# ATOMIC INTERSECTION OF $\sigma$-FIELDS AND SOME OF ITS CONSEQUENCES

PATRIZIA BERTI, LUCA PRATELLI, AND PIETRO RIGO

Abstract. Let $(\Omega, \mathcal{F}, P)$ be a probability space. For each $\mathcal{G} \subset \mathcal{F}$, define $\overline{\mathcal{G}}$ as the $\sigma$-field generated by $\mathcal{G}$ and those sets $F \in \mathcal{F}$ satisfying $P(F) \in \{0, 1\}$. Conditions for $P$ to be atomic on $\cap_{i=1}^k \overline{\mathcal{A}_i}$, with $\mathcal{A}_1, \ldots, \mathcal{A}_k \subset \mathcal{F}$ sub-$\sigma$-fields, are given. Conditions for $P$ to be 0-1-valued on $\cap_{i=1}^k \overline{\mathcal{A}_i}$ are given as well. These conditions are useful in various fields, including Gibbs sampling, iterated conditional expectations and the intersection property.

## 1. Introduction

Throughout, $(\Omega, \mathcal{F}, P)$ is a probability space, $\mathcal{A}_1, \ldots, \mathcal{A}_k \subset \mathcal{F}$ sub-$\sigma$-fields, $k \geq 2$, and we let

$$\mathcal{N} = \{F \in \mathcal{F} : P(F) \in \{0, 1\}\} \quad \text{and} \quad \overline{\mathcal{G}} = \sigma\big(\mathcal{G} \cup \mathcal{N}\big) \text{ for any } \mathcal{G} \subset \mathcal{F}.$$

As discussed in Section 3, the sub-$\sigma$-field

$$\mathcal{D} = \cap_{i=1}^k \overline{\mathcal{A}_i}$$

plays a role in various subjects, including Gibbs sampling, iterated conditional expectations and the intersection property. In a previous paper, in a Gibbs sampling framework, we investigated when $\cap_{i=1}^k \overline{\mathcal{A}_i} = \overline{\cap_{i=1}^k \mathcal{A}_i}$; see [3].

In this paper, instead, we focus on atomicity of $P$ on $\mathcal{D}$. In fact, atomicity of $P|\mathcal{D}$ (i.e., the restriction of $P$ to $\mathcal{D}$) has implications in each of the subjects mentioned above. It turns out that $P|\mathcal{D}$ is actually atomic under mild conditions.

An extreme form of atomicity for $P|\mathcal{D}$ is $\mathcal{D} = \mathcal{N}$, that is, $P$ 0-1-valued on $\mathcal{D}$. Indeed, $\mathcal{D} = \mathcal{N}$ is fundamental for Gibbs sampling and very useful for the intersection property; see [3], [7] and [10].

Our main results are in Sections 4 and 5. Section 4 gives general results on atomicity of $P|\mathcal{D}$. It includes a characterization (Theorem 2), a criterion for identifying the atoms (Theorem 3) and a sufficient condition (Theorem 4). Section 5, motivated by Gibbs sampling applications, concerns the particular case

$$\mathcal{A}_i = \sigma(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_k)$$

where $X_1, \ldots, X_k$ are any random variables on $(\Omega, \mathcal{F}, P)$. It contains working sufficient conditions for $\mathcal{D} = \mathcal{N}$ (Theorem 8) and for $P|\mathcal{D}$ to be atomic (Theorem 10). Indeed, $P|\mathcal{D}$ is atomic whenever the probability distribution of $(X_1, \ldots, X_k)$ is absolutely continuous with respect to a $\sigma$-finite product measure.

---

2000 *Mathematics Subject Classification.* 60A10, 60J22, 62B99.

*Key words and phrases.* Atomic probability measure – Gibbs sampling – Graphical models – Intersection property – Iterated conditional expectations.

Finally, it is worth noting that $\mathcal{D} = \mathcal{N}$ whenever $\mathcal{A}_r$ is independent of $\mathcal{A}_s$ for some $r, s$. Given $D \in \mathcal{D}$, in fact, one has $P(A_i \Delta D) = 0$ for some $A_i \in \mathcal{A}_i$, $i = 1, \ldots, k$. Hence, $P(D) = P(A_r \cap A_s) = P(A_r)P(A_s) = P(D)^2$.

## 2. Preliminaries

Let $(\mathcal{X}, \mathcal{E}, Q)$ be a probability space. A $Q$-*atom* is any set $H \in \mathcal{E}$ such that $Q(H) > 0$ and $Q(\cdot \mid H)$ is 0-1-valued. In general, there are three possible situations: (i) $Q$ is *nonatomic*, i.e., there are no $Q$-atoms; (ii) $Q$ is *atomic*, i.e., the $Q$-atoms form a (countable) partition of $\mathcal{X}$; (iii) there is $K \in \mathcal{E}$, $0 < Q(K) < 1$, such that $Q(\cdot \mid K)$ is nonatomic and $K^c$ is a (countable) disjoint union of $Q$-atoms.

Thus, $D \subset \Omega$ is an atom of $P|\mathcal{D}$ if and only if $D \in \mathcal{D}$, $P(D) > 0$ and $P(\cdot \mid D)$ is 0-1-valued on $\mathcal{D}$. In the sequel, when $P|\mathcal{D}$ is atomic, we also say that $\mathcal{D}$ *is atomic under* $P$.

For later purposes, we also note that $Q$ is nonatomic if and only if $(\mathcal{X}, \mathcal{E}, Q)$ supports a real random variable with uniform distribution on $(0, 1)$. In fact, if $U$ is a uniform random variable on $(\mathcal{X}, \mathcal{E}, Q)$, then $Q$ is nonatomic since $\sigma(U) \subset \mathcal{E}$ and $Q|\sigma(U)$ is nonatomic. Conversely, by Lyapunov's convexity theorem, if $Q$ is nonatomic the range of $Q$ is $[0, 1]$; see e.g. [8] or Theorem 5.1.6 of [4] for a proof (based on transfinite induction or Zorn's lemma, respectively). Since the range of $Q$ is $[0, 1]$, a uniform random variable on $(\mathcal{X}, \mathcal{E}, Q)$ can be obtained by arguing as in the proof of Lemma 2 of [1]; see also Theorem 3.1 of [2].

We finally recall that, for any sub-$\sigma$-field $\mathcal{G} \subset \mathcal{F}$,

$$\overline{\mathcal{G}} = \{F \in \mathcal{F} : P(F \Delta G) = 0 \text{ for some } G \in \mathcal{G}\}.$$

A straightforward consequence is that a real $\overline{\mathcal{G}}$-measurable function on $\Omega$ coincides a.s. with some $\mathcal{G}$-measurable function. Thus, if $U : \Omega \to \mathbb{R}$ is $\mathcal{D}$-measurable, then $U = U_i$ a.s. for some $\mathcal{A}_i$-measurable function $U_i : \Omega \to \mathbb{R}$, $i = 1, \ldots, k$.

## 3. Fields where $\mathcal{D}$ appears

We list some fields involving $\mathcal{D}$, by paying particular attention to the case where $P|\mathcal{D}$ is atomic. We stress by now that, for atomicity of $P|\mathcal{D}$ to be a real advantage, the atoms of $P|\mathcal{D}$ and their probabilities should be known.

Throughout, $X_i$ is a random variable on $(\Omega, \mathcal{F}, P)$ with values in the measurable space $(\mathcal{X}_i, \mathcal{B}_i)$, $i = 1, \ldots, k$, and

$$X_i^* = (X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_k).$$

3.1. **Intersection property.** Let $X$ be a random variable on $(\Omega, \mathcal{F}, P)$ with values in the measurable space $(\mathcal{X}, \mathcal{E})$. The intersection property (IP) is

$$X \perp X_i^* \mid X_i \text{ for } i = 1, \ldots, k \quad \Longrightarrow \quad X \perp (X_1, \ldots, X_k)$$

where the notation "$U \perp V \mid W$" stands for "$U$ conditionally independent of $V$ given $W$". It is well known that IP may fail. As a trivial example, take $X$ not independent of $X_1$ and $X_i = X_1$ for all $i$.

IP is involved in a number of arguments. It appears, for instance, in graphical models, zero entries in contingency tables, causal inference and estimation in Markov processes; see [10] and references therein.

The connections between IP and $\mathcal{D}$ are made clear by part (b) of the next (obvious) result. Part (a) is already known for $k = 2$ (see Proposition 2.2 of [10] and references therein) but we give a proof to make the paper self-contained.

**Theorem 1.** *Let $\mathcal{A}_i = \sigma(X_i)$ for all $i$. Then:*

**(a)** $X \perp X_i^* \mid X_i$ *for* $i = 1, \ldots, k \iff E(f(X) \mid X_1, \ldots, X_k) = E(f(X) \mid \mathcal{D})$ *a.s. for each real bounded measurable function $f$ on $(\mathcal{X}, \mathcal{E})$;*

**(b)** $X \perp X_i^* \mid X_i$ *for* $i = 1, \ldots, k$ *and* $X \perp \mathcal{D} \iff X \perp (X_1, \ldots, X_k)$*;*

**(c)** $X \perp \mathcal{D}$ *if and only if*

$$P(X \in A, X_1 \in B_1) = P(X \in A)\, P(X_1 \in B_1) \quad \textit{whenever}$$

$A \in \mathcal{E}, B_1 \in \mathcal{B}_1$ *and* $P\big(\{X_1 \in B_1\} \Delta \{X_i \in B_i\}\big) = 0$ *for some* $B_i \in \mathcal{B}_i,\ i = 2, \ldots, k$.

*Proof.* **(a)** Suppose $E\big(f(X) \mid X_1, \ldots, X_k\big) = E(f(X) \mid \mathcal{D})$ a.s. for all bounded measurable $f$ on $(\mathcal{X}, \mathcal{E})$. Given $i$, since $\mathcal{D} \subset \overline{\sigma(X_i)} \subset \overline{\sigma(X_1, \ldots, X_k)}$, then

$$E(f(X) \mid X_i) = E(f(X) \mid \mathcal{D}) = E\big(f(X) \mid X_1, \ldots, X_k\big) \text{ a.s.}$$

for all bounded measurable $f$, that is, $X \perp X_i^* \mid X_i$. Conversely, suppose $X \perp X_i^* \mid X_i$ for all $i$. Given a bounded measurable $f$,

$$E(f(X) \mid X_1) = E\big(f(X) \mid X_1, \ldots, X_k\big) = E(f(X) \mid X_i) \quad \text{a.s. for all } i.$$

Thus, $E(f(X) \mid X_1)$ is $\overline{\sigma(X_i)}$-measurable for all $i$, i.e., it is $\mathcal{D}$-measurable. Therefore, $E\big(f(X) \mid X_1, \ldots, X_k\big) = E(f(X) \mid X_1) = E(f(X) \mid \mathcal{D})$ a.s..

**(b)** "$\Longleftarrow$" is obvious. As to "$\Longrightarrow$", it suffices noting that $E\big(f(X) \mid X_1, \ldots, X_k\big) = E(f(X) \mid \mathcal{D}) = Ef(X)$ a.s. for all bounded measurable $f$, where the first equality is by part (a) and the second is because $X \perp \mathcal{D}$.

**(c)** Just note that

$$\mathcal{D} = \{F \in \mathcal{F} : P\big(F \Delta \{X_i \in B_i\}\big) = 0 \text{ for some } B_i \in \mathcal{B}_i,\ i = 1, \ldots, k\}.$$

□

Thus, $X \perp \mathcal{D}$ is a (natural) sufficient condition for IP. In a sense, it is necessary as well, since it is a consequence of $X \perp (X_1, \ldots, X_k)$. Heuristically, $X \perp \mathcal{D}$ means that $X$ is not affected by that part of information which is common to $X_1, \ldots, X_k$.

To test whether $X \perp \mathcal{D}$, atomicity of $\mathcal{D}$ under $P$ can help. If $P|\mathcal{D}$ is atomic, in fact, $X \perp \mathcal{D}$ reduces to

$$P(X \in A,\, D) = P(X \in A)\, P(D) \quad \text{for all } A \in \mathcal{E} \text{ and atoms } D \text{ of } P|\mathcal{D}.$$

As shown in Theorem 4, for $P|\mathcal{D}$ to be atomic, it is enough that the distribution of $(X_1, \ldots, X_k)$ is absolutely continuous with respect to a $\sigma$-finite product measure; see also Lemma 6.

A last note is that $X \perp \mathcal{D}$ is trivially true whenever $\mathcal{D} = \mathcal{N}$. In [10], a paper which inspired the present Subsection, $\mathcal{D} = \mathcal{N}$ was firstly viewed as a sufficient condition for IP. In [3], in a Gibbs sampling framework, $\mathcal{D} = \mathcal{N}$ was given a characterization and various sufficient conditions.

3.2. **Iterated conditional expectations.** Let $X$ be a real random variable on $(\Omega, \mathcal{F}, P)$ such that $EX^2 < \infty$ and

$$\mathcal{D}_{mk+i} = \mathcal{A}_i \quad \text{for all } m = 0, 1, \ldots \text{ and } i = 1, \ldots, k.$$

Define $Z_0 = X$ and $Z_n = E(Z_{n-1} \mid \mathcal{D}_n)$ for $n \geq 1$. Then,

$$Z_n \overset{a.s.}{\to} E(X \mid \mathcal{D}) \quad \text{as } n \to \infty. \tag{1}$$

This classical result was obtained by Burkholder-Chow [5] for $k = 2$ and Delyon-Delyon [6] for arbitrary $k$. See [7] for some historical notes.

Suppose now that $\mathcal{D}$ is atomic under $P$ and the goal is estimating $EX$. Then, $E(X \mid \mathcal{D}) = \sum_j I_{D_j} E(X \mid D_j)$ a.s. where $D_1, D_2, \ldots$ denote the (disjoint) atoms of $P|\mathcal{D}$. Thus, one should apply relation (1) on each atom $D_j$, so as to obtain an estimate for $E(X \mid D_j)$, and then use the formula $EX = \sum_j P(D_j) E(X \mid D_j)$.

3.3. **Gibbs sampling.** As noted in [7], the limit theorem of Burkholder-Chow and Delyon-Delyon (Subsection 3.2) is intrinsically connected to Gibbs sampling.

Let $X_1, \ldots, X_k$ be the canonical projections on

$$(\Omega, \mathcal{F}) = \Big(\prod_{i=1}^k \mathcal{X}_i, \prod_{i=1}^k \mathcal{B}_i\Big).$$

Each $X_i$ is assumed to admit a regular conditional distribution $\gamma_i$ given $X_i^*$. In the notation $u = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k)$, where $x_j \in \mathcal{X}_j$ for all $j \neq i$, this means that: (i) $\gamma_i(u)$ is a probability measure on $\mathcal{B}_i$ for every $u$; (ii) $u \mapsto \gamma_i(u)(A)$ is measurable for $A \in \mathcal{B}_i$; (iii) $P(B) = \int\int I_B(x_1, \ldots, x_k) \gamma_i(u)(dx_i)\, \gamma_i^*(du)$ for $B \in \prod_{i=1}^k \mathcal{B}_i$, where $\gamma_i^*$ denotes the marginal distribution of $X_i^*$.

The Gibbs-chain

$$Y_n = (Y_{1,n}, \ldots, Y_{k,n}), \quad n \geq 0,$$

can be informally described as follows. Starting from $\omega = (x_1, \ldots, x_k)$, the next state $\omega^* = (a_1, \ldots, a_k)$ is obtained by sequentially generating $a_k, a_{k-1}, \ldots, a_1$, each $a_i$ being selected from the conditional distribution of $X_i$ given $X_1 = x_1, \ldots, X_{i-1} = x_{i-1}, X_{i+1} = a_{i+1}, \ldots, X_k = a_k$. Formally, $(Y_n)$ is the homogeneous Markov chain with state space $(\Omega, \mathcal{F})$ and transition kernel

$$K(\omega, B) = K\big((x_1, \ldots, x_k),\, B\big)$$
$$= \int \ldots \int I_B(a_1, \ldots, a_k) \prod_{i=1}^k \gamma_i(x_1, \ldots, x_{i-1}, a_{i+1}, \ldots, a_k)(da_i).$$

Note that $P$ is a stationary distribution for $(Y_n)$, i.e., $P(\cdot) = \int K(\omega, \cdot) P(d\omega)$. Let $\mathbb{P}$ denote the law of $(Y_n)$ such that $Y_0 \sim P$.

The Gibbs chain is constructed mainly for sampling from $P$. To this end, the following SLLN is fundamental

$$\frac{1}{n} \sum_{i=0}^{n-1} \phi(Y_i) \to \int \phi\, dP, \ \mathbb{P}\text{-a.s., for all } \phi \in L_1(P). \tag{2}$$

Another basic requirement of $(Y_n)$, stronger than (2), is ergodicity on some set $S \in \mathcal{F}$, that is

$$P(S) = 1, \quad K(\omega, S) = 1 \quad \text{and} \quad \|K^n(\omega, \cdot) - P\| \to 0 \text{ for each } \omega \in S,$$

where $\|\cdot\|$ is total variation norm and $K^n$ the $n$-th iterate of $K$.

Now, letting $\mathcal{A}_i = \sigma(X_i^*)$ for all $i$, the SLLN under (2) is equivalent to

$$\mathcal{D} = \mathcal{N}.$$

In addition, in case $\mathcal{F}$ is countably generated and $P$ absolutely continuous with respect to a $\sigma$-finite product measure, $\mathcal{D} = \mathcal{N}$ if and only if $(Y_n)$ is ergodic on $S_0 = \{\omega \in \Omega : K(\omega, \cdot) \ll P\}$. See Theorems 4.2, 4.5 and Remark 4.7 of [3].

Conditions for $\mathcal{D} = \mathcal{N}$ (when $\mathcal{A}_i = \sigma(X_i^*)$ for all $i$) are given in Theorem 8; see also Lemma 6.

Strictly speaking, thus, Gibbs sampling is admissible only if $\mathcal{D} = \mathcal{N}$. At least in principle, however, it makes sense even if $\mathcal{D} \neq \mathcal{N}$, provided $\mathcal{D}$ is atomic under $P$. In fact, if $P|\mathcal{D}$ is atomic (with disjoint atoms $D_1, D_2, \ldots$), then (2) turns into

$$(2^*) \qquad \frac{1}{n}\sum_{i=0}^{n-1} \phi(Y_i) \to \int \phi(\omega)\, P(d\omega \mid D_j),$$

$$\mathbb{P}\text{-a.s. on } \{Y_0 \in D_j\}, \text{ for all } j \text{ and } \phi \in L_1(P).$$

The SLLN under (2*) can be proved by the same argument used for Theorem 4.2 of [3] plus the observation that $K(\cdot, B) = I_B(\cdot)$, $P$-a.s., for each $B \in \mathcal{D}$.

If the atoms $D_j$ are only a finite number, and they are known together with their probabilities $P(D_j)$, then (2*) can be used to evaluate $\int \phi\, dP$. The chain $(Y_n)$ should be started on each $D_j$, so as to obtain an estimate for $\int \phi(\omega)\, P(d\omega \mid D_j)$, and then the formula $\int \phi\, dP = \sum_j P(D_j) \int \phi(\omega)\, P(d\omega \mid D_j)$ should be applied.

As shown in Theorem 10, for $P|\mathcal{D}$ to be atomic, it is enough that the distribution of $(X_1, \ldots, X_k)$ is absolutely continuous with respect to a $\sigma$-finite product measure.

## 4. Atomicity of $\mathcal{D}$ under $P$

We begin with a definition. Say that $H \subset \Omega$ has the *trivial intersection property*, or briefly that $H$ is TIP, in case $H \in \mathcal{F}$, $P(H) > 0$, and

$$A_i \in \mathcal{A}_i \text{ and } P\big(A_i \Delta A_1 \mid H\big) = 0 \text{ for } i = 1, \ldots, k \implies P(A_1 \mid H) \in \{0, 1\}.$$

Here are some obvious consequences of the definition.

(i) If $H$ is TIP, $A_i \in \mathcal{A}_i$ and $P\big(A_i \Delta A_1 \mid H\big) = 0$ for all $i$, then either $P(A_i \mid H) = 0$ for all $i$ or $P(A_i \mid H) = 1$ for all $i$.

(ii) Let $H \in \mathcal{F}$ with $P(H) > 0$ and write $P_H = P(\cdot \mid H)$. Then, $H$ is TIP if and only if $\Omega$ is $P_H$-TIP (i.e., $\Omega$ is TIP under $P_H$). Moreover, $\Omega$ is TIP if and only if $\mathcal{D} = \mathcal{N}$. Therefore, the definition of TIP set may be rephrased as follows: $H$ is TIP if and only if

$$\mathcal{D}_{P_H} = \mathcal{N}_{P_H}$$

where $\mathcal{N}_{P_H} = \{F \in \mathcal{F} : P_H(F) \in \{0, 1\}\}$ and $\mathcal{D}_{P_H} = \cap_{i=1}^k \sigma(\mathcal{A}_i \cup \mathcal{N}_{P_H})$.

(iii) Let $Q$ be a probability on $\mathcal{F}$. If $P$ and $Q$ are equivalent (i.e., $P \ll Q$ and $Q \ll P$), then $H$ is $Q$-TIP if and only if it is $P$-TIP. If $P \ll Q$ and $H \subset \{\frac{dP}{dQ} > 0\}$, for some given version of $\frac{dP}{dQ}$, then $H$ is $Q$-TIP if and only if it is $P$-TIP.

The present notion of TIP set generalizes the one given in [3] for $k = 2$. Among other things, such a notion is basic for characterizing atomicity of $P|\mathcal{D}$.

**Theorem 2.** *Let $H \subset \Omega$. Then,*

- **(a)** *If $H$ is TIP, there is an atom $H^*$ of $P|\mathcal{D}$ satisfying $H^* \supset H$ and $P(D \mid H^*) = P(D \mid H)$ for all $D \in \mathcal{D}$;*
- **(b)** *For $H$ to be an atom of $P|\mathcal{D}$ it is necessary and sufficient that $H \in \mathcal{D}$ and $H$ is TIP;*
- **(c)** *$\mathcal{D}$ is atomic under $P$ if and only if $P(\cup_n H_n) = 1$ for some countable collection $H_1, H_2, \ldots$ of TIP sets.*

*Proof.* **(a)** Suppose $H$ is TIP. We first prove that $P(\cdot \mid H)$ is 0-1 on $\mathcal{D}$. Given $D \in \mathcal{D}$, for each $i$ there is $A_i \in \mathcal{A}_i$ such that $P(A_i \Delta D) = 0$. Hence, $P\big((A_i \Delta A_1) \cap H\big) \leq P(A_i \Delta A_1) = 0$ for all $i$, and $H$ TIP implies $P(D \mid H) = P(A_1 \mid H) \in \{0, 1\}$. Next, by a standard argument, there is $H^* \in \mathcal{D}$ such that $H^* \supset H$ and

$$P(H^*) = \inf\{P(D) : H \subset D \in \mathcal{D}\}.$$

Let $D \in \mathcal{D}$. If $P(D \mid H) = 1$, then

$$H \subset (D \cap H^*) \cup (D^c \cap H) \in \mathcal{D},$$

so that $P(H^*) \leq P\big((D \cap H^*) \cup (D^c \cap H)\big) = P(D \cap H^*)$ by definition of $H^*$. Hence, $P(D \mid H^*) = 1$. Taking complements, if $P(D \mid H) = 0$ then $P(D \mid H^*) = 0$. Thus, $H^*$ is an atom of $P|\mathcal{D}$ and $P(\cdot \mid H^*) = P(\cdot \mid H)$ on $\mathcal{D}$.

**(b)** If $H \in \mathcal{D}$ is TIP, then $H$ is an atom of $P|\mathcal{D}$ since $P(H \mid H^*) = P(H \mid H) = 1$, where $H^*$ is as in point (a). Conversely, suppose $H$ is an atom of $P|\mathcal{D}$. To prove $H$ TIP, we fix $A_i \in \mathcal{A}_i$ such that $P(A_i \Delta A_1 \mid H) = 0$ for $i = 1, \ldots, k$. For each $i$, since $H \in \mathcal{D} \subset \overline{\mathcal{A}_i}$, some $H_i \in \mathcal{A}_i$ meets $P(H \Delta H_i) = 0$. Moreover,

$$P\big((A_1 \cap H) \Delta (A_i \cap H_i)\big) \leq P(H \Delta H_i) + P\big((A_i \Delta A_1) \cap H\big) = 0.$$

Hence, $A_1 \cap H \in \overline{\mathcal{A}_i}$ for all $i$, that is, $A_1 \cap H \in \mathcal{D}$. Since $H$ is an atom of $P|\mathcal{D}$, it follows that $P(A_1 \mid H) = P(A_1 \cap H \mid H) \in \{0, 1\}$. Thus, $H$ is TIP.

**(c)** If $P|\mathcal{D}$ is atomic, it suffices to take the $H_n$ as the atoms of $P|\mathcal{D}$ and to apply point (b). Conversely, if $P(\cup_n H_n) = 1$ with the $H_n$ TIP, for each $n$ point (a) implies $H_n \subset H_n^*$ for some atom $H_n^*$ of $P|\mathcal{D}$. Then, $P|\mathcal{D}$ is atomic since $P(\cup_n H_n^*) \geq P(\cup_n H_n) = 1$. □

By Theorem 2, $P|\mathcal{D}$ is atomic provided $\Omega$ can be covered by countably many TIP sets $H_1, H_2, \ldots$. In this case, every atom $D$ admits the representation $D = \cup_{i \in I} H_i$ a.s. for some index set $I$ (by point (a)). The next issue, thus, is identifying such atoms using the $H_n$ as building blocks. Indeed, the atoms are maximal TIP sets, according to the following result.

**Theorem 3.** *Suppose $P(\cup_n H_n) = 1$, where $H_1, H_2, \ldots$ are TIP, and let $D \subset \Omega$. Then, $D$ is an atom of $P|\mathcal{D}$ if and only if $D$ is TIP and*

$$D \cup H_n \text{ fails to be TIP whenever } P(H_n \setminus D) > 0. \tag{3}$$

*Proof.* By Theorem 2, it can be assumed $D$ TIP, and we have to prove that condition (3) is equivalent to $D \in \mathcal{D}$. Suppose (3) holds. Let $N = \{n : P(H_n \setminus D) > 0\}$. If $N = \emptyset$, then $P(D^c) \leq \sum_n P(H_n \setminus D) = 0$, so that $D \in \mathcal{N} \subset \mathcal{D}$. If $N \neq \emptyset$, by (3), for each $n \in N$ there are $A_{i,n} \in \mathcal{A}_i$, $i = 1, \ldots, k$, satisfying

$$P\big(A_{i,n} \Delta A_{1,n} \mid D \cup H_n\big) = 0 \text{ and } P\big(A_{i,n} \mid D \cup H_n\big) \in (0, 1) \text{ for all } i.$$

Since $D$ and $H_n$ are TIP, one also has $P(A_{i,n} \mid D) \in \{0, 1\}$ and $P(A_{i,n} \mid H_n) \in \{0, 1\}$ for all $i$, and thus

$$P(A_{i,n} \mid D) = 1 - P(A_{i,n} \mid H_n) \text{ for all } i.$$

Define $F_{i,n} = A_{i,n}$ or $F_{i,n} = A_{i,n}^c$ as $P(A_{i,n} \mid D) = 1$ or $P(A_{i,n} \mid D) = 0$, and

$$A_i = \cap_{n \in N} F_{i,n}.$$

Then, $P(A_i \mid D) = 1$ and $P(A_i \mid H_n) = 0$ for all $i$ and $n \in N$. Hence, given $i$,

$$P(A_i \Delta D) = P(A_i \setminus D) \leq \sum_{n \in N} P\big(A_i \cap D^c \cap H_n\big) \leq \sum_{n \in N} P\big(A_i \cap H_n\big) = 0.$$

Since $A_i \in \mathcal{A}_i$, it follows that $D \in \overline{\mathcal{A}_i}$, that is, $D \in \mathcal{D}$. Conversely, suppose $D \in \mathcal{D}$ and $D \cup H_n$ is TIP for some $n$. Since $P(D \mid D \cup H_n) \in \{0, 1\}$ (by point (a) of Theorem 2) and $P(D) > 0$ (as $D$ is TIP), then $P(H_n \setminus D) = 0$. □

In real problems, it is not unusual that $P \ll Q$ for some probability $Q$ on $\mathcal{F}$ which makes $\mathcal{A}_1, \ldots, \mathcal{A}_k$ independent. This does not imply $\mathcal{D} = \mathcal{N}$ (see Examples 3.16 and 3.17 of [3]) but it suffices for atomicity of $P|\mathcal{D}$. Actually, it is enough that a couple of the $\mathcal{A}_i$ are independent under $Q$.

**Theorem 4.** *$P|\mathcal{D}$ is atomic provided $P \ll Q$ for some probability measure $Q$ on $\mathcal{F}$ which makes $\mathcal{A}_r$ and $\mathcal{A}_s$ independent for some $r$, $s$.*

*Proof.* Fix $H \in \mathcal{D}$ with $P(H) > 0$ and let $P_H = P(\cdot \mid H)$. If $P_H|\mathcal{D}$ is nonatomic, the probability space $(\Omega, \mathcal{D}, P_H|\mathcal{D})$ supports a real random variable with uniform distribution; see Section 2. Hence, it suffices to prove that each $\mathcal{D}$-measurable function $U : \Omega \to \mathbb{R}$ satisfies $P_H(U \in C) = 1$ for some countable set $C \subset \mathbb{R}$. Further, since $P_H \ll P$, it is enough to show that $P(U \in C) = 1$. Let $U : \Omega \to \mathbb{R}$ be $\mathcal{D}$-measurable. Then, $U = U_i$ a.s. for some $U_i : \Omega \to \mathbb{R}$ satisfying $\sigma(U_i) \subset \mathcal{A}_i$, $i = 1, \ldots, k$. Define the countable set $C = \{c \in \mathbb{R} : Q(U_s = c) > 0\}$. Since $U_r$ and $U_s$ are independent under $Q$,

$$Q(U_r \notin C,\, U_r = U_s) = \int_{\{U_r \notin C\}} Q\{x : U_s(x) = U_r(\omega)\}\, Q(d\omega) = 0.$$

Thus, $P \ll Q$ yields

$$P(U \in C) = 1 - P\big(U \notin C,\, U = U_r = U_s\big) = 1 - P\big(U_r \notin C,\, U_r = U_s\big) = 1.$$

□

For $k = 2$, Theorem 4 reduces to Theorem 3.10 of [3].

**Remark 5.** Let $\mathcal{A}_i = \sigma(X_i)$ for all $i$, where $X_i : \Omega \to \mathcal{X}_i$ is a random variable and $\mathcal{X}_i$ a separable metric space (equipped with its Borel $\sigma$-field $\mathcal{B}_i$). Then, $P|\mathcal{D}$ need not be atomic even though

$$P\big(X_i = f(X_j)\big) = 0 \tag{4}$$

for all $i \neq j$ and all measurable functions $f : \mathcal{X}_j \to \mathcal{X}_i$. We mention this fact since, for some time, we conjectured $P|\mathcal{D}$ atomic under (4).

As an example, let $k = 2$, $X_1 = (U, W)$ and $X_2 = (V, W)$, where $U$, $V$, $W$ are real independent random variables with nonatomic distributions. Take $\mathcal{X}_i = \mathbb{R}^2$ and $\mathcal{A}_i = \sigma(X_i)$ for $i = 1, 2$. Given a measurable function $f = (f_1, f_2) : \mathbb{R}^2 \to \mathbb{R}^2$, one obtains $P\big(X_2 = f(X_1)\big) \leq P\big(V = f_1(U, W)\big) = 0$ since $V$ has nonatomic distribution and is independent of $(U, W)$. Likewise, $P\big(X_1 = f(X_2)\big) = 0$. However, $P|\mathcal{D}$ is nonatomic, as $\sigma(W) \subset \mathcal{D}$ and $W$ has nonatomic distribution.

Finally, we state a simple but useful fact as a lemma. Let $Q$ be a probability measure on $\mathcal{F}$. Say that $P$ and $Q$ are *equivalent on rectangles* in case

$$P(A) = 0 \Leftrightarrow Q(A) = 0 \text{ for each } A \in \mathcal{R},$$

$$\text{where } \mathcal{R} = \{\cap_{i=1}^k A_i : A_i \in \mathcal{A}_i,\, i = 1, \ldots, k\}.$$

If $A \in \mathcal{R}$, then $A^c$ is a finite union of elements of $\mathcal{R}$. Hence, $P(A) = 1 \Leftrightarrow Q(A) = 1$ and $P(A \Delta B) = 0 \Leftrightarrow Q(A \Delta B) = 0$ whenever $A$, $B \in \mathcal{R}$ and $P$, $Q$ are equivalent on rectangles. Note that $\mathcal{A}_i \subset \mathcal{R}$ for all $i$. Note also that $P$ and $Q$ need not be equivalent on $\sigma(\mathcal{R})$ even though they are equivalent on rectangles.

**Lemma 6.** *Let $P$ and $Q$ be equivalent on rectangles. If $D$ is an atom of $Q|\mathcal{D}_Q$, there is $A \in \mathcal{R}$ such that $Q(A\Delta D) = 0$ and $A$ is an atom of $P|\mathcal{D}_P$. Moreover,*

$$\mathcal{D}_Q = \mathcal{N}_Q \Leftrightarrow \mathcal{D}_P = \mathcal{N}_P, \text{ and}$$
$$\mathcal{D}_Q \text{ is atomic under } Q \Leftrightarrow \mathcal{D}_P \text{ is atomic under } P.$$

*(Here, $\mathcal{N}_Q = \{F \in \mathcal{F} : Q(F) \in \{0,1\}\}$, $\mathcal{D}_Q = \cap_{i=1}^k \sigma(\mathcal{A}_i \cup \mathcal{N}_Q)$, $\mathcal{N}_P = \mathcal{N}$ and $\mathcal{D}_P = \mathcal{D}$).*

*Proof.* We first prove that, for each $D \in \mathcal{D}_Q$ with $Q(D) > 0$, there is $A = A(D)$ satisfying $A \in \mathcal{A}_1 \cap \mathcal{D}_P$, $Q(A\Delta D) = 0$ and $P(A) > 0$. Take in fact $A_i \in \mathcal{A}_i$ with $Q(A_i\Delta D) = 0$, $i = 1, \ldots, k$, and let $A = A_1$. Then $A \in \mathcal{A}_1$, $Q(A\Delta D) = 0$ and $P(A) > 0$. Since $Q(A_i\Delta A) = 0$ for all $i$, then $P(A_i\Delta A) = 0$ for all $i$, so that $A \in \mathcal{D}_P$. Next, let $D$ be an atom of $Q|\mathcal{D}_Q$ and $A = A(D)$. Then, $A \in \mathcal{A}_1 \subset \mathcal{R}$. Given $G \in \mathcal{D}_P$, for each $i$ there is $G_i \in \mathcal{A}_i$ such that $P(G\Delta G_i) = 0$. Again, $P(G_i\Delta G_1) = 0$ for all $i$ implies $Q(G_i\Delta G_1) = 0$ for all $i$, so that $G_1 \in \mathcal{D}_Q$. Since $A$ is an atom of $Q|\mathcal{D}_Q$ (as $Q(A\Delta D) = 0$), either $Q(A \cap G_1) = 0$ or $Q(A \cap G_1^c) = 0$. Accordingly, either $P(A \cap G) = P(A \cap G_1) = 0$ or $P(A \cap G^c) = P(A \cap G_1^c) = 0$, i.e., $A$ is an atom of $P|\mathcal{D}_P$. Next, if $\mathcal{D}_Q = \mathcal{N}_Q$, then $\Omega$ is an atom of $Q|\mathcal{D}_Q$. Thus, $A = A(\Omega)$ is an atom of $P|\mathcal{D}_P$ and $P(A) = 1$, i.e., $\mathcal{D}_P = \mathcal{N}_P$. Finally, suppose $Q|\mathcal{D}_Q$ is atomic with (disjoint) atoms $D_1, D_2, \ldots$. Let $A_j = A(D_j)$ and $A = \cup_j A_j$. Then, each $A_j$ is an atom of $P|\mathcal{D}_P$, and $P(A) = 1$ since $Q(A) = 1$ and $A \in \mathcal{A}_1 \subset \mathcal{R}$. Therefore, $P|\mathcal{D}_P$ is atomic. □

## 5. Applications to Gibbs sampling

As remarked in Subsection 3.3, in a Gibbs sampling framework it is fundamental that $\mathcal{D} = \mathcal{N}$, or at least that $\mathcal{D}$ is atomic under $P$, when the $\mathcal{A}_i$ are given by

$$\mathcal{A}_i = \sigma(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_k).$$

In this section, we let $\mathcal{A}_i = \sigma(X_i^*)$ for all $i$, where $X_i^* = (X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_k)$ and the $X_i$ are random variables on $(\Omega, \mathcal{F}, P)$ with values in the measurable spaces $(\mathcal{X}_i, \mathcal{B}_i)$, $i = 1, \ldots, k$. We also let $\mathcal{D}_0 = \cap_i \overline{\sigma(X_i)}$. Since $\mathcal{D}_0 \subset \mathcal{D}$, $P$ is 0-1-valued or atomic on $\mathcal{D}_0$ whenever it is so on $\mathcal{D}$.

Let $\mathcal{X} = \prod_{i=1}^k \mathcal{X}_i$ and let $\mathcal{B} = \prod_{i=1}^k \mathcal{B}_i$ denote the product $\sigma$-field on $\mathcal{X}$. Define two measures on $\mathcal{B}$ as

$$\gamma(\cdot) = P\big((X_1, \ldots, X_k) \in \cdot\big) \quad \text{and} \quad \mu = \mu_1 \times \ldots \times \mu_k$$

where each $\mu_i$ is a $\sigma$-finite measure on $\mathcal{B}_i$. Thus, $\gamma$ is the probability distribution of $(X_1, \ldots, X_k)$ and $\mu$ a $\sigma$-finite product measure.

By Theorem 4, it follows that $P|\mathcal{D}_0$ is atomic whenever $\gamma \ll \mu$. Whether or not $\gamma \ll \mu$ implies $P|\mathcal{D}$ atomic is a bit more delicate and is the main focus of this section. We start by noting that, in the independent case, things are as expected.

**Lemma 7.** *Let $\mathcal{A}_i = \sigma(X_i^*)$ for all $i$ and*

$$\mathcal{D}_j = \cap_{i=1}^j \overline{\sigma(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_j)}, \quad j = 2, \ldots, k.$$

*If $X_j$ is independent of $(X_1, \ldots, X_{j-1})$, then $\mathcal{D}_j = \mathcal{D}_{j-1}$. In particular, if $X_1, \ldots, X_k$ are independent, then $\mathcal{D} = \mathcal{N}$ and $H = \{X_1 \in B_1, \ldots, X_k \in B_k\}$ is TIP as far as $B_i \in \mathcal{B}_i$ for all $i$ and $P(H) > 0$.*

*Proof.* Since $\mathcal{D}_{j-1} \subset \mathcal{D}_j$, it suffices to prove $\mathcal{D}_{j-1} \supset \mathcal{D}_j$. Let $A \in \mathcal{D}_j$. Then, $I_A = f_i(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_j)$ a.s. for some bounded measurable function $f_i$, $i = 1, \ldots, j$. Let $\alpha_j$ denote the probability distribution of $X_j$. If $X_j$ is independent of $(X_1, \ldots, X_{j-1})$, then

$$\begin{aligned} I_A &= E\big(I_A \mid X_1, \ldots, X_{j-1}\big) \\ &= E\big(\, f_i(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_j) \mid X_1, \ldots, X_{j-1}\big) \\ &= \int f_i(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_{j-1}, t)\, \alpha_j(dt) \quad \text{a.s. for each } i < j. \end{aligned}$$

Thus, $A \in \mathcal{D}_{j-1}$. Next, suppose $X_1, \ldots, X_k$ are independent. By what already proved,

$$\mathcal{D} = \mathcal{D}_k = \mathcal{D}_{k-1} = \ldots = \mathcal{D}_2 = \overline{\sigma(X_1)} \cap \overline{\sigma(X_2)} = \mathcal{N},$$

or equivalently $\Omega$ is TIP. Since $X_1, \ldots, X_k$ are still independent under $P(\cdot \mid H)$, it follows that $\Omega$ is $P(\cdot \mid H)$-TIP, that is, $H$ is $P$-TIP. □

The independence assumption can be considerably relaxed. Next result is inspired to Corollary 3.7 of [3].

**Theorem 8.** *Suppose $\mathcal{A}_i = \sigma(X_i^*)$ for all $i$, $\gamma \ll \mu$ and $f$ is a version of $\frac{d\gamma}{d\mu}$. Let*

$$H = \{(X_1, \ldots, X_k) \in B\} \quad \textit{where } B \in \mathcal{B} \textit{ and } B \subset \{f > 0\}.$$

*Then, $H$ is TIP provided*

$$\cup_{i=1}^k \{X_i^* \in B_i^*\} \supset H \supset \cap_{i=1}^k \{X_i \in B_i\}, \quad \textit{where } B_i^* = \times_{j \neq i} B_j, \tag{5}$$
$$\textit{for some } B_i \in \mathcal{B}_i,\ i = 1, \ldots, k, \ \textit{with } P(X_1 \in B_1, \ldots, X_k \in B_k) > 0.$$

*Proof.* It can be assumed $(\Omega, \mathcal{F}, P) = (\mathcal{X}, \mathcal{B}, \gamma)$ and $X_1, \ldots, X_k$ the canonical projections (so that $H = B$). For each $i$, since $\mu_i$ is $\sigma$-finite, there is a probability $Q_i$ on $\mathcal{B}_i$ equivalent to $\mu_i$. Let $Q = Q_1 \times \ldots \times Q_k$ denote the corresponding product probability on $\mathcal{B}$. Since $P \ll Q$ and $P(H) > 0$, then $Q(H) > 0$. Since $f > 0$ on $H$, then $Q(\cdot \mid H)$ is equivalent to $P(\cdot \mid H)$. Thus, $H$ is $P$-TIP if and only if it is $Q$-TIP. We next prove that $H$ is $Q$-TIP. Let $K = \{X_1 \in B_1, \ldots, X_k \in B_k\}$. Since $Q(K) > 0$ (due to $P(K) > 0$) and $X_1, \ldots, X_k$ are independent under $Q$, Lemma 7 implies that $K$ is $Q$-TIP. Fix $A_i \in \mathcal{A}_i$ with $Q(A_i \Delta A_1 \mid H) = 0$, $i = 1, \ldots, k$. Since $K$ is $Q$-TIP and $K \subset H$, then $Q(A_1 \mid K) \in \{0, 1\}$, say $Q(A_1 \mid K) = 0$ (so that $Q(A_i \mid K) = 0$ for all $i$). Given $j$, since $Q(X_j \in B_j) > 0$ and

$$Q\big(A_j,\, X_j^* \in B_j^*\big)\, Q(X_j \in B_j) = Q\big(A_j,\, X_j^* \in B_j^*,\, X_j \in B_j\big) = Q(A_j \cap K) = 0,$$

then $Q(A_j,\, X_j^* \in B_j^*) = 0$. Also, $\{X_j^* \in B_j^*\} \cap \{X_r \notin B_r\} = \emptyset$ for $j \neq r$ and $H \subset \cup_{i=1}^k \{X_i^* \in B_i^*\}$. Thus, letting $A = \cap_i A_i$,

$$\begin{aligned} Q\big(A_1 \cap H\big) &= Q\big(A \cap H\big) = Q\big(A \cap H \cap K^c\big) \\ &\leq Q\big(A \cap (\cup_j \{X_j^* \in B_j^*\}) \cap (\cup_r \{X_r \notin B_r\})\big) \\ &\leq \sum_j \sum_r Q\big(A,\, X_j^* \in B_j^*,\, X_r \notin B_r\big) \\ &= \sum_j Q\big(A,\, X_j^* \in B_j^*,\, X_j \notin B_j\big) \\ &\leq \sum_j Q\big(A_j,\, X_j^* \in B_j^*\big) = 0. \end{aligned}$$

Thus, $H$ is $Q$-TIP, and this concludes the proof. □

By Theorem 8, $\mathcal{D} = \mathcal{N}$ in case $\gamma \ll \mu$ and condition (5) holds with $B = \{f > 0\}$.

**Example 9.** Let $X_3 = X_1 X_2$ where $X_1$ and $X_2$ are i.i.d. random variables with values in $\{-1, 1\}$ and $P(X_1 = -1) = P(X_1 = 1) = \frac{1}{2}$. Let $\mu_1 = \mu_2 = \mu_3$ be counting measure on $\{-1, 1\}$ and $\mathcal{D}_0 = \cap_i \overline{\sigma(X_i)}$. Since the $X_i$ are pairwise independent (even if not independent), $\mathcal{D}_0 = \mathcal{N}$. Since $P(X_3 = 1) = \frac{1}{2}$ and $\mathcal{D} \supset \sigma(X_3)$, then $\mathcal{D} \neq \mathcal{N}$. Thus, $\mathcal{D}_0 = \mathcal{N}$ and $\gamma \ll \mu$ do not imply $\mathcal{D} = \mathcal{N}$. Note also that $\cup_{i=1}^3 \{X_i = 1\} = \Omega$, $P(X_1 = X_2 = X_3 = 1) > 0$ while $H = \Omega$ is not TIP. Thus, condition (5) cannot be weakened into

$$\cup_{i=1}^k \{X_i \in B_i\} \supset H \supset \cap_{i=1}^k \{X_i \in B_i\} \quad \text{with} \quad P(X_1 \in B_1, \ldots, X_k \in B_k) > 0.$$

Our last and main result is that $\mathcal{D}$ is atomic under $P$ as far as $\gamma \ll \mu$.

**Theorem 10.** *Let $\mathcal{A}_i = \sigma(X_i^*)$ for all $i$. If $\gamma \ll \mu$, then $P|\mathcal{D}$ is atomic.*

*Proof.* Let $Q$ be a probability measure on $\mathcal{F}$ which makes $X_1, \ldots, X_k$ independent. Denote $M_Q$ the class of those probabilities $\mathbb{P}$ on $\mathcal{F}$ such that $\mathbb{P} \ll Q$ and

$$\mathcal{N}_{\mathbb{P}} = \{F \in \mathcal{F} : \mathbb{P}(F) \in \{0, 1\}\}, \quad \mathcal{D}_{\mathbb{P}} = \cap_{i=1}^k \sigma\big(\sigma(X_i^*) \cup \mathcal{N}_{\mathbb{P}}\big), \quad \text{with } \mathbb{P} \in M_Q.$$

Arguing by induction on $k$, we now prove that each $\mathbb{P} \in M_Q$ is atomic on $\mathcal{D}_{\mathbb{P}}$.

Let $k = 2$ and $\mathbb{P} \in M_Q$. Since $X_1^* = X_2$ and $X_2^* = X_1$, then $\mathbb{P}$ is atomic on $\mathcal{D}_{\mathbb{P}}$ by Theorem 4.

Given $k \geq 3$, define $V_i = (X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_{k-1})$. By induction, suppose that each $\mathbb{P} \in M_Q$ is atomic on

$$\mathcal{V}_{\mathbb{P}} = \cap_{i=1}^{k-1} \sigma\big(\sigma(V_i) \cup \mathcal{N}_{\mathbb{P}}\big).$$

We have to prove that each $\mathbb{P} \in M_Q$ is atomic on $\mathcal{D}_{\mathbb{P}}$. Accordingly, we fix $\mathbb{P} \in M_Q$ and a $\mathcal{D}_{\mathbb{P}}$-measurable function $U : \Omega \to \mathbb{R}$. Arguing as in the proof of Theorem 4, it suffices to show that $\mathbb{P}(U \in C) = 1$ for some countable set $C \subset \mathbb{R}$.

Since $\sigma(U) \subset \mathcal{D}_{\mathbb{P}}$ and $\mathcal{A}_i = \sigma(X_i^*)$, one obtains $\mathbb{P}\big(U = f_i(X_i^*)\big) = 1$, $i = 1, \ldots, k$, for some real measurable function $f_i$ on $(\prod_{j \neq i} \mathcal{X}_j, \prod_{j \neq i} \mathcal{B}_j)$. Let

$$A_x = \{f_1(V_1, x) = \ldots = f_{k-1}(V_{k-1}, x)\} \quad \text{for } x \in \mathcal{X}_k,$$
$$F(t, x) = Q\big(A_x \cap \{f_1(V_1, x) \leq t\}\big) \quad \text{for } t \in \mathbb{R} \text{ and } x \in \mathcal{X}_k.$$

Since $F(t, \cdot)$ is $\mathcal{B}_k$-measurable for fixed $t$, $F$ is a real cadlag process on the measurable space $(\mathcal{X}_k, \mathcal{B}_k)$. Let $J = \{(t, x) : F(t, x) > F(t-, x)\}$ be the jump set of $F$. By a well known result (see e.g. [9], Proposition 2.26), $J$ is contained in a countable union of graphs, that is,

$$J \subset \cup_n \{(g_n(x), x) : x \in \mathcal{X}_k\}$$

for suitable $\mathcal{B}_k$-measurable functions $g_n : \mathcal{X}_k \to \mathbb{R}$, $n = 1, 2, \ldots$.

Fix $x \in \mathcal{X}_k$ with $Q(A_x) > 0$ and define $Q_x(\cdot) = Q(\cdot \mid A_x)$. Then, $Q_x$ is atomic on $\mathcal{V}_{Q_x}$ (since $Q_x \in M_Q$) and $f_1(V_1, x)$ is $\mathcal{V}_{Q_x}$-measurable (since $Q_x(A_x) = 1$). Thus, $F(\cdot, x)$ is a purely jump function, that is, $Q\big(A_x \cap \{f_1(V_1, x) \notin J_x\}\big) = 0$ where $J_x = \{t : F(t, x) > F(t-, x)\}$. Integrating over $x$ yields

$$Q\big(f_1(X_1^*) = \ldots = f_{k-1}(X_{k-1}^*)\big)$$
$$= Q\big(f_1(X_1^*) = \ldots = f_{k-1}(X_{k-1}^*) = g_n(X_k) \text{ for some } n\big).$$

Let $C = \{c \in \mathbb{R} : Q(f_k(X_k^*) = c) > 0\}$. Since $f_k(X_k^*)$ and $g_n(X_k)$ are independent under $Q$, then $Q\big(f_k(X_k^*) \notin C,\, f_k(X_k^*) = g_n(X_k)\big) = 0$ for all $n$. Hence,

$$\begin{aligned} &Q\big(f_k(X_k^*) \notin C \text{ and } f_1(X_1^*) = \ldots = f_k(X_k^*)\big) \\ &\leq Q\big(f_k(X_k^*) \notin C \text{ and } f_k(X_k^*) = g_n(X_k) \text{ for some } n\big) = 0. \end{aligned}$$

Therefore, $\mathbb{P} \ll Q$ and $\mathbb{P}\big(U = f_i(X_i^*)\big) = 1$ for all $i = 1, \ldots, k$ imply

$$\mathbb{P}(U \in C) = 1 - \mathbb{P}\big(f_k(X_k^*) \notin C \text{ and } f_1(X_1^*) = \ldots = f_k(X_k^*)\big) = 1.$$

Since $C$ is countable, $\mathbb{P}$ is atomic on $\mathcal{D}_{\mathbb{P}}$. This concludes the induction argument and proves that each $\mathbb{P} \in M_Q$ is atomic on $\mathcal{D}_{\mathbb{P}}$.

Finally, to prove $P$ atomic on $\mathcal{D} = \mathcal{D}_P$, it can be assumed $(\Omega, \mathcal{F}, P) = (\mathcal{X}, \mathcal{B}, \gamma)$ and $X_1, \ldots, X_k$ the canonical projections. Also, since $\mu$ is a $\sigma$-finite product measure, $\mu$ is equivalent to some probability $Q$ on $\mathcal{B} = \mathcal{F}$ which makes $X_1, \ldots, X_k$ independent. Hence, $\gamma \ll \mu$ implies $P = \gamma \ll Q$. This concludes the proof. □

Note that Theorem 4 could be obtained as a corollary of previous Theorem 10. However, Theorem 4 has been stated as an autonomous result, since it is a useful preliminary step toward Theorem 10.

Finally, the scope of Theorems 8 and 10 can be enlarged via Lemma 6. Following this route, sometimes, the assumption $\gamma \ll \mu$ can be circumvented. Let $Z_i : \mathcal{X} \to \mathcal{X}_i$ denote the $i$-th canonical projection and $Z_i^* = (Z_1, \ldots, Z_{i-1}, Z_{i+1}, \ldots, Z_k)$. Moreover, suppose $\gamma$ is equivalent on rectangles to some probability $\nu$ on $\mathcal{B}$, i.e., $\gamma(A) = 0 \Leftrightarrow \nu(A) = 0$ for each set $A$ of the form $A = \{Z_1^* \in C_1, \ldots, Z_k^* \in C_k\}$ with $C_i \in \prod_{j \neq i} \mathcal{B}_j$ for all $i$. Then, $\mathcal{D}$ is atomic under $P$ provided $\nu \ll \mu$; cfr. Lemma 6 and Theorem 10. Or else, $\mathcal{D} = \mathcal{N}$ whenever $\nu \ll \mu$ and

$$\cup_{i=1}^k \{Z_i^* \in B_i^*\} \supset \{\frac{d\nu}{d\mu} > 0\} \supset \cap_{i=1}^k \{Z_i \in B_i\}$$

for some $B_1, \ldots, B_k$ such that $\nu\big(Z_1 \in B_1, \ldots, Z_k \in B_k\big) > 0$; cfr. Lemma 6 and Theorem 8. Note also that, for $k = 2$, equivalence on rectangles reduces to

$$\gamma(B_1 \times B_2) = 0 \Leftrightarrow \nu(B_1 \times B_2) = 0 \quad \text{whenever } B_1 \in \mathcal{B}_1,\, B_2 \in \mathcal{B}_2.$$

As an example (suggested by an anonymous referee) suppose $(X_n : n \geq 1)$ is an exchangeable sequence of real random variables with Ferguson-Dirichlet mixing measure. For $k = 2$, $(X_1, X_2)$ is distributed as

$$\gamma(B_1 \times B_2) = a\,\beta(B_1)\,\beta(B_2) + (1 - a)\,\beta(B_1 \cap B_2)$$

where $0 < a < 1$ and $\beta$ is a probability on the real Borel sets. Then $\mathcal{D} = \mathcal{N}$, as $\gamma$ is equivalent on rectangles to $\beta \times \beta$. However, if $\beta$ is nonatomic, $\gamma$ fails to be absolutely continuous with respect to any $\sigma$-finite product measure. It can be shown that, for every $k \geq 2$, one obtains $\mathcal{D} = \mathcal{N}$ for $(X_1, \ldots, X_k)$ as well.

*Acknowledgment:* This paper benefited from the helpful suggestions of an anonymous referee.

## References


[1] Berti P., Pratelli L., Rigo P. (2006) Asymptotic behaviour of the empirical process for exchangeable data, *Stoch. Proc. Appl.*, 116, 337-344.

[2] Berti P., Pratelli L., Rigo P. (2007) Skorohod representation on a given probability space, *Prob. Theo. Rel. Fields*, 137, 277-288.

[3] Berti P., Pratelli L., Rigo P. (2008) Trivial intersection of $\sigma$-fields and Gibbs sampling, *Ann. Probab.*, 36, 2215-2234.

[4] Bhaskara Rao K.P.S., Bhaskara Rao M. (1983) *Theory of charges*, Academic Press.

[5] Burkholder D.L., Chow Y.S. (1961) Iterates of conditional expectation operators, *Proc. Amer. Math. Soc.*, 12, 490-495.

[6] Delyon B., Delyon F. (1999) Generalization of von Neumann's spectral sets and integral representation of operators, *Bull. Soc. Math. Fr.*, 127, 25-41.

[7] Diaconis P., Khare K., Saloff-Coste L. (2007) Stochastic alternating projections, *Preprint*, Dept. of Statistics, Stanford University, currently available at: `http://www-stat.stanford.edu/~cgates/PERSI/papers/altproject-2.pdf`

[8] Halmos P.R. (1947) On the set of values of a finite measure, *Bull. Amer. Math. Soc.*, 53, 138-141.

[9] Karatzas I., Shreve S.E. (1991) *Brownian motion and stochastic calculus* (second edition), Springer.

[10] San Martin E., Mouchart M., Rolin J.M. (2005) Ignorable common information, null sets and Basu's first theorem, *Sankhya*, 67, 674-698.



Patrizia Berti, Dipartimento di Matematica Pura ed Applicata "G. Vitali", Universita' di Modena e Reggio-Emilia, via Campi 213/B, 41100 Modena, Italy
*E-mail address*: patrizia.berti@unimore.it

Luca Pratelli, Accademia Navale, viale Italia 72, 57100 Livorno, Italy
*E-mail address*: pratel@mail.dm.unipi.it

Pietro Rigo (corresponding author), Dipartimento di Economia Politica e Metodi Quantitativi, Universita' di Pavia, via S. Felice 5, 27100 Pavia, Italy
*E-mail address*: prigo@eco.unipv.it