Probab. Theory Relat. Fields **112**, 425–456 (1998)

Probability Theory and Related Fields

The almost equivalence of pairwise and mutual independence and the duality with exchangeability

Yeneng Sun[1,2]

[1] Department of Mathematics, National University of Singapore, Singapore 119260. e-mail: matsuny@math.nus.edu.sg
[2] Cowles Foundation, Yale University, New Haven, CT 06520, USA

Received: 16 April 1997 / Revised version: 18 May 1998

Abstract. For a large collection of random variables in an ideal setting, pairwise independence is shown to be almost equivalent to mutual independence. An asymptotic interpretation of this fact shows the equivalence of asymptotic pairwise independence and asymptotic mutual independence for a triangular array (or a sequence) of random variables. Similar equivalence is also presented for uncorrelatedness and orthogonality as well as for the constancy of joint moment functions and exchangeability. General unification of multiplicative properties for random variables are obtained. The duality between independence and exchangeability is established through the random variables and sample functions in a process. Implications in other areas are also discussed, which include a justification for the use of mutually independent random variables derived from sequential draws where the underlying population only satisfies a version of weak dependence. Macroscopic stability of some mass phenomena in economics is also characterized via almost mutual independence. It is also pointed out that the unit interval can be used to index random

The author is indebted to Donald Brown and Peter Loeb for stimulating discussions and help. He is also grateful to Donald Andrews, Andrew Barron, Martin Barlow, Xuming He, Ward Henson, Ali Khan, Heracles Polemarchakis, David Pollard, Qiman Shao, and Marten Wegkamp for useful questions and comments. This work was done while the author was visiting the Cowles Foundation for Research in Economics at Yale university in 1996-1997; the support from the Director, John Geanakoplos, and the Staff of the institute is gratefully acknowledged.

variables in the ideal setting, provided that it is endowed together with some sample space a suitable *larger* measure structure.

Mathematics Subject Classification (1991): Primary: 60G05, 60G09; Secondary: 28E05, 60G07, 90A46.

1. Introduction

Independence has long been a primary focus of probability theory, and is often considered the most important concept in the subject which helped foster an independent development of the theory of probability beyond its measure-theoretic context. The definition of independence is the abstraction of a highly intuitive and empirical notion, linking relevant mathematical theorems to experimentally observable results in the real world.

There are several versions of independence in the probabilistic literature, such as pairwise independence and various versions of independence involving a multiple number of random variables. For a finite collection of random variables, these notions are all very different. However, as demonstrated by the subject of probability theory itself, distinctive new properties arise whenever mass phenomena are studied. One might ask if a large number of random variables are considered, how much difference is still there among those different notions of independence?

The first aim of this paper is to show that all the notions of independence are, in fact, almost identical to their pairwise counterpart in an ideal setting. An asymptotic interpretation shows that in the usual sequential setting, these notions are asymptoticly equivalent even though they are still different. As a consequence of this type of equivalence result, we are also able to unify multiple versions of various, seemingly unrelated, multiplicative properties of random variables, such as those involving generating functions, characteristic functions, and maximum of random variables. To illustrate this general unification, consider a large collection of real-valued random variables; if for essentially every pair of random variables in the collection, the distribution function of the maximum of the pair is the product of the individual distribution functions of the two random variables, then for essentially every tuple of n random variables chosen from the original collection, the characteristic function of the sum of these n random variables is the product of the individual characteristic functions of the random variables in the tuple. This shows that some notions, though apparently having no relation at all in the finite case, are in fact essentially equivalent in the ideal setting. Note that the possibility of deriving results involving multiple versions of some notion from a pairwise counterpart is a major underlying theme of this paper. In the sequel, such a general principle will often be summarized as "*Two implies many*".

Another basic concept in probability theory is exchangeability. Our second aim is to show that the notions of independence and exchangeability are dual to each other in the sense that almost mutual independence (almost exchangeability) of the random variables in a process in an ideal setting is equivalent to almost exchangeability (almost mutual independence) of the sample functions of the process. Such a duality result can also be interpreted in the asymptotic setting by routine techniques. As another instance of the phenomenon "*Two implies many*", pairwise and multiple versions of exchangeability are also shown to be almost equivalent.

Since the continuum is commonly used to model a large number of entities, it is natural to explore the possibility of studying the various types of independence in the setting of a continuum of random variables. However, it is well known that the usual mathematical framework does not permit a meaningful study of a continuum of random variables with low intercorrelation (see [D], [FG]). In particular, it was pointed out by Doob (see [D], p.67) that if the random variables of a continuous parameter process are independent and have a common distribution (not concentrated at a single point), then the process is not jointly measurable and even has no measurable standard modification with respect to the relevant product measure. The following proposition provides a more general result in the same spirit.

Proposition 1.1. *Let* $(I, \mathscr{I}, \mu)$ *and* $(X, \mathscr{X}, \nu)$ *be any two probability spaces with a complete product probability space* $(I \times X, \mathscr{I} \otimes \mathscr{X}, \mu \otimes \nu)$, *and* f *a function from* $I \times X$ *to a separable metric space. If* f *is jointly measurable on the product probability space, and for* $\mu \otimes \mu$*-almost all* $(i_1, i_2) \in I \times I$, f_{i_1} *and* f_{i_2} *are independent* (*this condition is called almost sure pairwise independence*), *then, for* μ*-almost all* $i \in I$, f_i *is a constant random variable, where* f_i *is the function on* X *defined by* $f(i, \cdot)$.

Proof. We only prove the case that f is real-valued and bounded. As in the proof of Theorem 6.5 in [S2], the general case can be proven by using the compositions of the indicator functions of some open sets with f which still satisfy the condition of almost sure pairwise independence.

Assume that f is real-valued and bounded. Let A be any measurable set in I. By the Fubini Theorem, it is easy to establish the following identities:

$$\int_X \left[\int_A (f(i,x) - Ef_i)\, d\mu(i) \right]^2 d\nu(x)$$
$$= \int_X \int_A (f(i_1,x) - Ef_{i_1})\, d\mu(i_1) \int_A (f(i_2,x) - Ef_{i_2})\, d\mu(i_2)\, d\nu(x)$$
$$= \iint_{A\times A} \int_X (f(i_1,x) - Ef_{i_1})(f(i_2,x) - Ef_{i_2})\, d\nu(x)\, d\mu \otimes \mu(i_1,i_2) \ ,$$

which is zero by the condition of almost sure pairwise independence. Hence, for ν-almost all $x \in X$, $\int_A (f(i,x) - Ef_i)\, d\mu(i) = 0$.

Thus, for any measurable set B in X, $\iint_{A\times B}(f(i,x) - Ef_i)\, d\mu \otimes \nu = 0$. This means that the singed measure τ defined on $(I \times X, \mathscr{I} \otimes \mathscr{X})$ by integrating $f(i,x) - Ef_i$ on sets in $\mathscr{I} \otimes \mathscr{X}$ agrees with the zero measure on the rectangles. Note that the product algebra $\mathscr{I} \otimes \mathscr{X}$ is generated by all the rectangles $A \times B$, and the collection of rectangles is also closed under finite intersections, i.e., a π-system. By applying Dynkin's $\pi - \lambda$ theorem (see [C], p. 44 and [Du], p. 404), we obtain that the singed measure τ is equal to the zero measure. Thus, both $f(i,x) - Ef_i$ and 0 are Radon-Nikodym derivatives of the same measure. By the uniqueness of the Radon-Nikodym derivatives, we have $f(i,x) = Ef_i$ for $\mu \otimes \nu$-almost all $(i,x) \in I \times X$. Therefore, for μ-almost all $i \in I$, f_i is the constant random variable Ef_i. □

Note that if μ is atomless and the process f has mutually independent random variables, then the condition of almost sure pairwise independence is obviously satisfied. The previous result is still valid when μ has an atom A; one can simply observe that the almost sure pairwise independence condition implies the essential constancy of the random variables f_i for almost all $i \in A$. The above proposition simply says that no matter what kind of measure spaces are taken, independence and joint measurability with respect to the usual measure-theoretic product are never compatible with each other except for the trivial case. Thus, to study independence in a continuum setting, one *has to go beyond* the usual measure-theoretic framework.

The essential idea in the approach used here as well as in the earlier papers [S1] and [S2] (further results are in [S3]) is to use a larger measure-theoretic framework to conduct various simple measure-theoretic and probabilistic operations which are not applicable in the traditional framework due to the incompatibility of joint measurability and independence. Except for trivial cases, the processes considered here are measurable with respect to a σ-algebra (called Loeb product algebra) but not measurable with respect to the strictly smaller product σ-algebra in the usual sense. A key feature is the

Fubini property, as first shown by Keisler in [K1] (see also [AFHL], [Cu] and [L2]), for the bigger Loeb product algebra based on a type of measure spaces introduced by Loeb in [L1]. As shown in the appendix of [S2], there is no hope of developing a similar framework when Lebesgue spaces are used to index random variables in a process. On the other hand, the universality result in Theorem 6.2 of [S2] shows that one can construct processes on the relevant Loeb product algebra of *any* two atomless Loeb spaces whose random variables are almost surely pairwise independent and may take any variety of distributions in a well-defined sense. These processes are not jointly measurable in the usual sense as guaranteed by Proposition 1.1. As argued in [S1] and [S2], another significant advantage of using this larger framework is that the study of processes in the ideal setting is simply a way of studying general triangular arrays or sequences of random variables through the systematic applications of existing measure-theoretic techniques to a new setting. As demonstrated in this paper and in [S1] and [S2], distinctive new phenomena do arise naturally in this context.

The rest of the paper is organized as follows. Section 2 contains the main mathematical results of this paper. Some consequences and implications of the main results are discussed in Section 3. Based on some additional general results which are in a certain sense the best possible, the proofs of all the theorems as stated in Section 2 are given in Section 4. Section 5 concerns with asymptotic interpretation of the results in the ideal setting. Before moving to the next section, we note that if one prefers to use the unit interval $I = [0, 1]$ to index random variables in a process, then one can work with a measure μ on I induced by a bijection between I and a hyperfinite set in an ultrapower construction based on $\mathbb{N}$ (see [AFHL]). This new measure μ on I cannot be the Lebesgue measure. In fact, based on Theorem 7.16 in [S2] and Proposition 4.9 below it is easy to show that for a nontrival almost iid process, almost all the sample functions are μ-measurable but *not* Lebesgue measurable (for details, see [S5]). All the results and proofs in this paper can be reproduced in terms of the new measure μ, provided that the unit interval I with measure μ is to be endowed together with some sample space a suitable *larger* product measure structure as above. The point is that the particular choice of a parameter space itself is not an issue; what is really relevant is the associated measure structure.

2. The main results

In this section, we present four different instances of the phenomenon "*Two implies many*" through the notions of uncorrelatedness, constancy of joint moment functions, independence and exchangeability. In addition, we show that exchangeability is, in fact, the dual notion of independence. The proofs of these results are given in Section 4.

We shall now fix some notation. Let T be a hyperfinite set, $\mathscr{T}$ the internal algebra of all internal subsets of T, and λ an internal finitely additive probability measure on $(T, \mathscr{T})$. Let $(T, L(\mathscr{T}), L(\lambda))$ be the standardization, i.e., the Loeb space, formed from $(T, \mathscr{T}, \lambda)$; this standard probability space will be the hyperfinite parameter space for the processes to be considered here. Starting with another internal probability space $(\Omega, \mathscr{A}, P)$, we let the Loeb space $(\Omega, L(\mathscr{A}), L(P))$ be our sample probability space. For basic properties of Loeb spaces, see [L1], [AFHL] and [HL].

Note that the internal product space $(T \times \Omega, \mathscr{T} \otimes \mathscr{A}, \lambda \otimes P)$ is also an internal probability space. The corresponding Loeb space is denoted by $(T \times \Omega, L(\mathscr{T} \otimes \mathscr{A}), L(\lambda \otimes P))$; it will be referred to as the Loeb product space. Similarly, for a positive integer n, let $(T^n, \mathscr{T}^n, \lambda^n)$ and $(\Omega^n, \mathscr{A}^n, P^n)$ be the n-fold internal product spaces of $(T, \mathscr{T}, \lambda)$ and $(\Omega, \mathscr{A}, P)$ respectively; the associated n-fold Loeb product spaces are denoted by $(T^n, L(\mathscr{T}^n), L(\lambda^n))$ and $(\Omega^n, L(\mathscr{A}^n), L(P^n))$. Other types of product spaces for Loeb spaces will appear in Section 4. As usual, a measurable function of two variables is called a process. Given a process f, for each $t \in T$, and $\omega \in \Omega$, f_t denotes the function $f(t, \cdot)$ on Ω and f_ω denotes the function $f(\cdot, \omega)$ on T. Since the Fubuni type property is satisfied by internal processes on the internal product space $(T \times \Omega, \mathscr{T} \otimes \mathscr{A}, \lambda \otimes P)$, one can also obtain the same type of property for processes on the corresponding Loeb product space. The latter Fubini Theorem is often called the Fubini Theorem for Loeb measures or Keisler's Fubini Theorem (see [AFHL], [Cu], [K1] and [L2]). The functions f_t are usually called the random variables of the process f, while the f_ω form the sample functions of the process. Note that the measurability of f_t and f_ω is guaranteed by the relevant Fubini Theorem.

The following theorem shows that the usual notion of uncorrelatedness can be used to deduce its multiple versions for a hyperfinite collection of random variables. The result becomes trivial in the usual finite setting since the almost sure uncorrelatedness condition implies the relevant random variables to be essentially constant.

Theorem 1. *Let f be a real-valued process on $(T \times \Omega, L(\mathscr{T} \otimes \mathscr{A}), L(\lambda \otimes P))$ with a finite m-th moment for some $m \geq 2$, i.e., $|f|^m$ is $L(\lambda \otimes P)$-integrable. If the random variables f_t are almost surely uncorrelated, i.e., for $L(\lambda \otimes \lambda)$-almost all $(t_1, t_2) \in T \times T$, f_{t_1} and f_{t_2} are uncorrelated, then for each $2 \leq n \leq m$, the random variables f_t are almost surely uncorrelated in n-tuple, i.e., for $L(\lambda^n)$-almost all $(t_1, t_2, \ldots, t_n) \in T^n$,*

$$E(f_{t_1}f_{t_2}\dots f_{t_n}) = Ef_{t_1}Ef_{t_2}\dots Ef_{t_n} \ .$$

A classical example of Bernstein shows that there are three events which are pairwise independent but not mutually independent (see [F], p. 126). It is also well known that there are m non-independent events among which any $m-1$ of them are mutually independent (see, for example, [WSS]). The above theorem shows that if one considers a hyperfinite collection $\{C_t : t \in T\}$ of events, then almost sure pairwise independence implies mutual independence for almost all n- tuples of events $C_{t_1}, C_{t_2}, \dots, C_{t_n}$, where $C \in L(\mathscr{T} \otimes \mathscr{A})$ and n is any positive integer greater than two. One can simply take f to be the indicator function χ_C of C and observe that $E(f_{t_1}f_{t_2}\dots f_{t_n}) = L(P)(C_{t_1} \cap C_{t_2} \cap \dots \cap C_{t_n})$ and $E(f_{t_i}) = L(P)(C_{t_i})$.

For a real-valued process f, one often calls the second joint moment function $Ef_{t_1}f_{t_2}$ the autocorrelation function of the process. The next theorem says that the essential constancy of the second joint moment function implies that of higher order joint moment functions.

Theorem 2. *Let f be a real-valued process on $(T \times \Omega, L(\mathscr{T} \otimes \mathscr{A}), L(\lambda \otimes P))$ with a finite m-th moment for some $m \geq 2$. If the autocorrelation function of the random variables f_t is essentially constant, then for each $1 \leq n \leq m$, the n-th joint moment of random variables f_t are essentially the constant $\int_\Omega (Ef_\omega)^n \, dL(P)(\omega)$, i.e., for $L(\lambda^n)$-almost all $(t_1, t_2, \dots, t_n) \in T^n$, $Ef_{t_1}f_{t_2}\dots f_{t_n} = \int_\Omega (Ef_\omega)^n \, dL(P)(\omega)$. In particular, the covariance function $cov(f_{t_1}, f_{t_2})$ is essentially equal to the nonnegative constant $\int_\Omega (Ef_\omega)^2 \, dL(P)(\omega) - \left(\int_\Omega Ef_\omega \, dL(P)(\omega)\right)^2$.*

Theorem 1 covers the almost equivalence of pairwise and multiple versions of independence for events. We shall now move to the case of random variables. Note that pairwise independence is also weaker than mutual independence for a finite collection of random variables (see [F], p. 220).

Theorem 3. *Let f be a process from $(T \times \Omega, L(\mathscr{T} \otimes \mathscr{A}), L(\lambda \otimes P))$ to a separable metric space X and n be an integer greater than or equal to two. Then the following are equivalent:*

(1) *the random variables f_t are almost surely pairwise independent, i.e., for $L(\lambda \otimes \lambda)$-almost all $(t_1, t_2) \in T \times T$, f_{t_1} and f_{t_2} are independent;*
(2) *the random variables f_t are almost surely independent in n-tuple, i.e., $f_{t_1}, f_{t_2}, \dots, f_{t_n}$ are mutually independent for $L(\lambda^n)$-almost all $(t_1, t_2, \dots, t_n) \in T^n$.*

The idea of exchangeability has a wide range of applications in both pure and applied probability (see, for example, [A], [Bi], [CT] and

[K]), which is commonly studied in the sequential setting. A sequence $\mathscr{C}$ of random variables is said to be exchangeable if for any $n \geq 1$, the joint distribution of any n random variables from $\mathscr{C}$ depends only on n but not on the particular choice and order of these n random variables. As for the case of independence, here we also consider the almost equivalence of pairwise and multiple versions of exchangeability in the ideal setting.

Theorem 4. *Let f be a process from $(T \times \Omega, L(\mathscr{T} \otimes \mathscr{A}), L(\lambda \otimes P))$ to a separable metric space X and n be an integer greater than or equal to two. Then the following are equivalent*:

(1) *the random variables f_t are almost surely pairwise exchangeable, i.e., there is a distribution v on $X \times X$ such that for $L(\lambda \otimes \lambda)$-almost all $(t_1, t_2) \in T \times T$, f_{t_1} and f_{t_2} have a joint distribution v*;

(2) *the random variables f_t are almost surely exchangeable in n-tuple, i.e., there is a distribution v_n on X^n such that the joint distribution of* $f_{t_1}, f_{t_2}, \dots, f_{t_n}$ is v_n *for $L(\lambda^n)$-almost all $(t_1, t_2, \dots, t_n) \in T^n$. In addition, the v_n must have the form*

$$v_n(C_1 \times C_2 \times \cdots \times C_n) = \int_{\Omega} \tau_\omega(C_1) \times \tau_\omega(C_2) \times \cdots \times \tau_\omega(C_n)\, dL(P)(\omega)$$

for all Borel sets $C_1, C_2, \dots, C_n$ in X, where τ_ω is the distribution on X induced by f_ω. Moreover, for $L(\lambda)$-almost all $t \in T$, the distribution of f_t is simply the distribution induced by f viewed as a random variable on the Loeb product space.

The notion of exchangeability has already been linked to conditional independence through the classical de Finetti theorem (see [CT], p. 222, and [K]). The final result of this section relates exchangeability to *unconditional* independence. It shows that the two notions are in fact dual to each other in the sense that almost exchangeability of the random variables in a hyperfinite process is equivalent to almost independence of the sample functions of the process in corresponding settings. One can also view this type of duality the other way around by interchanging the index and sample variables.

Theorem 5. *Let f be a process from $(T \times \Omega, L(\mathscr{T} \otimes \mathscr{A}), L(\lambda \otimes P))$ to a separable metric space X and n be an integer greater than or equal to two. Then the following are equivalent*:

(1) *the random variables f_t are almost surely exchangeable in n-tuple*;

(2) *the sample functions f_ω, as random variables on $(T, L(\mathscr{T}), L(\lambda))$, are almost surely independent in n-tuple.*

We say a collection of random variables *almost mutually independent* (*almost exchangeable*) if they are almost surely independent (exchangeable) in n-tuple for all $n \geq 2$. Thus, Theorem 3 shows the equivalence of almost mutual independence and almost sure pairwise independence, while Theorem 5 reveals the duality of mutual independence and exchangeability in the almost sense.

3. Some implications of the main results

We first observe that some additional characterizations can be easily obtained for those processes satisfying various versions of consistent law of large numbers (or simply the consistency law) as studied in [S1] and [S2]. This law can be viewed as a formal version of the intuitive observation, as characterized with the aphorism "*No betting system can beat the house*", which simply means that a gambler cannot change the expectation of his return by betting at a particular subsequence. It is also documented in [FG] that if the special case of a continuum of iid random variables is used to model individual risks, then one should require sample averages of any nonnegligible subcollection of the random variables to be constant, though this is shown in [FG] to be impossible in the usual setting. Such type of condition which requires stability not only for a whole system but also for the large subsystems will be referred to as *macroscopic stability*.

For a formal definition of the consistency law in the setting of sample means, see Definition 1 in [S1]. Theorem 2 in [S1] (see also Theorem 4.6 in [S2]) characterizes those square integrable processes satisfying the consistency law by almost sure uncorrelatedness. Thus, by that characterization and Theorem 1 here, it is easy to see that multiple versions of uncorrelatedness are not only sufficient but also necessary for the satisfiability of the consistency law.

Theorem 4 in [S1] (see also Theorem 7.6 in [S2]) says that almost sure pairwise independence is necessary and sufficient for the satisfiability of the consistency law in distribution. It also includes the unexpected result that the almost sure versions of various multiplicative properties in the pairwise setting are all equivalent to almost sure pairwise independence. Since Theorem 3 here shows the almost equivalence of pairwise and multiple versions of independence, it is clear that the multiple versions of various multiplicative properties are also equivalent to each other as well as to the satisfiability of the consistency law in distribution (for a formal definition of this concept, see Definition 2 in [S1]) as indicated by the following proposition. This proposition relies on the notion of a separating class for some distributions whose formal definition in our context can be found in Definition 3 in [S1].

Proposition 3.1. *Let $\mathscr{E}$ be a class of real or complex valued Borel functions on a separable metric space X, and f a process from $(T \times \Omega, L(\mathscr{T} \otimes \mathscr{A}), L(\lambda \otimes P))$ to X. Assume that $\mathscr{E}$ is a separating class for all the distributions induced by the sample functions f^A_ω as well as the distribution of the process f^A viewed as a random variable on $A \times \Omega$, where $A \in \mathscr{T}$ with $L(\lambda)(A) > 0$, and f^A is the restriction of f to $A \times \Omega$ with the rescaled Loeb product measure. Let m be a positive integer greater than or equal to two. Assume that $\varphi(f)$ is a process with a finite m-th moment for each $\varphi \in \mathscr{E}$. Then the following are equivalent:*

(1) *f satisfies the consistency law in distribution;*
(2) *the random variables f_t are almost mutually independent;*
(3) *for each $\varphi \in \mathscr{E}$, the $(\varphi(f))_t$ are almost surely uncorrelated, i.e., for $L(\lambda \otimes \lambda)$-almost all $(t_1, t_2) \in T \times T$, $E((\varphi(f))_{t_1} \overline{(\varphi(f))_{t_2}}) = E(\varphi(f))_{t_1} E\overline{(\varphi(f))_{t_2}}$;*
(4) *if the functions in $\mathscr{E}$ are all real-valued, then for $L(\lambda^n)$-almost all $(t_1, t_2, \ldots, t_n) \in T^n$,*
$$E(\varphi(f_{t_1})\varphi(f_{t_2}) \ldots \varphi(f_{t_n})) = E\varphi(f_{t_1})E\varphi(f_{t_2}) \ldots E\varphi(f_{t_n})$$
holds for all $\varphi \in \mathscr{E}$, where n is any integer between 2 and m;
(5) *if the functions in $\mathscr{E}$ are complex-valued, then for $L(\lambda^n)$-almost all $(t_1, t_2, \ldots, t_n) \in T^n$,*
$$E\big(\varphi^{i_1}(f_{t_1})\varphi^{i_2}(f_{t_2}) \ldots \varphi^{i_n}(f_{t_n})\big) = E\varphi^{i_1}(f_{t_1})E\varphi^{i_2}(f_{t_2}) \ldots E\varphi^{i_n}(f_{t_n})$$
holds for all $\varphi \in \mathscr{E}$, where n is any integer between 2 and *m, $i_1, i_2, \ldots, i_n = 0$ or 1, and $\varphi^0 = \varphi, \varphi^1 = \bar{\varphi}$, the complex conjugate of φ.*

To see the above general unification of multiplicative properties in more concrete situations, we note that both the complex exponentials e^{iux} and the indicator functions of the intervals $(-\infty, u]$ form separating classes for all distributions on $\mathbb{R}$, and the class of functions $\{z^x : z \in (-1, 1)\}$ is separating for distributions of random variables taking values in natural numbers (see Section 7.4 in [S2]). Applying the above proposition to these separating classes leads to the general equivalence of independence and the presence of multiplicative properties involving characteristic functions, maximum of random variables, and generating functions, which include the type of equivalence involving maximum of random variables in pairs and product of characteristic functions in n-tuples, as discussed in the third paragraph of Section 1.

Many economic models consider the case where individual agents face idiosyncratic risks, i.e., risks of gains or losses that are non-negligible on the individual level but can be exactly predicted in the aggregate on the macroscopic level (see [A1], [FG] and [S2] for many references). Then, a natural question arises: how to characterize uncertainty or risks on the individual level so that there is no randomness from macroscopic point of view. Since the satisfiability of consistency law, as a formalization of the concept of macroscopic stability, is characterized by almost mutual independence or uncorrelatedness, the answer to this question is thus obvious. Note that the phenomenon of macroscopic stability is observable while the independence condition is only a theoretical assumption based on intuition. It is interesting that equivalence can still be established between them. One may view from the fact that large scale insurance systems usually do not fail to claim that the insured risks must satisfy a sort of mutual independence or uncorrelatedness conditions.

Note that even though some models may only require the validity of law of large numbers (in an ideal or a discrete setting) explicitly for a whole system, the sufficient conditions used to imply the law usually also apply to the subsystems, and thus the consistency law is, in fact, satisfied implicitly. For example, the usual mixing conditions (see [Bi]) are such sufficient conditions. One can observe that the indexes among the terms in a subsequence of a sequence of mixing random variables are in some sense further away than those in the original sequence, and thus the subsequence must also satisfy the mixing conditions. Thus, *almost mutual independence* and almost sure uncorrelatedness (or their equivalent notions) may remain to be the best general conditions to ensure the stability of a whole system. For some applications of the exact law to particular economic contexts, see [KS] and [S4].

Proposition 3.1 shows how to characterize almost mutual independence by using characteristic functions, maximum of random variables, and generating functions. The following proposition provides an analog for the case of exchangeability.

Proposition 3.2. *Let $\mathscr{D}$ be a collection of distributions on a separable metric space X, $\mathscr{E}$ a separating class for $\mathscr{D}$ of real or complex valued Borel functions, and f a process from $(T \times \Omega, L(\mathscr{T} \otimes \mathscr{A}), L(\lambda \otimes P))$ to X such that for all $B \in \mathscr{A}$ with $L(P)(B) > 0$, the distributions of the random variables f_t^B as well as the distribution of the process f^B viewed as a random variable on $T \times B$ are all in $\mathscr{D}$, where f^B is the restriction of f to $T \times B$ with the rescaled Loeb product measure. Let m be a positive* *integer greater than or equal to two. Assume that $\varphi(f)$ has a finite m-th moment for each $\varphi \in \mathscr{E}$. Then the following are equivalent*:

(1) *the random variables f_t are almost surely pairwise exchangeable*;
(2) *for any $n \geq 2$, the random variables f_t are almost surely exchangeable in n-tuple*;
(3) *for each $\varphi \in \mathscr{E}$, $E((\varphi(f))_{t_1}\overline{(\varphi(f))_{t_2}}) = E|E\varphi(f_\omega)|^2$ for $L(\lambda \otimes \lambda)$-almost all $(t_1, t_2) \in T \times T$*;
(4) *if the functions in $\mathscr{E}$ are all real-valued, then, for $L(\lambda^n)$-almost all $(t_1, t_2, \ldots, t_n) \in T^n$,*
$$E(\varphi(f_{t_1})\varphi(f_{t_2})\ldots\varphi(f_{t_n})) = E[E\varphi(f_\omega)]^n$$
holds for all $\varphi \in \mathscr{E}$, where n is any integer between 2 and m;
(5) *if the functions in $\mathscr{E}$ are complex-valued, then, for $L(\lambda^n)$-almost all $(t_1, t_2, \ldots, t_n) \in T^n$,*
$$E\big(\varphi^{i_1}(f_{t_1})\varphi^{i_1}(f_{t_2})\ldots\varphi^{i_n}(f_{t_n})\big) = \int_\Omega E\varphi^{i_1}(f_\omega)E\varphi^{i_2}(f_\omega)\ldots E\varphi^{i_n}(f_\omega)\,dL(P)(\omega)$$
holds for all $\varphi \in \mathscr{E}$, where n is any integer between 2 and m, $i_1, \ldots, i_n = 0$ or 1, and $\varphi^0 = \varphi, \varphi^1 = \bar{\varphi}$, the complex conjugate of φ.

Proof. (1) $\Longleftrightarrow$ (2) is shown in Theorem 4. By Theorem 5, (1) is equivalent to the almost sure pairwise independence of the sample functions f_ω. By interchanging the index and sample variables in (2) and (4) of Theorem 7.6 in [S2], we know that the almost sure pairwise independence of f_ω is equivalent to (3). Thus (1), (2) and (3) are all equivalent.

The implications (4) $\Longrightarrow$ (3) and (5) $\Longrightarrow$ (3) are clear. (3) $\Longrightarrow$ (4) follows from Theorem 2 by regrouping countably many null sets together. Finally, if we assume (2), then the essential constancy of the n-th order joint distributions implies the essential constancy of the function $E(\varphi^{i_1}(f_{t_1})\varphi^{i_1}(f_{t_2})\ldots\varphi^{i_n}(f_{t_n}))$; by integrating the function with respect to $t_1, t_2, \ldots, t_n$ and changing the relevant iterated integrals, (5) can be obtained. One can also obtain (5) by using the formula for v_n in Theorem 4 (2). □

To understand the above proposition in a more specific setting, consider a natural number valued process f on the Loeb product space. Assume that the characteristic function of the difference $f_{t_1} - f_{t_2}$ is essentially independent of the choices of t_1 and t_2. Then the generating function of the sum $f_{t_1} + f_{t_2} + \cdots + f_{t_n}$ is also essentially unrelated to particular choices of $t_1, t_2, \ldots, t_n$, where n is any positive integer. This is also a version of "*Two implies many*" relating different notions.

The notion of weak dependence is often used to mean that if any random variable in a given collection (a sequence or a triangular array) of random variables is approximately independent in some sense to most other random variables in the collection, then this collection is said to be weakly dependent. In our idealized setting, this notion simply means that any random variable is independent of others outside a negligible set, which is precisely the notion of almost sure pairwise independence. It is easy to check that the transferred version of the pervasive mixing conditions (see Section 20, [Bi]) does lead to almost sure pairwise independence, which also means that these mixing conditions are indeed stronger than the asymptotic version of almost sure pairwise independence. Now consider a large population modelled by a hyperfinite process with weakly dependent random variables in the idealized sense. Then, by randomly drawing a sequence of random variables from the underlying hyperfinite population, one can certainly expect to obtain a pairwise independent sequence since the underlying population are almost so. However, it is rather surprising that the resulting sequence is, in fact, *mutually independent* (see Proposition 3.4 below). That is, sequential draws derive *mutual independence* from a version of weak dependence.

To give a rigorous formulation of the type of result in Proposition 3.4, we need a suitable σ-algebra on the countable product T^{∞} together with a measure. We use $\mathscr{T}_m$ to denote the collection of all subsets of T^{∞} in the form $A_m \times T^{\infty}$, where A_m is some internal set in the internal product algebra $\mathscr{T}^m$. Define a set function $L(\lambda_m)$ on $\mathscr{T}_m$ by letting $L(\lambda_m)(A_m \times T^{\infty}) = L(\lambda^m)(A_m)$. Let $\mathscr{T}_{\infty}$ be the union of all the $\mathscr{T}_m$ and $L(\lambda_{\infty})$ the set function on $\mathscr{T}_{\infty}$ such that $L(\lambda_{\infty})(C) = L(\lambda_m)(C)$ if $C \in \mathscr{T}_m$. By the Fubini property, $L(\lambda_{\infty})$ is a well defined finitely additive measure on the algebra $\mathscr{T}_{\infty}$.

The following proposition shows that $(T^{\infty}, \mathscr{T}_{\infty}, L(\lambda_{\infty}))$ can be extended to a countably additive complete measure space. Note that $\mathscr{T}_{\infty}$ itself is not internal, and the usual result on Loeb extension (see [L1] and [AFHL]) thus does not cover this case.

Proposition 3.3. *The finitely additive measure space* $(T^{\infty}, \mathscr{T}_{\infty}, L(\lambda_{\infty}))$ *can be extended to a countably additive complete measure space* $(T^{\infty}, L(\mathscr{T}_{\infty}), L(\lambda_{\infty}))$.

Proof. Let $\{C_n\}_{n=1}^{\infty}$ be a decreasing sequence of sets in $\mathscr{T}_{\infty}$ with empty intersection. By the construction of $\mathscr{T}_{\infty}$, one can find a sequence of internal sets $\{A_n\}_{n=1}^{\infty}$ and a non-decreasing sequence $\{m_n\}_{n=1}^{\infty}$ of positive integers such that $C_n = A_n \times T^{\infty}$ and $A_n \in \mathscr{T}^{m_n}$. For $\ell \leq n$, let π_{ℓ}^n be the mapping from T^{m_n} to $T^{m_{\ell}}$ by projecting a tuple in T^{m_n} to its first

m_ℓ coordinates; then $\pi^n_\ell(A_n) \subseteq A_\ell$. Take the transfer $\{m_n\}_{n\in {}^*\mathbb{N}}$ of the sequence $\{m_n\}_{n=1}^\infty$ and an internal extension $\{A_n\}_{n\in {}^*\mathbb{N}}$ of the sequence $\{A_n\}_{n=1}^\infty$ of internal sets. By spillover and $\aleph_1$-saturation, one can obtain $h \in {}^*\mathbb{N}_\infty$ such that for all $n \leq h$, $A_n \in \mathscr{T}^{m_n}$ and $\pi^n_\ell(A_n) \subseteq A_\ell$ for all $\ell \in \mathbb{N}$, where π^n_ℓ is defined in exactly the same way as in the case when n is finite.

We claim that $A_n = \emptyset$ for all $n \in {}^*\mathbb{N}_\infty$ with $n \leq h$; if not, one can find such an n with $(t_1, t_2, \ldots, t_{m_n}) \in A_n$. Then $(t_1, t_2, \ldots, t_{m_\ell}) \in A_\ell$ for any $\ell \in \mathbb{N}$. If $m_n \in {}^*\mathbb{N}_\infty$, then it is obvious that $\{t_p\}_{p=1}^\infty$ is in C_ℓ for all $\ell \in \mathbb{N}$, which contradicts the assumption that the intersection of all the C_ℓ is empty. If $m_n \in \mathbb{N}$, one can choose t_p arbitrarily for any $p > m_n$ to obtain the same contradiction. Hence the claim is proven.

By spillover, we know that for some $n \in \mathbb{N}$, $A_n = \emptyset$, and so is C_n. Thus, we obtain a trivial limit, $\lim_{n\to\infty} L(\lambda_\infty)(C_n) = 0$. This means that $L(\lambda_\infty)$ is indeed countably additive on $\mathscr{T}_\infty$. As in [L1], the Caratheodory extension theorem implies that $L(\lambda_\infty)$ can be extended to the σ-algebra $\sigma(\mathscr{T}_\infty)$ generated by $\mathscr{T}_\infty$. Let $(T^\infty, L(\mathscr{T}_\infty), L(\lambda_\infty))$ be the measure space obtained by completing the measure space $(T^\infty, \sigma(\mathscr{T}_\infty), L(\lambda_\infty))$, and we are done. □

We are now ready to present Proposition 3.4.

Proposition 3.4. *Let f be a process from $(T \times \Omega, L(\mathscr{T} \otimes \mathscr{A}), L(\lambda \otimes P))$ to a separable metric space X. If the random variables f_t are almost surely pairwise independent, then for $L(\lambda_\infty)$*-almost all $(t_1, t_2, \ldots, t_n, \ldots) \in T^\infty$, *the sequence $\{f_{t_n}\}_{n=1}^\infty$ of random variables are mutually independent.*

Proof. For each $n \geq 2$, let A_n be the collection of all the $(t_1, t_2, \ldots, t_n) \in T^n$ such that $f_{t_1}, f_{t_2}, \ldots, f_{t_n}$ are mutually independent. Then $L(\lambda^n)(A_n) = 1$ by Theorem 3. Let $C_n = A_n \times T^\infty$ and $C = \cap_{n=1}^\infty C_n$. By the fact that one can find an internal set whose symmetric difference with A_n is null, we can obtain that $C_n \in L(\mathscr{T}_\infty)$ with $L(\lambda_\infty)(C_n) = 1$. Hence $C \in L(\mathscr{T}_\infty)$ and $L(\lambda_\infty)(C) = 1$. It is clear that for any $(t_1, t_2, \ldots, t_n, \ldots) \in C$, the random variables in the sequence $\{f_{t_n}\}_{n=1}^\infty$ are mutually independent. □

The following is an analog of Proposition 3.4 in the setting of exchangeability.

Proposition 3.5. *Let f be a process from $(T \times \Omega, L(\mathscr{T} \otimes \mathscr{A}), L(\lambda \otimes P))$ to a separable metric space X. If the random variables f_t are almost surely* *pairwise exchangeable, then for* $L(\lambda_\infty)$-almost all $(t_1, t_2, \ldots, t_n, \ldots) \in T^\infty$, *the sequence* $\{f_{t_n}\}_{n=1}^\infty$ *is exchangeable.*

4. Proof of the theorems and additional results

First note that we allow both the index set and sample space of a process to be possibly non-hyperfinite in this section. The main reason for this is to work with the transfer of a common sample space (which is usually not hyperfinite) of some triangular array of random variables in Section 5. Since the Fubini type result holds no matter the sets are hyperfinite or not, there is no additional burden at all by relaxing the hyperfiniteness restriction on the index sets also. This allows us to give the index and sample variables symmetric treatments in some situations. We shall also consider processes with the same sample space but possibly different index sets or with the same index set but possibly different sample spaces. Corresponding Loeb product spaces are also defined accordingly. In this section, we shall formulate and prove some additional results which are then used to prove the theorems in Section 2.

To prove Proposition 4.2 below, we need the following technical lemma on the integrability of some relevant functions. The proof is based on the Tonelli theorem for Loeb measures (see [HL], p. 204). Similar integrability problems will arise quite often in the proof of other results. They can usually be solved by using the same idea and the proof will be omitted. It is important to note that the f^i in Lemma 4.1 and other places is a process itself, not the i-th power of a process f.

Lemma 4.1. L*et* m *be a positive integer greater than or equal to* 2. *For each* $i = 1, 2, \ldots, m$, let f^i *be a real-valued process on a Loeb product space* $(T_i \times \Omega, L(\mathcal{T}_i \otimes \mathcal{A}), L(\lambda_i \otimes P))$ *with a finite m-th moment*; *that is,* $\int_{T_i \times \Omega} |f^i(t, \omega)|^m \, dL(\lambda_i \otimes P) < \infty$. *Then*

(1) *the function* G *on* $\left(\Pi_{i=1}^m T_i\right) \times \Omega^m$ *defined by* $G(t_1, \ldots, t_m, \omega_1, \ldots, \omega_m) = \Pi_{i=1}^m \Pi_{j=1}^m f^i(t_i, \omega_j)$ *is* $L\left(\left(\Pi_{i=1}^m \lambda_i\right) \otimes P^m\right)$*-integrable*;
(2) *for* $L\left(\Pi_{i=1}^m \lambda_i\right)$*-almost all* $(t_1, \ldots, t_m) \in \Pi_{i=1}^m T_i$, $f_{t_1}^1 f_{t_2}^2 \ldots f_{t_m}^m$ *is integrable over* $(\Omega, L(\mathcal{A}), L(P))$;
(3) *the function* $E(f_{t_1}^1 f_{t_2}^2 \ldots f_{t_m}^m)$ *on* $\Pi_{i=1}^m T_i$ *has a finite m-th moment with respect to the measure* $L\left(\Pi_{i=1}^m \lambda_i\right)$.

Proof. The Tonelli theorem for Loeb measures implies the following:

$$\int_{(\Pi_{i=1}^m T_i)\times\Omega^m} |G(t_1,\ldots,t_m,\omega_1,\ldots,\omega_m)|\, dL\big(\big(\Pi_{i=1}^m \lambda_i\big)\otimes P^m\big)$$

$$= \int_{\Pi_{i=1}^m T_i} \Pi_{j=1}^m \int_\Omega \Pi_{i=1}^m |f^i(t_i,\omega_j)|\, dL(P)(\omega_j)\, dL(\Pi_{i=1}^m \lambda_i)$$

$$= \int_{\Pi_{i=1}^m T_i} (E|f_{t_1}^1 f_{t_2}^2 \ldots f_{t_m}^m|)^m\, dL(\Pi_{i=1}^m \lambda_i)$$

$$\leq \int_{\Pi_{i=1}^m T_i} \Pi_{i=1}^m \int_\Omega |f^i(t_i,\omega)|^m\, dL(P)(\omega)\, dL(\Pi_{i=1}^m \lambda_i)$$

$$= \Pi_{i=1}^m \int_{T_i\times\Omega} |f^i(t_i,\omega)|^m\, dL(\lambda_i\otimes P) < \infty \ ,$$

where the inequality is obtained by applying the usual Hölder inequality (see, for example, [Lo], p. 158) repeatedly. The rest of the proof is clear. □

The following proposition shows that for a given finite collection of processes with possibly different index sets, if the second joint moment function of the sample functions is essentially constant for each process (except one of them), then the processes are uncorrelated in m-tuple in the sense specified below. Note that in order to work with the multiple versions of independence in Theorem 3, one has to take the inverse images of *different* sets in the target space of a process. The resulting processes are different, even though they are constructed from the same source process. Thus, it is necessary to work with a finite collection of different processes.

Proposition 4.2. *Let $m \geq 2$ be a positive integer. For each $i = 1, 2, \ldots, m$, let f^i be a real-valued process on a Loeb product space $(T_i \times \Omega, L(\mathcal{T}_i \otimes \mathcal{A}), L(\lambda_i \otimes P))$ with a finite m-th moment. If for each $1 \leq i \leq m-1$, the autocorrelation function of the sample functions f_ω^i is essentially constant on $\Omega\times\Omega$, then for $L(\Pi_{i=1}^m \lambda_i)$-almost all $(t_1, t_2, \ldots, t_m) \in \Pi_{i=1}^m T_i$, $f_{t_1}^1, f_{t_2}^2, \ldots, f_{t_m}^m$ are uncorrelated in m-tuple, i.e., $E(f_{t_1}^1 f_{t_2}^2 \ldots f_{t_m}^m) = Ef_{t_1}^1 Ef_{t_2}^2 \ldots Ef_{t_m}^m$.*

Proof. Lemma 4.1 implies that the function $Ef_{t_1}^1 f_{t_2}^2 \ldots f_{t_m}^m$ on $\Pi_{i=1}^m T_i$ has a finite m-th moment, and hence a finite second moment with respect to the measure $L(\Pi_{i=1}^m \lambda_i)$. By Keisler's Fubini theorem for Loeb measures, we can obtain

$$
\begin{aligned}
I = &\int_{\Pi_{i=1}^m T_i} (Ef^1_{t_1} f^2_{t_2} \dots f^m_{t_m} - Ef^1_{t_1} Ef^2_{t_2} \dots Ef^m_{t_m})^2 \, dL(\Pi_{i=1}^m \lambda_i) \\
= &\int_{\Pi_{i=1}^m T_i} \int_\Omega \Pi_{i=1}^m f^i(t_i, \omega_1) \, dL(P)(\omega_1) \\
&\quad \times \int_\Omega \Pi_{i=1}^m f^i(t_i, \omega_2) \, dL(P)(\omega_2) \, dL(\Pi_{i=1}^m \lambda_i) \\
&\quad - 2 \int_{\Pi_{i=1}^m T_i} Ef^1_{t_1} Ef^2_{t_2} \dots Ef^m_{t_m} Ef^1_{t_1} f^2_{t_2} \dots f^m_{t_m} \, dL(\Pi_{i=1}^m \lambda_i) \\
&\quad + \int_{\Pi_{i=1}^m T_i} (Ef^1_{t_1} Ef^2_{t_2} \dots Ef^m_{t_m})^2 \, dL(\Pi_{i=1}^m \lambda_i) \\
= &\int_{\Omega \times \Omega} \Pi_{i=1}^m \int_{T_i} f^i(t_i, \omega_1) f^i(t_i, \omega_2) \, dL(\lambda_i)(t_i) \, dL(P \otimes P)(\omega_1, \omega_2) \\
&\quad - 2 \int_{(\omega_1, \dots, \omega_m) \in \Omega^m} \int_{\omega \in \Omega} \left[\Pi_{i=1}^m \int_{t_i \in T_i} f^i(t_i, \omega) f^i(t_i, \omega_i) \, dL(\lambda_i)(t_i) \right] \\
&\quad \times dL(P)(\omega) \, dL(P^m)(\omega_1, \dots, \omega_m) + \Pi_{i=1}^m \int_{t_i \in T_i} (Ef^i_{t_i})^2 \, dL(\lambda_i)(t_i) \ .
\end{aligned}
$$

Now, for each $1 \le i \le m-1$, since the autocorrelation function of f^i_ω is essentially constant, it is easy to check by the Fubini theorem that $Ef^i_{\omega_1} f^i_{\omega_2} = \int_{t_i \in T_i} (Ef^i_{t_i})^2 \, dL(\lambda_i)$ for $L(P \otimes P)$-almost all $(\omega_1, \omega_2) \in \Omega \times \Omega$ (see Theorem 4.6 in [S2]). Hence, the Fubini theorem implies that

$$
\begin{aligned}
&\int_{\Omega \times \Omega} \Pi_{i=1}^m \int_{T_i} f^i(t_i, \omega_1) f^i(t_i, \omega_2) \, dL(\lambda_i)(t_i) \, dL(P \otimes P)(\omega_1, \omega_2) \\
&\quad = \int_{\Omega \times \Omega} \int_{T_m} f^m(t_m, \omega_1) f^m(t_m, \omega_2) \, dL(\lambda_m) \Pi_{i=1}^{m-1} E\big(Ef^i_{t_i}\big)^2 \\
&\qquad\qquad dL(P \otimes P)(\omega_1, \omega_2) \\
&\quad = \Pi_{i=1}^{m-1} E\big(Ef^i_{t_i}\big)^2 \int_{\Omega \times \Omega} \int_{T_m} f^m(t_m, \omega_1) f^m(t_m, \omega_2) \, dL(\lambda_m) \\
&\qquad\qquad dL(P \otimes P)(\omega_1, \omega_2) \\
&\quad = \Pi_{i=1}^m \int_{t_i \in T_i} \big(Ef^i_{t_i}\big)^2 \, dL(\lambda_i)(t_i) \ .
\end{aligned}
$$

Next, we note that for each $1 \le i \le m-1$, $\int_{t_i \in T_i} f^i_\omega(t_i) f^i_{\omega_i}(t_i) \, dL(\lambda_i)(t_i)$ is also essentially equal to $E(Ef^i_{t_i})^2$. Therefore, by the Fubini theorem again,

$$\int_{(\omega_1,\ldots,\omega_m)\in\Omega^m}\int_{\omega\in\Omega}\left[\Pi_{i=1}^m\int_{t_i\in T_i} f^i(t_i,\omega)f^i(t_i,\omega_i)\,dL(\lambda_i)(t_i)\right] \times dL(P)(\omega)\,dL(P^m)(\omega_1,\ldots,\omega_m)$$
$$= \Pi_{i=1}^{m-1}E(Ef^i_{t_i})^2\int_{\Omega\times\Omega}\int_{T_m} f^m(t_m,\omega)f^m(t_m,\omega_m)\,dL(\lambda_m)(t_m) \times dL(P\otimes P)(\omega,\omega_m)$$
$$= \Pi_{i=1}^m\int_{t_i\in T_i}(Ef^i_{t_i})^2\,dL(\lambda_i)(t_i)\ .$$

Replacing the relevant formulas in the expansion of I by $\Pi_{i=1}^m\int_{t_i\in T_i}(Ef^i_{t_i})^2\,dL(\lambda_i)$, we obtain $I=0$, and hence the proposition follows. □

Next, we consider a converse of the previous proposition in the following proposition. It shows that uncorrelatedness implies the constancy of joint moment functions for the relevant sample functions. Note that the case for $m=1$ in this proposition is Theorem 3.8 in [S2].

Proposition 4.3. *Let m be a positive integer. For each $i=1,2,\ldots,m$, let f^i be a real-valued process on a Loeb product space $(T\times\Omega_i, L(\mathscr{T}\otimes\mathscr{A}_i), L(\lambda\otimes P_i))$ with a finite m-th moment (with a finite second moment when $m=1$). If for each $1\le i\le m$, the random variables f^i_t are almost surely uncorrelated, i.e., for $L(\lambda\otimes\lambda)$-almost all $(t_1,t_2)\in T\times T$, the random variables $f^i_{t_1}$ and $f^i_{t_2}$ on Ω_i are uncorrelated, then for $L(\Pi_{i=1}^m P_i)$-almost all $(\omega_1,\omega_2,\ldots,\omega_m)\in\Pi_{i=1}^m\Omega_i$,*

$$Ef^1_{\omega_1}f^2_{\omega_2}\ldots f^m_{\omega_m} = \int_T\left[Ef^1_tEf^2_t\ldots Ef^m_t\right]dL(\lambda)\ .$$

Proof. Denote $\int_T\left[Ef^1_tEf^2_t\ldots Ef^m_t\right]dL(\lambda)$ by c. By the Fubini theorem,

$$\int_{\Pi_{i=1}^m\Omega_i}[Ef^1_{\omega_1}f^2_{\omega_2}\ldots f^m_{\omega_m}-c]^2\,dL(\Pi_{i=1}^mP_i)(\omega_1,\ldots,\omega_m)$$
$$= c^2-2c\int_{\Pi_{i=1}^m\Omega_i}\int_T f^1_{\omega_1}(t)f^2_{\omega_2}(t)\ldots f^m_{\omega_m}(t)\,dL(\lambda)(t)\, dL(\Pi_{i=1}^mP_i)(\omega_1,\ldots,\omega_m)$$
$$+\int_{\Pi_{i=1}^m\Omega_i}\int_T\Pi_{i=1}^m f^i(t_1,\omega_i)\,dL(\lambda)(t_1) \times\int_T\Pi_{i=1}^m f^i(t_2,\omega_i)\,dL(\lambda)(t_2)\,dL(\Pi_{i=1}^mP_i)$$
$$= c^2-2c\int_T\Pi_{i=1}^m\int_{\Omega_i}f^i_t(\omega_i)\,dL(P_i)(\omega_i)\,dL(\lambda)$$

$$+\int_{(t_1,t_2)\in T\times T} \Pi_{i=1}^{m} \int_{\Omega_i} f_{t_1}^i(\omega_i) f_{t_2}^i(\omega_i)\, dL(P_i)\, dL(\lambda\otimes\lambda)$$
$$= c^2 - 2c^2 + c^2 = 0 \ .$$

The rest is obvious. □

In Proposition 4.2, we only require $m-1$ of the m processes to have essentially constant autocorrelation functions for the relevant sample functions. On the other hand, all the m processes are assumed to have almost surely uncorrelated random variables in Proposition 4.3. The following example shows that these conditions are optimal. That is, the numbers $m-1$ and m in respective settings cannot be reduced.

Example 4.4. Choose $B \in L(\mathscr{A})$ with $0 < L(P)(B) < 1$. Define $f^i \equiv 1$ for $1 \le i \le m-2$, and $f^{m-1} = f^m = \chi_B$, where χ_B is the indicator function of B in Ω. Then, it is obvious that for each $1 \le i \le m-2$, the process f^i has essentially constant autocorrelation function for its sample functions, while the conclusion of Proposition 4.2 is not valid.

For the case of Proposition 4.3, let $f^i \equiv 1$ for $1 \le i \le m-1$, and $f^m = \chi_B$. Then, each of the first $m-1$ processes has almost surely uncorrelated random variables; but $Ef^1_{\omega_1} f^2_{\omega_2} \dots f^m_{\omega_m} = \chi_B(\omega_m)$ is certainly not an essentially constant function. □

Part of Theorem 4.6 in [S2] shows the duality of the notions of uncorrelatedness and constancy of the autocorrelation functions in an almost sense. We restate this basic fact in the following corollary and give a new proof based on Propositions 4.2 and 4.3.

Corollary 4.5. *Let f be a square integrable real-valued process on a Loeb product space $(T\times\Omega, L(\mathscr{T}\otimes\mathscr{A}), L(\lambda\otimes P))$. Then the following are equivalent*:

(1) *the random variables f_t are almost surely uncorrelated*;
(2) *the autocorrelation function of the sample functions f_ω is essentially constant.*

Proof. By taking $f^1 = f^2 = f$, we observe that $(2) \Longrightarrow (1)$ follows from Proposition 4.2 while $(1) \Longrightarrow (2)$ is a special case of Proposition 4.3. □

By Corollary 4.5 and Proposition 4.2, one can obtain that for a given collection of m processes, if, except possibly one of them, all have almost surely uncorrelated random variables, then the processes are uncorrelated in m-tuple among themselves. The case $m = 2$ is already shown in Theorem 2 in [S1].

Corollary 4.6. *Let $m \geq 2$ be a positive integer. For each $i = 1, 2, \ldots, m$, let f^i be a real-valued process on a Loeb product space $(T_i \times \Omega, L(\mathcal{T}_i \otimes \mathcal{A}), L(\lambda_i \otimes P))$ with a finite m-th moment. If for each $1 \leq i \leq m-1$, the random variables f_t^i in the process f^i are almost surely uncorrelated, then for $L(\Pi_{i=1}^m \lambda_i)$-almost all $(t_1, t_2, \ldots, t_m) \in \Pi_{i=1}^m T_i$, $f_{t_1}^1, f_{t_2}^2, \ldots, f_{t_m}^m$ are uncorrelated in m-tuple, i.e., $E(f_{t_1}^1 f_{t_2}^2 \ldots f_{t_m}^m) = Ef_{t_1}^1 Ef_{t_2}^2 \ldots Ef_{t_m}^m$.*

Now, we consider an analog of the above corollary in terms of the constancy of joint moment functions.

Corollary 4.7. *Let m be a positive integer. For each $i = 1, 2, \ldots, m$, let f^i be a real-valued process on a Loeb product space $(T \times \Omega_i, L(\mathcal{T} \otimes \mathcal{A}_i)$, $L(\lambda \otimes P_i))$* with a finite m-th moment (*with a finite second moment when $m = 1$*). *If for each $1 \leq i \leq m$, the autocorrelation function of the sample functions f_ω^i in the process f^i is essentially constant, then for $L(\Pi_{i=1}^m P_i)$-almost all $(\omega_1, \omega_2, \ldots, \omega_m) \in \Pi_{i=1}^m \Omega_i$,*

$$Ef_{\omega_1}^1 f_{\omega_2}^2 \ldots f_{\omega_m}^m = \int_T \left[Ef_t^1 Ef_t^2 \ldots Ef_t^m \right] dL(\lambda) \ .$$

Proof. By the equivalence result in Corollary 4.5, we know that for each $1 \leq i \leq m$, the random variables f_t^i in the process f_i are almost surely uncorrelated. Hence the result follows from Proposition 4.3. □

We are now ready to prove Theorems 1 and 2.

Proof of Theorem 1. One can simply take $f^i = f$ for all $1 \leq i \leq n$. Then each f^i still has a finite n-th moment. The rest follows from Corollary 4.6. □

Proof of Theorem 2. One can simply take $f^i = f$ for all $1 \leq i \leq n$. Then each f^i still has a finite n-th moment. The rest follows from Corollary 4.7 by renaming the variables ω_i to t_i and t to ω. □

For the two Loeb spaces $(T, L(\mathcal{T}), L(\lambda))$ and $(\Omega, L(\mathcal{A}), L(P))$, one can also take their product in the usual sense to obtain a measure space $(T \times \Omega, L(\mathcal{T}) \otimes L(\mathcal{A}), L(\lambda) \otimes L(P))$, which is clearly contained in the corresponding Loeb product space $(T \times \Omega, L(\mathcal{T} \otimes \mathcal{A}), L(\lambda \otimes P))$ (noted first by Anderson in [A1]). For simplicity, we will not distinguish the difference between $(T \times \Omega, L(\mathcal{T}) \otimes L(\mathcal{A}), L(\lambda) \otimes L(P))$ and its completion. Both the product σ-algebra $L(\mathcal{T}) \otimes L(\mathcal{A})$ and its completion with respect to $L(\lambda) \otimes L(P)$) will be denoted by $\mathscr{U}$. It is interesting to note that except the trivial case that one of the Loeb spaces is purely atomic $\mathscr{U}$ is always strictly contained in the Loeb product algebra $L(\mathscr{T} \otimes \mathscr{A})$ (see Proposition 6.6 in [S2]). For the case that T is a hyperfinite set, Ω the internal power set on T, and both endowed with the Loeb counting probability measures, such a proper inclusion was first observed by Hoover (see [AFHL]). The intimate connection between the law of large numbers and $\mathscr{U}$ is also established in [S1] and [S2]. In particular, it is shown that an integrable real-valued process f on the Loeb product space satisfies the consistency law if and only if the conditional expectation $E(f|_{\mathscr{U}})$ is essentially a function of t (see Theorem 1 in [S1] and Theorem 3.17 in [S2]).

Remark 4.8. Theorem 4.6 in [S2] also points out that the autocorrelation function of the sample functions of a process is essentially constant if and only if the conditional expectation of the process with respect to $\mathscr{U}$ depends only on the indices of the random variables. By symmetry, it is obvious that the autocorrelation function of the random variables f_t is essentially constant if and only if $E(f|_{\mathscr{U}})$ is essentially a function on Ω; and in this case $E(f|_{\mathscr{U}})(t, \omega) = Ef_\omega$ for $L(\lambda \otimes P)$-almost all $(t, \omega) \in T \times \Omega$, i.e., $f(t, \omega)$ can be expressed as the sum $Ef_\omega + e(t, \omega)$, where the process e has almost surely orthogonal random variables.

As noted in the paragraph above Proposition 4.2, even if one works with a fixed process as in Theorem 3, one still needs to consider uncorrelatedness involving possibly different processes to prove some relevant results. Thus, it might be helpful to prove an analog of Corollary 4.6 for the case of independence.

Proposition 4.9. *Let $n \geq 2$ and $f^i, i = 1, 2, \ldots, n$ be processes from a Loeb product space $(T \times \Omega, L(\mathscr{T} \otimes \mathscr{A}), L(\lambda \otimes P))$ to a separable metric space X. If for each $1 \leq i \leq n-1$ the random variables f_t^i are almost surely pairwise independent, i.e., for $L(\lambda \otimes \lambda)$-almost all $(t_1, t_2) \in T \times T$, $f_{t_1}^i$ and $f_{t_2}^i$ are independent, then for $L(\lambda^n)$-almost all $(t_1, t_2, \ldots, t_n) \in T^n$, $f_{t_1}^1, f_{t_2}^2, \ldots, f_{t_n}^n$ are mutually independent.*

Proof. Fix a countable open base $\{Q_j\}_{j=1}^\infty$ for X. Let $\{O_k\}_{k=1}^\infty$ be a list of all the finite intersections of the sets Q_j. Now, fix an n-tuple $(\ell_1, \ell_2, \ldots, \ell_n)$ of positive integers (not necessarily different). Then, take the n sets $O_{\ell_1}, O_{\ell_2}, \ldots, O_{\ell_n}$ from the sequence $\{O_k\}_{k=1}^\infty$. For each $1 \leq i \leq n-1$, by the assumption of almost sure pairwise independence of the random variables f_t^i, we know that the process $\chi_{O_{\ell_i}}(f^i)$ has almost surely uncorrelated random variables, where $\chi_{O_{\ell_i}}$ is the indicator function of the set O_{ℓ_i}. Then, Corollary 4.6 implies that there is an $L(\lambda^n)$-null set $\mathcal{N}_{\ell_1\ell_2\ldots\ell_n}$ such that for all $(t_1, t_2, \ldots, t_n) \notin \mathcal{N}_{\ell_1\ell_2\ldots\ell_n}$, the random variables $\chi_{O_{\ell_1}}(f^1_{t_1})$, $\chi_{O_{\ell_2}}(f^2_{t_2}), \ldots,$ $\chi_{O_{\ell_n}}(f^n_{t_n})$ are uncorrelated in n-tuple. Let $\mathcal{N}$ be the union of all the null sets $\mathcal{N}_{\ell_1\ell_2\ldots\ell_n}$. Then $\mathcal{N}$ is still an $L(\lambda^n)$-null set, and for all $(t_1, t_2, \ldots, t_n) \notin \mathcal{N}$,

$$L(P)((f^1_{t_1})^{-1}(O_{\ell_1}) \cap (f^2_{t_2})^{-1}(O_{\ell_2}) \cap \ldots \cap (f^n_{t_n})^{-1}(O_{\ell_n}))$$

$$= L(P)((f^1_{t_1})^{-1}(O_{\ell_1}))L(P)\big((f^2_{t_2})^{-1}(O_{\ell_2})\big) \ldots L(P)((f^n_{t_n})^{-1}(O_{\ell_n}))$$

holds for all n-tuples $(\ell_1, \ell_2, \ldots, \ell_n)$ of positive integers. Since the collection $\{O_k\}_{k=1}^{\infty}$ generates the Borel σ-algebra of X and is closed under the formation of finite intersections, the Extension Theorem in [Lo] (see p. 237) implies that for all $(t_1, t_2, \ldots, t_n) \notin \mathcal{N}$, $f^1_{t_1}, f^2_{t_2}, \ldots, f^n_{t_n}$ are mutually independent. One can also view this fact on mutual independence directly through a consequence of Dynkin's $\pi - \lambda$ theorem which guarantees that two probability measures agree on a π-system are the same (see [C], p. 45 or [Du], p. 404). Note that for $(t_1, t_2, \ldots, t_n) \notin \mathcal{N}$, the joint distribution of $f^1_{t_1}, f^2_{t_2}, \ldots, f^n_{t_n}$ agrees with the product of its marginal distributions on all the sets from the collection

$$\{O_{\ell_1} \times O_{\ell_2} \times \cdots \times O_{\ell_n} : 1 \le \ell_1, \ell_2, \ldots, \ell_n < \infty\} \ .$$

Since this collection is still closed under finite intersections and also generates the Borel σ-algebra of X^n, i.e., a π-system, the joint distribution of $f^1_{t_1}, f^2_{t_2}, \ldots, f^n_{t_n}$ is thus equal to the product of its marginals, and hence $f^1_{t_1}, f^2_{t_2}, \ldots, f^n_{t_n}$ are mutually independent. □

For any fixed $n \ge 2$, the implication (1) $\Longrightarrow$ (2) in Theorem 3 clearly follows from Proposition 4.9 by taking $f^i = f$ for $1 \le i \le n$. It remains to prove the other half of the equivalence in the theorem.

Proof of Theorem 3. Assume (2) is valid. Let A be the set of all the n-tuples $(t_1, t_2, \ldots, t_n) \in T^n$ such that $f_{t_1}, f_{t_2}, \ldots, f_{t_n}$ are mutually independent. Then (2) says that $L(\lambda^n)(A) = 1$. By the Fubini theorem, for $L(\lambda^{n-2})$-almost all $(t_3, t_4, \ldots, t_n) \in T^{n-2}$, the set

$$A_{(t_3,t_4,\ldots,t_n)} = \{(t_1, t_2) : (t_1, t_2, t_3, t_4, \ldots, t_n) \in A\}$$

is of probability 1, and hence we can choose a particular $(n-2)$-tuple $(s_3, s_4, \ldots, s_n)$ with the property, i.e., $L(\lambda \otimes \lambda)(A_{(s_3,s_4,\ldots,s_n)}) = 1$. Thus, for any $(t_1, t_2) \in A_{(s_3,s_4,\ldots,s_n)}$, $f_{t_1}, f_{t_2}, f_{s_3}, \ldots, f_{s_n}$ are mutually independent, and so are f_{t_1}, f_{t_2}. Therefore (1) follows. □

The following proposition is an analog of Proposition 4.9 for the case of exchangeability. The consideration of different processes below might be useful elsewhere.

Proposition 4.10. *Let $n \geq 2$ and $f^i, i = 1, 2, \ldots, n$ be processes from a Loeb product space $(T \times \Omega, L(\mathcal{T} \otimes \mathcal{A}), L(\lambda \otimes P))$ to a separable metric space X. If for each $1 \leq i \leq n$ the random variables f^i_t are almost surely pairwise exchangeable, i.e., there is a distribution v^i on $X \times X$ such that for $L(\lambda \otimes \lambda)$-almost all $(t_1, t_2) \in T \times T$, the joint distribution of $f^i_{t_1}$ and $f^i_{t_2}$ is v^i, then there is a distribution v_n on X^n such that for $L(\lambda^n)$-almost all $(t_1, t_2, \ldots, t_n) \in T^n$, the joint distribution of $f^1_{t_1}, f^2_{t_2}, \ldots, f^n_{t_n}$ is v_n.*

Proof. As in the proof of Proposition 4.9, we can obtain a countable open base $\{O_k\}_{k=1}^{\infty}$ which is closed under the formation of finite intersections. Fix an n-tuple $(\ell_1, \ell_2, \ldots, \ell_n)$ of positive integers. For each $1 \leq i \leq n$, by the assumption of almost sure pairwise exchangeability of the random variables f^i_t, we obtain that the autocorrelation function of the random variables in the process $\chi_{O_{\ell_i}}(f^i)$ is essentially constant. By renaming the variables ω_i to t_i and t to ω in Corollary 4.7, we know that there is an $L(\lambda^n)$-null set $\mathcal{N}_{\ell_1\ell_2\ldots\ell_n}$ such that for all $(t_1, t_2, \ldots, t_n) \notin \mathcal{N}_{\ell_1\ell_2\ldots\ell_n}$, the joint moment of the random variables $\chi_{O_{\ell_1}}(f^1_{t_1}), \chi_{O_{\ell_2}}(f^2_{t_2}), \ldots, \chi_{O_{\ell_n}}(f^n_{t_n})$ is

$$\int_{\Omega} [E\chi_{O_{\ell_1}}(f^1_{\omega}) E\chi_{O_{\ell_2}}(f^2_{\omega}) \ldots E\chi_{O_{\ell_n}}(f^n_{\omega})] \, dL(P)(\omega) \ .$$

Let $\mathcal{N}$ be the union of all the null sets $\mathcal{N}_{\ell_1\ell_2\ldots\ell_n}$. Then $\mathcal{N}$ is still an $L(\lambda^n)$-null set, and for all $(t_1, t_2, \ldots, t_n) \notin \mathcal{N}$,

$$\begin{aligned} &L(P)((f^1_{t_1})^{-1}(O_{\ell_1}) \cap \ldots \cap (f^n_{t_n})^{-1}(O_{\ell_n})) \\ &= \int_{\Omega} \Pi_{i=1}^n L(\lambda)((f^i_{\omega})^{-1}(O_{\ell_i})) \, dL(P)(\omega) \end{aligned}$$

holds for all n-tuples $(\ell_1, \ell_2, \ldots, \ell_n)$ of positive integers. Define a distribution v_n on X^n by letting $v_n(B) = \int_{\Omega} \left[\Pi_{i=1}^n L(\lambda)(f^i_{\omega})^{-1} \right](B) \, dL(P)(\omega)$ for any Borel set B in X^n. Then the previous identity shows that for all $(t_1, t_2, \ldots, t_n) \notin \mathcal{N}$, the joint distribution of $f^1_{t_1}, f^2_{t_2}, \ldots, f^n_{t_n}$ and v_n agree on all the sets in the π-system for the Borel σ-algebra of X^n

$$\{O_{\ell_1} \times O_{\ell_2} \times \cdots \times O_{\ell_n} : 1 \leq \ell_1, \ell_2, \ldots, \ell_n < \infty\} \ ,$$

and hence are equal by Dynkin's $\pi - \lambda$ theorem as quoted earlier. □

Note that Example 4.4 can still be used to show that the number $m - 1$ in Corollary 4.6 and Proposition 4.9 as well as the number m in Corollary 4.7 and Proposition 4.10 cannot be reduced in the same context. We also note that exchangeability has been studied by using model-theoretic methods in [H] and [Ho]. It is noted in [H] that one can always find an exchangeable sequence among a given continuum of random variables (for some discrete versions, see [A] and [K]). [Ho] is concerned with a special notion of partial exchangeability for multiply indexed arrays $\{X_{i_1\dots i_n}\}_{i_1\dots i_n\in\mathbb{N}}$, called dissociated row-column exchangeability (DRCE). It shows how DRCE may be related by a certain sampling procedure to random variables on product spaces in a graded probability space which could be constructed from a sequence of Loeb product spaces.

We are now ready to prove Theorems 4 and 5.

Proof of Theorem 4. The same argument as in the proof of Theorem 3 can be used to prove (2) $\Longrightarrow$ (1) here.

Next, if we assume (1), then by taking $f^i = f$ for $1 \le i \le n$, it follows from Proposition 4.10 that the joint distribution of $f_{t_1}, f_{t_2}, \dots, f_{t_n}$ is essentially independent of $t_1, t_2, \dots, t_n$; the essentially common joint distribution is ν_n with $\nu_n(B) = \int_\Omega \tau^n_\omega(B)\, dL(P)(\omega)$ for any Borel set B in X^n, where τ_ω is the distribution on X induced by f_ω and τ^n_ω its n-fold product.

Finally, for any $\ell \ge 1$, (1) also implies the essential constancy of the autocorrelation function of the random variables in the process $\chi_{O_\ell}(f)$. By Theorem 2, we obtain that

$$E\chi_{O_\ell}(f_t) = \int_\Omega E\chi_{O_\ell}(f_\omega)\, dL(P) = \int_{T\times\Omega} \chi_{O_\ell}(f)\, dL(\lambda\otimes P)$$

for $L(\lambda)$-almost all $t \in T$. By taking the union of countably many $L(\lambda)$-null sets, we can obtain that for $L(\lambda)$-almost all $t \in T$, the respective distributions of f_t and f agree on all the O_ℓ, and hence are identical by Dynkin's $\pi - \lambda$ theorem. Therefore, (2) holds. □

Proof of Theorem 5. By Theorems 3 and 4, we only have to show the equivalence for the case $n = 2$. Let $\{O_k\}_{k=1}^\infty$ be a countable open base of X which is closed under the formation of finite intersections. Then, by interchanging the index and sample variables in Corollary 4.5, we know that for each $k \ge 1$, the essential constancy of the autocorrelation function of the random variables $\chi_{O_k}(f_t)$ is equivalent to the uncorrelatedness of the sample functions $\chi_{O_k}(f_\omega)$.

If (1) holds, then the autocorrelation function of the random variables $\chi_{O_k}(f_t)$ is indeed essentially constant for each k. By the above equivalence and by taking the union of countably many $L(P\otimes P)$-null sets, we can obtain that for $L(P\otimes P)$-almost all $(\omega_1, \omega_2) \in \Omega\times\Omega$, the sample functions $\chi_{O_k}(f_{\omega_1})$ and $\chi_{O_k}(f_{\omega_2})$ are uncorrelated for all $k \geq 1$. The same argument in the end of the proof of Proposition 4.9 can be used to claim almost sure pairwise independence for the sample functions f_ω, i.e., (2) holds.

Next, note that (2) implies that for any $k \geq 1$, the sample functions in the process $\chi_{O_k}(f)$ are uncorrelated, and hence the autocorrelation function of the random variables $\chi_{O_k}(f_t)$ is essentially constant by the equivalence in the first paragraph. By interchanging the index and sample variables t and ω in Corollary 4.7, we know that for $L(\lambda \otimes \lambda)$-almost all $(t_1, t_2) \in T \times T$, the joint moment

$$E\chi_{O_{\ell_1}}(f_{t_1})\chi_{O_{\ell_2}}(f_{t_2}) = \int_\Omega E\chi_{O_{\ell_1}}(f_\omega)E\chi_{O_{\ell_2}}(f_\omega)\, dL(P)(\omega) \ ,$$

for any given $\ell_1, \ell_2 \geq 1$. By taking the union of countably many $L(\lambda \otimes \lambda)$-null sets, we can find an $L(\lambda \otimes \lambda)$-null set $\mathcal{N}$ and a distribution μ on $X \times X$ such that for any $(t_1, t_2) \in \mathcal{N}$, μ and the joint distribution of f_{t_1} and f_{t_2} agree on all the $O_{\ell_1} \times O_{\ell_2}$, and hence are identical by Dynkin's $\pi - \lambda$ theorem as in the proof of Proposition 4.10. Therefore, (1) follows. □

In the previous part of this section, we have considered uncorrelatedness, constancy of joint moment functions, independence and exchangeability. We shall now move to the study of orthogonality in the same context. The surprising point is that when one works with a finite collection of processes, one can only require almost sure orthogonality for one of these processes. This is in contrast with the previous cases where all the processes in the collection (or except one of them) must satisfy the relevant conditions.

The following lemma is an analog of Corollary 4.5 in the setting of orthogonality. It shows that the almost sure orthogonality of the random variables and sample functions in a process are equivalent, which can be proven by establishing the equality of $\iint_{\Omega\times\Omega} \left(\int_T f_{\omega_1}(t) f_{\omega_2}(t)\, dL(\lambda)\right)^2 dL(P \otimes P)$ with $\iint_{T\times T}\left(\int_\Omega f_{t_1}(\omega) f_{t_2}(\omega)\, dL(P)\right)^2 dL(\lambda \otimes \lambda)$ (see Theorem 4.5 in [S2]).

Lemma 4.11. *Let f be a square integrable real-valued process on a Loeb product space $(T \times \Omega, L(\mathcal{T} \otimes \mathcal{A}), L(\lambda \otimes P))$. Then the following are equivalent*:

(1) *the random variables f_t are almost surely orthogonal, i.e., for $L(\lambda \otimes \lambda)$-almost all $(t_1, t_2) \in T \times T$, f_{t_1} and f_{t_2} are orthogonal*;
(2) *the sample functions f_ω are almost surely orthogonal.*

The following proposition is an analog of Corollary 4.6 in the setting of orthogonality. It can be proven by showing $\int_{\Pi_{i=1}^m T_i}\left(Ee_{t_1}^1 e_{t_2}^2 \ldots e_{t_m}^m\right)^2 dL(\Pi_{i=1}^m \lambda_i)$ to be zero.

Proposition 4.12. *Let m be a positive integer. For each $i = 1, 2, \ldots m$, let e^i be a real-valued process on a Loeb product space $(T_i \times \Omega, L(\mathscr{T}_i \otimes \mathscr{A}), L(\lambda_i \otimes P))$ with a finite m-th moment (with a finite second moment when $m = 1$). If there is a j between 1 and m such that the random variables $e_{t_j}^j$ are almost surely orthogonal, then for $L(\Pi_{i=1}^m \lambda_i)$-almost all $(t_1, t_2, \ldots, t_m) \in \Pi_{i=1}^m T_i$, $Ee_{t_1}^1 e_{t_2}^2 \ldots e_{t_m}^m = 0$.*

The following corollary is an analog of Proposition 4.3.

Corollary 4.13. *For $i = 1, 2, \ldots, m$, let e^i be a real-valued process on a Loeb product space $(T \times \Omega_i, L(\mathscr{T} \otimes \mathscr{A}_i), L(\lambda \otimes P_i))$ with a finite m-th moment (with a finite second moment when $m = 1$). If there is a j between 1 and m such that the random variables e_t^j are almost surely orthogonal, then for $L(\Pi_{i=1}^m P_i)$-almost all $(\omega_1, \omega_2, \ldots, \omega_m) \in \Pi_{i=1}^m \Omega_i$, $Ee_{\omega_1}^1 e_{\omega_2}^2 \ldots e_{\omega_m}^m = 0$.*

Proof. By Lemma 4.11, the sample functions $e_{\omega_j}^j$ are almost surely orthogonal. The result can then be proven by interchanging the index and sample variables in Proposition 4.12. □

In Proposition 4.12 and Corollary 4.13, we allow the processes under consideration to be different. In the following corollary, we work with a fixed process. We can again obtain a sort of results in the style of "*Two implies many*", i.e., almost sure orthogonality implies its corresponding multiple versions for the random variables as well as for sample functions. The proof is obvious.

Corollary 4.14. *Let e be a real-valued process on $(T \times \Omega, L(\mathscr{T} \otimes \mathscr{A}), L(\lambda \otimes P))$ which has almost orthogonal random variables and also a finite m-th moment for some $m \geq 2$. Then the random variables e_t are almost orthogonal in n-tuple for any $1 \leq n \leq m$, i.e., for $L(\lambda^n)$-almost all $(t_1, t_2, \ldots, t_n) \in T^n$, $Ee_{t_1} e_{t_2} \ldots e_{t_n} = 0$, and so are the sample functions.*

The previous corollary shows that the usual orthogonality (second order) implies its higher order counterpart. It is natural to ask when the converse holds. The following result indicates that an even order orthogonality also implies the usual orthogonality (second order). On the other hand, it seems unlikely that a similar result holds for the odd case.

Proposition 4.15. *Let e be a real-valued process on* $(T \times \Omega, L(\mathscr{T} \otimes \mathscr{A}), L(\lambda \otimes P))$ *with a finite m-th moment for some positive even integer m. If the random variables* e_t *are almost surely orthogonal in m-tuple, then they are also almost surely orthogonal.*

Proof. By the Fubini Theorem, we can establish the following identities

$$\begin{aligned}
&\int_{\Omega\times\Omega}\left[\int_T e_{\omega_1}(t)e_{\omega_2}(t)\,dL(\lambda)(t)\right]^m dL(P\otimes P)\\
&\quad=\int_{\Omega\times\Omega}\left(\Pi_{i=1}^m\int_T e_{\omega_1}(t_i)e_{\omega_2}(t_i)\,dL(\lambda)(t_i)\right)dL(P\otimes P)\\
&\quad=\int_{T^m}\int_\Omega \Pi_{i=1}^m e(t_i,\omega_1)\,dL(P)(\omega_1)\int_\Omega \Pi_{i=1}^m e(t_i,\omega_2)\,dL(P)(\omega_2)\,dL(\lambda^m)\\
&\quad=\int_{T^m}\left[\int_\Omega e_{t_1}(\omega)e_{t_2}(\omega)\dots e_{t_m}(\omega)\,dL(P)(\omega)\right]^2 dL(\lambda^m)\ .
\end{aligned}$$

Since the random variables e_t are almost surely orthogonal in m-tuple, we have

$$\int_\Omega e_{t_1}(\omega)e_{t_2}(\omega)\dots e_{t_m}(\omega)\,dL(P)=0$$

for $L(\lambda^m)$-almost all $(t_1,t_2,\dots,t_m)\in T^m$, and hence

$$\int_{\Omega\times\Omega}\left[\int_T e_{\omega_1}(t)e_{\omega_2}(t)\,dL(\lambda)(t)\right]^m dL(P\otimes P)=0\ .$$

Since m is even, we obtain that $\int_T e_{\omega_1}(t)e_{\omega_2}(t)\,dL(\lambda)=0$ for $L(P\otimes P)$-almost all $(\omega_1,\omega_2)\in\Omega\times\Omega$. The rest follows from Lemma 4.11. □

Remark 4.16. Since many applied probabilistic models involve not only uncertainty and large number of entities but also time constraints, one is naturally led to use a hyperfinite set to index a large collection of stochastic processes with time and sample parameters. As illustrated in Section 8 of [S2], many results in the previous sections can be routinely extended to such "hyperprocesses". In particular, Theorems 3–5 can be restated to the case of general hyperprocesses with continuous time parameters. Note that the various notions of independence and exchangeability for hyperprocesses should be defined in terms of the finite dimensional distributions of stochastic processes.

5. Asymptotic equivalence of pairwise and mutual independence and the duality with exchangeability

In this section, we shall translate Theorem 3 and the dual version of Theorem 5 to obtain information about triangular arrays or sequences of random variables. In particular, some versions of pairwise and mutual independence are shown to be asymptotically equivalent. The duality between independence and exchangeability is also established in the asymptotic setting.

We shall now fix some notation for the asymptotic case. Let $(\Omega, \mathscr{A}, P)$ be a fixed probability space which will be used as the common sample space of the triangular array of random variables to be considered. For each $n \geq 1$, let $(T_n, \mathscr{T}_n, \lambda_n)$ be a finite probability space, which will be the n-th index space, and we shall assume that the number of points in T_n goes to infinity as $n \to \infty$; let g_n be a process from $T_n \times \Omega$ to a separable metric space X such that $g_n(t, \cdot)$ is a random variable on Ω for each $t \in T_n$. The measure λ_n provides the weight for each point in the finite index set T_n and $\mathscr{T}_n$ is the power set of T_n. Such a sequence of processes $g = \{g_n\}_{n \geq 1}$ will be called a triangular array of random variables. For a positive integer m and a separable metric space X, ρ_m denotes a Prohorov distance on the space of distributions on X^m.

Note that the usual definition of a triangular array of random variables is set in the form, $x_1^n, x_2^n, \ldots, x_n^n$, $n = 1, 2, \ldots$, which corresponds to our case when $T_n = \{1, 2, \ldots, n\}$ and λ_n is the uniform probability measure on T_n. So integrals on T_n are just the arithmetic averages. In this section, we consider general weighted averages rather than the special arithmetic averages, since the proofs are the same.

The following proposition shows that for a triangular array of random variables, asymptotic pairwise independence implies its asymptotic multiple versions; the other implication is clear and omitted here. As noted earlier in Section 3, the asymptotic pairwise independence as presented here is simply a version of the usual notion of weak dependence. In particular, it covers the type of mixing conditions as discussed in [Bi].

Proposition 5.1. *For any $t_n^1, t_n^2, \ldots, t_n^m \in T_n$, let $\mu_{t_n^i}$ be the distribution of the random variable $f_n(t_n^i, \cdot)$ and $\mu_{t_n^1 t_n^2 \ldots t_n^m}$ the joint distribution of the random variables $f_n(t_n^1, \cdot), f_n(t_n^2, \cdot), \ldots, f_n(t_n^m, \cdot)$. Assume that the collection of distributions induced by all the f_n* on *X (viewed as random variables on $T_n \times \Omega$) is tight. For any $\varepsilon > 0$ and $n \geq 1$, define*

$$T_n^m(\varepsilon) = \{(t_n^1, t_n^2, \ldots, t_n^m) \in (T_n)^m : \rho_m(\mu_{t_n^1 t_n^2 \ldots t_n^m}, \Pi_{i=1}^m \mu_{t_n^i}) \leq \varepsilon\} \ .$$

If $\lim_{n\to\infty}(\lambda_n \otimes \lambda_n)(T_n^2(\delta)) = 1$ *for any* $\delta > 0$, *then* $\lim_{n\to\infty}(\lambda_n)^m (T_n^m(\varepsilon)) = 1$ *for any* $\varepsilon > 0$.

Proof. We transfer the sequence to the nonstandard universe to obtain a sequence $\{f_n\}_{n\in{}^*\mathbb{N}}$ of internal processes on the associated sequence $\{(T_n \times {}^*\Omega, \mathscr{T}_n \otimes {}^*\mathscr{A}, \lambda_n \otimes {}^*P) : n \in {}^*\mathbb{N}\}$ of internal probability spaces. The tightness assumption on the processes f_n implies that for each $n \in \mathbb{N}_\infty$, the standard part of the $f_n(t_n, \omega)$ exists for almost all $(t_n, \omega) \in T_n \times {}^*\Omega$, and hence for almost all $t_n \in T_n$, $f_n(t_n, \cdot)$ has a standard part.

Next, fix $n \in {}^*\mathbb{N}_\infty$ and omit the subindex n in the rest of this paragraph for simplicity. By spillover, we can obtain that $\lambda \otimes \lambda(T^2(h)) = 1$ for some positive infinitesimal h. Thus, for $L(\lambda \otimes \lambda)$-almost all $(t^1, t^2) \in T \times T$, $\rho_2(\mu_{t^1 t^2}, \mu_{t^1} \otimes \mu_{t^2}) \leq h$. It is easy to see that $\mu_{t^1 t^2} \simeq L(P)({}^\circ f_{t^1}, {}^\circ f_{t^2})^{-1}$ and $\mu_{t^1} \otimes \mu_{t^2} \simeq L(P)({}^\circ f_{t^1})^{-1} \otimes L(P)({}^\circ f_{t^2})^{-1}$, and hence

$$L(P)({}^\circ f_{t^1}, {}^\circ f_{t^2})^{-1} = L(P)({}^\circ f_{t^1})^{-1} \otimes L(P)({}^\circ f_{t^2})^{-1} \ .$$

Therefore, the random variables ${}^\circ f_t$ are almost surely pairwise independent, and by Theorem 3, also almost surely independent in m-tuple. By the fact that the topology of weak convergence of distributions on X^m restricted to the product measures is simply the m-fold product topology of the topology of weak convergence of distributions on X (see Theorem 3.2 in [Bi]; p.21), we can obtain for almost all m-tuples $(t^1, t^2, \ldots, t^m)$, $\rho_m(\mu_{t^1 t^2 \ldots t^m}, \Pi_{i=1}^m \mu_{t^i}) \simeq 0$.

Now we resume the index n and also fix an $\varepsilon \in \mathbb{R}^+$. The previous paragraph shows that $(\lambda_n)^m (T_n^m(\varepsilon)) \simeq 1$ for any $n \in {}^*\mathbb{N}_\infty$ and for any positive standard real number ε. Hence $\lim_{n\to\infty}(\lambda_n)^m (T_n^m(\varepsilon)) = 1$. □

Note that if we start from three random variables which are pairwise independent but not mutually independent, then, by taking independent replicas of the three random variables, we can obtain a sequence in which pairs of random variables are independent but some triples are not mutually independent. This is the usual way of constructing a sequence of pairwise independent but not mutually independent random variables (see, for example, [F], p.220). Note that most triples in such a sequence are still mutually independent. Proposition 5.1 says that this is approximately the general case in the sense that even if one works on a sequence with approximate pairwise independence, then "almost all" triples are still approximately mutually independent. That is, one cannot expect the pairs in a sequence to be independent but the triples to be "highly" non-independent.

Next, we transfer the result on the duality of independence and exchangeability to the asymptotic setting. The proof is omitted.

Proposition 5.2. *Let* $m \geq 2$. *For any* $t_n^1, t_n^2, \ldots, t_n^m \in T_n$, *let* $\mu_{t_n^i}$ *be the distribution of the random variable* $f_n(t_n^i, \cdot)$ *for each* $1 \leq i \leq m$, *and* $\mu_{t_n^1 t_n^2 \ldots t_n^m}$ *the joint distribution of the random variables* $f_n(t_n^1, \cdot), f_n(t_n^2, \cdot), \ldots, f_n(t_n^m, \cdot)$. *For any* $\omega_n^1, \omega_n^2, \ldots, \omega_n^m \in \Omega$, *let* $\tau_{\omega_n^1 \omega_n^2 \ldots \omega_n^m}$ *be the joint distribution of the sample functions* $f_n(\cdot, \omega_n^1), f_n(\cdot, \omega_n^2), \ldots, f_n(\cdot, \omega_n^m)$. *Define a distribution* ν_n^m *on* X^m *by letting* $\nu_n^m(B) = \int_{T_n} [Pf_{t_n}^{-1}]^m (B)\, d\lambda_n(t_n)$ *for any Borel set B in* X^m, *where* $[Pf_{t_n}^{-1}]^m$ *is the m-fold product distribution of* $Pf_{t_n}^{-1}$. *Assume that the collection of distributions induced by all the* f_n *on* X (*viewed as random variables on* $T_n \times \Omega$) *is tight. Then the following are equivalent.*

(1) *for any* $\delta > 0$ *and* $n \geq 1$, *let*
$$T_n^m(\delta) = \{(t_n^1, t_n^2, \ldots, t_n^m) \in (T_n)^m : \rho_m(\mu_{t_n^1 t_n^2 \ldots t_n^m}, \Pi_{i=1}^m \mu_{t_n^i}) \leq \delta\} \ ;$$
then $\lim_{n\to\infty} (\lambda_n)^m (T_n^m(\delta)) = 1$;

(2) *for any* $\varepsilon > 0$, $\lim_{n\to\infty}(P^m)(W_n^m(\varepsilon)) = 1$, *where*
$$W_n^m(\varepsilon) = \{(\omega_n^1, \omega_n^2, \ldots, \omega_n^m) \in \Omega^m : \rho_m(\tau_{\omega_n^1 \omega_n^2 \ldots \omega_n^m}, \nu_n^m) \leq \varepsilon\} \ .$$

Since a nonstandard model is *elementarily equivalent* to the corresponding standard model, it is routine to interpret results from one model to the other. The general possibility of such a translation was already demonstrated by Brown and Robinson in [BR], where large finite results on the cores of exchange economies are obtained from some internal counterparts. The procedure usually involves the so-called *lifting*, *pushing-down* and *transfer*, and by now is standard. Thus, there is no point to transfer all the results in earlier sections to the large finite setting, since we know that this is always possible and no additional scientific significance is added. We may also point out that when an exact result in the measure-theoretic setting is reinterpreted into the discrete case, much mathematical elegance may be lost in this process of translation (though the scientific meaning is still retained). This may partially explain why the type of discrete results in this paper were not considered before in the literature.

To conclude this paper, we note that earlier applications of nonstandard methods and their different variations (see, for example, [AFHL], [HL], [K1], and [K2]) usually focus on obtaining exact results via either internal or large finite approximations, where the main concern is often on existence issues. On the contrary, our focus here is to obtain qualitative properties of general processes and then to pass the properties to triangular arrays of random variables automatically. To put it in a different way, we derive "difficult" approximate results for the large finite case from "easy" exact results in the limit model rather than the other way around. In some sense, this is the opposite to the usual approach of using Loeb measures.

References

[AFHL] Albeverio, S., Fenstad, J.E., Hoegh-Krohn, R. and Lindstrom, T.L. *Nonstandard Methods in Stochastic Analysis and Mathematical Physics*, Academic Press, Orlando, Florida, 1986.

[A] Aldous, D.J. Exchangeability and related topics, in *École d'Été de Probabilités de Saint-Flour XIII – 1983* (P. L. Hennequin ed.), Lectures Notes in Mathematics, Vol. 1117, Springer-Verlag, Berlin, 1985, pp.1–198.

[A1] Anderson, R.M. A nonstandard representation for Brownian motion and Ito integration, *Israel Journal of Mathematics* **25** (1976), 15–46.

[A2] Anderson, R.M. Non-standard analysis with applications to economics, in *Handbook of Mathematical Economics* **IV** (W. Hildenbrand and H. Sonnenschein eds.), North-Holland, New York, 1991.

[Bi] Billingsley, P. *Convergence of Probability Measures*, Wiley, New York, 1968.

[BR] Brown, D.J. and Robinson, A. A limit theorem on the cores of large standard exchange economies, *Proceedings of the National Academy of Sciences USA* **69** (1972), 1258–1260.

[CT] Chow, Y.H. and H. Teicher, H. *Probability Theory: independence, interchangeability and martingales*, Springer-Verlag, New York, 1978.

[C] Cohn, D.J. *Measure Theory*, Birkhäuser, Boston, 1980.

[Cu] Cutland, N.J. Nonstandard measure theory and its applications, *Bulletin of London Mathematical Society* **15** (1983), 529–589.

[D] Doob, J.L. *Stochastic Processes*, Wiley, New York, 1953.

[Du] Durrett, R. *Probability: theory and examples*, Wadsworth, Belmont, California, 1991.

[FG] Feldman, M. and Gilles, C. An expository note on individual risk without aggregate uncertainty, *Journal of Economic Theory* **35** (1985), 26–32.

[F] Feller, W. *An Introduction to Probability Theory and Its Applications*, 3rd Edition, Wiley, New York, 1968.

[H] Henson, C.W. The existence of an exchangeable sequence among a continuum of random variables, private communication.

[Ho] Hoover, D.N. Row-column exchangeability and a generalized model for probability, in *Exchangeability in Probability and Statistics* (G. Koch and F. Spizzichino eds.), North-Holland, 1982.

[HL] Hurd, A.E. and Loeb, P.A. *An Introduction to Nonstandard Real Analysis*, Academic Press, Orlando, Florida, 1985.

[K1] Keisler, H.J. Hyperfinite model theory, in *Logic Colloquium 76* (R. O. Gandy and J. M. E. Hyland eds.) North-Holland, Amsterdam, 1977.

[K2] Keisler, H.J. Neometric spaces, *Advances in Mathematics* **118** (1996), 134–175.

[KS] Khan, M.A. and Sun, Y.N. The capital asset pricing model and arbitrage pricing theory: a unification, *Proceedings of the National Academy of Sciences USA* **94** (1997), 4229–4232.

[K] Kingman, J.F.C. Uses of exchangeability, *Annals of Probability* **6** (1978), 183–197.

[L1] Loeb, P.A. Conversion from nonstandard to standard measure spaces and applications in probability theory, *Transactions of the American Mathematical Society* **211** (1975), 113–122.

[L2] Loeb, P.A. A nonstandard functional approach to Fubini's theorem, *Proceedings of the American Mathematical Society* **93** (1985), 343–346.

[Lo] Loéve, M. *Probability Theory I*, 4th Edition, Springer-Verlag, New York, 1977.

[S1] Sun, Y.N. Hyperfinite law of large numbers, *Bulletin of Symbolic Logic* **2** (1996), 189–198.

[S2] Sun, Y.N. A theory of hyperfinite processes: the complete removal of individual uncertainty via exact LLN, *Journal of Mathematical Economics* **29** (1998), 419–503.

[S3] Sun, Y.N. Conditional distributions for hyperfinite processes, preliminary draft, National University of Singapore and Cowles Foundation, 1997.

[S4] Sun, Y.N. The complete removal of individual uncertainty: multiple optimal choices and random exchange economies, preprint, National University of Singapore and Cowles Foundation, 1997

[S5] Sun, Y.N. On the Lebesgue measurability of sample functions for a continuum of iid random variables (to be published).

[WSS] Wang, Y.H., Stoyanov, J. and Shao, Q.M. On independence and dependence properties of a set of random events, *The American Statistician* **47** (1993), 112–115.