Probab. Theory Relat. Fields 118, 211– 250 (2000)
Digital Object Identifier (DOI) 10.1007/s004400000086

Olav Kallenberg

Spreading-invariant sequences and processes on bounded index sets

Received: 9 November 1999 / Revised version: 16 March 2000 /
Published online: 18 September 2000 –

Abstract. We say that a random sequence is *spreadable* if all subsequences of equal length have the same distribution. For infinite sequences the notion is equivalent to exchangeability but for finite sequences it is more general. The present paper is devoted to a systematic study of finite spreadable sequences and of processes on [0, 1] with spreadable increments. In particular, we show how many basic results in the exchangeable case—notably the predictable sampling theorem, the Wald-type identities, and various martingale and weak convergence results—admit extensions to a spreadable setting. We also identify some additional conditions that ensure the exchangeability of a spreadable sequence or process.

1. Introduction

A finite sequence of random elements $\xi = (\xi_1, \dots, \xi_n)$ in some measurable space $(S, \mathscr{S})$ is said to be *spreading-invariant in distribution* or simply *spreadable* if for any $m < n$ we have

$$(\xi_{k_1}, \dots, \xi_{k_m}) \stackrel{d}{=} (\xi_1, \dots, \xi_m), \quad 1 \le k_1 < \dots < k_m \le n. \tag{1}$$

This should be compared with the stronger condition of *exchangeability*, where (1) is required for all sets of distinct (but not necessarily increasing) indices $k_1, \dots, k_m \in \{1, \dots, n\}$. Note that (1) follows by induction from the more primitive condition

$$(\xi_1, \dots, \xi_{k-1}, \xi_{k+1}, \dots, \xi_n) \stackrel{d}{=} (\xi_1, \dots, \xi_{n-1}), \quad k = 1, \dots, n. \tag{2}$$

An infinite random sequence $\xi = (\xi_1, \xi_2, \dots)$ is said to be exchangeable or spreadable if every finite subsequence has this property. Ryll-Nardzewski (1957) showed that, in the infinite case, the two notions are in fact equivalent, so that by de Finetti's theorem an infinite spreadable sequence in a Borel space is mixed i.i.d. (cf. Kallenberg (1997), Theorem 9.16). The mentioned equivalence fails the

O. Kallenberg: Department of Mathematics, Auburn University, 218 Parker Hall, Auburn, AL 36849-5310, USA. e-mail: `olavk@mail.auburn.edu`

Mathematics Subject Classification (2000): Primary 60G09; Secondary 28D05, 60F17, 60G40, 60G44, 60G57, 60H05

Key words and phrases: Exchangeable or spreadable sequences and processes – Extremality and integral representations – Martingale characteristics – Regularization – Weak convergence – Predictable sampling – Wald-type identities

finite case, simply because (as Kingman (1978) first observed) the spreadability of a distribution on S^n puts restrictions only on the $n-1$-dimensional marginals, which do not determine the full distribution in general. Some interesting examples are given below.

The situation in continuous time is similar. Here the exchangeability and spreadability are defined in terms of the increments. Thus, we say that an $\mathbb{R}^d$-valued process X on $[0, 1]$ has *exchangeable or spreadable increments* if the sequence $\xi_{nj} = X_{j/n} - X_{(j-1)/n}$, $j = 1, \ldots, n$, is exchangeable or spreadable for every $n \in \mathbb{N}$. If we assume in addition that X is continuous in probability, then by Theorem 4.7 below it has a version that is *right-continuous with left-hand limits* (*rcll* for short). This justifies our definition of an exchangeable or spreadable process on $[0, 1]$ as one with exchangeable or spreadable increments, rcll paths, and initial value 0. A process on $\mathbb{R}_+$ is said to be exchangeable or spreadable if the same properties hold on every finite subinterval $[0, t]$.

An alternative approach to spreadability in continuous time is suggested by (2). Here we consider, for any times $a < b$ in $[0, 1]$, the process

$$\widehat{X}_{a,b}(t) = X_{t\wedge a} + X_{b+(t-a)_+} - X_b, \quad t \in [0, 1-b+a], \tag{3}$$

and note that X is spreadable iff it is continuous in probability and satisfies $\widehat{X}_{a,b} \overset{d}{=} X$ on $[0, 1-b+a]$ for arbitrary $a < b$. Iterating (3) in finitely many steps yields an equivalent condition corresponding to (1).

For a more explicit statement of the latter condition, we may identify X with the associated finitely additive set function on the class $\mathcal{U}$ of finite interval unions $\bigcup_i (s_i, t_i]$, given by $X(s, t] = X_t - X_s$. For any $U \in \mathcal{U}$ we put $X_U(t) = X(U_t)$ and $\lambda_U(t) = \lambda(U_t)$ where $U_t = U \cap (0, t]$ and λ denotes Lebesgue measure on $[0, 1]$. Next we define $\widehat{X}_U = X_U \circ \lambda_U^{-1}$ or, in integral notation,

$$\widehat{X}_U(t) = \int_U 1\{\lambda_U(s) \le t\}\, dX_s, \quad 0 \le t \le \lambda U, \tag{4}$$

where $1\{\cdot\}$ denotes the indicator function of the set within brackets. (Note that the integral in (4) is elementary since $U \cap 1\{\lambda_U \le t\} \in \mathcal{U}$ for all t.) In particular, $\widehat{X}_U = \widehat{X}_{a,b}$ when $U = (0, a] \cup (b, 1]$. We may now state the continuous-time counterpart of (1) in the form $\widehat{X}_U \overset{d}{=} X$ on $[0, \lambda U]$, where $U \in \mathcal{U}$ is arbitrary.

The present paper is devoted to a systematic study of finite spreadable sequences and of spreadable processes on $[0, 1]$. Such a study is interesting for several reasons: 1) Many results, previously known in a more special context, are best understood in the present generality. This is especially true for some basic martingale properties, the optional skipping property, and the Wald-type identities, all known for exchangeable processes. 2) The exchangeability of a sequence or process can be deduced from the weaker hypothesis of spreadability together with a variety of additional constraints. Thus, the present theory contributes to our understanding of exchangeable objects (whose importance is more generally acknowledged). 3) Many problems associated with spreadable sequences and processes are more challenging than their exchangeble counterparts (basically because fewer symmetries are available in the spreadable case), and their solution often leads to results that are both attractive and surprising. (Here the weak convergence theory of Section 4 might qualify as an example.) Thus, the present theory has arguably a considerable intrinsic interest. 4) Often the hard problems of the area force us to develop new tools and techniques that may be of some independent interest. A case in point is our construction in Section 5 of a new stochastic integral, which is needed already for the *formulation* of the predictable sampling theorem in continuous time.

Let us summarize briefly some highlights of the paper. We begin in Section 2 with an integral representation of the general spreadable distribution in terms of extreme points. In the same section we examine the relationship between spreadable sequences and processes. In Section 3 we show, under a moment condition, that spreadable processes are special semimartingales whose local characteristics are themselves martingales. Here we also consider some additional conditions that ensure the exchangeability of a spreadable process. Section 4 is devoted to the weak convergence theory of spreadable sequences and processes. Here our first key result guarantees the existence, for every spreadable process, of an exchangeable process with the same characteristics. This fundamental correspondence allows us in the next step to extend the basic tightness criteria for exchangeable processes to a spreadable context. With those two results as our main tools, we proceed to derive the basic regularization theorem and various convergence criteria and norm relations. Our final Sections 5 and 6 contain the ultimate versions of the predictable sampling (or optional skipping) theorem and the Wald-type identities for spreadable sequences and processes. The latter are both interesting and powerful results, whose history goes back to some classical statements for i.i.d. sequences due to Doob (1936) and Wald (1945), respectively.

Since there is no general representation formula in the spreadable case, the construction of nontrivial and interesting examples requires both ingenuity and some calculation. To indicate the possibilities, we give two examples of extreme, spreadable (but not exchangeable) distributions on $\{0, 1, 2\}^3$. In other words, these are distributions of spreadable sequences of length 3 with values in the set $\{0, 1, 2\}$:

001	012	020	102	120	122	200	202	210	221
2	2	2	1	1	1	2	1	1	1

012	021	101	102	110	120	122	202	210	211	221
2	2	2	1	2	1	1	1	1	2	1

In each table, the first row gives the possible configurations and the second one shows the corresponding probabilities, up to a normalization. (Thus, the probabilities in the first example are $1/7$ or $1/14$ and in the second example $1/8$ or $1/16$.) To verify the extremality, we note that every component, in the sense of convex combinations, is supported by the same set of configurations. The associated probabilities are then uniquely determined by the spreadability constraints. Other examples are obtained by reversing the sequences or by permuting the symbols 0, 1, and 2.

We are immediately struck by the peculiar lack of symmetry or simple pattern. It is indeed remarkable that so many wondrous properties are hidden behind such an apparent complexity and disorder. The present irregularity is in sharp contrast to the trite symmetry in the exchangeable case. Thus, the only ergodic exchangeable distributions on $\{0,1,2\}^3$ are of the form (apart from permutations of the digits)

000
1

001	010	100
1	1	1

012	021	102	120	201	210
1	1	1	1	1	1

Much of the present material is related to some earlier ideas and results of the author, especially from Kallenberg (1973, 1982, 1988a, 1989). We may also call attention to the crucial role of the spreadability concept in various other areas, such as for the subsequence principle (cf. Aldous (1985), Section 8) and for certain higher-dimensional symmetries (cf. Kallenberg (1992)). For a general introduction to exchangeability theory, we recommend Aldous' (1985) lecture notes, supplemented by the relevant portions of Kallenberg (1997), especially the concluding pages of Chapters 9 and 14. Constant use will be made of some basic notions and results on weak convergence, semimartingales, and stochastic integration, for which we refer to the relevant chapters in Jacod and Shiryaev (1987) and Kallenberg (1997).

For the reader's convenience, we review the basic representations in the exchangeable case (cf. Kallenberg (1997), Theorems 9.16–17, 9.21, 14.25). We have already quoted the de Finetti–Ryll-Nardzewski theorem, the fact that any infinite exchangeable or spreadable sequence in a Borel space is mixed i.i.d. The continuous-time analogue is *Bühlmann's theorem*, which states that any exchangeable or spreadable process on $\mathbb{R}_+$ is a mixture of Lévy processes. Next we note that any finite exchangeable sequence is a mixture of so-called *urn sequences*, which can be generated by drawing without replacement from an urn with finitely many tickets.

The final and most difficult case is that of exchangeable processes X on $[0,1]$ taking values in $\mathbb{R}^d$. Here the general distribution may be described through the representation formula

$$X_t = \alpha t + \sigma B_t + \sum\nolimits_j \beta_j (1\{\tau_j \le t\} - t), \quad t \in [0,1], \tag{5}$$

where B is a d-dimensional Brownian bridge, the variables $\tau_1, \tau_2, \ldots$ are independent of B and i.i.d. $U(0,1)$ (uniformly distributed on $[0,1]$), and the vector- or matrix-valued coefficients α, σ, and $\beta_1, \beta_2, \ldots$ are independent of $(B, \{\tau_j\})$ and such that $\sum_j |\beta_j|^2 < \infty$ a.s. The series in (5) then converges a.s., uniformly on $[0,1]$, which ensures that X will have a.s. rcll paths.

The distribution of the process X in (5) (often written as $\mathscr{L}(X)$) determines (and is determined by) that of the triple (α, β, γ), where the point process β on $\mathbb{R}^d \setminus \{0\}$ and the random $d \times d$ matrix γ are given by

$$\beta = \sum\nolimits_j \delta_{\beta_j}, \qquad \gamma = \sigma\sigma' + \sum\nolimits_j \beta_j \beta_j'.$$

(Here δ_x denotes the unit mass at x and the prime denotes transposition.) The correspondence $\mathscr{L}(X) \leftrightarrow \mathscr{L}(\alpha, \beta, \gamma)$ is even a homeomorphism with respect to Skorohod's J_1 topology on $D(\mathbb{R}_+, \mathbb{R}^d)$ and the vague topology on $\mathscr{M}(\mathbb{R}^d \setminus \{0\})$. (Those are the spaces of rcll functions $[0,1] \to \mathbb{R}^d$ and of locally finite measures on

$\mathbb{R}^d \setminus \{0\}$, respectively.) We finally note that X is a semimartingale with covariation matrix $[X]_1 = \gamma$.

We conclude this section with some remarks on notation. All random objects are assumed to be defined on some abstract probability space Ω with probability measure P and associated expectation E, often also equipped with a discrete or continuous filtration $\mathscr{F} = (\mathscr{F}_k)$ or $(\mathscr{F}_t)$. Independence is expressed by the symbol $\perp\!\!\!\perp$ and equality in distribution by $\stackrel{d}{=}$. The arrows $\stackrel{P}{\to}$ and $\stackrel{d}{\to}$ indicate convergence in probability or distribution, respectively, and we write $\stackrel{fd}{\longrightarrow}$ for convergence of the finite-dimensional distributions. The space of all rcll maps $T \to S$ is denoted by $D(T, S)$, and we write $\mathscr{M}(S)$ for the space of locally finite measures on S and $\mathscr{M}_1(S)$ for the subspace of probability measures on S. If nothing else is said, these spaces are endowed with the Skorohod J_1 topology and the vague and weak topologies, respectively. We also write $\mathscr{B}(S)$ for the class of Borel sets in S.

For any $\mathbb{R}^d$-valued process X, we define $X_t^* = \sup_{s \le t} |X_s|$ and $X^* = \sup_t |X_t| = \sup_t X_t^*$. If X is a semimartingale, then $[X]_t$ denotes the matrix-valued covariation process with components $[X^i, X^j]_t$ for $i, j \le d$. The symbol Δ will be used both for symmetric differences of sets and for jumps of processes. We shall often write $V \cdot X$ for the integral process $\int_0^t V dX$ and put $\mu f = \int f d\mu$ when μ is a measure. In Lebesgue integrals we may omit the integrator from our notation and write $\int V$ as short for $\int V_s ds$ or $V \cdot \lambda$. The symbol $\otimes$ is used for both product σ-fields and product measures. Shift operators in discrete or continuous time are written as θ_k or θ_t.

Superscripts will often be employed as indices, rarely as exponents, and we may occasionally write $X_n(t)$ as X_t^n for convenience. The relation $x \lesssim y$ or $y \gtrsim x$ means by definition that $x \le cy$ for some constant c; if even $x \gtrsim y$ we may write $x \asymp y$. Either relation is said to be *uniform* in a parameter t if the relevant constants can be chosen to be independent of t. Finally, we adopt the conventions $\mathbb{R}_+ = [0, \infty)$, $\mathbb{N} = \{1, 2, \ldots\}$, and $\mathbb{Z}_+ = \{0, 1, \ldots\}$. Some more special notation will be explained when it first occurs.

2. Extremality and jump structure

In this section we prove some general integral representations of spreadable distributions in terms of extreme points, examine the relationship between spreadable sequences and processes, and consider some simple cases where the spreadability of a sequence or process implies its exchangeability.

To motivate our first topic, we note that any mixture of spreadable distributions is again spreadable. Here we address the reverse problem of decomposing a spreadable distribution into extreme distributions of the same type. Recall that a spreadable distribution μ is said to be *extreme* if any relation $\mu = c\mu_1 + (1-c)\mu_2$ with $c \in (0, 1)$ and spreadable μ_1 and μ_2 implies $\mu_1 = \mu_2 = \mu$. For *infinite* spreadable sequences, a unique representation of the mentioned type is given by the de Finetti–Ryll-Nardzewski theorem, where the i.i.d. distributions play the role of extremal elements. Similar representations for finite exchangeable sequences and for exchangeable processes on $[0, 1]$ or $\mathbb{R}_+$ are implicitly given by the characterizations described in Section 1.

In each of those four basic cases, the exchangeability may be characterized as invariance in distribution under a suitable group of measurable transformations of the underlying space. In such a situation, a general theory is available that guarantees the existence of a unique integral representation (cf. Dynkin (1978) and Aldous (1985), Section 12). Though there is no such description in the general spreadable case, we can still prove the existence of a general integral representation over extreme points. However, the corresponding uniqueness fails in general, as we shall see in Corollary 2.6 below.

Theorem 2.1. *In each of the following cases, the distribution of ξ or X is a mixture of extreme distributions of the same type:*

(i) *ξ is a finite, spreadable sequence in some Borel space S;*
(ii) *X is an $\mathbb{R}^d$-valued, spreadable process on* $[0, 1]$.

In the proof we shall refer to Lemma 4.5 below, which is permissible since no subsequent results depend on the present theorem. A similar remark applies to the use of Theorem 3.5 in the proof of Lemma 2.3 below.

Proof. (i) Embedding S as a Borel set in $[0, 1]$, we may regard ξ as a spreadable random sequence in $[0, 1]$. The space $\mathcal{M}_1([0, 1]^n)$ of probability measures on $[0, 1]^n$ is compact and metrizable (cf. Rogers and Williams (1994), Theorem 81.3), and we also note that the subset K of spreadable distributions on $[0, 1]^n$ is convex and closed, hence compact. By a standard form of Choquet's theorem (cf. Alfsen (1971)), the element $\mathcal{L}(\xi)$ has then an integral representation

$$P\{\xi \in B\} = \int \mu(B)\, \nu(d\mu), \quad B \in \mathcal{B}([0, 1]^n), \tag{1}$$

in terms of some probability measure ν on the set $\operatorname{ex} K$ of extreme elements of K. In particular, we obtain $\nu\{\mu;\ \mu S^n < 1\} = 0$. Letting $\tilde{\nu}$ denote the image of ν under the restriction map $\mu \mapsto \mu|_{S^n}$, we get for any $B \in \mathcal{B}(S^n)$

$$P\{\xi \in B\} = \int \mu|_{S^n}(B)\, \nu(d\mu) = \int \mu'(B)\, \tilde{\nu}(d\mu'). \tag{2}$$

It remains to note that, if $\mu \in \operatorname{ex} K$ is restricted to S^n, then $\mu|_{S^n}$ is an extreme, spreadable distribution on S^n.

(ii) The space $S = D([0, 1], \mathbb{R}^d)$ is Polish in the Skorohod topology (cf. Jacod and Shiryaev (1987), Theorem VI.1.14), and so it may be embedded as a Borel subset of a compact metric space J (cf. Rogers and Williams (1994), Theorem 82.5). The space $\mathcal{M}_1(J)$ is again compact and metrizable, and $\mathcal{M}_1(S)$ can be identified with the subset $\{\mu \in \mathcal{M}_1(J);\ \mu S = 1\}$ (op. cit., Theorem 83.7).

Now let K denote the convex set of all spreadable distributions on S, and note that K remains convex as a subset of $\mathcal{M}_1(J)$. The closure $\overline{K}$ in $\mathcal{M}_1(J)$ is again convex and also compact. Thus, Choquet's theorem yields an integral representation as in (1), where B is now an arbitrary Borel set in J and ν is a probability measure on $\operatorname{ex} \overline{K}$.

Since $\mathscr{L}(X)$ is restricted to S, we may proceed as before to derive a representation as in (2), where $\tilde{\nu}$ is now the image of ν under the restriction map $\mu \mapsto \mu|_S$. It remains to show that, if $\mu \in \text{ex}\,\overline{K}$ with $\mu S = 1$, then $\mu|_S \in \text{ex}\,K$. But this is clear, since K is closed in $\mathscr{M}_1(S)$ by Lemma 4.5 below and therefore $\overline{K} \cap \mathscr{M}_1(S) = K$. □

In Section 1 we saw that a finite, spreadable sequence need not be exchangeable. The equivalence of the two notions will now be established under an additional assumption. Further result of this type are given in Lemma 2.4 and Theorem 3.5 below. In the exchangeable case, we note that a distribution is extreme iff it is *ergodic*, in the sense that the invariant σ-field is trivial.

Lemma 2.2. *Let $\xi = (\xi_1, \ldots, \xi_n)$ be a spreadable sequence in some measurable space S such that the measure $\beta = \sum_j \delta_{\xi_j}$ is a.s. nonrandom. Then ξ is ergodic exchangeable.*

Proof. Introduce an exchangeable permutation $\tilde{\xi}_1, \ldots, \tilde{\xi}_n$ of $\xi_1, \ldots, \xi_n$, let $(\mathscr{F}_k)$ and $(\tilde{\mathscr{F}}_k)$ be the filtrations induced by the two sequences, and put $\beta_k = \sum_{j \le k} \delta_{\xi_j}$ and $\tilde{\beta}_k = \sum_{j \le k} \delta_{\tilde{\xi}_j}$. We shall prove by induction that

$$(\xi_1, \ldots, \xi_k) \stackrel{d}{=} (\tilde{\xi}_1, \ldots, \tilde{\xi}_k), \quad k = 0, \ldots, n. \tag{3}$$

This is vacuously true for $k = 0$, and for $k = n$ it implies the asserted statement.

Now assume that (3) holds for some fixed $k < n$. Since $(\xi_1, \ldots, \xi_k, \xi_m)$ has the same distribution for all $m > k$, we get for any measurable function $g \ge 0$ on S

$$\begin{aligned} E[g(\xi_{k+1})|\mathscr{F}_k] &= \cdots = E[g(\xi_n)|\mathscr{F}_k] \\ &= (n-k)^{-1} E\Big[\sum\nolimits_{m>k} g(\xi_m)\Big|\, \mathscr{F}_k\Big] \\ &= (n-k)^{-1}(\beta - \beta_k)g. \end{aligned}$$

A similar relation holds for $\tilde{\xi}_{k+1}$, $\tilde{\mathscr{F}}_k$, and $\tilde{\beta}_k$, since the basic hypotheses remain fulfilled for the sequence $(\tilde{\xi}_j)$ with the same nonrandom measure β. Using the induction hypothesis, we get for any measurable function $f \ge 0$ on S^k

$$\begin{aligned} Ef(\xi_1, \ldots, \xi_k) g(\xi_{k+1}) &= E(f(\xi_1, \ldots, \xi_k) E[g(\xi_{k+1})|\mathscr{F}_k]) \\ &= (n-k)^{-1} Ef(\xi_1, \ldots, \xi_k)(\beta - \beta_k)g \\ &= (n-k)^{-1} Ef(\tilde{\xi}_1, \ldots, \tilde{\xi}_k)(\beta - \tilde{\beta}_k)g \\ &= E(f(\tilde{\xi}_1, \ldots, \tilde{\xi}_k) E[g(\tilde{\xi}_{k+1})|\tilde{\mathscr{F}}_k]) \\ &= Ef(\tilde{\xi}_1, \ldots, \tilde{\xi}_k) g(\tilde{\xi}_{k+1}), \end{aligned}$$

which proves (3) with k replaced by $k + 1$. This completes the induction, and the assertion follows. □

The notion of extremality clearly depends on the underlying symmetry. The next result relates the exchangeable and spreadable versions.

Lemma 2.3. *The distribution of an exchangeable sequence or process is ergodic iff it is extreme in the spreadable sense.*

Proof. The sufficiency is obvious, since every exchangeable sequence or process is also spreadable. Now consider the distribution μ of some ergodic, exchangeable sequence $\xi_1, \dots, \xi_n$ in some space S, and note that μ is a.s. restricted to the permutations of some fixed sequence $a_1, \dots, a_n \in S$. If μ is a convex combination of some spreadable distributions μ_1 and μ_2, then even the latter have the stated property. From Lemma 2.2 it follows that μ_1 and μ_2 are exchangeable, and since μ is extreme in the exchangeable sense, we get $\mu_1 = \mu_2$. This shows that μ remains extreme in the spreadable sense.

Next consider the distribution μ of some ergodic, exchangeable process X on $[0, 1]$, and note that X has finite first moments and a.s. fixed jump sizes and terminal value. If μ is a convex combination of some spreadable distributions μ_1 and μ_2, then even the latter measures have the stated properties. Hence, the μ_i are exchangeable by Theorem 3.5 below, and since μ is extreme in the exchangeable sense, it follows that $\mu_1 = \mu_2$. Thus, μ is again extreme in the spreadable sense. □

A random measure ξ on $[0, 1]$ is said to be spreadable if the corresponding *distribution function* $X_t = \xi[0, t]$ is a spreadable process on $[0, 1]$. By a *simple point process* we mean a purely atomic random measure ξ such that $\xi\{s\} = 0$ or 1 for all s. The next result shows that the notions of spreadability and exchangeability are equivalent for simple point processes and diffuse random measures. This prepares for our study of more general processes in Theorem 3.5. Parts (ii) and (iii) below are in fact equivalent to Lemma 3.4 in Kallenberg (1982); they are restated here with a short direct proof, for the convenience of the reader.

Lemma 2.4. *In each of the following cases, a spreadable random sequence or measure ξ is exchangeable:*

(i) *ξ is a finite sequence in* $\{0, 1\}$*;*
(ii) *ξ is a simple point process on* $[0, 1]$*;*
(iii) *ξ is a diffuse random measure on* $[0, 1]$.

With subsequent applications in mind, we note that a diffuse random measure ξ on $[0, 1]$ is exchangeable iff $\xi = \alpha\lambda$ a.s. for some random variable $\alpha \geq 0$. For a simple point process ξ on $[0, 1]$, exchangeability means that ξ is a (homogeneous) *mixed binomial* (or *sample) process,* so that $\xi = \sum_{k \leq \kappa} \delta_{\tau_k}$ for some i.i.d. $U(0, 1)$ random variables $\tau_1, \tau_2, \dots$ and some independent, $\mathbb{Z}_+$-valued random variable κ. The revised terminology is motivated by the fact that, for any $B \in \mathscr{B}([0, 1])$ and given κ, the random variable ξB is conditionally binomially distributed with parameters κ and λB.

Proof. (i) Let $\xi = (\xi_1, \dots, \xi_n)$ be spreadable in $\{0, 1\}$, fix any permutation $p = (p_1, \dots, p_n)$ of $1, \dots, n$, and put $\eta = (\xi_{p_1}, \dots, \xi_{p_n})$. Using the spreadability of ξ, we get for any subset $A \subset \{1, \dots, n\}$

$$\sum\nolimits_{j \in A} \eta_j = \sum\nolimits_{j \in p^{-1}A} \xi_j \overset{d}{=} \sum\nolimits_{j \leq |A|} \xi_j \overset{d}{=} \sum\nolimits_{j \in A} \xi_j .$$

Hence, $\xi \stackrel{d}{=} \eta$ by Theorem 10.9 in Kallenberg (1997), and the exchangeability follows since p was arbitrary.

(ii)–(iii) Let ξ be a spreadable simple point process or diffuse random measure on $[0, 1]$. Fix any $n \in \mathbb{N}$ and a permutation $p = (p_1, \ldots, p_n)$ of $1, \ldots, n$. Define a measure-preserving transformation f on $(0, 1]$ by

$$f(x) = x + n^{-1}(p_j - j), \quad x \in I_{nj} \equiv n^{-1}(j-1, j], \quad j = 1, \ldots, n,$$

and introduce the random measure $\eta = \xi \circ f^{-1}$ on $(0, 1]$. Using the spreadability of ξ, we get

$$\eta U = \xi(f^{-1}U) \stackrel{d}{=} \xi[0, \lambda U] \stackrel{d}{=} \xi U, \quad U \in \mathscr{U},$$

and so $\xi \stackrel{d}{=} \eta$ on $(0, 1]$ by a version of Theorem 10.9 in Kallenberg (1997). In particular, ξ has exchangeable increments over $I_{n,1}, \ldots, I_{n,n}$, and the asserted exchangeability follows since $\xi\{0\} = 0$ a.s. □

For a spreadable process on $[0, 1]$, it is not clear whether the jump structure can be described in terms of spreadable sequences in general. The following result gives a useful connection in a special case.

Theorem 2.5. *Let X be a step process in $\mathbb{R}^d$ with a fixed number of jumps and let ξ and η denote the associated jump size sequence and jump time process. Then*

(i) *X is spreadable iff ξ and η are independent and spreadable.*

In that case

(ii) *X and ξ are simultaneously extreme;*
(iii) *X and ξ are simultaneously exchangeable.*

Proof. (i) Assume that ξ and η are independent and spreadable. Fix any $U, V \in \mathscr{U}$ with $\lambda U = \lambda V$, and let $\xi'_1, \ldots, \xi'_{\eta U}$ and $\xi''_1, \ldots, \xi''_{\eta V}$ be the jump sizes of $\widehat{X}_U$ and $\widehat{X}_V$, respectively, enumerated from left to right. Since ξ is spreadable and independent of η, Fubini's theorem yields

$$P[(\xi'_1, \ldots, \xi'_k) \in \cdot \,|\, \eta] = P\{(\xi_1, \ldots, \xi_k) \in \cdot\} \text{ a.s. on } \{\eta U = k\}.$$

Combining with the same relation for the ξ''_j and noting that $\hat{\eta}_U \stackrel{d}{=} \hat{\eta}_V$ since η is spreadable, we obtain

$$(\hat{\eta}_U, \xi'_1, \ldots, \xi'_{\eta U}) \stackrel{d}{=} (\hat{\eta}_V, \xi''_1, \ldots, \xi''_{\eta V}),$$

which implies $\widehat{X}_U \stackrel{d}{=} \widehat{X}_V$. Thus, X is spreadable.

Conversely, assume that X is spreadable with n jumps, and let $U, V \in \mathscr{U}$ with $\lambda U = \lambda V$. Since $\widehat{X}_U \stackrel{d}{=} \widehat{X}_V$ and $\eta[0, 1] = n$, we get for any $B \in \mathscr{B}(\mathbb{R}^{nd})$

$$\begin{aligned} P\{\xi \in B,\ \eta U^c = 0\} &= P\{\xi \in B,\ \eta U = n\} \\ &= P\{\xi \in B,\ \eta V = n\} \\ &= P\{\xi \in B,\ \eta V^c = 0\}. \end{aligned}$$

Keeping B fixed with $P\{\xi \in B\} > 0$ and using a version of Theorem 10.9 in Kallenberg (1997), we conclude that η is conditionally spreadable, given the event $\{\xi \in B\}$. By Lemma 2.4, η is then exchangeable and hence a binomial process with n points. Since the conditional distribution is the same for all B, it follows that η is spreadable and independent of ξ.

To show that even ξ is spreadable, we may assume that $n > 0$. Fix any $a < b$ in $(0, 1)$ and put

$$I = (0, a], \qquad U = (0, a] \cup (b, 1], \qquad V = (0, 1 - b + a],$$

so that $\lambda U = \lambda V$. By the spreadability of X we have $\widehat{X}_U \stackrel{d}{=} \widehat{X}_V$, and then also

$$(\widehat{X}_U, \eta I, \eta U) \stackrel{d}{=} (\widehat{X}_V, \eta I, \eta V),$$

In particular, we get for any $k \in \{0, \ldots, n-1\}$

$$(\xi_1, \ldots, \xi_k, \xi_{k+2}, \ldots, \xi_n)\, 1\{\eta I = k, \eta U = n-1\}$$
$$\stackrel{d}{=} (\xi_1, \ldots, \xi_{n-1})\, 1\{\eta I = k, \eta V = n-1\}.$$

Since $\xi \perp\!\!\!\perp \eta$, and also

$$P\{\eta I = k, \eta U = n-1\} = P\{\eta I = k, \eta V = n-1\} > 0$$

by the spreadability of η, we obtain

$$(\xi_1, \ldots, \xi_k, \xi_{k+2}, \ldots, \xi_n) \stackrel{d}{=} (\xi_1, \ldots, \xi_{n-1}), \quad 0 \le k < n,$$

which implies the required spreadability of ξ.

(ii) Every distribution $\mu = \mathscr{L}(\xi)$ on $\mathbb{R}^{nd}$ determines uniquely a corresponding distribution $\tilde{\mu} = \mathscr{L}(X)$ on $D([0, 1], \mathbb{R}^d)$. The mapping $\mu \mapsto \tilde{\mu}$ is clearly linear and injective, and the measures μ and $\tilde{\mu}$ are simultaneously spreadable. Now assume that $\tilde{\mu}$ is extreme, and let $\mu = c\mu_1 + (1-c)\mu_2$ for some spreadable probabilities μ_1, μ_2 and some constant $c \in (0, 1)$. Then $\tilde{\mu} = c\tilde{\mu}_1 + (1-c)\tilde{\mu}_2$, and since $\tilde{\mu}$ is extreme and $\tilde{\mu}_1$ and $\tilde{\mu}_2$ are spreadable, we get $\tilde{\mu}_1 = \tilde{\mu}_2$. Thus, $\mu_1 = \mu_2$, which shows that even μ is extreme. The converse implication follows by the same argument since the inverse mapping $\tilde{\mu} \mapsto \mu$ is again linear.

(iii) Let $\xi = (\xi_1, \ldots, \xi_n)$ be exchangeable. Since η is a binomial process independent of ξ, the transfer Theorem 5.10 in Kallenberg (1997) ensures the existence of some i.i.d. $U(0, 1)$ random variables $\sigma_1, \ldots, \sigma_n \perp\!\!\!\perp \xi$ with $\eta = \sum_j \delta_{\sigma_j}$. Writing $\tau_1, \ldots, \tau_n$ for the increasing enumeration of the σ_j, we have $\sigma_j \equiv \tau_{\pi_j}$ for some (σ_j)-measurable permutation $(\pi_1, \ldots, \pi_n)$ of $1, \ldots, n$. By exchangeability and independence, the sequences ξ and $(\xi_{\pi_1}, \ldots, \xi_{\pi_n})$ are equally distributed, conditionally on (σ_j), and therefore $(\xi_{\pi_j}) \perp\!\!\!\perp (\sigma_j)$. To see that X is exchangeable, it remains to write

$$X_t = \sum\nolimits_{k \le n} \xi_k 1\{\tau_k \le t\} = \sum\nolimits_{j \le n} \xi_{\pi_j} 1\{\sigma_j \le t\}, \quad t \in [0, 1].$$

Conversely, assume that X is exchangeable. To show that this is also true for ξ, we may reduce by conditioning to the case when $\sum_k \delta_{\xi_k}$ is nonrandom. But then ξ is exchangeable by Lemma 2.2. □

The last result can be used to extend some counterexamples for spreadable sequences to the continuous-time case.

Corollary 2.6. *There exist some spreadable sequences* ξ *and processes* X *such that*

(i) ξ *and* X *are not exchangeable;*

(ii) *the extremal representations of* $\mathscr{L}(\xi)$ *and* $\mathscr{L}(X)$ *are not unique.*

Proof. Consider on $\{1, 2, 3\}^2$ the extreme, spreadable distributions

$$\mu_1 = \boxed{\begin{array}{ccc} 12 & 23 & 31 \\ 1 & 1 & 1 \end{array}}, \qquad \mu_2 = \boxed{\begin{array}{ccc} 32 & 21 & 13 \\ 1 & 1 & 1 \end{array}},$$

$$\nu_1 = \boxed{\begin{array}{cc} 12 & 21 \\ 1 & 1 \end{array}}, \qquad \nu_2 = \boxed{\begin{array}{cc} 23 & 32 \\ 1 & 1 \end{array}}, \qquad \nu_3 = \boxed{\begin{array}{cc} 31 & 13 \\ 1 & 1 \end{array}},$$

and note that μ_1 and μ_2 are not exchangeable. Furthermore, $\frac{1}{2}(\mu_1 + \mu_2) = \frac{1}{3}(\nu_1 + \nu_2 + \nu_3)$. To get a distribution of a spreadable sequence that satisfies both (i) and (ii), we may take $\mu = \frac{1}{3}(\mu_1 + 2\mu_2)$. By Theorem 2.5, conditions (i) and (ii) remain fulfilled for the corresponding continuous-time distribution $\tilde{\mu}$. □

When the space S is finite, the distributions of all spreadable sequences in S of length n form a convex polyhedron in an appropriate affine subspace of S^n. In fact, the sample space is then finite and the spreadability condition yields finitely many linear constraints on the corresponding probabilities. Hence, in this case there are finitely many extreme distributions, each corresponding to a vertex of the mentioned polyhedron.

In the special case of random pairs, we note that a distribution is extreme, spreadable iff it gives equal weight m^{-1} to all pairs (a_k, a_{k+1}) for some distinct elements $a_1, \ldots, a_m \in S$, where a_{m+1} is interpreted as a_1. These distributions are clearly exchangeable only for $m = 0$ and 1.

3. Martingale methods

In this section we examine some basic martingale properties of spreadable processes and exhibit conditions that ensure the exchangeability of a spreadable process. Our results extend and improve some statements for exchangeable processes in Kallenberg (1982, 1988a).

We may relate the spreadability of a finite or infinite sequence $\xi = (\xi_1, \xi_2, \ldots)$ to a discrete filtration $\mathscr{F} = (\mathscr{F}_0, \mathscr{F}_1, \ldots)$. Then say that ξ is $\mathscr{F}$*-spreadable* if it is adapted to $\mathscr{F}$ and such that the shifted sequence $\theta_k \xi = (\xi_{k+1}, \xi_{k+2}, \ldots)$ is conditionally spreadable given $\mathscr{F}_k$ for every $k \geq 0$. To avoid requiring the existence of conditional distributions, we may state our condition in terms of elementary conditional probabilities, given any set $A \in \mathscr{F}_k$ with $PA > 0$. Note that any spreadable sequence ξ is spreadable for the induced filtration $\mathscr{F}_k = \sigma\{\xi_j;\ j \leq k\}$.

An infinite sequence ξ is said to be *strongly stationary* or *$\mathscr{F}$-stationary* if $\theta_\tau \xi \overset{d}{=} \xi$ for every finite optional (or stopping) time $\tau \geq 0$. For finite sequences $\xi = (\xi_1, \ldots, \xi_n)$, we interpret the condition as

$$(\xi_{\tau+1}, \ldots, \xi_{\tau+k}) \overset{d}{=} (\xi_1, \ldots, \xi_k) \text{ whenever } \tau + k \leq n \text{ a.s.}$$

We finally consider the martingale property of the so-called *prediction sequence* $\mu_k = P[\theta_k \xi \in \cdot | \mathscr{F}_k]$, $k \geq 0$. Again we may avoid regularity requirements by stating our condition in terms of elementary conditional probabilities. Thus, for infinite sequences, we say that the μ_k form an *$\mathscr{F}$-martingale* if $\theta_{k+1}\xi \overset{d}{=} \theta_k \xi$ *over* $\mathscr{F}_k$ for all $k \geq 0$, in the sense that

$$P[\theta_{k+1}\xi \in \cdot; A] = P[\theta_k \xi \in \cdot; A], \quad A \in \mathscr{F}_k, \ k \geq 0.$$

For finite sequences $\xi = (\xi_1, \ldots, \xi_n)$, the martingale condition is interpreted as

$$(\xi_{k+2}, \ldots, \xi_n) \overset{d}{=} (\xi_{k+1}, \ldots, \xi_{n-1}) \text{ over } \mathscr{F}_k, \quad k = 0, \ldots, n-2.$$

The mentioned conditions are related by the following result, which extends the corresponding statement for infinite exchangeable sequences in Kallenberg (1982, 1988a) (cf. Kallenberg (1997), Proposition 9.18).

Lemma 3.1. *Let $\xi = (\xi_1, \xi_2, \ldots)$ be a finite or infinite, $\mathscr{F}$-adapted random sequence in some measurable space S. Then these conditions are equivalent:*

(i) *ξ is $\mathscr{F}$-spreadable;*
(ii) *ξ is $\mathscr{F}$-stationary;*
(iii) *$\mu_k = P[\theta_k \xi \in \cdot | \mathscr{F}_k]$ forms an $\mathscr{F}$-martingale.*

Proof. The proof for infinite sequences carries over with obvious changes. □

We turn to the continuous-time case. An $\mathbb{R}^d$-valued process X on $[0, 1]$ or $\mathbb{R}_+$ is said to be spreadable with respect to a filtration $\mathscr{F} = (\mathscr{F}_t)$ or simply *$\mathscr{F}$-spreadable* if it is $\mathscr{F}$-adapted with $X_0 = 0$ and such that the shifted process $\theta_s X - X_s$ is conditionally spreadable given $\mathscr{F}_s$ for every $s \geq 0$. To justify our use of martingale theory and stochastic calculus, we may pass to a new filtration $\mathscr{G}$ that is both *right-continuous* and *complete*, in the sense that $\mathscr{G}_t = \mathscr{G}_{t+} \equiv \bigcap_{u>t} \mathscr{G}_u$ for all $t \geq 0$ and every $\mathscr{G}_t$ contains the null sets in the P-completion of $\sigma\{\mathscr{G}_u;\ u \geq 0\}$. Recall that every filtration $\mathscr{F}$ has a smallest right-continuous and complete extension, the so-called *usual augmentation* of $\mathscr{F}$ (cf. Kallenberg (1997), Lemma 6.8).

Lemma 3.2. *If a process X on $[0, 1]$ or $\mathbb{R}_+$ is spreadable for some filtration $\mathscr{F}$, then it remains so for the usual augmentation of $\mathscr{F}$.*

Proof. For every t, the shifted process $\theta_t X - X_t$ is conditionally spreadable given $\mathscr{F}_t$. By the chain rule for conditional expectations, this remains true with $\mathscr{F}_t$ replaced by $\mathscr{F}_{s+}$ for any $s < t$. Since X is right-continuous, the latter version extends for fixed s to $t = s$, which means that X is spreadable with respect to the right-continuous filtration $(\mathscr{F}_{t+})$. It remains to note that any conditional probability is unaffected by completion of the associated σ-field. □

For later needs, we record a continuous-time analogue of Lemma 3.1, which follows from the earlier result by a straightforward approximation.

Lemma 3.3. *Let X be an $\mathbb{R}^d$-valued, $\mathscr{F}$-adapted process on $[0,1]$ or $\mathbb{R}_+$ with $X_0 = 0$ and rcll paths. Then these conditions are equivalent:*

(i) *X is $\mathscr{F}$-spreadable;*
(ii) *X has $\mathscr{F}$-stationary increments;*
(iii) *$\mu_t = P[\theta_t X - X_t \in \cdot|\mathscr{F}_t]$ forms an $\mathscr{F}$-martingale.*

We turn to the semimartingale properties of spreadable processes. Given an rcll process X on $[0,1]$ or $\mathbb{R}_+$, we define the associated *jump point process* β by $\beta_t A = \sum_{s\le t} 1_A(\Delta X_s)$ for any $A \in \mathscr{B}(\mathbb{R}^d \setminus \{0\})$ and let $\hat\beta$ denote the compensator of β. When X is a semimartingale, we write X^c for the continuous local martingale component of X. Finally, if X is a special semimartingale, we define the *compensator* $\hat X$ of X as the a.s. unique predictable process of locally finite variation and initial value 0 such that $X - \hat X$ is a local martingale. The processes $\hat X$, $[X^c]$, and $\hat\beta$ are collectively referred to as the *local characteristics* of X.

Given a filtration $\mathscr{F}$, we say that a process X is *$\mathscr{F}_s$-integrable* if $E[|X_t||\mathscr{F}_s] < \infty$ a.s. for all t. (If X is $\mathscr{F}$-spreadable on $[0,1]$, then by Theorem 4.10 it suffices to assume this condition for a fixed $t \in (s,1)$.) The stronger notion of *uniform $\mathscr{F}_0$-integrability* may be defined in the obvious way in terms of the conditional distributions $P[X \in \cdot\,|\mathscr{F}_0]$. We say that X is a *conditional $\mathscr{F}$-martingale* on some interval I if X is $\mathscr{F}_s$-integrable for all $s \in I$ and satisfies $X_s = E[X_t|\mathscr{F}_s]$ a.s. for all $s < t$ in I. Note that any conditional martingale on $[0,1)$ or $\mathbb{R}_+$ is a local martingale.

We show under a moment condition that a spreadable process is a semimartingale and describe the associated local characteristics. This extends and improves a result for exchangeable processes in Theorem 4.1 of Kallenberg (1988a). We conjecture that no moment condition is actually needed for the semimartingale property in part (ii).

Proposition 3.4. *Let X be an $\mathbb{R}^d$-valued, $\mathscr{F}$-spreadable process on $[0,1]$ with jump point process β. Then*

(i) *$\hat\beta$ admits a conditional martingale density on $(0,1)$;*
(ii) *if X is $\mathscr{F}_0$-integrable, it is a uniformly $\mathscr{F}_0$-integrable special semimartingale on $[0,1]$, such that $[X^c]$ is a.s. linear and $\hat X$ admits a conditional martingale density on $[0,1)$.*

Proof. (ii) Define

$$M_t = \frac{E[X_1 - X_t|\mathscr{F}_t]}{1-t}, \quad t \in [0,1). \tag{1}$$

By the spreadability of X, we get for any times $s \le t < 1$ with rationally dependent residuals $1-s$ and $1-t$

$$E[M_t|\mathscr{F}_s] = \frac{E[X_1 - X_t|\mathscr{F}_s]}{1-t} = \frac{E[X_1 - X_s|\mathscr{F}_s]}{1-s} = M_s. \tag{2}$$

Approximating from the right in s, we may extend the formula to arbitrary $s \le t < 1$, which shows that M is a conditional martingale on $[0, 1)$. Writing $N_t = E[X_1|\mathscr{F}_t]$, we have

$$X_t = N_t - (1-t)M_t, \quad t \in [0, 1), \tag{3}$$

and integrating by parts gives

$$dX_t = dN_t - (1-t)dM_t + M_t dt. \tag{4}$$

This shows that X is a special semimartingale on $[0, 1)$ with compensator

$$\hat{X}_t = \int_0^t M_s ds, \quad t \in [0, 1).$$

From (3) we note that X is uniformly $\mathscr{F}_0$-integrable on $[0, \frac{1}{2}]$, and the spreadability of X allows us to extend this property to all of [0, 1]. The matrix-valued process $[X^c] = [X]^c$ is again spreadable, by the approximation property in Theorem I.4.47 of Jacod and Shiryaev (1987). Since it is also continuous and of finite variation, it is a.s. linear by Lemma 2.4.

To extend the semimartingale property to the closed interval [0, 1], we may use (1), Jensen's inequality, the spreadability of X, and Theorem 4.10 below (whose proof depends only on the semimartingale property on $[0, 1)$) to write for any $t < 1$

$$E^0|M_t| \le \frac{E^0|X_1 - X_t|}{1-t} = \frac{E^0|X_{1-t}|}{1-t} \lesssim \frac{E^0|X_{1/2}|}{(1-t)^{1/2}},$$

where E^0 is short for $E[\,\cdot\,|\mathscr{F}_0]$. This gives

$$E^0 \int_0^1 |d\hat{X}_t| = \int_0^1 E^0|M_t|\, dt \lesssim E^0|X_{1/2}| \int_0^1 (1-t)^{-1/2}\, dt < \infty,$$

and shows that $\hat{X}$ has $\mathscr{F}_0$-integrable variation whereas $X - \hat{X}$ is a uniformly $\mathscr{F}_0$-integrable local martingale on [0, 1].

(i) Combining Lemma 2.4 above with Lemma 4.2 in Kallenberg (1988a), we note that $E[\beta\{|x| > \varepsilon\}|\mathscr{F}_t] < \infty$ a.s. for all $t \in (0, 1]$ and $\varepsilon > 0$. We may then define a measure-valued process μ on $[0, 1)$ by

$$\mu_t A = \frac{E[\beta_1 A - \beta_t A|\mathscr{F}_t]}{1-t}, \quad t \in [0, 1),\ A \in \mathscr{B}^d. \tag{5}$$

The conditional martingale property of μ follows as in (2), and proceeding as in (3) and (4), we see that μ is a.s. a density of $\hat{\beta}$. □

In Corollary 2.6 we saw that a spreadable process need not be exchangeable. Here we consider some additional conditions that ensure the equivalence of the two properties. Our results extend and improve Theorem 3.3 in Kallenberg (1982) and Lemmas 2.2 and 2.4 above.

Theorem 3.5. *Let X be an $\mathbb{R}^d$-valued, $\mathscr{F}$-spreadable process on* $[0, 1]$ *with jump point process β. Then X is $\mathscr{F}$-exchangeable under each of these conditions:*

(i) *X is $\mathscr{F}_0$-integrable and the pair (X_1, β_1) is $\mathscr{F}_0$-measurable;*
(ii) *X has a.s. finite variation and β_1 is $\mathscr{F}_0$-measurable.*

We conjecture that every *continuous* spreadable process is exchangeable. The proof of part (i) requires two lemmas.

Lemma 3.6. *Let X be such as in Theorem* 3.5 (i). *Then X is a special semimartingale with $[X^c]_t \equiv t[X]_1^c$ a.s., and for any $t \in [0,1)$ we have a.s.*

$$\hat{\beta}_t = \int_0^t \frac{\beta_1 - \beta_s}{1-s}\, ds, \tag{6}$$

$$\hat{X}_t = tX_1 - \int_0^t ds \int_0^s \frac{d(X_r - \hat{X}_r)}{1-r}. \tag{7}$$

Proof. From Proposition 3.4 and its proof we note that X is a special semimartingale such that $[X^c]$ is a.s. linear, and also that $\hat{\beta}$ has the conditional martingale density $\mu_t = (\beta_1 - \beta_t)/(1-t)$. The latter statement implies (6). Furthermore, $\hat{X}$ has a conditional martingale density M satisfying

$$X_1 - X_t = (1-t)M_t, \quad t \in [0,1]. \tag{8}$$

Integration by parts gives

$$-(1-t)dM_t + M_t dt = dX_t = d(X_t - \hat{X}_t) + d\hat{X}_t,$$

and so, by the uniqueness of the canonical decomposition,

$$d\hat{X}_t = M_t dt, \qquad dM_t = -\frac{d(X_t - \hat{X}_t)}{1-t}.$$

Equation (7) follows as we integrate the latter relations and note that $M_0 = X_1$ a.s. in view of (8). □

Lemma 3.7. *For processes X as in Theorem* 3.5 (i), *the conditional distribution $P[X \in \cdot\, | \mathscr{F}_0]$ is a unique, measurable function of X_1, $[X]_1^c$, and β_1.*

Proof. Let $A_1, \ldots, A_n \in \mathscr{B}(\mathbb{R}^d)$ be disjoint and bounded away from 0, put $\kappa_r = \beta_1 A_r$, and let $\tau_1^r < \cdots < \tau_{\kappa_r}^r$ be the points of the process βA_r. Write $\hat{\tau}_0^r = 0$ and $\hat{\tau}_j^r = \hat{\beta} A_r(\tau_j^r)$, and put $\gamma_j^r = \hat{\tau}_j^r - \hat{\tau}_{j-1}^r$. By (6) we have

$$\gamma_j^r = (\kappa_r - j + 1)\, \log\left\{\frac{1-\tau_{j-1}^r}{1-\tau_j^r}\right\}, \quad j \le \kappa_r,\ r \le n.$$

Solving recursively for τ_j^r gives

$$\tau_j^r = 1 - \exp\left\{-\sum\nolimits_{i \le j} \frac{\gamma_i^r}{\kappa_r - i + 1}\right\}, \quad j \le \kappa_r,\ r \le n. \tag{9}$$

By a version of Theorem II.6.6 in Jacod and Shiryaev (1987), the continuous martingale component X^c is conditionally a Brownian motion with covariance matrix

$[X]_1^c$, given $\mathscr{F}_0$, whereas the γ_j^r are i.i.d. standard exponential random variables independent of $(\mathscr{F}_0, X^c)$. Furthermore, (9) exhibits the processes βA_r as measurable functions of the variables κ_r and γ_j^r, where the former are $\mathscr{F}_0$-measurable by hypothesis. This specifies the joint conditional distribution of X^c and $\beta A_1, \ldots, \beta A_n$ given $\mathscr{F}_0$. Since the sets A_r were arbitrary and since (6) expresses $\hat{\beta}$ as a measurable function of β, the conditional distribution is then a.s. unique for the martingale component $X - \hat{X}$. In view of (7), this is also true for the process X itself. □

Proof of Theorem 3.5. (i) By the transfer Theorem 5.10 in Kallenberg (1997), we may choose the process Y to be conditionally exchangeable, given $\mathscr{F}_0$, with directing triple $(X_1, [X]_1^c, \beta_1)$. Let $\mathscr{G}$ be the right-continuous, complete filtration induced by Y and $\mathscr{F}_0$, and note that even Y satisfies the hypotheses of the theorem, but now relative to the filtration $\mathscr{G}$. By Lemma 3.7, X and Y have then the same conditional distribution given $\mathscr{F}_0$, and in particular X is conditionally exchangeable on $[0, 1]$. Applying this result to the shifted process $\theta_s X - X_s$ for arbitrary $s \in [0, 1)$, we see more generally that X is $\mathscr{F}$-exchangeable.

(ii) The continuous component X^c of X—in the sense of functions of bounded variation—is again spreadable, and so are the monotone components in the Jordan decomposition of each coordinate process. By Lemma 2.4 it follows that $X_t^c \equiv t\alpha$ a.s. for some random vector α in $\mathbb{R}^d$. If $\mathscr{F}$ is right-continuous, as we may assume by Lemma 3.2, then α is clearly $\mathscr{F}_0$-measurable. Since also a.s.

$$X_1 = \alpha + \int_0^1 x\, \beta_1(dx),$$
$$|X_t| \le \int_0^1 |dX_s| = |\alpha| + \int_0^1 |x|\, \beta_1(dx) < \infty,$$

we conclude that X_1 is $\mathscr{F}_0$-measurable and X is $\mathscr{F}_0$-integrable. The asserted exchangeability of X now follows by part (i). □

We can also give the following alternative proof of part (ii), based on the more elementary but less intuitive Theorem 2.5.

Second proof of (ii). For any $\varepsilon > 0$, let X^ε denote the sum of all jumps of modulus $> \varepsilon$, and denote the corresponding jump sizes by $\xi_1, \ldots, \xi_n$. Then X^ε is again spreadable, and so the sequence (ξ_k) is spreadable by Theorem 2.5 (i). Since $\sum_k \delta_{\xi_k}$ is nonrandom by hypothesis, Lemma 2.2 shows that (ξ_k) is even exchangeable. Then so is the process X^ε by Theorem 2.5 (iii), and the same thing is true for the jump component X' of X since ε was arbitrary.

Next Lemma 2.4 gives $X_t - X_t' \equiv \alpha t$ a.s. for some random vector α in $\mathbb{R}^d$, and by Lemma 3.2 we may assume that α is $\mathscr{F}_0$-measurable. Applying the previous argument to the conditional distributions, we note that X' is conditionally exchangeable on $[t, 1]$ given $\mathscr{F}_t$ for every $t \in [0, 1]$. The same thing is then true for X itself, since $X - X'$ is a.s. linear and $\mathscr{F}_t$-measurable. □

4. Distributional properties

In this section we study some weak convergence and related properties of spreadable sequences and processes. A comprehensive weak convergence theory for exchangeable processes was developed in Kallenberg (1973, 1975, 1982, 1988b), and we refer to Chapter 14 of Kallenberg (1997) for some basic ideas and results. Since there is no general representation formula in the spreadable case, only a partial extension of the exchangeable theory is possible.

We begin with a limit theorem for finite spreadable sequences, which extends an elementary result for exchangeable sequences from Kallenberg (1973). The present statement also contains Ryll-Nardzewski's (1957) version of de Finetti's theorem (cf. Kallenberg (1997), Theorem 9.16), the fact that any *infinite* spreadable sequence is mixed i.i.d.

Given a finite random sequence $\xi = (\xi_1, \dots, \xi_m)$ in some Polish space S, we define the associated *empirical distribution* as the random probability measure $\nu = m^{-1} \sum_j \delta_{\xi_j}$ on S. Here ν is regarded as a random element in the space $\mathscr{M}_1(S)$ endowed with the weak topology. For sequences ξ_n of lengths $m_n \xrightarrow{d} \infty$, the convergence $\xi_n \xrightarrow{d} \xi$ is defined by the corresponding set of finite-dimensional conditions.

Theorem 4.1. *Let $\xi_1, \xi_2, \dots$ be spreadable sequences of finite lengths $m_1, m_2, \dots \to \infty$ in a Polish space S and let $\nu_1, \nu_2, \dots$ denote the associated empirical distributions. Then $\xi_n \xrightarrow{d}$ some ξ in S^∞ iff $\nu_n \xrightarrow{d}$ some ν in $\mathscr{M}_1(S)$, in which case $\mathscr{L}(\xi) = E\nu^\infty$.*

We give a direct proof based on the following simple moment estimate. Using the Ryll-Nardzewski theorem, one can also give a slightly shorter but more sophisticated proof along the lines of Theorem 4.8 below.

Lemma 4.2. *Let $\xi_1, \dots, \xi_n$ be square-integrable random variables with constant mean m, variance σ^2, and covariance $\sigma^2\rho$, and fix any distributions (p_j) and (q_j) on $\{1, \dots, n\}$. Then*

$$E\Big(\sum\nolimits_j p_j\xi_j - \sum\nolimits_j q_j\xi_j\Big)^2 \le 2\sigma^2(1-\rho)\sup\nolimits_j |p_j - q_j|.$$

Proof. Write $d_j = p_j - q_j$. Noting that $\sum_j d_j = 0$ and $\sum_j |d_j| \le 2$, we get

$$\begin{aligned} E\Big(\sum\nolimits_j d_j\xi_j\Big)^2 &= E\Big(\sum\nolimits_j d_j(\xi_j - m)\Big)^2 \\ &= \sum\nolimits_i \sum\nolimits_j d_i d_j \operatorname{cov}(\xi_i, \xi_j) \\ &= \sigma^2\rho\,\Big(\sum\nolimits_j d_j\Big)^2 + \sigma^2(1-\rho)\sum\nolimits_j d_j^2 \\ &\le \sigma^2(1-\rho)\,\sup\nolimits_i |d_i| \sum\nolimits_j |d_j| \\ &\le 2\sigma^2(1-\rho)\sup\nolimits_j |d_j|. \end{aligned}$$

□

Proof of Theorem 4.1. First assume that $\nu_n \xrightarrow{d} \nu$. Fix any continuous functions $f_1, \ldots, f_k : S \to [0,1]$. Write $[m_n/k] = r_n$ and $I_{nj} = \{(j-1)r_n + 1, \ldots, jr_n\}$. Using the spreadability of ξ_n, Jensen's inequality, and Lemma 4.2, we get

$$\begin{aligned}
\left| E \prod_{j \le k} f_j(\xi_{nj}) - E \prod_{j \le k} \nu_n f_j \right|
&= \left| r_n^{-k} E \prod_{j \le k} \sum_{i \in I_{nj}} f_j(\xi_{ni}) - E \prod_{j \le k} \nu_n f_j \right| \\
&\le \sum_{j \le k} E \left| r_n^{-1} \sum_{i \in I_{nj}} f_j(\xi_{ni}) - \nu_n f_j \right| \\
&\le \sum_{j \le k} \left\| r_n^{-1} \sum_{i \in I_{nj}} f_j(\xi_{ni}) - \nu_n f_j \right\|_2 \\
&\le 2k r_n^{-1/2} \sim 2k^{3/2} m_n^{-1/2} \to 0.
\end{aligned}$$

Since also

$$E \prod_{j \le k} \nu_n f_j \to E \prod_{j \le k} \nu f_j,$$

by the continuity of the mappings $\mu \mapsto \mu f_j$ on $\mathscr{M}_1(S)$, we obtain

$$E \prod_{j \le k} f_j(\xi_{nj}) \to E \prod_{j \le k} \nu f_j = E \prod_{j \le k} f_j(\xi_{0j}), \tag{1}$$

where $\xi = (\xi_{0j})$ has distribution $E\nu^\infty$.

To deduce the required convergence $\xi_n \xrightarrow{d} \xi$, we note in particular that $\xi_{nj} \xrightarrow{d} \xi_{0j}$ for each j. Hence, by Prohorov's theorem, (ξ_n) is tight and any subsequence $N' \subset \mathbb{N}$ has a further subsequence $N'' \subset N'$ such that $\xi_n \xrightarrow{d}$ some ξ' along N''. To see that $\xi' \stackrel{d}{=} \xi$, we note that (1) remains true along N'' with ξ_{0j} replaced by ξ'_{0j}, and therefore

$$E \prod_{j \le k} f_j(\xi_{0j}) = E \prod_{j \le k} f_j(\xi'_{0j})$$

for any k and $f_1, \ldots, f_k$. By a simple approximation, this extends to the indicator functions of any open sets $G_1, \ldots, G_k \subset S$, and by a monotone class argument it follows that $\xi \stackrel{d}{=} \xi'$. Hence, $\xi_n \xrightarrow{d} \xi$ along N'' and then also along $\mathbb{N}$, since the limit is independent of the choice of subsequence.

Conversely, assume that $\xi_n \xrightarrow{d} \xi$. The spreadability of the sequences ξ_n implies the weak convergence $E\nu_n = \mathscr{L}(\xi_{n1}) \xrightarrow{w} \mathscr{L}(\xi_{01})$, and by Prohorov's theorem it follows that the sequence $(E\nu_n)$ is tight. By Lemma 7.14 in Aldous (1985), the tightness carries over to the sequence of random probability measures ν_n. Now let $\nu_n \xrightarrow{d} \nu$ along some subsequence $N' \subset \mathbb{N}$. The direct assertion gives $\xi_n \xrightarrow{d} \xi'$ along N' where $\mathscr{L}(\xi') = E\nu^\infty$. Since also $\xi_n \xrightarrow{d} \xi$, we get $\mathscr{L}(\xi) = E\nu^\infty$.

To see that $\mathscr{L}(\nu)$ is uniquely determined by $\mathscr{L}(\xi)$, fix any bounded, measurable function f on S and a constant $r \in \mathbb{R}$. By the law of large numbers,

$$\begin{aligned}
P\{\nu f \le r\} &= E\nu^\infty \left\{ \limsup_n n^{-1} \sum_{j \le n} f(x_j) \le r \right\} \\
&= P\left\{ \limsup_n n^{-1} \sum_{j \le n} f(\xi_{0j}) \le r \right\},
\end{aligned}$$

and the required uniqueness follows by the Cramér–Wold theorem. By Prohorov's theorem, the convergence $\nu_n \xrightarrow{d} \nu$ then extends to the original sequence. □

The continuous-time case is more difficult. Here our treatment is based on the following key result, where for any spreadable sequence or process we associate an exchangeable sequence or process with the same one-dimensional distributions. This correspondence allows us to extend many statements for exchangeable processes to the more general context of spreadability.

For any random sequence $\xi = (\xi_1, \dots, \xi_n)$, we introduce the associated *occupation sequence* $\beta_k = \sum_{j \le k} \delta_{\xi_j}$, $k = 1, \dots, n$. When X is an $\mathbb{R}^d$-valued semimartingale on $[0, 1]$, the associated *characteristics* α_t, β_t, and γ_t are given by

$$\alpha_t = X_t, \quad \beta_t = \sum\nolimits_{s \le t} \delta_{\Delta X_s}, \quad \gamma_t^{ij} = [X^i, X^j]_t, \quad t \in [0, 1]. \tag{2}$$

The definitions of α_t and β_t continue to make sense without the semimartingale property, as long as the paths of X are rcll.

Proposition 4.3.

(i) *For any spreadable sequence $\xi = (\xi_1, \dots, \xi_n)$ in some measurable space S, there exists an exchangeable sequence $\tilde{\xi} = (\tilde{\xi}_1, \dots, \tilde{\xi}_n)$ such that the associated occupation sequences satisfy*

$$\beta_k \stackrel{d}{=} \tilde{\beta}_k, \quad k = 1, \dots, n.$$

(ii) *For any $\mathbb{R}^d$-valued, spreadable process X on $[0, 1]$, there exists an exchangeable process $\tilde{X}$ such that the associated processes of characteristics satisfy*

$$(\alpha_t, \beta_t) \stackrel{d}{=} (\tilde{\alpha}_t, \tilde{\beta}_t), \quad t \in [0, 1]. \tag{3}$$

If X is a semimartingale, then (3) can be strengthened to

$$(\alpha_t, \beta_t, \gamma_t) \stackrel{d}{=} (\tilde{\alpha}_t, \tilde{\beta}_t, \tilde{\gamma}_t), \quad t \in [0, 1]. \tag{4}$$

In particular, we note that $\sum_t |\Delta X_t|^2 < \infty$ a.s. for any $\mathbb{R}^d$-valued, spreadable process X on $[0, 1]$. In the semimartingale case we may take $t = 1$ in (4) to see that $\mathscr{L}(\tilde{X})$ is unique. It is not clear whether uniqueness still holds under the weaker condition (3). In the following proof and throughout the remainder of the section, we consider *summation processes* of the form

$$X_t = \sum\nolimits_{j \le mt} \xi_j, \quad t \in [0, 1], \tag{5}$$

where $\xi_1, \dots, \xi_m$ are random vectors in $\mathbb{R}^d$.

Proof. (i) Let $\tau_1, \dots, \tau_n$ form an exchangeable permutation of $1, \dots, n$ independent of ξ and consider the exchangeable sequence $\tilde{\xi}$ with elements $\tilde{\xi}_k = \xi_{\tau_k}$ for $k = 1, \dots, n$. For each $k \in \{1, \dots, n\}$ we enumerate the variables $\tau_1, \dots, \tau_k$ in increasing order as $\sigma_{k1}, \dots, \sigma_{kk}$. The spreadability of ξ and exchangeability of $\tilde{\xi}$

carry over to the random measures $\mu_j = \delta_{\xi_j}$ and $\tilde{\mu}_j = \delta_{\tilde{\xi}_j}$, respectively, and for any $k \le n$ we get

$$\beta_k = \sum\nolimits_{j \le k} \mu_j \overset{d}{=} \sum\nolimits_{j \le k} \mu_{\sigma_{kj}} = \sum\nolimits_{j \le k} \mu_{\tau_j} = \sum\nolimits_{j \le k} \tilde{\mu}_j = \tilde{\beta}_k.$$

(ii) Here we put $t_{nj} = j2^{-n}$ and $\xi_{nj} = X_{t_{nj}} - X_{t_{n,k-1}}$ for all $n \in \mathbb{N}$ and $j \le 2^n$. For each $n \in \mathbb{N}$ we introduce an exchangeable permutation $(\tilde{\xi}_{nj})$ of the increments ξ_{nj} and write X^n and $\tilde{X}^n$ for the corresponding summation processes on $[0, 1]$. Since $\tilde{X}^n_t \overset{d}{=} X^n_t = X_t$ for all n-dyadic times $t = t_{nj}$, the sequence $(\tilde{X}^n)$ is tight in $D([0, 1], \mathbb{R}^d)$ by the exchangeable version of Proposition 4.4 below, and therefore $\tilde{X}^n \overset{d}{\to} \tilde{X}$ along a subsequence for some exchangeable process $\tilde{X}$ on $[0, 1]$. By a d-dimensional version of a convergence criterion in Kallenberg (1973) (cf. Kallenberg (1997), Proposition 14.24), the corresponding characteristic triples satisfy

$$(\tilde{\alpha}^n_t, \tilde{\beta}^n_t, \tilde{\gamma}^n_t) \overset{d}{\to} (\tilde{\alpha}_t, \tilde{\beta}_t, \tilde{\gamma}_t), \quad t \in [0, 1], \tag{6}$$

where $\tilde{\beta}_t$ and $\tilde{\beta}^n_t$ are regarded as random elements in $\mathcal{M}(\mathbb{R}^d \setminus \{0\})$ with the vague topology.

If X is a semimartingale, we may use part (i) together with the approximation property for the quadratic variation (cf. Jacod and Shiryaev (1987), Theorem I.4.47) to write for dyadic times $t \in [0, 1]$

$$(\tilde{\alpha}^n_t, \tilde{\beta}^n_t, \tilde{\gamma}^n_t) \overset{d}{=} (\alpha^n_t, \beta^n_t, \gamma^n_t) \overset{P}{\to} (\alpha_t, \beta_t, \gamma_t). \tag{7}$$

Relation (4) follows for dyadic t by combination of (6) and (7), and then in general by the right continuity of both sides. Without the semimartingale hypothesis, we can only assert convergence in (7) for the first two components. This still guarantees the truth of (3). □

Our next key step is to establish some tightness criteria for spreadable processes, originally stated for exchangeable processes in Kallenberg (1973) (though proved more convincingly in Kallenberg (1997), Proposition 14.24 and Theorem 14.25).

Proposition 4.4. *Let $X_1, X_2, \ldots$ be $\mathbb{R}^d$-valued, spreadable processes on $[0, 1]$ or summation processes on $[0, 1]$ based on spreadable sequences in $\mathbb{R}^d$ of lengths $m_n \to \infty$. Fix any relatively compact sequence $t_1, t_2, \ldots \in (0, 1)$. Then the following two conditions are equivalent:*

(i) *$\{X_n\}$ is tight in $D([0, 1], \mathbb{R}^d)$;*
(ii) *$\{X_n(t_n)\}$ is tight in $\mathbb{R}^d$.*

If the X_n are semimartingales, it is also equivalent that

(iii) *$\{(\alpha_n, \gamma_n)\}$ is tight in $\mathbb{R}^d \times \mathbb{R}^{d^2}$.*

Proof. The implication (i) ⇒ (ii) holds generally in $D([0, 1], \mathbb{R}^d)$ by the continuity of the mapping $x \mapsto \sup_t |x_t|$. To prove the remaining assertions, we choose some associated exchangeable processes $\tilde{X}_n$ with characteristic triples $(\tilde{\alpha}_n, \tilde{\beta}_n, \tilde{\gamma}_n)$ as in

Proposition 4.3. By the criteria for exchangeable processes, the sequences $\{\tilde{X}_n\}$, $\{\tilde{X}_n(t_n)\}$, and $\{(\tilde{\alpha}_n, \tilde{\gamma}_n)\}$ are simultaneously tight. In particular, Proposition 4.3 shows that (ii) ⇔ (iii) whenever each process X_n is a semimartingale.

In the general case, (ii) implies that $\{\tilde{X}_n\}$ is tight in $D([0,1], \mathbb{R}_+)$. Noting that $x_n(h_n) \to x(0)$ whenever $x_n \to x$ in $D([0,1], \mathbb{R}^d)$ and $h_n \to 0$ and using Lemma 3.3 and Proposition 4.3, we get in the continuous-time case

$$X_n(\tau_n + h_n) - X_n(\tau_n) \overset{d}{=} X_n(h_n) \overset{d}{=} \tilde{X}_n(h_n) \overset{P}{\to} 0,$$

for any X_n-optional times τ_n and positive constants $h_n \to 0$ with $\tau_n + h_n \leq 1$ a.s. Since the sequence $X_n(t) \overset{d}{=} \tilde{X}_n(t)$ is tight in $\mathbb{R}^d$ for every $t \in [0,1]$, the tightness of $\{X_n\}$ in $D([0,1], \mathbb{R}^d)$ follows by Aldous' criterion (cf. Kallenberg (1997), Theorem 14.11). Thus, (ii) ⇒ (i). The argument applies with obvious modifications to the case of summation processes. □

We also need the following closure properties for spreadable processes, which again extend the exchangeable versions from Kallenberg (1973).

Lemma 4.5. *Let $X_1, X_2, \ldots$ be $\mathbb{R}^d$-valued, spreadable processes on $[0,1]$ or summation processes based on spreadable sequences in $\mathbb{R}^d$ of lengths $m_n \to \infty$, and assume that $X_n \overset{d}{\to} X$ in $D([0,1], \mathbb{R}^d)$. Then X is again spreadable. A similar result holds for spreadable random measures on $[0,1]$.*

Proof. Writing X_n as X^n, we have in the spreadable case $X^n_v - X^n_u \overset{d}{=} X^n_{v-u}$ for any $u < v$. If u, v, and $v-u$ are a.s. continuity points for X, we obtain $X_v - X_u \overset{d}{=} X_{v-u}$. Similarly, $X_1 - X_u \overset{d}{=} X_{1-u}$ whenever u and $1-u$ are a.s. continuity points. Fixing any $t \in (0,1]$, we may let $u \uparrow t$ and $v \downarrow t$ to get $\Delta X_t \overset{d}{=} 0$. This shows that X has no fixed discontinuities, and therefore $X^n \overset{fd}{\longrightarrow} X$. In particular, the spreadability of the X^n carries over to X. A slightly modified argument applies to the case of summation processes.

Next let $\xi_1, \xi_2, \ldots$ be spreadable random measures on $[0,1]$ with $\xi_n \overset{d}{\to} \xi$. The corresponding random distribution functions $X^n_t = \xi_n[0,t]$ are then tight in $D([0,1], \mathbb{R}^d)$ by Proposition 4.4, and we get convergence $X^n \overset{d}{\to} \tilde{X}$ along a subsequence $N' \subset \mathbb{N}$ for some spreadable process $\tilde{X}$. As before, it follows that $X^n \overset{fd}{\longrightarrow} \tilde{X}$ along N', and so $\xi_n \overset{d}{\to} \tilde{\xi}$ where $\tilde{\xi}[0,t] = \tilde{X}_t$. But then $\tilde{\xi} \overset{d}{=} \xi$, and so the spreadability of $\tilde{\xi}$, inherited from $\tilde{X}$, carries over to ξ. □

The preceding results allow us to show that the functional and finite-dimensional modes of convergence are equivalent for spreadable processes on $[0,1]$. This again extends a statement for exchangeable processes in Kallenberg (1973).

Corollary 4.6. *Let $X_1, X_2, \ldots$ be $\mathbb{R}^d$-valued, spreadable processes on $[0,1]$ or summation processes on $[0,1]$ based on spreadable sequences in $\mathbb{R}^d$ of lengths $m_n \to \infty$. Then for any process X in $D([0,1], \mathbb{R}^d)$ we have $X_n \overset{d}{\to} X$ iff $X_n \overset{fd}{\longrightarrow} X$.*

Proof. If $X_n \xrightarrow{d} X$, then X is spreadable by Lemma 4.5 and hence continuous in probability. Since $x_n \to x$ in $D([0,1], \mathbb{R}^d)$ implies $x_n(t) \to x(t)$ at every continuity point t of x, we conclude that $X_n \xrightarrow{fd} X$. Conversely, $X_n \xrightarrow{fd} X$ implies that $\{X_n\}$ is tight by Proposition 4.4, and so by Prohorov's theorem $X_n \xrightarrow{d} X$ in $D([0,1], \mathbb{R}^d)$. □

The necessary tools are now available to prove our fundamental regularization theorem for spreadable processes. It shows in particular that a process with spreadable increments has a right-continuous version with left-hand limits iff it is continuous in probability. In this weaker form, the result is known for exchangeable processes (cf. Kallenberg (1997), Theorem 14.25). The present statement justifies our definition of spreadable processes in Section 1.

Theorem 4.7. *Let X be an $\mathbb{R}^d$-valued process with spreadable increments and $X_0 = 0$, defined on the set of dyadic rationals in $[0,1]$. Then X extends a.s. to a spreadable process on $[0,1]$ with rcll paths.*

Proof. For each $n \in \mathbb{N}$ we introduce a summation process X^n based on the increments of X on the n-dyadic set D_n. Since $X_t^m = X_t$ for all $m \geq n$ when $t \in D_n$, the sequence (X^n) is tight in $D([0,1], \mathbb{R}^d)$ by Proposition 4.4. By Prohorov's theorem, we have convergence $X^n \xrightarrow{d} Y$ in $D([0,1], \mathbb{R}^d)$ along a subsequence, and the limiting process Y is spreadable by Lemma 4.5. But then $X^n \xrightarrow{fd} Y$ by Corollary 4.6, and so $X \stackrel{d}{=} Y$ on $D = \bigcup_n D_n$. Finally, we may use the transfer Theorem 5.10 in Kallenberg (1997) to construct a spreadable process $\tilde{X} \stackrel{d}{=} Y$ with $X = \tilde{X}$ a.s. on D. □

We turn to a partial continuous-time extension of Theorem 4.1. For any sequences (ξ_n) and (η_n) of random elements in a Polish space S, we write $\xi_n \stackrel{d}{\sim} \eta_n$ to mean that $\xi_n \xrightarrow{d} \xi$ iff $\eta_n \xrightarrow{d} \xi$ for any random element ξ in S. When ξ is a spreadable random measure on some interval $[0,u]$, we define the processes α_t and β_t as in (2) from the associated distribution function $X_t = \xi[0,t]$. Let us say that $\tilde{\xi}$ is a *corresponding* exchangeable random measure on $[0,u]$ if the terminal values α_u and β_u agree for ξ and $\tilde{\xi}$. If ξ is instead an exchangeable random measure on $\mathbb{R}_+$, so that the associated process X is a mixture of subordinators, we write ν for the random Lévy measure of X and ρ for the rate of increase of the linear drift component.

Theorem 4.8. *Let $\xi_1, \xi_2, \ldots$ be spreadable random measures on some intervals $[0, u_n] \to \mathbb{R}_+$ and let $\tilde{\xi}_1, \tilde{\xi}_2, \ldots$ be corresponding exchangeable random measures on the same intervals. Then $\xi_n \stackrel{d}{\sim} \tilde{\xi}_n$ in both $\mathscr{M}(\mathbb{R}_+)$ and $D(\mathbb{R}_+, \mathbb{R}_+)$. A similar result holds for summation processes based on spreadable sequences in $\mathbb{R}_+$.*

The result fails for processes on a common finite interval, since by Corollary 2.6 and Proposition 4.3 there exist some spreadable random measures on $[0,1]$ with different distributions but the same characteristics at 1. We conjecture that the statement remains true, possibly under a moment condition, for any $\mathbb{R}^d$-valued, spreadable processes on some intervals $[0, t_n] \to \mathbb{R}_+$.

Proof. First assume that $\tilde{\xi}_n \xrightarrow{d} \tilde{\xi}$ in $\mathscr{M}(\mathbb{R}_+)$, and note that the convergence extends to $D(\mathbb{R}_+, \mathbb{R}_+)$ by Corollary 4.6. By Proposition 4.3 and the continuity in $D(\mathbb{R}_+, \mathbb{R}_+)$, the associated characteristics satisfy

$$(\alpha_t^n, \beta_t^n) \stackrel{d}{=} (\tilde{\alpha}_t^n, \tilde{\beta}_t^n) \xrightarrow{d} (\tilde{\alpha}_t, \tilde{\beta}_t), \quad t \geq 0, \tag{8}$$

under the vague topology on $\mathscr{M}((0, \infty))$. In particular, (ξ_n) is tight in $D(\mathbb{R}_+, \mathbb{R}_+)$ by Proposition 4.4. Now assume that $\xi_n \xrightarrow{d} \xi$ in $D(\mathbb{R}_+, \mathbb{R}_+)$ along a subsequence. Then ξ is spreadable by Lemma 4.5 and hence exchangeable by Ryll-Nardzewski's theorem. By continuity, we have

$$(\alpha_t^n, \beta_t^n) \xrightarrow{d} (\alpha_t, \beta_t), \quad t \geq 0,$$

and comparing with (8) gives $(\alpha_t, \beta_t) \stackrel{d}{=} (\tilde{\alpha}_t, \tilde{\beta}_t)$ for all $t \geq 0$. Hence, the characteristics of the limiting processes satisfy $(\rho, \nu) \stackrel{d}{=} (\tilde{\rho}, \tilde{\nu})$, by the law of large numbers, and so $\xi \stackrel{d}{=} \tilde{\xi}$. Since the limiting distribution is independent of subsequence, Prohorov's theorem yields $\xi_n \xrightarrow{d} \tilde{\xi}$ along $\mathbb{N}$.

Conversely, assume that $\xi_n \xrightarrow{d} \xi$ in $\mathscr{M}(\mathbb{R}_+)$. Then Proposition 4.3 yields

$$\tilde{\alpha}_t^n \stackrel{d}{=} \alpha_t^n \xrightarrow{d} \alpha_t, \quad t \geq 0,$$

and so $(\tilde{\xi}_n)$ is tight in $D(\mathbb{R}_+, \mathbb{R}_+)$. If $\tilde{\xi}_n \xrightarrow{d} \tilde{\xi}$ along a subsequence, then $\tilde{\xi} \stackrel{d}{=} \xi$ holds as before, and Prohorov's theorem yields $\tilde{\xi}_n \xrightarrow{d} \xi$ along $\mathbb{N}$. □

We may next extend some one-dimensional convergence criteria from Kallenberg (1988b).

Proposition 4.9. *Let $X^1, X^2, \ldots$ be $\mathbb{R}^d$-valued, spreadable processes or summation processes on some intervals $[0, u_n] \to \mathbb{R}_+$, where the latter are based on some spreadable sequences of lengths m_n such that $m_n/u_n \to \infty$. Consider a mixed Lévy process X in $\mathbb{R}^d$ that is either continuous with finite means or ergodic with a finite exponential moment. Then $X^n \xrightarrow{d} X$ in $D(\mathbb{R}_+, \mathbb{R}^d)$ iff $X_t^n \xrightarrow{d} X_t$ for all $t \geq 0$.*

Again the statement fails for processes on a fixed interval, since by Proposition 4.3 the distribution of a spreadable process on $[0, 1]$ may not be determined by its one-dimensional projections.

Proof. Assume that $X_t^n \xrightarrow{d} X_t$ for all $t \geq 0$. Then, by Proposition 4.4, the sequence (X^n) is tight in $D([0, r], \mathbb{R}^d)$ for every $r > 0$ and hence also in $D(\mathbb{R}_+, \mathbb{R}^d)$. Now assume that $X^n \xrightarrow{d} Y$ in $D(\mathbb{R}_+, \mathbb{R}^d)$ along a subsequence $N' \subset \mathbb{N}$. Then Y is spreadable by Lemma 4.5, and by Ryll-Nardzewski's theorem it is even exchangeable. For any $t \geq 0$ we get $X_t^n \xrightarrow{d} Y_t$ along N', and so $X_t \stackrel{d}{=} Y_t$ for all t. Under the stated conditions on X, we may conclude from Theorem 4.1 in Kallenberg (1988b) that $X \stackrel{d}{=} Y$. Since N' was arbitrary, we obtain $X^n \xrightarrow{d} X$. □

We conclude this section with some basic norm relations for spreadable processes, which appear to be new even in the exchangeable case. Our proofs depend on both the present distributional methods and the martingale methods of Section 3.

Theorem 4.10. *For any real, spreadable processes X on $[0,1]$ we have*

$$\|X_t^*\|_p \asymp \|X_t\|_p \lesssim t^{1/(p\vee 2)}\|X_s\|_p, \quad s\in(0,1),\ p\geq 1, \tag{9}$$

uniformly in $t\in[0,\frac{1}{2}]$ and $\mathscr{L}(X)$. The second relation remains valid for all $p>0$, and both relations hold for $p>0$ when X is exchangeable.

In particular, we note that

$$\|X_s\|_p \asymp \|X_t\|_p, \quad s,t\in(0,1),\ p>0, \tag{10}$$

uniformly in $\mathscr{L}(X)$. Our proof of Theorem 4.10 relies on the following preliminary result for exchangeable processes.

Lemma 4.11. *For any real, ergodic, exchangeable processes X on $[0,1]$ with characteristics α and γ we have, uniformly in $\mathscr{L}(X)$,*

$$\|X_t\|_p \asymp \|X^*\|_p \asymp |\alpha|+\gamma^{1/2}, \quad t\in(0,1),\ p>0.$$

Proof. Here $M_t=(X_t-\alpha t)/(1-t)$ is a martingale on $[0,1)$. Noting that X has quadratic variation γ and using the BDG inequalities and the symmetry of $X_t-\alpha t$ under reflection of $[0,1]$, we get for any $t\in[0,1]$ and $p\geq 1$

$$\|X_t\|_p \leq \|X^*\|_p \lesssim |\alpha|+\gamma^{1/2}, \tag{11}$$

which extends by Jensen's inequality to arbitrary $p>0$.

To prove the reverse relations, suppose that instead $\|X_t^n\|_p/(|\alpha_n|+\gamma_n^{1/2})\to 0$ for some ergodic exchangeable processes X^n with characteristics $(\alpha_n,\gamma_n)\neq 0$ and some $t\in(0,1)$. By scaling we may assume that $\|X_t^n\|_p\leq 1$ and $|\alpha_n|+\gamma_n^{1/2}\to\infty$. But these conditions are mutually contradictory by Proposition 4.4. Thus, $\|X_t\|_p\gtrsim|\alpha|+\gamma^{1/2}$, and the assertion follows by combination with (11). □

Proof of Theorem 4.10. Relation (10) holds by Proposition 4.3 and Lemma 4.11. Assuming $\|X_s\|_p<\infty$ for some fixed $s\in(0,1)$ and $p\geq 1$, we conclude that X is integrable. By the proof of Proposition 3.4, we may write $X_t=(X_t-\hat{X}_t)+\hat{X}_t$ where the compensator $\hat{X}$ is absolutely continuous and admits the martingale density $M_t=E[X_1-X_t|\mathscr{F}_t]/(1-t)$. By (10) and Jensen's inequality, we have $\|M_t\|_p\lesssim\|X_s\|_p$ for fixed $s,t\in(0,1)$. Using the continuous-time version of Minkowski's inequality, we get for any $t\leq\frac{1}{2}$

$$\|\hat{X}_t^*\|_p \leq \left\|\int_0^t |M_r|\,dr\right\|_p \leq \int_0^t \|M_r\|_p\,dr \leq t\|M_{1/2}\|_p \lesssim t\|X_s\|_p. \tag{12}$$

As an alternative for $p>1$, we may use Doob's inequality to write

$$\|\hat{X}_t^*\|_p \leq t\|M_{1/2}^*\|_p \lesssim t\|M_{1/2}\|_p \lesssim t\|X_s\|_p.$$

Next we conclude from the BDG inequalities that

$$\|(X-\hat{X})^*_t\|_p \lesssim \|\gamma_t^{1/2}\|_p, \quad p \geq 1, \tag{13}$$

uniformly in $t \in [0,1]$. To estimate the right-hand side in terms of $\|X_s\|_p$, we may assume by Proposition 4.3 that X is ergodic exchangeable. Using Jensen's inequality, the spreadability of γ_t, and Lemma 4.11, we get for $p \leq 2$

$$\|\gamma_t^{1/2}\|_p^2 \leq \|\gamma_t^{1/2}\|_2^2 = E\gamma_t = t\gamma \lesssim t\|X_s\|_p^2. \tag{14}$$

For $p > 2$ we have instead

$$\|\gamma_t^{1/2}\|_p^p = E\gamma_t^{p/2} \leq \gamma^{p/2-1}E\gamma_t = t\gamma^{p/2} \lesssim t\|X_s\|_p^p. \tag{15}$$

Combining (12)–(15) gives

$$\begin{aligned}\|X_t^*\|_p &\leq \|(X-\hat{X})_t^*\|_p + \|\hat{X}_t^*\|_p \\ &\lesssim t\|X_s\|_p + t^{1/(p\vee 2)}\|X_s\|_p \\ &\lesssim t^{1/(p\vee 2)}\|X_s\|_p.\end{aligned}$$

In particular, $\|X^*_{1/2}\|_p \asymp \|X_{1/2}\|_p$ uniformly in $\mathscr{L}(X)$. Applying this for each $t \in [0, \frac{1}{2}]$ to the spreadable process $Y(s) = X(2st)$ $s \in [0,1]$, we obtain $\|X_t^*\|_p \asymp \|X_t\|_p$ uniformly in $t \in [0, \frac{1}{2}]$ and $\mathscr{L}(X)$. This completes the proof of (9).

To extend the second relation in (9) to arbitrary $p > 0$, we may assume by Proposition 4.3 that X is ergodic exchangeable. Then a simple calculation gives $EX_t^2 = \alpha^2 t^2 + t(1-t)\gamma$, and so for $p \leq 2$ we get by Jensen's inequality and Lemma 4.11

$$\|X_t\|_p^2 \leq EX_t^2 \leq t(\alpha^2+\gamma) \asymp t\|X_s\|_p^2.$$

In the exchangeable case, Lemma 4.11 yields $\|X^*_{1/2}\|_p \asymp \|X_{1/2}\|_p$ for every $p > 0$, which extends as before to $\|X_t^*\|_p \asymp \|X_t\|_p$, uniformly in $t \in [0, \frac{1}{2}]$ and $\mathscr{L}(X)$. □

5. Predictable sampling

The main purpose of this section is to extend the *optional skipping* or *predictable sampling* property to spreadable sequences and processes. The result was originally proved by Doob (1936) for sequences of i.i.d. random variables (cf. Doob (1953), Theorem III.5.2, and the historical remarks in Halmos (1985), pp. 74–76). It was extended in Kallenberg (1982) to exchangeable sequences and processes on bounded or unbounded index sets.

We begin with the quite elementary discrete-time result.

Proposition 5.1. *Let $\xi = (\xi_1, \ldots, \xi_n)$ be an $\mathscr{F}$-spreadable sequence in some measurable space S and let $\tau_1 < \cdots < \tau_k$ be $\mathscr{F}$-predictable times in $\{1, \ldots, n\}$. Then*

$$(\xi_{\tau_1}, \ldots, \xi_{\tau_k}) \stackrel{d}{=} (\xi_1, \ldots, \xi_k). \tag{1}$$

This represents the ultimate extension for *monotone* sampling, since the stated property trivially implies spreadability of the underlying sequence or process. If ξ is exchangeable, then (1) remains true for any a.s. distinct, predictable times $\tau_1, \ldots, \tau_k$ in $\{1, \ldots, n\}$, *regardless of order* (cf. Kallenberg (1988a) or (1997), Theorem 9.19). The stronger version implies that ξ is exchangeable and is therefore false in the general spreadable case.

Proof. We proceed by induction on k, starting from the triviality for $k = 0$. Assuming the statement to hold for less than k predictable times, we turn to the case of k such times $\tau_1, \ldots, \tau_k$. For any measurable function $f : S^k \to \mathbb{R}_+$ and index $j \le n$ we get

$$\begin{aligned} E[f(\xi_{\tau_1}, \ldots, \xi_{\tau_k});\ \tau_1 = j] &= E[f(\xi_j, \xi_{\tau_2}, \ldots, \xi_{\tau_k});\ \tau_1 = j] \\ &= E[f(\xi_j, \xi_{j+1}, \ldots, \xi_{j+k-1});\ \tau_1 = j] \\ &= E[f(\xi_{n-k+1}, \ldots, \xi_n);\ \tau_1 = j], \end{aligned}$$

where the second equality holds by the induction hypothesis, applied to sequences of the form $(\theta_j \xi, \eta)$ with η an $\mathscr{F}_j$-measurable random element, and the last equality holds by the $\mathscr{F}$-spreadability of ξ and the predictability of τ_1. Summing over j and using once more the spreadability of ξ, we obtain

$$(\xi_{\tau_1}, \ldots, \xi_{\tau_k}) \overset{d}{=} (\xi_{n-k+1}, \ldots, \xi_n) \overset{d}{=} (\xi_1, \ldots, \xi_k),$$

which completes the induction. □

To state the corresponding result in continuous time, consider any $\mathbb{R}^d$-valued, $\mathscr{F}$-spreadable process X on $[0, 1]$. If X is a semimartingale (which holds by Proposition 3.4 when X has finite first moments), then for any predictable set $A \subset [0, 1]$ we may form the processes $\lambda_A = 1_A \cdot \lambda$ and $X_A = 1_A \cdot X$, or more explicitly,

$$\lambda_A(t) = \lambda(A \cap [0, t]),\ X_A(t) = \int_0^{t+} 1_A(s)\, dX_s, \quad t \in [0, 1], \tag{2}$$

where the second formula is understood in the sense of component-wise stochastic integration (cf. Kallenberg (1997), Chapter 23). Introducing the right-continuous inverse

$$\tau_s = \inf\{t \in [0, 1];\ \lambda_A(t) > s\}, \quad s \in [0, \lambda A], \tag{3}$$

we may define a process $\widehat{X}_A$ on $[0, \lambda A]$ by

$$\widehat{X}_A(s) = X_A(\tau_s), \quad s \in [0, \lambda A]. \tag{4}$$

For A in $\mathscr{U}$—the class of finite, nonrandom interval unions—this agrees with our definition from Section 1. Thus, for spreadable processes X we have $\widehat{X}_A \overset{d}{=} X$ on $[0, \lambda A]$ for all $A \in \mathscr{U}$. Our aim is to extend this relation to any predictable set $A \subset [0, 1]$.

Before we can state the general result, we need to make sense of $\widehat{X}_A$ for arbitrary X and A. This requires us to extend the integral $X_A = 1_A \cdot X$ in (2) to any spreadable processes X and predictable sets A. Though the stochastic integration $V \cdot X$ of an arbitrary bounded, predictable process V requires X to be a semimartingale (cf. Kallenberg (1997), Theorem 23.21), less may be needed when the integrand V takes values in $\{0, 1\}$. In any case, we shall see how the restriction map $A \mapsto X_A$ can be defined with appropriate additivity and continuity properties when X is an arbitrary spreadable process on $[0, 1]$.

If A is a.s. a finite union of intervals $(s, t]$, we define X_A as the usual Stieltjes integral $1_A \cdot X$. Any predictable set of this form can be written as

$$A = \bigcup\nolimits_{j \le m} (\sigma_j, \tau_j], \tag{5}$$

where the interval endpoints are optional times in $[0, 1]$ satisfying

$$\sigma_1 \le \tau_1 \le \sigma_2 \le \cdots \le \sigma_m \le \tau_m.$$

The associated *elementary predictable integral* X_A is given by

$$X_A(t) = \sum\nolimits_{j \le m} (X_{t \wedge \tau_j} - X_{t \wedge \sigma_j}), \quad t \in [0, 1]. \tag{6}$$

The following result gives the required extension to the predictable σ-field $\mathscr{P}$ on $[0, 1]$. Here we say that a mapping $A \mapsto X_A$ on $\mathscr{P}$ is *additive* if $X_{A \cup B} = X_A + X_B$ a.s. for any disjoint sets $A, B \in \mathscr{P}$.

Theorem 5.2. *Let X be an $\mathscr{F}$-spreadable process on* $[0, 1]$. *Then the elementary predictable integral in* (6) *extends a.s. uniquely to an additive map $A \mapsto X_A$ on $\mathscr{P}$ such that $\lambda A_n \xrightarrow{P} 0$ implies $(X_{A_n})^*_t \xrightarrow{P} 0$ for all $t \in [0, 1)$. The process X_A is a.s. rcll on* $[0, 1)$ *with $\Delta X_A = 1_A \Delta X$, and we have $X_A = 1_A \cdot X$ a.s. whenever X is a semimartingale on* $[0, 1)$.

Our proof is based on two lemmas. We say that A is a *simple predictable set* if it can be written as in (5) for some optional times σ_j and τ_j taking values in a fixed dyadic set $D_n = \{k2^{-n};\ k = 0, \ldots, 2^n\}$.

Lemma 5.3. *Let A be a simple predictable set in* $[0, 1]$ *with $\lambda A \ge t \ge 0$ a.s. Then $\widehat{X}_A \stackrel{d}{=} X$ on* $[0, t]$.

Proof. Let A be given by (5), where the interval endpoints σ_j and τ_j take values in the dyadic set D_n. Fixing a $t \in [0, 1]$ with $t \le \lambda A$ a.s., we may assume that $\lambda A = m2^{-n}$ a.s. for some integer $m \le 2^n$. Dividing A into m disjoint intervals $(a, b]$ of length 2^{-n}, we may next assume that $\tau_j = \sigma_j + 2^{-n}$ for all j. By Lemma 3.2 we may take $\mathscr{F}$ to be right-continuous, in which case $\sigma_1, \ldots, \sigma_m$ become $\mathscr{F}$-optional.

Let us now put $t_k = k2^{-n}$ for $k \le 2^n$ and introduce the processes

$$Y_k(s) = X_{t_{k-1}+s} - X_{t_{k-1}}, \quad s \in [0, 2^{-n}],\ \ k = 1, \ldots, 2^n.$$

The sequence $(Y_1, \ldots, Y_{2^n})$ is clearly spreadable with respect to the discrete filtration $\mathscr{G}_k = \mathscr{F}_{t_k}$, $k = 0, \ldots, 2^n$, and we note that the times $\kappa_j = 2^n \tau_j =$

$2^n\sigma_j + 1$ are integer-valued and $\mathscr{G}$-predictable with $\kappa_1 < \cdots < \kappa_m$. Hence, by Proposition 5.1,

$$(Y_{\kappa_1}, \ldots, Y_{\kappa_m}) \stackrel{d}{=} (Y_1, \ldots, Y_m),$$

which implies $\widehat{X}_A \stackrel{d}{=} X$ on $[0, m2^{-n}] \supset [0, t]$. □

Lemma 5.4. *Let* $A_1, A_2, \ldots$ *be simple predictable sets in* $[0, 1]$ *with* $\lambda A_n \stackrel{P}{\to} 0$. *Then* $(X_{A_n})_t^* \stackrel{P}{\to} 0$ *for all* $t \in [0, 1)$.

Proof. For any simple predictable set $A \subset [0, 1]$ we have

$$(X_A)_t^* = (\widehat{X}_A \circ \lambda_A)_t^* = (\widehat{X}_A)_{\rho_t}^*, \quad t \in [0, 1), \tag{7}$$

where $\rho_t = \lambda_A(t)$. Fixing any dyadic time $t \in [0, 1)$, we may choose for every $n \in \mathbb{N}$ some simple predictable set $A'_n \subset [0, 1]$ with $\lambda A'_n = 1 - t$ such that $A_n = A'_n$ on $[0, t]$ when $\lambda A_n \le 1 - t$. Letting $\varepsilon \in (0, 1 - t)$, we get by (7) and Lemma 5.3

$$\begin{aligned} E[(X_{A_n})_t^* \wedge 1] &\le E[(\widehat{X}_{A'_n})_\varepsilon^* \wedge 1] + P\{\lambda A_n > \varepsilon\} \\ &= E[X_\varepsilon^* \wedge 1] + P\{\lambda A_n > \varepsilon\}. \end{aligned}$$

Here the right-hand side tends to 0 as $n \to \infty$ and then $\varepsilon \to 0$, since $\lambda A_n \stackrel{P}{\to} 0$ and X is right-continuous with $X_0 = 0$. Thus, $(X_{A_n})_t^* \stackrel{P}{\to} 0$. □

Proof of Theorem 5.2. Fix any predictable set $A \subset [0, 1]$. By a monotone class argument, we may choose some simple predictable sets $A_1, A_2, \ldots \subset [0, 1]$ such that $\lambda(A \Delta A_n) \stackrel{P}{\to} 0$. Then $\lambda(A_m \Delta A_n) \stackrel{P}{\to} 0$ as $m, n \to \infty$, and so by Lemma 5.4 we have for any $t \in [0, 1)$

$$\begin{aligned} (X_{A_m} - X_{A_n})_t^* &= (X_{A_m \setminus A_n} - X_{A_n \setminus A_m})_t^* \\ &\le (X_{A_m \setminus A_n})_t^* + (X_{A_n \setminus A_m})_t^* \stackrel{P}{\to} 0. \end{aligned}$$

Thus, there exists a process X_A satisfying

$$(X_{A_n} - X_A)_t^* \stackrel{P}{\to} 0, \quad t \in [0, 1). \tag{8}$$

Note that X_A is a.s. rcll on $[0, 1)$, since this property holds trivially for each process X_{A_n}.

To see that the limit X_A is a.s. independent of approximating sequence (A_n), assume that also $\lambda(A \Delta A'_n) \stackrel{P}{\to} 0$ for some simple predictable sets A'_n. Then

$$\lambda(A_n \Delta A'_n) \le \lambda(A \Delta A_n) + \lambda(A \Delta A'_n) \stackrel{P}{\to} 0,$$

and so by (8) and Lemma 5.4,

$$(X_A - X_{A'_n})_t^* \le (X_A - X_{A_n})_t^* + (X_{A_n} - X_{A'_n})_t^* \stackrel{P}{\to} 0.$$

To prove the general continuity, let $A_1, A_2, \ldots \subset [0,1]$ be predictable with $\lambda A_n \xrightarrow{P} 0$ and fix any $t \in [0,1)$. Choose some simple predictable sets $A'_1, A'_2, \ldots \subset [0,1]$ such that

$$E\lambda(A_n \Delta A'_n) + E[(X_{A_n} - X_{A'_n})^*_t \wedge 1] \le n^{-1}, \quad n \in \mathbb{N}.$$

Then

$$\lambda A'_n \le \lambda A_n + \lambda(A_n \Delta A'_n) \xrightarrow{P} 0,$$

and so by Lemma 5.4

$$(X_{A_n})^*_t \le (X_{A'_n})^*_t + (X_{A_n} - X_{A'_n})^*_t \xrightarrow{P} 0.$$

The additivity of the mapping $A \mapsto X_A$ is obvious for simple predictable sets A. To prove the general result, consider any disjoint, predictable sets $A, B \subset [0,1]$. Choose some simple predictable sets A_n and B_n with $\lambda(A\Delta A_n) \xrightarrow{P} 0$ and $\lambda(B\Delta B_n) \xrightarrow{P} 0$, and note that the differences $B'_n = B_n \setminus A_n$ satisfy

$$\lambda(B\Delta B'_n) \le \lambda(A\Delta A_n) + \lambda(B\Delta B_n) + \lambda(A \cap B) \xrightarrow{P} 0.$$

We may then assume that $A_n \cap B_n = \emptyset$ for all n. Also note that

$$\lambda((A \cup B)\Delta(A_n \cup B_n)) \le \lambda(A\Delta A_n) + \lambda(B\Delta B_n) \xrightarrow{P} 0.$$

Since $X_{A_n \cup B_n} = X_{A_n} + X_{B_n}$ for every n, we get for $t \in [0,1)$

$$\begin{aligned}&(X_{A\cup B} - X_A - X_B)^*_t \\ &\quad \le (X_{A\cup B} - X_{A_n \cup B_n})^*_t + (X_{A_n} - X_A)^*_t + (X_{B_n} - X_B)^*_t \xrightarrow{P} 0,\end{aligned}$$

which shows that $X_{A\cup B} = X_A + X_B$ a.s.

The relation $\Delta X_A = 1_A \Delta X$ is clearly true for simple predictable sets A. To extend the formula to the general case, we choose some simple predictable sets A_n with $\lambda(A\Delta A_n) \xrightarrow{P} 0$ and note that

$$(\Delta X_{A_n} - \Delta X_A)^*_t \xrightarrow{P} 0, \quad t \in [0,1),$$

since $(X_{A_n} - X_A)^*_t \xrightarrow{P} 0$. It remains to verify that

$$(1_{A_n}\Delta X - 1_A \Delta X)^*_t \xrightarrow{P} 0, \quad t \in [0,1).$$

An equivalent claim is that $(1_{A_n}\Delta X)^*_t \xrightarrow{P} 0$ on $[0,1)$ for any predictable sets A_n with $\lambda A_n \xrightarrow{P} 0$. This follows if we can show that the jump point process ξ of X satisfies

$$\xi(A^t_n \times B_\varepsilon) \xrightarrow{P} 0, \quad t < 1, \ \varepsilon > 0,$$

where $A^t_n = A_n \cap [0,t]$ and $B_\varepsilon = \mathbb{R} \setminus (-\varepsilon, \varepsilon)$. To this aim, put

$$\tau_k = \inf\{t \in [0,1];\ \xi([0,t] \times B_\varepsilon) \ge k\}, \quad k \in \mathbb{N}.$$

By dual predictable projection followed by dominated convergence as $n \to \infty$, we have

$$E \int_0^{\tau_k +} 1_{A_n^t \times B_\varepsilon}(s, x)\, \xi(ds\, dx) = E \int_0^{\tau_k} 1_{A_n^t \times B_\varepsilon}(s, x)\, \hat{\hat{\xi}}(ds\, dx)$$
$$= E \int_0^{\tau_k} 1_{A_n^t}(s)\, \mu_s B_\varepsilon\, ds \to 0,$$

where (μ_s) denotes the martingale density in Proposition 3.4 (i). As $n \to \infty$, the inner integral on the left tends in probability to 0, and the assertion follows since k is arbitrary.

If X is a semimartingale with decomposition $M + V$, we may choose the approximating sets A_n in (8) such that

$$\int_0^t 1_{A \Delta A_n}\, d[M] + \int_0^t 1_{A \Delta A_n}\, |dV| \xrightarrow{P} 0, \quad t \in [0, 1).$$

For $t \in [0, 1)$ we get

$$(X_A - 1_A \cdot X)_t^* \le (X_A - X_{A_n})_t^* + (1_A \cdot X - X_{A_n})_t^* \xrightarrow{P} 0,$$

which implies $X_A = 1_A \cdot X$ a.s. □

We may now state the continuous-time version of our predictable sampling theorem. Given any spreadable process X and predictable set A, we define the process $\widehat{X}_A$ by (3) and (4), where X_A is given by Theorem 5.2.

Theorem 5.5. *Let X be an $\mathbb{R}^d$-valued, $\mathscr{F}$-spreadable process on $[0, 1]$ and let $A \subset [0, 1]$ be $\mathscr{F}$-predictable. Then*

$$\widehat{X}_A \stackrel{d}{=} X \ \text{on}\ [0, t), \quad t \in [0, 1] \ \text{with}\ \lambda A \ge t \ \text{a.s.}$$

Again a much more general version—essentially invariance in distribution under arbitrary predictable and measure-preserving transformations—holds for exchangeable processes on $[0, 1]$ or $\mathbb{R}_+$ (cf. Kallenberg (1988a) and the special case in Kallenberg (1997), Proposition 16.9).

Proof. By a monotone class argument, we may choose some simple predictable sets $A_1, A_2, \ldots$ in $[0, 1]$ with $\lambda(A \Delta A_n) \xrightarrow{P} 0$. Fixing any dyadic numbers $t_n \uparrow t$, we note that the times

$$\sigma_n = \inf\{s \le 1;\ \lambda_{A_n^c}(s) > 1 - t_n\}, \quad n \in \mathbb{N},$$

are optional and take values in some fixed dyadic sets D_m. Hence, the sets $A_n' = A_n \cup (\sigma_n, 1]$ are again simple predictable, and we have

$$\lambda A_n' = \lambda_{A_n}(\sigma_n) + 1 - \sigma_n \ge t_n,$$
$$\lambda(A_n \Delta A_n') = \lambda(A_n^c \cap (\sigma_n, 1]) = (t_n - \lambda A_n) \vee 0$$
$$\le \lambda(A \Delta A_n) \xrightarrow{P} 0,$$

which shows that $\lambda(A \Delta A'_n) \xrightarrow{P} 0$. Dropping the primes, we may henceforth assume that $\lambda A_n \geq t_n$ a.s. for all n. Then Lemma 5.3 yields

$$\widehat{X}_{A_n} \stackrel{d}{=} X \text{ on } [0, t_n], \quad n \in \mathbb{N}. \tag{9}$$

Now introduce as in (3) the right-continuous inverses $\tau = (\tau_s)$ and $\tau^n = (\tau_s^n)$ of the processes λ_A and λ_{A_n}, respectively. As $n \to \infty$,

$$(\lambda_{A_n} - \lambda_A)^* \leq \lambda(A_n \Delta A) \xrightarrow{P} 0,$$

and so $(\lambda_{A_n} - \lambda_A)^* \to 0$ a.s. along a subsequence $N' \subset \mathbb{N}$. Outside the same P-null set, we obtain $\tau_s^n \to \tau_s$ along N' for every continuity point $s < t$ of τ, and so by Fubini's theorem, with convergence along N',

$$\int_0^t P\{\tau_s^n \not\to \tau_s\}\, ds = E\lambda\{s < t;\ \tau_s^n \not\to \tau_s\} = 0,$$

which implies

$$P\{\tau_s^n \to \tau_s\} = 1, \quad s < t \text{ a.e. } \lambda. \tag{10}$$

We also note that a.s. $\Delta X_A(\tau_s) = 0$ for all but countably many $s < t$ since X_A is rcll and τ is strictly increasing. By Fubini's theorem, we obtain

$$\int_0^t P\{\Delta X_A(\tau_s) \neq 0\}\, ds = E\lambda\{s < t;\ \Delta X_A(\tau_s) \neq 0\} = 0,$$

which gives

$$P\{\Delta X_A(\tau_s) = 0\} = 1, \quad s < t \text{ a.e. } \lambda. \tag{11}$$

By Theorem 5.2 we have $(X_{A_n} - X_A)_r^* \xrightarrow{P} 0$ for every $r < 1$, and so a.s. $(X_{A_n} - X_A)_r^* \to 0$ for all $r < 1$ along some further subsequence $N'' \subset N'$. Thus, with probability 1, $X_{A_n}(r_n) \to X_A(r)$ for every sequence $r_n \to r < 1$ with $\Delta X_A(r) = 0$. Noting that $\tau_s < 1$ a.s. for all $s < t$, we get by (4), (10), and (11) for almost every $s < t$

$$\widehat{X}_{A_n}(s) = X_{A_n}(\tau_s^n) \to X_A(\tau_s) = \widehat{X}_A(s),$$

a.s. along N''. Using (9) gives

$$\begin{aligned}(X_{s_1}, \ldots, X_{s_m}) &\stackrel{d}{=} \big(\widehat{X}_{A_n}(s_1), \ldots, \widehat{X}_{A_n}(s_m)\big)\\ &\xrightarrow{d} \big(\widehat{X}_A(s_1), \ldots, \widehat{X}_A(s_m)\big)\end{aligned}$$

for $s_1, \ldots, s_m < t$ outside some λ-null set. Thus,

$$\big(\widehat{X}_A(s_1), \ldots, \widehat{X}_A(s_m)\big) \stackrel{d}{=} (X_{s_1}, \ldots, X_{s_m}),$$

which extends to arbitrary $s_1, \ldots, s_m < t$ since both X and $\widehat{X}_A$ are right-continuous, the latter by the right continuity of X_A and τ. This shows that indeed $\widehat{X}_A \stackrel{d}{=} X$ on $[0, t)$. □

6. Wald-type identities

This final section is devoted to some general *Wald-type* or *decoupling identities* for spreadable sequences and processes. The topic originated with some elementary relations by Wald (1945) for sums of i.i.d. random variables. In continuous time, the earliest results are equivalent to some basic martingale properties of stochastic integrals. General versions for exchangeable sequences and processes on bounded or unbounded index sets were established in Kallenberg (1989). All these results are based on the remarkable observation that certain product moments involving stochastic sums or integrals can be computed as if the integrands and integrators were independent. The general identities are powerful enough to imply the corresponding versions of the predictable sampling theorem.

In the spreadable case, the basic results are identities involving certain *tetrahedral moments*, from which identities for ordinary product moments can be derived as easy corollaries. We begin with the basic relation in discrete time. To avoid distracting technicalities, we consider only bounded random variables. Binomial coefficients are denoted by $c_{n,k} = n!/k!(n-k)!$, and we shall often write $(xy)^j_k = x^j_k y^j_k$ for convenience.

Proposition 6.1. *Let $\xi = (\xi^1, \dots, \xi^d)$ and $\eta = (\eta^1, \dots, \eta^d)$ be bounded random sequences in $\mathbb{R}^d$ of length $n \geq d$ such that ξ is $\mathscr{F}$-spreadable and η is $\mathscr{F}$-predictable, and assume that the sums*

$$S_j = \sum_{k_j < \cdots < k_d} \cdots \sum \eta^j_{k_j} \cdots \eta^d_{k_d}, \quad j = 1, \dots, d, \tag{1}$$

are $\mathscr{F}_0$-measurable. Then

$$E \sum_{k_1 < \cdots < k_d} \cdots \sum (\xi\eta)^1_{k_1} \cdots (\xi\eta)^d_{k_d}$$

$$= c^{-1}_{n,d}\, E \sum_{h_1 < \cdots < h_d} \cdots \sum \xi^1_{h_1} \cdots \xi^d_{h_d} \sum_{k_1 < \cdots < k_d} \cdots \sum \eta^1_{k_1} \cdots \eta^d_{k_d}.$$

Before providing a proof, we state the corresponding identity for product moments. Here we write $x^J = \prod_{j \in J} x^j$ for convenience. By an *ordered partition* of a set J we mean a partition of J into disjoint, nonempty subsets $J_1, \dots, J_m$, listed in a specified order.

Corollary 6.2. *Let $\xi = (\xi^1, \dots, \xi^d)$ and $\eta = (\eta^1, \dots, \eta^d)$ be finite, bounded sequences in $\mathbb{R}^d$ such that ξ is $\mathscr{F}$-spreadable and η is $\mathscr{F}$-predictable, and assume that the sums*

$$\sum_{k_1 < \cdots < k_m} \cdots \sum \eta^{J_1}_{k_1} \cdots \eta^{J_m}_{k_m}, \quad J_1, \dots, J_m \text{ disjoint in } \{1, \dots, d\},$$

are $\mathscr{F}_0$-measurable. Then

$$E \prod_j \sum_k (\xi\eta)^j_k$$

$$= \sum_{J_1,\dots,J_m} c^{-1}_{n,m}\, E \sum_{h_1 < \cdots < h_m} \cdots \sum \xi^{J_1}_{h_1} \cdots \xi^{J_m}_{h_m} \sum_{k_1 < \cdots < k_m} \cdots \sum \eta^{J_1}_{k_1} \cdots \eta^{J_m}_{k_m},$$

where the outer summation on the right extends over all ordered partitions $J_1, \ldots,$ J_m *of* $\{1, \ldots, d\}$.

This extends a result for exchangeable sequences in Theorem 3.1 of Kallenberg (1989). Note, however, that stronger hypotheses are needed in the spreadable case, since we could otherwise proceed as in Section 6 of the same paper to show that the sequence is in fact exchangeable.

Proof. Combine Proposition 6.1 with the elementary decomposition

$$\prod_{j \le d} \sum\nolimits_{k \le n} x_k^j = \sum\nolimits_{J_1, \ldots, J_m} \sum_{k_1 < \cdots < k_m} \!\!\cdots\!\! \sum x_{k_1}^{J_1} \cdots x_{k_m}^{J_m},$$

where the outer summation on the right extends over all ordered partitions $J_1, \ldots,$ J_m of $\{1, \ldots, d\}$. □

For the proof of Proposition 6.1 we need a simple lemma.

Lemma 6.3. *Let* $\eta = (\eta^1, \ldots, \eta^d)$ *be an* $\mathscr{F}$*-predictable sequence in* $\mathbb{R}^d$ *of length* $n \ge d$ *such that the sums* S_j *in* (1) *are* $\mathscr{F}_0$*-measurable. Then the sequence*

$$T_r = \sum_{r < k_1 < \cdots < k_d} \!\!\cdots\!\! \sum \eta_{k_1}^1 \cdots \eta_{k_d}^d, \quad r = 0, \ldots, n-d,$$

is again $\mathscr{F}$*-predictable.*

Proof. The statement is obvious for $d = 1$. Proceeding by induction, we assume that the statement is true with d replaced by $d - 1$. Turning to the case of d, we note that

$$T_r = S_1 - \sum_{k \le r} \eta_k^1 \sum_{k < k_2 < \cdots < k_d} \!\!\cdots\!\! \sum \eta_{k_2}^2 \cdots \eta_{k_d}^d, \quad r = 1, \ldots, n-d.$$

Applying the induction hypothesis to the $d - 1$ -fold inner sum, we see that the kth term on the right is $\mathscr{F}_{k-1}$-measurable. The assertion now follows since $\mathscr{F}_h \subset \mathscr{F}_k$ for $h \le k$. □

Proof of Proposition 6.1. Using repeatedly the $\mathscr{F}$-spreadability of ξ, the $\mathscr{F}$-predictability of η, and Lemma 6.3, we get

$$\begin{aligned}
&E \sum_{k_1 < \cdots < k_d} \!\!\cdots\!\! \sum (\xi\eta)_{k_1}^1 \cdots (\xi\eta)_{k_d}^d \\
&\quad = E\, \xi_n^d \sum_{k_1 < \cdots < k_d} \!\!\cdots\!\! \sum (\xi\eta)_{k_1}^1 \cdots (\xi\eta)_{k_{d-1}}^{d-1} \eta_{k_d}^d \\
&\quad = E\, \xi_n^d \sum_{k_1 < \cdots < k_{d-1}} \!\!\cdots\!\! \sum (\xi\eta)_{k_1}^1 \cdots (\xi\eta)_{k_{d-1}}^{d-1} \sum_{k_d > k_{d-1}} \eta_{k_d}^d \\
&\quad = E\, \xi_{n-1}^{d-1} \xi_n^d \sum_{k_1 < \cdots < k_{d-1}} \!\!\cdots\!\! \sum (\xi\eta)_{k_1}^1 \cdots (\xi\eta)_{k_{d-2}}^{d-2} \eta_{k_{d-1}}^{d-1} \sum_{k_d > k_{d-1}} \eta_{k_d}^d
\end{aligned}$$

$$= E\, \xi_{n-1}^{d-1} \xi_n^d \sum_{k_1 < \cdots < k_{d-2}} \cdots \sum (\xi\eta)_{k_1}^1 \cdots (\xi\eta)_{k_{d-2}}^{d-2} \sum_{k_d > k_{d-1} > k_{d-2}} \sum \eta_{k_{d-1}}^{d-1} \eta_{k_d}^d$$

$$= \cdots = E\, \xi_{n-d+1}^1 \cdots \xi_n^d \sum_{k_1 < \cdots < k_d} \cdots \sum \eta_{k_1}^1 \cdots \eta_{k_d}^d$$

$$= c_{n,d}^{-1}\, E \sum_{h_1 < \cdots < h_d} \cdots \sum \xi_{h_1}^1 \cdots \xi_{h_d}^d \sum_{k_1 < \cdots < k_d} \cdots \sum \eta_{k_1}^1 \cdots \eta_{k_d}^d. \qquad \square$$

We turn to the basic tetrahedral identity in continuous time. Again we may avoid some technical complications by considering only bounded integrands and integrators of bounded variation. The *tetrahedral regions* Δ_k are given by

$$\Delta_k = \{(s_1, \ldots, s_k) \in [0,1]^k;\ s_1 < \cdots < s_k\}, \quad k \in \mathbb{N}.$$

Theorem 6.4. *Consider on* $[0,1]$ *an* $\mathscr{F}$*-spreadable process* $X = (X^1, \ldots, X^d)$ *of bounded variation and some bounded,* $\mathscr{F}$*-predictable processes* $V^1, \ldots, V^d$ *such that the integrals*

$$\eta_k = \int \cdots \int_{\Delta_{d-k+1}} V^k \cdots V^d, \quad k = 1, \ldots, d, \tag{2}$$

are $\mathscr{F}_0$*-measurable. Then*

$$E \int \cdots \int_{\Delta_d} V^1 dX^1 \cdots V^d dX^d$$
$$= d!\, E \int \cdots \int_{\Delta_d} dX^1 \cdots dX^d \int \cdots \int_{\Delta_d} V^1 \cdots V^d. \tag{3}$$

As a consequence, we obtain the following product moment identity, which extends Theorem 4.1 in Kallenberg (1989) for exchangeable processes. Again we note that stronger hypotheses are needed in the spreadable case, since we could otherwise use the result to prove that X is exchangeable.

For any two semimartingales X and Y we write $dXdY = d[X,Y]$. Differentials of higher order are defined recursively, so that for any semimartingales $X^1, \ldots, X^d$ we have

$$\int_0^t dX_s^1 \cdots dX_s^d = \sum_{s \le t} \Delta X_s^1 \cdots \Delta X_s^d, \quad t \ge 0,\ d \ge 3.$$

If $X_0^1 = \cdots = X_0^d = 0$, we put

$$X_t^J = \int_0^t \prod_{j \in J} dX_s^j, \quad \emptyset \ne J \subset \{1, \ldots, d\}.$$

Finally, we use V_s^J to denote the product $\prod_{j \in J} V_s^j$.

Corollary 6.5. *Consider on* $[0, 1]$ *an* $\mathscr{F}$*-spreadable process* $X = (X^1, \ldots, X^d)$ *of bounded variation and some bounded,* $\mathscr{F}$*-predictable processes* $V^1, \ldots, V^d$ *such that the integrals*

$$\int \cdots \int_{\Delta_m} V^{J_1} \cdots V^{J_m}, \quad J_1, \ldots, J_m \textit{ disjoint in } \{1, \ldots, d\},$$

are $\mathscr{F}_0$*-measurable. Then*

$$E \prod\nolimits_{j \le d} \int_0^1 V^j dX^j$$
$$= \sum\nolimits_{J_1,\ldots,J_k} k!\, E \int \cdots \int_{\Delta_k} dX^{J_1} \cdots dX^{J_k} \int \cdots \int_{\Delta_k} V^{J_1} \cdots V^{J_k},$$

where the summation extends over all ordered partitions $J_1, \ldots, J_k$ *of the set* $\{1, \ldots, d\}$.

This follows from Theorem 6.4 by means of the following tetrahedral decomposition, which generalizes the integration-by-parts formula for general semimartingales (cf. Kallenberg (1997), Theorem 23.6).

Lemma 6.6. *Let* $X^1, \ldots, X^d$ *be real semimartingales starting at* 0. *Then*

$$X_t^1 \cdots X_t^d = \sum_{J_1,\ldots,J_k} \int \cdots \int_{s_1 < \cdots < s_k \le t} dX_{s_1}^{J_1} \cdots dX_{s_k}^{J_k}, \quad t \ge 0,$$

where the summation extends over all ordered partitions $J_1, \ldots, J_k$ *of the set* $\{1, \ldots, d\}$.

Proof. By the substitution rule for general semimartingales (cf. Kallenberg (1997), Theorem 23.7), we have

$$X_t^1 \cdots X_t^d = \sum\nolimits_J \int_0^t dX_s^J \prod\nolimits_{k \notin J} X_{s-}^k, \quad t \ge 0,$$

where the summation extends over all nonempty subsets $J \subset \{1, \ldots, d\}$. The assertion now follows by iteration in finitely many steps. □

Several lemmas are needed for the proof of Theorem 6.4. We begin with a continuous-time version of Lemma 6.3, which can be proved by a similar argument.

Lemma 6.7. *Let* $V^1, \ldots, V^d$ *be bounded,* $\mathscr{F}$*-predictable processes on* $[0, 1]$ *such that the integrals in* (2) *are* $\mathscr{F}_0$*-measurable. Then the process*

$$Y_t = \int_t^1 V_{s_1}^1 ds_1 \int_{s_1}^1 V_{s_2}^2 ds_2 \int_{s_2}^1 \cdots \int_{s_{d-1}}^1 V_{s_d}^d ds_d, \quad t \in [0, 1],$$

is again $\mathscr{F}$*-predictable.*

Theorem 6.4 will first be proved under a simplifying assumption.

Lemma 6.8. *The statement of Theorem* 6.4 *holds when* $V^1, \dots, V^d$ *are supported by some interval* $[0, 1-\varepsilon]$ *with* $\varepsilon > 0$.

Proof. Though our formal argument is similar to the proof of Proposition 6.1, the justification requires a lot more care. To explain the key steps, consider a spreadable process X of bounded variation and a bounded, predictable process V. Recall from Proposition 3.4 that X is a special semimartingale, whose compensator $\hat{X}$ admits a martingale density M on $[0, 1)$. Using repeatedly the definition of M, the martingale properties of $X - \hat{X}$ and M, and Fubini's theorem, we get for $0 < t < t + h \le 1$

$$\begin{aligned} E\int_0^t V_s dX_s &= E\int_0^t V_s d\hat{X}_s = E\int_0^t V_s M_s ds \\ &= \int_0^t E(V_s M_s) ds = \int_0^t E(V_s M_t) ds \\ &= E\, M_t \int_0^t V_s ds = h^{-1} E(X_{t+h} - X_t) \int_0^t V_s ds. \end{aligned}$$

Let us now put $t_k = 1 - \varepsilon(d-k)/d$ for $k = 0, \dots, d$ and define

$$\rho_k = \frac{X^k_{t_k} - X^k_{t_{k-1}}}{t_k - t_{k-1}}, \quad k = 1, \dots, d.$$

Since X remains $\mathscr{F}$-spreadable on $[0, 1-\varepsilon]$ under the "probability" measures $E[\rho_d \cdots \rho_k; \cdot]$, we obtain more generally

$$E\,\rho_d \cdots \rho_{k+1} \int_0^{1-\varepsilon} V_s dX_s = E\,\rho_d \cdots \rho_k \int_0^{1-\varepsilon} V_s ds, \quad k = 1, \dots, d. \tag{4}$$

To explain our second key step, consider a semimartingale X and two bounded, predictable processes U and V where $\int_0^1 U_t dt$ is $\mathscr{F}_0$-measurable. Write $A_t = \int_0^t U_s ds$ and $Y_t = \int_0^t V_s dX_s$, and note that $[A, Y] = 0$. Integrating by parts gives

$$\int_0^1 Y_{t-} dA_t = A_1 Y_1 - \int_0^1 A_t dY_t = \int_0^1 (A_1 - A_t) dY_t$$

or

$$\int_0^1 U_t dt \int_0^{t-} V_s dX_s = \int_0^1 V_t dX_t \int_t^1 U_s ds. \tag{5}$$

Now to prove (3), we may imitate the discrete-time argument, alternating the use of (4) and (5) as follows:

$$\begin{aligned} &E\int \cdots \int_{\Delta_d} V^1_{s_1} dX^1_{s_1} \cdots V^d_{s_d} dX^d_{s_d} \\ &= E\int_0^1 V^d_{s_d} dX^d_{s_d} \int_0^{s_d-} V^{d-1}_{s_{d-1}} dX^{d-1}_{s_{d-1}} \cdots \int_0^{s_2-} V^1_{s_1} dX^1_{s_1} \\ &= E\,\rho_d \int_0^1 V^d_{s_d} ds_d \int_0^{s_d-} V^{d-1}_{s_{d-1}} dX^{d-1}_{s_{d-1}} \cdots \int_0^{s_2-} V^1_{s_1} dX^1_{s_1} \end{aligned}$$

$$
\begin{aligned}
&= E\,\rho_d \int_0^1 V_{s_{d-1}}^{d-1} dX_{s_{d-1}}^{d-1} \int_{s_{d-1}}^1 V_{s_d}^d ds_d \int_0^{s_{d-1}-} V_{s_{d-2}}^{d-2} dX_{s_{d-2}}^{d-2} \cdots \\
&= E\,\rho_d \rho_{d-1} \int_0^1 V_{s_{d-1}}^{d-1} ds_{d-1} \int_{s_{d-1}}^1 V_{s_d}^d ds_d \int_0^{s_{d-1}-} V_{s_{d-2}}^{d-2} dX_{s_{d-2}}^{d-2} \cdots \\
&= E\,\rho_d \rho_{d-1} \int_0^1 V_{s_{d-2}}^{d-2} dX_{s_{d-2}}^{d-2} \int_{s_{d-2}}^1 V_{s_{d-1}}^{d-1} ds_{d-1} \int_{s_{d-1}}^1 V_{s_d}^d ds_d \cdots \\
&= E\,\rho_d \rho_{d-1} \rho_{d-2} \int_0^1 V_{s_{d-2}}^{d-2} ds_{d-2} \int_{s_{d-2}}^1 V_{s_{d-1}}^{d-1} ds_{d-1} \int_{s_{d-1}}^1 V_{s_d}^d ds_d \cdots \\
&= \cdots = E\,\rho_d \cdots \rho_1 \int_0^1 V_{s_1}^1 ds_1 \int_{s_1}^1 V_{s_2}^2 ds_2 \int_{s_2}^1 \cdots \int_{s_{d-1}}^1 V_{s_d}^d ds_d \\
&= E\,\rho_1 \cdots \rho_d \int_0^1 V_{s_d}^d ds_d \int_0^{s_d} V_{s_{d-1}}^{d-1} ds_{d-1} \int_0^{s_{d-1}} \cdots \int_0^{s_2} V_{s_1}^1 ds_1 \\
&= d!\, E \int \cdots \int_{\Delta_d} dX_{r_1}^1 \ldots dX_{r_d}^d \int \cdots \int_{\Delta_d} V_{s_1}^1 ds_1 \cdots V_{s_d}^d ds_d .
\end{aligned}
$$

Here the first equality holds by definitions, equalities 2, 4, 6, ... hold by (4), and equalities 3, 5, 7, ... hold by (5) together with Lemma 6.7. The process is continued recursively until all stochastic integrals are converted into associated Lebesgue integrals.

The second relation from the end holds by Fubini's theorem. Finally, the last equality follows by the same computations, in the special case when $V_s^1, \ldots, V_s^n$ are $\mathscr{F}_0$-measurable and independent of s. □

For the extension to the general case, we need to employ a measurable selection based on the following lemma.

Lemma 6.9. *Let ξ and η be random elements in some spaces S and T where T is Borel, and assume that $f(\xi, \eta) = 0$ a.s. for some measurable function $f: S \times T \to \mathbb{R}$. Then there exists a ξ-measurable random element $\hat{\eta}$ of T such that $f(\xi, \hat{\eta}) = 0$ a.s.*

Proof. Put $A = f^{-1}\{0\}$ and let πA denote the projection of A on S. By the general section theorem (cf. Dellacherie (1972), Theorem T37) there exists a measurable function $g: S \to T$ such that

$$(\xi, g(\xi)) \in A \;\text{ a.s. on }\; \{\xi \in \pi A\}.$$

Since $(\xi, \eta) \in A$ implies $\xi \in \pi A$, we also note that

$$P\{\xi \in \pi A\} \geq P\{(\xi, \eta) \in A\} = 1.$$

Thus, the assertion holds with $\hat{\eta} = g(\xi)$. □

The following truncation lemma will be needed to reduce the proof of Theorem 6.4 to the special case of Lemma 6.8.

Lemma 6.10. *Let $V^1, \ldots, V^d$ be $\mathscr{F}$-predictable processes on* $[0, 1]$ *such that* $|V^k| \le 1$ *and the integrals η_k in* (2) *are $\mathscr{F}_0$-measurable. Then for every $\varepsilon \in (0, \frac{1}{2}]$ there exist some predictable processes $\tilde{V}^1, \ldots, \tilde{V}^d$ with a.s. the same values of the integrals in* (2)*, such that for any $k \le d$ we have $\tilde{V}^k = V^k$ on* $[0, 1-2\varepsilon]$*,* $|\tilde{V}^k| \le 2$ *on* $(1-2\varepsilon, 1-\varepsilon]$*, and* $\tilde{V}^k = 0$ *on* $(1-\varepsilon, 1]$.

Proof. Define the random signed measures $\xi_1, \ldots, \xi_n$ on $[0, 1]$ by

$$\xi_k B = \int_B V_s^k ds, \quad B \in \mathscr{B}([0, 1]), \quad 1 \le k \le d.$$

Since the V^k are predictable and hence progressively measurable, we note that the ξ_k are adapted to $\mathscr{F}$. Equation (2) and the conditions $|V^k| \le 1$ translate into

$$(\xi_k \otimes \cdots \otimes \xi_d)\Delta_{d-k+1} = \eta_k, \qquad 1 \le k \le d, \tag{6}$$

$$|\xi_k[a, b]| \le b - a, \quad 0 \le a \le b \le 1, \quad 1 \le k \le d. \tag{7}$$

Approximating the region Δ_{d-k+1} by finite unions of rectangles, we see by dominated convergence that (6) defines a measurable constraint on the random measures $\xi_1, \ldots, \xi_d$ and variables $\eta_1, \ldots, \eta_d$. The same thing is true for (7) since it suffices to consider rational a and b. The whole collection of conditions (6) and (7) may then be summarized by an equation

$$F(\xi_1, \ldots, \xi_d;\ \eta_1, \ldots, \eta_d) = 0 \text{ a.s.}$$

for some measurable function F.

Now fix any $\varepsilon \in (0, \frac{1}{2}]$ and let $\xi_1', \ldots, \xi_d'$ denote the restrictions of $\xi_1, \ldots, \xi_d$ to $[0, 1-2\varepsilon]$. By Lemma 6.9 there exist some signed random measures $\hat{\xi}_1, \ldots, \hat{\xi}_d$, each measurable with respect to $\xi_1', \ldots, \xi_d'$ and $\eta_1, \ldots, \eta_d$, such that $\hat{\xi}_k = \xi_k$ on $[0, 1-2\varepsilon]$ for all k and

$$F(\hat{\xi}_1, \ldots, \hat{\xi}_d;\ \eta_1, \ldots, \eta_d) = 0 \text{ a.s.}$$

In other words, the $\hat{\xi}_k$ are $\mathscr{F}_{1-2\varepsilon}$-measurable and satisfy (6) and (7) a.s.

From the version of (7) for $\hat{\xi}_1, \ldots, \hat{\xi}_d$ we note that these measures are a.s. absolutely continuous with densities $\hat{V}^1, \ldots, \hat{V}^d$ bounded by ± 1. By the martingale approach to the Radon–Nikodym theorem (cf. Doob (1953), Section VII.8), we may choose the $\hat{V}^k$ to be $\mathscr{F}_{1-2\varepsilon} \otimes \mathscr{B}$ -measurable on $(1-2\varepsilon, 1]$. Completing the definition by taking $\hat{V}^k = V^k$ on $[0, 1-2\varepsilon]$ for all k, we note that the $\hat{V}^k$ become a.s. bounded by ± 1 and satisfy (2).

We now introduce the function

$$f(t) = t - \tfrac{1}{2}(t - 1 + 2\varepsilon)_+, \quad t \in [0, 1], \tag{8}$$

and consider for each $k \in \{1, \ldots, d\}$ the signed random measure $\tilde{\xi}_k = \hat{\xi}_k \circ f^{-1}$ on $[0, 1-\varepsilon]$. Since f is strictly increasing, we have $f(t_1) < \cdots < f(t_k)$ iff $t_1 < \cdots < t_k$, and so

$$(f^{\otimes k})^{-1}\Delta_k = \Delta_k, \quad k \in \mathbb{N}.$$

Thus, (6) remains a.s. fulfilled for $\tilde{\xi}_1, \ldots, \tilde{\xi}_d$. Also note that $\tilde{\xi}_k = \xi_k$ on $[0, 1-2\varepsilon]$ for all k.

Inverting (8), we note that on $(1-2\varepsilon, 1-\varepsilon]$ the random measures $\tilde{\xi}_1, \ldots, \tilde{\xi}_d$ have a.s. the densities

$$\tilde{V}_t^k = 2\hat{V}^k(2t - 1 + 2\varepsilon), \quad \varepsilon \le 1 - t < 2\varepsilon, \ \ 1 \le k \le d. \tag{9}$$

We may complete the construction by setting $\tilde{V}^k = V^k$ on $[0, 1-2\varepsilon]$ and $\tilde{V}^k = 0$ on $(1-\varepsilon, 1]$. The $\tilde{V}^k$ are again $\mathscr{F}_{1-2\varepsilon} \otimes \mathscr{B}$ -measurable and hence predictable on $(1-2\varepsilon, 1-\varepsilon]$. The predictability also holds trivially on the intervals $[0, 1-2\varepsilon]$ and $(1-\varepsilon, 1]$. Furthermore, (2) remains a.s. true for $\tilde{V}^1, \ldots, \tilde{V}^d$ since the measures $\tilde{\xi}_1, \ldots, \tilde{\xi}_d$ satisfy (6) a.s. Finally, (9) shows that the processes $\tilde{V}^k$ are a.s. bounded by ± 2. □

Proof of Theorem 6.4. For any $\varepsilon \in (0, \frac{1}{2}]$, let $V_\varepsilon^1, \ldots, V_\varepsilon^d$ denote the truncated processes constructed in Lemma 6.10, and conclude from Lemma 6.8 that (3) holds with $V^1, \ldots, V^d$ replaced by $V_\varepsilon^1, \ldots, V_\varepsilon^d$. As $\varepsilon \to 0$ we have $V_\varepsilon^k \to V^k$ on $[0, 1)$ for all k, and (3) follows in the stated form by dominated convergence. □

We may use the present results to give alternative proofs of the predictable sampling theorems of Section 5, at least under some simplifying assumptions.

Second proof of Proposition 5.1. We may assume that the ξ_j take values in $[0, 1]$. By the Cramér–Wold theorem, it suffices to show for any $c_1, \ldots, c_k \in \mathbb{R}$ that $\sum_j c_j \xi_{\tau_j} \overset{d}{=} \sum_j c_j \xi_j$. Since both sides are bounded, it is equivalent that

$$E\Big(\sum\nolimits_j c_j \xi_{\tau_j}\Big)^n = E\Big(\sum\nolimits_j c_j \xi_j\Big)^n, \quad n \in \mathbb{N}. \tag{10}$$

To see this, we introduce the predictable sequence

$$\alpha_i = \inf\{j;\ \tau_j = i\}, \quad i = 1, \ldots, l,$$

and note that $\sum_j c_j \xi_{\tau_j} = \sum_i c_{\alpha_i} \xi_i$ where $c_\infty = 0$ by convention. Equations (10) now follow from Corollary 6.2 if we can only show that

$$\sum_{i_1 < \cdots < i_h} \!\!\cdots \sum c_{\alpha_{i_1}}^{r_1} \cdots c_{\alpha_{i_h}}^{r_h} = \sum_{j_1 < \cdots < j_h} \!\!\cdots \sum c_{j_1}^{r_1} \cdots c_{j_h}^{r_h}$$

for any positive integers $h \le k$ and $r_1, \ldots, r_h$. Here the product on the left vanishes unless $(i_1, \ldots, i_h) = (\tau_{j_1}, \ldots, \tau_{j_h})$ for some $j_1 < \ldots < j_h$. Applying the corresponding substitution and noting that $\alpha_{\tau_j} = j$ for all j, we see that the two sums agree. □

Second (partial) proof of Theorem 5.5. We consider only the case when X has bounded variation. As in the case of (10), we need to show that

$$E\Big(\int f\, d\widehat{X}_A\Big)^n = E\Big(\int f\, dX\Big)^n, \quad n \in \mathbb{N},$$

for any function $f : [0, 1] \to \mathbb{R}$ of the form

$$f(s) = \sum\nolimits_{j \le m} c_j 1\{s \le t_j\}, \quad s \in [0, 1],$$

where $c_1, \ldots, c_m \in \mathbb{R}$ and $0 \le t_1 < \cdots < t_m \le \lambda A$ a.s. By the definition of $\widehat{X}_A$ we have $\int f \, d\widehat{X}_A = \int V dX$ where V denotes the predictable process

$$V_s = 1_A(s) \sum\nolimits_{j \le m} c_j 1\{s \le \tau_{t_j}\}, \quad s \in [0, 1],$$

defined in terms of the right-continuous inverse τ of λ_A. By Corollary 6.5 it is then enough, for any k and $r_1, \ldots, r_k$ in $\mathbb{N}$, to show that

$$\int \cdots \int_{\Delta_k} V_{s_1}^{r_1} \cdots V_{s_k}^{r_k} = \int \cdots \int_{\Delta_k} f_{s_1}^{r_1} \cdots f_{s_k}^{r_k}.$$

But this follows easily from the substitution rule $\int (f \circ g) \, d\mu = \int f \, d(\mu \circ g^{-1})$ for Lebesgue–Stieltjes integrals. □

References

Aldous, D.J.: Exchangeability and related topics.In: École d'Été de Probabilités de Saint-Flour XIII—1983. Lect. Notes in Math. Springer, Berlin. **1117**, 1–198 (1985)

Alfsen, E.M.: Compact Convex Sets and Boundary Integrals. Springer, Berlin (1971)

Dellacherie, C.: Capacités et Processus Stochastiques. Springer, Berlin (1972)

Doob, J.L.: Note on probability. Ann. Math. (2) **37**, 363–367 (1936)

Doob, J.L.: Stochastic Processes. Wiley, New York (1953)

Dynkin, E.B.: Sufficient statistics and extreme points. Ann. Probab. **6**, 705–730 (1978)

Halmos, P.R.: I want to be a Mathematician. Springer, New York (1985)

Jacod, J., Shiryaev, A.N.: Limit Theorems for Stochastic Processes. Springer, Berlin (1987)

Kallenberg, O.: Canonical representations and convergence criteria for processes with interchangeable increments. Z. Wahrscheinlichkeitstheorie verw. Gebiete **27**, 23–36 (1973)

Kallenberg, O.: On symmetrically distributed random measures. Trans. Amer. Math. Soc. **202**, 105–121 (1975)

Kallenberg, O.: Characterizations and embedding properties in exchangeability. Z. Wahrscheinlichkeitstheorie verw. Gebiete **60**, 249–281 (1982)

Kallenberg, O.: Spreading and predictable sampling in exchangeable sequences and processes. Ann. Probab. **16**, 508–534 (1988a)

Kallenberg, O.: One-dimensional uniqueness and convergence criteria for exchangeable processes. Stoch. Proc. Appl. **28**, 159–183 (1988b)

Kallenberg, O.: General Wald-type identities for exchangeable sequences and processes. Probab. Th. Rel. Fields **83**, 447–487 (1989)

Kallenberg, O.: Symmetries on random arrays and set-indexed processes. J. Theor. Probab. **5**, 727–765 (1992)

Kallenberg, O.: Foundations of Modern Probability. Springer, New York (1997)

Kingman, J.F.C.: Uses of exchangeability. Ann. Probab. **6**, 183–197 (1978)

Rogers, L.C.G., Williams, D.: Diffusions, Markov Processes, and Martingales, Vol. I., Wiley, Chichester (1994)

Ryll-Nardzewski, C.: On stationary sequences of random variables and the de Finetti's equivalence. Colloq. Math. **4**, 149–156 (1957)

Wald, A.: Sequential tests of statistical hypotheses. Ann. Math. Statist. **16**, 117–186 (1945)