a.s. $P_{C_n}(A_n^0 \cap A_n^1) = P_{C_n}(A_0^0)P_{C_n}(A_1^1)$,
- whilst using the elementary equality $|1_C - 1_D| = 1_{C \Delta D}$ for sets $C$ and $D$, we find that

$$P|P_{C_n}(A_n^0 \cap A_n^1) - P_{C_n}(A_n^0)\cap P_{C_n}(A_n^1)| \leq P(A_0^1 \Delta A_n^0) + P(A_0^0 \Delta A_n^1)$$

and likewise

$$P|P_{C_n}(A_n^0)P_{C_n}(A_n^1) - P_{C_n}(A_0^0)\cap P_{C_n}(A_1^1)| \leq P(A_0^1 \Delta A_n^0) + P(A_0^0 \Delta A_n^1).$$

In particular, $P_{C_n}(A_n^0)P_{C_n}(A_n^1) - P_{C_n}(A_0^0)\cap P_{C_n}(A_1^1) \to 0$ and $P_{C_n}(A_0^0 \cap A_0^1) - P_{C_n}(A_0^0)\cap P_{C_n}(A_0^1) \to 0$ in $\mathbb{P}$-probability. Since convergence in probability is preserved under addition and multiplication and since the limit in probability is a.s. unique, letting $n \to \infty$, yields that $\mathbb{P}$-a.s. $P_{C}(A_0^0 \cap A_1^1) = P_{C}(A_0^0)\cap P_{C}(A_1^1)$, as required. \qed

**Example 3.3.** We show that conditional independence is generally not preserved under $\text{(WC)}$ to the $\mathbb{P}$-$\lim \inf$ (see (4), second bullet point, for the notation). — Without the latter insistence, a counterexample is trivial: take any $A \in \mathcal{F}$ with $\mathbb{P}(A) \in (0, 1)$, for $n \in \mathbb{N}$ let $B_n = H := \sigma(A)^\mathbb{P}$ be the $\mathbb{P}$-complete $\sigma$-field generated by $A$. Then $B_n$ is conditionally independent of $B_n$ given $B_n$ for each $n$, but the strong (indeed, in every sense) limit $H$ is not conditionally independent of itself given the trivial $\sigma$-subfield, to which the $B_n$ converge weakly.) — Take $\Omega = [0, 1]$, $\mathcal{F} = \mathcal{B}(\Omega)$ the Borel $\sigma$-field, $\mathbb{P} = \text{Lebesgue measure.}$ For $n \in \mathbb{N}$ set

$$B_n := \bigcup_{k=0}^{2^n - 2 - 1} \left[ \frac{2k}{2^n} , \frac{2k + 1}{2^n} \right),$$

and let $B_n = \overline{\sigma(B_n)}^\mathbb{P}$ be the $\mathbb{P}$-complete $\sigma$-field generated by $B_n$. Finally let $\mathcal{B}_0 = \{\emptyset, \Omega\}^\mathbb{P}$ be the trivial $\sigma$-subfield. Then $B_n$ converges weakly to $\mathcal{B}_0 = \mathcal{B}_p = \mathbb{P}$-$\lim \inf_{n \to \infty} B_n$, for
but not strongly [15, Example 3.4]. Finally, denote \( x = (\sqrt{6} - 1)/8 \); notice that \( 1/2 < 1 - 2x + 1/16 < 1 - x < 1 - x + 1/16 < 1 \); and consider the events
\[
A := \left[ \frac{3}{16}, \frac{7}{16} \right] \cup [1 - x, 1) \quad \text{and} \quad B := \left[ \frac{1}{16}, \frac{5}{16} \right) \cup \left( 1 - 2x + \frac{1}{16}, 1 - x + \frac{1}{16} \right).
\]
so that
\[
A \cap B = \left[ \frac{3}{16}, \frac{5}{16} \right) \cup \left( 1 - x, 1 - x + \frac{1}{16} \right).
\]
A simple calculation yields that \( \mathbb{P}\)-a.s. for each \( n \in \mathbb{N}_{\geq 4} \),
\[
\mathbb{P}(A|\mathcal{B}_n) = \mathbb{P}(B|\mathcal{B}_n) = \frac{1/8}{1/4} \mathbbm{1}_{B_n} + \frac{1/8 + x}{3/4} \mathbbm{1}_{\Omega \setminus B_n}
\]
whilst
\[
\mathbb{P}(A \cap B|\mathcal{B}_n) = \frac{1/16}{1/4} \mathbbm{1}_{B_n} + \frac{1/16 + 1/16}{3/4} \mathbbm{1}_{\Omega \setminus B_n}.
\]
This renders the \( \mathbb{P}\)-a.s. equality \( \mathbb{P}(\mathcal{B}_n(A)) \mathbb{P}(\mathcal{B}_n(B)) = \mathbb{P}(\mathcal{B}_n(A \cap B)) \). Thus \( A \) and \( B \), equivalently the respective \( \mathbb{P}\)-complete \( \sigma \)-fields generated by them, are conditionally independent given \( \mathcal{B}_n \) for each \( n \in \mathbb{N}_{\geq 4} \). But \( A \) and \( B \) are not conditionally independent given the weak limit \( \mathcal{B}_0 = \mathcal{B}_\mathbb{P} \) of the \( \mathcal{B}_n \), for they are not independent: \( (1/4 + x)^2 \neq 1/8 + 1/16 \), as is readily verified.  

4. Invariance under passage to equivalent probability measure

In what follows, for a convergence mode (iC), by saying that it is invariant under passage to an equivalent probability measure, we mean that, whenever \( (\mathcal{B}_n)_{n \in \mathbb{N}_0} \subset \mathfrak{F} \),
\[
\mathcal{B}_n \to \mathcal{B}_0 \quad \text{in the sense of (iC) under} \quad \mathbb{P} \iff \mathcal{B}_n \to \mathcal{B}_0 \quad \text{in the sense of (iC) under} \quad \mathbb{Q},
\]
provided \( \mathbb{Q} \approx \mathbb{P} \). (Note that, given \( (\Omega, \mathcal{F}), \mathfrak{F} \) depends on \( \mathbb{P} \) only through \( \mathcal{N} = \mathbb{P}^{-1}(\{0\}) \).)

Recall that for finite measures \( \mu \) and \( \nu \) on \( (\Omega, \mathcal{F}) \), \( \mu \ll \nu \) is equivalent to
\[
\forall \epsilon \in (0, \infty) \exists \delta \in (0, \infty) \forall A \in \mathcal{F} : \nu(A) < \delta \Rightarrow \mu(A) < \epsilon.
\]

**Proposition 4.1.** \([\mathrm{HC}]\) is invariant under passage to an equivalent probability measure. Indeed the distance \( D \), up to equivalence, depends on \( \mathbb{P} \) only up to equivalence.

**Remark 4.2.** By \([\mathrm{2}]\), \([\mathrm{ONC}_p^q]\) \( q < p \), is equivalent to \([\mathrm{HC}]\), so that \([\mathrm{ONC}_p^q]\) \( q < p \), too is invariant under passage to an equivalent probability measure. The case \( p = q \) remains open.

**Proof.** Let \( \mathbb{Q} \approx \mathbb{P} \), denote by \( D_\mathbb{P} \) and \( D_\mathbb{Q} \) the metrics associated to \( \mathbb{P} \) and \( \mathbb{Q} \), respectively. Let \( \epsilon \in (0, \infty) \), \( A \in \mathfrak{F} \). Since \( \mathbb{Q} \ll \mathbb{P} \), there is a \( \delta \in (0, \infty) \) such that for all \( A \in \mathcal{F} \),
\[
\mathbb{P}(A) < \delta \Rightarrow \mathbb{Q}(A) < \epsilon.
\]
Then for all \( B \in \mathfrak{F} \), \( D_\mathbb{P}(A, B) < \delta \) implies \( D_\mathbb{Q}(A, B) \leq 2\epsilon \). \( \square \)

**Proposition 4.3.** \([\mathrm{WC}]\) is invariant under passage to an equivalent probability measure.

**Proof.** By \([\mathrm{1}]\), second bullet point, it is enough to verify that \( \mathcal{B}_\mathbb{P} = \mathcal{B}_\mathbb{Q} \) whenever \( \mathbb{P} \sim \mathbb{Q} \). But assuming \( \mathbb{Q} \ll \mathbb{P} \), thanks to the equivalent condition for absolute continuity noted above, if for some \( A \in \mathcal{F} \) there exist \( A_n \in \mathcal{B}_n \) for \( n \in \mathbb{N} \) with \( \lim_{n \to \infty} \mathbb{P}(A_n \Delta A) = 0 \), then also \( \lim_{n \to \infty} \mathbb{Q}(A_n \Delta A) = 0 \). \( \square \)
Recall now the statement of abstract Bayes’ theorem. Letting $Q$ be another probability measure on $\mathcal{F}$, equivalent to $P$:

For any $\mathcal{F}/\mathcal{B}([-\infty, \infty])$-measurable $X$ and any sub-$\sigma$-field $\mathcal{G}$ of $\mathcal{F}$, $Q_\mathcal{G}(X)P_\mathcal{G}(\frac{dQ}{dP}) = P_\mathcal{G}(\frac{dQ}{dP}X)$ a.s., in the sense that the left hand-side is well-defined iff the right hand-side is so, whence they are equal.

**Proposition 4.4.** ([SC]) is invariant under passage to an equivalent probability measure.

**Proof.** Assume $\mathcal{B}_n \to \mathcal{B}_0$ strongly under $P$. Let $Q \sim P$. Note $dQ/dP$ and then all the $P_{\mathcal{B}_n}(\frac{dQ}{dP})$, $n \in \mathbb{N}_0$, may be chosen from their equivalence classes to be strictly positive everywhere. Let $A \in \mathcal{F}$. Then by Bayes’ rule a.s.

$$Q_{\mathcal{B}_n}1_A = P_{\mathcal{B}_n}\left(\frac{dQ}{dP}1_A\right) / P_{\mathcal{B}_n}\left(\frac{dQ}{dP}\right).$$

By (5) the numerator and denominator both converge as $n \to \infty$ in $L^1(P)$, hence in $P$- (equivalently, $Q$-) probability, to the respective expressions in which $\mathcal{B}_0$ replaces $\mathcal{B}_n$. Convergence in probability is preserved under taking quotients (assuming the denominators are non-zero; e.g. from the characterization through the a.s. convergence of subsequences) and the claim follows by another application of Bayes’ rule. □

**Remark 4.5.** According to [2, Proposition 3.3, Lemma 1.3] ([SC]) is equivalent to the conjunction of ([WC]) and

(\perp C) Orthogonal convergence. $\mathcal{B}_n$ converges to $\mathcal{B}_0$ orthogonally if $1_{A_n} - P_{\mathcal{B}_n}1_{A_n} \to 0$ weakly in $L^2(P)$, whenever $A_n \in \mathcal{B}_n$ for each $n \in \mathbb{N}$.

It remains open whether (\perp C) too is invariant under passage to an equivalent probability measure.

**Proposition 4.6.** ([ASC]) is invariant under passage to an equivalent probability measure.

**Proof.** Assume $\mathcal{B}_n \to \mathcal{B}_0$ in the almost-sure sense under $P$. Let $Q \sim P$. Then by Bayes’ rule, for any $f \in L^1(Q)$, a.s.

$$Q_{\mathcal{B}_n}f = P_{\mathcal{B}_n}\left(\frac{dQ}{dP}f\right) / P_{\mathcal{B}_n}\left(\frac{dQ}{dP}\right).$$

Since $\frac{dQ}{dP}f$ and $\frac{dQ}{dP}$ both belong to $L^1(P)$, by the very definition of ([ASC]) the numerator and denominator both converge $P$- (equivalently $Q$-) a.s. as $n \to \infty$ to the respective expressions in which $\mathcal{B}_0$ replaces $\mathcal{B}_n$. Another application of Bayes’ rule concludes the argument. □

**Question 4.7.** Given this invariance of the various convergence modes, can something akin to the characterization of convergence in probability through the a.s. convergence of subsequences, be offered? I.e. can the convergence modes be characterized in such a way as to make manifest the invariance under passage to an equivalent probability measure?
References

1. A. Alonso, A Countereexample on the Continuity of Conditional Expectations, Journal of Mathematical Analysis and Applications 129 (1988), no. 1, 1 – 5.
2. A. Alonso and F. Brambila-Paz, $L^p$-Continuity of Conditional Expectations, Journal of Mathematical Analysis and Applications 221 (1998), no. 1, 161 – 176.
3. Z. Artstein, Compact Convergence of $\sigma$-fields and Relaxed Conditional Expectation, Probability Theory and Related Fields 120 (2001), no. 3, 369–394.
4. E. S. Boylan, Equiconvergence of Martingales, The Annals of Mathematical Statistics 42 (1971), no. 2, 552–559.
5. H. Fetter, On the Continuity of Conditional Expectations, Journal of Mathematical Analysis and Applications 61 (1977), no. 1, 227 – 231.
6. S. Hu and N. S. Papageorgiou, Handbook of Multivalued Analysis: Volume II: Applications, Mathematics and Its Applications, Springer US, 2013.
7. O. Kallenberg, Foundations of Modern Probability, Probability and Its Applications, Springer, New York Berlin Heidelberg, 1997.
8. H. Kudo, On an Approximation to a Sufficient Statistic Including a Concept of Asymptotic Sufficiency, Journal of the Faculty of Science, University of Tokyo (1970), no. 17, 273–290.
9. ______, A Note on the Strong Convergence of $\Sigma$-Algebras, The Annals of Probability 2 (1974), no. 1, 76–83.
10. D. Landers and L. Rogge, An Inequality for the Hausdorff-Metric of $\sigma$-Fields, The Annals of Probability 14 (1986), no. 2, 724–730.
11. V. Mackevicius, J. Memin, and F. Coquet, Some Examples and Countereexamples of Convergence of $\sigma$-algebras and Filtrations, Lithuanian Mathematical Journal 40 (2000), no. 3.
12. H. G. Mukerjee, Almost Sure Equiconvergence of Conditional Expectations, The Annals of Probability 12 (1984), no. 3, 733–741.
13. J. Neveu, Mathematical Foundations of the Calculus of Probability, Holden-Day Series in Probability and Statistics, Holden-Day, 1965.
14. ______, Note on the Tightness of the Metric on the Set of Complete Sub $\sigma$-Algebras of a Probability Space, The Annals of Mathematical Statistics 43 (1972), no. 4, 1369–1371.
15. L. Piccinini, Convergence of Nonmonotone Sequences of Sub-$\sigma$-fields and Convergence of Associated Subspaces $L^p(B_n) (p \in [1, +\infty])$, Journal of Mathematical Analysis and Applications 225 (1998), no. 1, 73–90.
16. M. M. Rao, Foundations of stochastic analysis, Dover books on mathematics, Dover Publications, 2011.
17. L. Rogge, Uniform Inequalities for Conditional Expectations, The Annals of Probability 2 (1974), no. 3, 486–489.
18. M. Stinchcombe, A Further Note on Bayesian Information Topologies, Journal of Mathematical Economics 22 (1993), no. 2, 189–193.
19. B. Tsirelson, Noise as a Boolean algebra of $\sigma$-fields, The Annals of Probability 42 (2014), no. 1, 311–353.
20. T. V. Zandt, The Hausdorff Metric of $\sigma$-Fields and the Value of Information, The Annals of Probability 21 (1993), no. 1, 161–167.

Department of Mathematics, University of Ljubljana and Institute of Mathematics, Physics and Mechanics, Slovenia
E-mail address: matija.vidmar@fmf.uni-lj.si