The Annals of Probability
2000, Vol. 28, No. 2, 685–712

WEAK CONVERGENCE OF SOME CLASSES OF MARTINGALES WITH JUMPS

BY YOICHI NISHIYAMA

Institute of Statistical Mathematics

This paper deals with weak convergence of stochastic integrals with respect to multivariate point processes. The results are given in terms of an entropy condition for partitioning of the index set of the integrands, which is a sort of L^2-bracketing. We also consider ℓ^∞-valued martingale difference arrays, and present natural generalizations of Jain–Marcus's and Ossiander's central limit theorems. As an application, the asymptotic behavior of log-likelihood ratio random fields in general statistical experiments with abstract parameters is derived.

1. Introduction. Let $(E, \mathscr{E})$ be a Blackwell space. For every $n \in \mathbb{N}$, let μ^n be an E-valued multivariate point process defined on a stochastic basis $\mathbf{B}^n = (\Omega^n, \mathscr{F}^n, \mathbf{F}^n = (\mathscr{F}_t^n)_{t\in\mathbb{R}_+}, P^n)$, and let ν^n be the predictable compensator of μ^n. Let $\mathscr{W}^n = \{W^{n,\psi}: \psi \in \Psi\}$ be a class of predictable functions on $\Omega^n \times \mathbb{R}_+ \times E$, indexed by an arbitrary set Ψ. [Throughout this paper, we follow the standard definitions and notations of martingale theory, which can be found in the book by Jacod and Shiryaev (1987).] The main goal of this paper is to present some sufficient conditions for the weak convergence of the sequence of processes $(t, \psi) \rightsquigarrow X_t^{n,\psi}$, given by

$$X_t^{n,\psi} = W^{n,\psi} * (\mu^n - \nu^n)_t \qquad \forall\, t \in \mathbb{R}_+ \ \forall\, \psi \in \Psi,$$

as $n \to \infty$.

Our result has its roots in the modern theory of empirical processes for i.i.d. random sequences indexed by classes of sets or functions, which was initiated by the prominent work by Dudley (1978), and has been well developed by many authors in the 80s including Ossiander (1987) and Andersen, Giné, Ossiander and Zinn (1988). The recent book by van der Vaart and Wellner (1996) gives a comprehensive exposition of such results up to row-independent cases as well as a lot of applications to statistics. On the other hand, it is also important from a practical point of view to remove the assumption of independence. This problem has been considered in the last few years by several authors: see Arcones and Yu (1994), Doukhan, Massart and Rio (1995), Bae and Levental (1995) and Nishiyama (1997).

In order to explain the crucial point of our work, let us quote here a general criterion for the weak convergence of $\ell^\infty(T)$-valued random elements, which we shall use in Sections 3 and 4. See Theorem 1.5.4 and 1.5.6 of van der Vaart

Received November 1997; revised February 1999.
AMS 1991 *subject classifications*. 60F05, 60F17, 62F20.
Key words and phrases. Weak convergence, central limit theorem, point process, martingale, likelihood, Markov chain.

and Wellner (1996) for the proof. In what follows, the notation $\Rightarrow_{P^n}$ means the weak convergence under the sequence of probability measures P^n [see, e.g., Definition 1.3.3 of van der Vaart and Wellner (1996)]; we denote by P^* and E^* the outer probability and expectation with respect to the probability measure P, respectively.

THEOREM 1.1. *Let T be an arbitrary set. For every $n \in \mathbb{N}$, let $(\Omega^n, \mathscr{F}^n, P^n)$ be a probability space and X^n a mapping from Ω^n to $\ell^\infty(T)$. Consider the following statements*:

(i) *X^n converges weakly in $\ell^\infty(T)$ to a tight, Borel law.*

(ii) *Every finite-dimensional marginal of X^n converges weakly to a (tight,) Borel law.*

(iii) *For every $\varepsilon, \eta > 0$ there exists a finite partition $\{T_k : 1 \le k \le N\}$ of T such that*

$$\limsup_{n\to\infty} P^{n*}\left(\max_{1\le k\le N} \sup_{t,\, s\in T_k} |X^n(t) - X^n(s)| > \varepsilon\right) \le \eta.$$

Then, there is the equivalence (i) $\Leftrightarrow$ (ii) + (iii). *Furthermore, if the marginals of a process $X = (X(t)|t \in T)$ have the same laws as those of the limits in* (ii), *then there exists a version $\tilde{X}$ of X such that $X^n \Rightarrow_{P^n} \tilde{X}$.*

In our situation, the finite-dimensional convergence can be shown by using some well-known results for finite-dimensional local martingales [see, e.g., Jacod and Shiryaev (1987)]. Hence the crucial point is to check the condition (iii), and this is accomplished by means of a maximal inequality given in Section 2, which is described in terms of a certain entropy of series of finite partitions of the set Ψ. The inequality is proved by combining a *bracketing and chaining argument* which has been developed mainly for empirical processes and an *exponential inequality* for local martingales with bounded jumps. The former is originally due to Ossiander (1987) who established the central limit theorem for i.i.d. sequences under the metric entropy condition for L^2-bracketing, and is refined by van der Vaart and Wellner (1996) who showed no metric is necessary to formulate a certain entropy condition for L^2-bracketing. On the other hand, Bae and Levental (1995) have already shown that Bernstein–Freedman's inequality [Freedman (1975)] works well instead of the classical inequality of Bernstein for i.i.d. sequences, in their study on a central limit theorem for ergodic Markov chains. Van de Geer (1995) and Nishiyama (1997a, b) have taken such approaches in some situations of continuous-time martingales. It should be noted that the idea of the *partitioning entropy condition* introduced in Section 3 comes from those of Theorem 2.11.9 of van der Vaart and Wellner (1996) and Proposition 1.1 of Bae and Levental (1995). Although the results in Section 3 do not contain Theorem 2.2 of Nishiyama (1997), the conditions have been considerably refined. The refinement is partly due to the use of the tightness criterion in terms of partitioning [i.e., (iii) of Theorem 1.1

above] rather than the well-known stochastic ρ-equicontinuity criterion. Van der Vaart and Wellner (1996) are apparently the first to present the partitioning criterion.

Section 4 is devoted to the case of discrete-time martingales. Let Ψ be an arbitrary set. Let $\mathbf{B}^n = (\Omega^n, \mathscr{F}^n, \mathbf{F}^n, P^n)$ be a discrete-time stochastic basis, where $(\Omega^n, \mathscr{F}^n, P^n)$ is a probability space and $\mathbf{F}^n = \{\mathscr{F}_i^n\}_{i\in\mathbb{N}_0}$ is a nondecreasing sequence of sub-σ-fields of $\mathscr{F}^n$ indexed by $\mathbb{N}_0 = \{0\} \cup \mathbb{N}$.

DEFINITION 1.2. $\{\xi_i^n\}_{i\in\mathbb{N}} = \{(\xi_i^{n,\psi} | \psi \in \Psi)\}_{i\in\mathbb{N}}$ is called an $\ell^\infty(\Psi)$-valued martingale difference array on $\mathbf{B}^n$ if:

(i) ξ_i^n is a mapping from Ω^n to $\ell^\infty(\Psi)$ for every $i \in \mathbb{N}$;

(ii) $\{\xi_i^{n,\psi}\}_{i\in\mathbb{N}}$ is an $\mathbb{R}$-valued martingale difference array on $\mathbf{B}^n$ for every $\psi \in \Psi$.

It is required in (ii) that $\xi_i^{n,\psi}$ is $\mathscr{F}_i^n$-measurable and $E_{i-1}^n \xi_i^{n,\psi} = 0$ almost surely, for every $\psi \in \Psi$, where E_{i-1}^n denotes the $\mathscr{F}_{i-1}^n$-conditional expectation; the exceptional sets may depend on ψ. Notice also that we do not require any measurability of the $\ell^\infty(\Psi)$-valued random element ξ_i^n. In Section 4, starting from a maximal inequality again, we give some sufficient conditions to ensure the weak convergence of sequences of $\ell^\infty(\Psi)$-valued random elements

$$\sum_{i=1}^{\sigma^n} \xi_i^n = \left(\sum_{i=1}^{\sigma^n} \xi_i^{n,\psi} \,\middle|\, \psi \in \Psi \right),$$

where σ^n is a finite stopping time on $\mathbf{B}^n$. Our result generalizes that of Ossiander (1987). A natural generalization of Jain and Marcus's (1975) central limit theorem is also given. A major part of this section, as well as Section 5, was originally presented in Nishiyama (1996).

The asymptotic behavior of log-likelihood ratio random fields have been studied by many authors. The work by Vostrikova (1987), who considered a continuous-time semimartingale model, seems the most general result for cases of Euclidean parameters. In Section 5, we get a result for a discrete-time statistical experiment with general parameters in terms of the partitioning entropy of the parameter space. An application to ergodic Markov chains is also presented.

As mentioned above, van de Geer (1995) has successfully taken a bracketing entropy approach, which has the same nature as that in the present paper, to nonparametric maximum likelihood estimation for counting processes. She derived a probability inequality based on her generalization of Bernstein's inequality for martingales under a higher order moment condition on the size of jumps. Although our maximal inequalities presented in Sections 2 and 4 require that the jumps be uniformly bounded, this assumption can be replaced by a higher order moment based on Bernstein–van de Geer's inequality. See Nishiyama (1998) for the details.

Throughout this paper, the notation "$\lesssim$" means that the left-hand side is not bigger than the right up to a universal multiplicative constant. The notation "$\to_{P^n}$" means the convergence in P^n-probability.

2. Maximal inequality. Let us begin by preparing two definitions.

DEFINITION 2.1. Let $(\mathscr{X}, \mathscr{A}, \lambda)$ be a σ-finite measure space. For a given mapping $Z : \mathscr{X} \to \mathbb{R} \cup \{\infty\}$, we denote by $[Z]_{\mathscr{A},\lambda}$ any $\mathscr{A}$-measurable function $U : \mathscr{X} \to \mathbb{R} \cup \{\infty\}$ such that:

(i) $U \geq Z$ holds identically.

(ii) $\tilde{U} \geq U$ holds λ-almost everywhere, for every $\mathscr{A}$-measurable function $\tilde{U}$ such that $\tilde{U} \geq Z$ holds λ-almost everywhere.

The existence of such a random variable $[Z]_{\mathscr{A},\lambda}$ and its uniqueness up to a λ-negligible set follow from Lemma 1.2.1 of van der Vaart and Wellner (1996). They showed those facts when $(\mathscr{X}, \mathscr{A}, \lambda)$ is a probability space, but it is clear from their proof that λ may be replaced by a σ-finite measure.

DEFINITION 2.2. Let Ψ be an arbitrary set. $\Pi = \{\Pi(\varepsilon)\}_{\varepsilon \in (0, \Delta_\Pi]}$, where $\Delta_\Pi \in (0, \infty) \cap \mathbb{Q}$, is called a decreasing series of finite partitions (abb. DFP) [resp., nested series of finite partitions (abb. NFP)] of Ψ if it satisfies the following (i), (ii) and (iii) [resp., (i), (ii) and (iii$'$)]:

(i) Each $\Pi(\varepsilon) = \{\Psi(\varepsilon; k) : 1 \leq k \leq N_\Pi(\varepsilon)\}$ is a finite partition of Ψ; that is, $\Psi = \bigcup_{k=1}^{N_\Pi(\varepsilon)} \Psi(\varepsilon; k)$.

(ii) $N_\Pi(\Delta_\Pi) = 1$ and $\lim_{\varepsilon \downarrow 0} N_\Pi(\varepsilon) = \infty$.

(iii) $N_\Pi(\varepsilon) \geq N_\Pi(\varepsilon')$ whenever $\varepsilon \leq \varepsilon'$.

(iii$'$) $\Pi(\varepsilon) \supset \Pi(\varepsilon')$ whenever $\varepsilon \leq \varepsilon'$.

The ε-entropy $H_\Pi(\varepsilon)$ and the modified ε-entropy $\tilde{H}_\Pi(\varepsilon)$ of a DFP Π are defined by:

$$H_\Pi(\varepsilon) = \sqrt{\log N_\Pi(\varepsilon)};$$

$$\tilde{H}_\Pi(\varepsilon) = \sqrt{\log(1 + N_\Pi(\varepsilon))}.$$

Notice that any NFP is a DFP. Although the converse is not true, we can sometimes construct a new NFP from a given DFP, due to Lemma 2.4, given later, without loss of generality for our purpose.

Let us now turn to the context of multivariate point processes. Let $(E, \mathscr{E})$ be a Blackwell space. Let μ be an E-valued multivariate point process defined on a stochastic basis $\mathbf{B} = (\Omega, \mathscr{F}, \mathbf{F} = (\mathscr{F}_t)_{t \in \mathbb{R}_+}, P)$, and ν a "good" version of the predictable compensator of μ. We put $\tilde{\Omega} = \Omega \times \mathbb{R}_+ \times E$ and $\tilde{\mathscr{P}} = \mathscr{P} \otimes \mathscr{E}$, where $\mathscr{P}$ is the predictable σ-field on $\Omega \times \mathbb{R}_+$. We introduce the Doléans measure M_ν^P on $(\tilde{\Omega}, \tilde{\mathscr{P}})$, which is $\tilde{\mathscr{P}}$-σ-finite, given by

$$M_\nu^P(d\omega, dt, dx) = P(d\omega)\nu(\omega; dt, dx)$$

[see Theorem II.1.8 and III.3.15 of Jacod and Shiryaev (1987)].

Let $\mathscr{W} = \{W^\psi\colon \psi \in \Psi\}$ be a family of predictable functions on $\tilde{\Omega}$ indexed by an arbitrary set Ψ. We give a definition, using the notation of Definition 2.1, which plays the key role in our context.

DEFINITION 2.3. The predictable envelope $\overline{W}$ of $\mathscr{W} = \{W^\psi\colon \psi \in \Psi\}$ is defined by

$$\overline{W} = \left[\sup_{\psi\in\Psi} |W^\psi|\right]_{\tilde{\mathscr{P}},\, M_\nu^P}.$$

For a given DFP Π of Ψ, the quadratic Π-modulus $\|\mathscr{W}\|_\Pi$ of $\mathscr{W} = \{W^\psi\colon \psi \in \Psi\}$ is defined as the $\mathbb{R}_+ \cup \{\infty\}$-valued predictable process $t \rightsquigarrow \|\mathscr{W}\|_{\Pi,\,t}$ given by

$$\|\mathscr{W}\|_{\Pi,\,t} = \sup_{\varepsilon\in(0,\,\Delta_\Pi]\cap\mathbb{Q}} \max_{1\le k\le N_\Pi(\varepsilon)} \frac{\sqrt{|W(\Psi(\varepsilon; k))|^2 * \nu_t}}{\varepsilon} \qquad \forall\, t \in \mathbb{R}_+,$$

where

$$W(\Psi') = \left[\sup_{\psi,\,\phi\in\Psi'} |W^\psi - W^\phi|\right]_{\tilde{\mathscr{P}},\, M_\nu^P} \qquad \forall\, \Psi' \subset \Psi. \tag{2.1}$$

Here, and in the sequel, the notation of the stochastic integral "$W * \mu$" always means the pathwise Lebesgue–Stieltjes integral [see II.1.5 of Jacod and Shiryaev (1987)]. The stochastic integral in the L^2-sense does not appear in this paper. Notice that both $\overline{W}$ and $\|\mathscr{W}\|_\Pi$ depend on $\mathbf{F}$, P and ν, through $\tilde{\mathscr{P}}$ and M_ν^P.

LEMMA 2.4. *For any DFP* Π *such that* $\int_0^{\Delta_\Pi} H_\Pi(\varepsilon)\,d\varepsilon < \infty$, *there exists a NFP* Π' *such that*

$$\Delta_{\Pi'} = \Delta_\Pi;$$
$$\int_0^{\Delta_{\Pi'}} H_{\Pi'}(\varepsilon)\,d\varepsilon \le 4\int_0^{\Delta_\Pi} H_\Pi(\varepsilon)\,d\varepsilon;$$
$$\int_0^{\Delta_{\Pi'}} \tilde{H}_{\Pi'}(\varepsilon)\,d\varepsilon \le 4\int_0^{\Delta_\Pi} \tilde{H}_\Pi(\varepsilon)\,d\varepsilon;$$
$$\|\mathscr{W}\|_{\Pi',\,t} \le \|\mathscr{W}\|_{\Pi,\,t} \qquad \forall\, t \in \mathbb{R}_+.$$

PROOF. For every $\varepsilon \in (0, \Delta_\Pi]$, let us define

$$\Pi'(\varepsilon) = \bigvee_{i_0\le j\le i} \Pi(2^{-j}) \quad \text{if} \quad \varepsilon \in [2^{-i}, 2^{-i+1}) \cap (0, \Delta_\Pi], \qquad i \ge i_0,$$

where $i_0 = \min\{i \in \mathbb{Z}\colon 2^{-i} \le \Delta_\Pi\}$. Then, the constructed $\Pi' = \{\Pi'(\varepsilon)\}_{\varepsilon\in(0,\,\Delta_{\Pi'}]}$ is a NFP such that $\Delta_{\Pi'} = \Delta_\Pi$. The two inequalities for the integrals can be shown by a standard way [see, e.g., Lemma 2.4 of Andersen, Giné, Ossiander and Zinn (1988), Lemma 3.6 of Nishiyama (1997) or Lemma 2.2.2 of Nishiyama (1998)]. The last assertion is trivial from the construction of Π'. □

Suppose that the increasing process $t \rightsquigarrow \overline{W} * \nu_t$ is locally integrable, and define the random variables X_t^{ψ} and $X_t^{a,\psi}$ by

$$X_t^{\psi} = W^{\psi} * (\mu - \nu)_t \qquad \forall\, t \in \mathbb{R}_+ \;\forall\, \psi \in \Psi \tag{2.2}$$

and

$$X_t^{a,\psi} = W^{\psi} 1_{\{\overline{W} \le a\}} * (\mu - \nu)_t \qquad \forall\, t \in \mathbb{R}_+ \;\forall\, \psi \in \Psi \;\forall\, a > 0, \tag{2.3}$$

respectively. Then, the process $t \rightsquigarrow X_t^{\psi}$ is a local martingale and the process $t \rightsquigarrow X_t^{a,\psi}$ is a locally square-integrable martingale on $\mathbf{B}$, both of which have finite variation [see Proposition II.1.28 of Jacod and Shiryaev (1987)]. The following theorem gives some maximal inequalities for these processes in terms of $\|\mathscr{W}\|_{\Pi}$.

THEOREM 2.5. *Let μ be an E-valued multivariate point process defined on a stochastic basis $\mathbf{B}$, and ν a "good" version of the predictable compensator of μ. Let $\mathscr{W} = \{W^{\psi}: \psi \in \Psi\}$ be a family of predictable functions on $\tilde{\Omega}$, indexed by an arbitrary set Ψ, such that the increasing process $t \rightsquigarrow \overline{W} * \nu_t$ is locally integrable. Then, the following* (i) *and* (ii) *hold for any stopping time τ such that $\nu([0,\tau] \times E) < \infty$ almost surely*:

(i) *For any NFP Π of Ψ and any constants $\delta \in (0, \Delta_{\Pi}]$ and $K > 0$,*

$$E^* \sup_{t \in [0,\tau]} \sup_{\substack{1 \le k \le N_{\Pi}(\delta) \\ \psi,\phi \in \Psi(\delta;k)}} |X_t^{a,\psi} - X_t^{a,\phi}| 1_{\{\|\mathscr{W}\|_{\Pi,\tau} \le K\}} \lesssim K \int_0^{\delta} \tilde{H}_{\Pi}(\varepsilon)\, d\varepsilon,$$

where the random variables $X_t^{a,\psi}$ are defined by (2.3) *with $a = a(\delta, K) = \delta K / \tilde{H}_{\Pi}(\delta/2)$.*

(ii) *For any DFP Π of Ψ and any constants $K, L > 0$,*

$$E^* \sup_{t \in [0,\tau]} \sup_{\psi,\phi \in \Psi} |X_t^{\psi} - X_t^{\phi}| 1_{\{\|\mathscr{W}\|_{\Pi,\tau} \le K,\ |\overline{W}|^2 * \nu_\tau \le L\}} \lesssim K \int_0^{\Delta_{\Pi}} \tilde{H}_{\Pi}(\varepsilon)\, d\varepsilon + \frac{L}{\Delta_{\Pi} K},$$

where the random variables X_t^{ψ} are defined by (2.2).

REMARK. As stated in Section 1, the notation "$\lesssim$" means that the left-hand side is not bigger than the right up to a multiplicative universal constant.

PROOF OF THEOREM 2.5(I). Fix any $\delta, K > 0$; we may assume $\delta \in \mathbb{Q}$ without loss of generality. For every integer $p \ge 0$, we set

$$a_p = 2^{-p+1} \delta K / \tilde{H}_{\Pi}(2^{-p-1}\delta).$$

Next, choosing an element $\psi_{p,k}$ from each partitioning set $\Psi(2^{-p}\delta; k)$ such that

$$\{\psi_{p,k}: 1 \le k \le N_{\Pi}(2^{-p}\delta)\} \subset \{\psi_{p+1,k}: 1 \le k \le N_{\Pi}(2^{-(p+1)}\delta)\},$$

we define for every $\psi \in \Psi$

$$\begin{cases} \pi_p \psi = \psi_{p,k}, \\ \Pi_p \psi = \Psi(2^{-p}\delta; k), \end{cases} \qquad \text{if} \quad \psi \in \Psi(2^{-p}\delta; k).$$

For every integer $q \geq 1$, we introduce the stopping time

$$\tau_q = \inf\left\{ t \in \mathbb{R}_+ : \nu([0,t]\times E) > \frac{|\tilde{H}_\Pi(2^{-q-2}\delta)|^2}{16} - 1 \right\} \wedge \tau.$$

Since $\nu([0,\tau]\times E) < \infty$ almost surely and $\lim_{\varepsilon\downarrow 0} N_\Pi(\varepsilon) = \infty$, it holds that $\tau_q \uparrow \tau$ as $q \to \infty$ almost surely. Hence it is enough to show that

$$E^* \sup_{t\in[0,\tau_q]} \sup_{\psi\in\Psi} |X_t^{a,\psi} - X_t^{a,\pi_0\psi}| 1_{\{\|\mathscr{W}\|_{\Pi,\tau} \leq K\}} \lesssim K \int_0^\delta \tilde{H}_\Pi(\varepsilon)\,d\varepsilon \qquad \forall\, q \geq 1, \tag{2.4}$$

where $a = a(\delta, K)$.

Let us now fix any integer $q \geq 1$, and denote $\tau = \tau_q$ (there should be no risk of confusion). For every $p = 0, 1, \ldots, q$, we consider the predictable functions $W(\Pi_p \psi)$ on $\tilde{\Omega}$ defined by (2.1). Since $\Pi = \{\Pi(\varepsilon)\}_{\varepsilon\in(0,\Delta_\Pi]}$ is nested, it follows from Definition 2.1 that

$$2\overline{W} \geq W(\Pi_0\psi) \geq W(\Pi_1\psi) \geq \cdots \geq W(\Pi_q\psi), \tag{2.5}$$

M_ν^P-almost everywhere. Defining the values on the exceptional sets as zero, we can choose some versions such that the above inequality holds *identically*. Notice also that $W(\Pi_p\psi) = W(\Pi_p\phi)$ holds identically, whenever $\psi, \phi \in \Psi(2^{-q}\delta; k)$ for some k. Next, let us introduce the following predictable functions on $\tilde{\Omega}$:

$$A_p(\psi) = 1_{\{W(\Pi_0\psi)\leq a_0, \ldots, W(\Pi_{p-1}\psi)\leq a_{p-1},\, W(\Pi_p\psi)\leq a_p\}}, \qquad p = 0, 1, \ldots, q;$$
$$B_p(\psi) = 1_{\{W(\Pi_0\psi)\leq a_0, \ldots, W(\Pi_{p-1}\psi)\leq a_{p-1},\, W(\Pi_p\psi)> a_p\}}, \qquad p = 1, \ldots, q;$$
$$B_0(\psi) = 1_{\{W(\Pi_0\psi) > a_0\}}.$$

It is important that $A_p(\psi)$ and $B_p(\psi)$ depend on ψ only through the subsets $\Pi_0\psi, \ldots, \Pi_p\psi$ of Ψ. Next observe the identity

$$\begin{aligned} W^\psi - W^{\pi_0\psi} &= (W^\psi - W^{\pi_0\psi})B_0(\psi) \\ &\quad + \sum_{p=1}^{q} (W^\psi - W^{\pi_p\psi})B_p(\psi) \\ &\quad + (W^\psi - W^{\pi_q\psi})A_q(\psi) \\ &\quad + \sum_{p=1}^{q} (W^{\pi_p\psi} - W^{\pi_{p-1}\psi})A_{p-1}(\psi). \end{aligned}$$

Since $a_0 = 2a(\delta,\ K)$, we have $B_0(\psi) \leq 1_{\{\overline{W} > a(\delta, K)\}}$. Hence we obtain

$$\sup_{t\in[0,\tau]} \sup_{\psi\in\Psi} |X_t^{a(\delta,K),\psi} - X_t^{a(\delta,K),\pi_0\psi}| \leq (I_1) + (I_2) + (II_1) + (II_2) + (III),$$

where

$$(I_1) = \sup_{\psi \in \Psi} \sum_{p=1}^{q} W(\Pi_p \psi) B_p(\psi) * \mu_\tau,$$

$$(I_2) = \sup_{\psi \in \Psi} \sum_{p=1}^{q} W(\Pi_p \psi) B_p(\psi) * \nu_\tau,$$

$$(II_1) = \sup_{\psi \in \Psi} W(\Pi_q \psi) A_q(\psi) * \mu_\tau,$$

$$(II_2) = \sup_{\psi \in \Psi} W(\Pi_q \psi) A_q(\psi) * \nu_\tau,$$

$$(III) = \sup_{t \in [0,\tau]} \sup_{\psi \in \Psi} \sum_{p=1}^{q} \left| (W^{\pi_p \psi} - W^{\pi_{p-1}\psi}) A_{p-1}(\psi) * (\mu - \nu)_t \right|.$$

Further, it holds that $(I_1) \le (I_1') + (I_2)$ where

$$(I_1') = \sup_{\psi \in \Psi} \sum_{p=1}^{q} \left| W(\Pi_p \psi) B_p(\psi) * (\mu - \nu)_\tau \right|,$$

and that $(II_1) \le (II_1') + (II_2)$ where

$$(II_1') = \sup_{\psi \in \Psi} \left| W(\Pi_q \psi) A_q(\psi) * (\mu - \nu)_\tau \right|.$$

Hereafter we will obtain bounds for the terms (I_1'), (I_2), (II_1'), (II_2) and (III).

[Estimation of (I_2) and (II_2).] We can easily see that

$$\begin{aligned}
(I_2) &\le \sup_{\psi \in \Psi} \sum_{p=1}^{q} \frac{1}{a_p} |W(\Pi_p \psi)|^2 B_p(\psi) * \nu_\tau \\
&\le \max_{1 \le p \le q} \sup_{\psi \in \Psi} \frac{|W(\Pi_p \psi)|^2 B_p(\psi) * \nu_\tau}{|2^{-p}\delta|^2} \sum_{p=1}^{q} \frac{|2^{-p}\delta|^2}{a_p} \\
&\le K \sum_{p=1}^{q} 2^{-p-1} \delta \tilde{H}_\Pi(2^{-p-1}\delta) \quad \text{on the set } \{\|\mathscr{W}\|_{\Pi,\tau} \le K\}.
\end{aligned}$$

On the other hand, it follows from Schwarz's inequality that

$$\begin{aligned}
(II_2) &\le \sup_{\psi \in \Psi} \sqrt{|W(\Pi_q \psi)|^2 * \nu_\tau} \sqrt{\nu([0,\tau] \times E)} \\
&\le 2^{-q} \delta K \frac{\tilde{H}_\Pi(2^{-q-2}\delta)}{4}.
\end{aligned}$$

Hence we have

$$\begin{aligned}
E|(I_2) + (II_2)| 1_{\{\|\mathscr{W}\|_{\Pi,\tau} \le K\}} &\le K \sum_{p=1}^{q+1} 2^{-p-1} \delta \tilde{H}_\Pi(2^{-p-1}\delta) \\
&\le 2K \int_0^\delta \tilde{H}_\Pi(\varepsilon)\, d\varepsilon.
\end{aligned}$$

[Estimation of (I_1'), (II_1') and (III).] Let us consider the term (I_1'). We will apply Bernstein–Freedman's inequality for local martingales with bounded jumps [see, e.g., Section 4.13 of Liptser and Shiryaev (1989) or Corollary 3.3 (a) of Nishiyama (1997)] to the processes

$$t \rightsquigarrow M_t = W(\Pi_p \psi) B_p(\psi) * (\mu - \nu)_t.$$

It follows from

$$0 \le W(\Pi_p \psi) B_p(\psi) \le W(\Pi_{p-1}\psi) B_p(\psi) \le a_{p-1}$$

that $|\Delta M| \le a_{p-1}$; it is also clear that

$$\begin{aligned}\langle M, M\rangle_\tau &\le |W(\Pi_p \psi)|^2 B_p(\psi) * \nu_\tau \\ &\le |2^{-p}\delta K|^2 \quad \text{on the set} \quad \{\|\mathscr{W}\|_{\Pi,\tau} \le K\}.\end{aligned}$$

Thus we have

$$\begin{aligned}&P\left(\sup_{t\in[0,\tau]} \left|W(\Pi_p \psi) B_p(\psi) * (\mu-\nu)_t\right| > \varepsilon,\ \|\mathscr{W}\|_{\Pi,\tau} \le K\right) \\ &\qquad \le 2\exp\left(-\frac{\varepsilon^2}{2[a_{p-1}\varepsilon + |2^{-p}\delta K|^2]}\right) \qquad \forall\, \varepsilon > 0.\end{aligned}$$

Hence it follows from Lemma 2.2.10 of van der Vaart and Wellner (1996) with an appropriate truncation that

$$\begin{aligned}&E \sup_{\psi\in\Psi} \sup_{t\in[0,\tau]} \left|W(\Pi_p \psi) B_p(\psi) * (\mu-\nu)_t\right| 1_{\{\|\mathscr{W}\|_{\Pi,\tau}\le K\}} \\ &\qquad \lesssim a_{p-1} |\tilde{H}_\Pi(2^{-p}\delta)|^2 + 2^{-p}\delta K \tilde{H}_\Pi(2^{-p}\delta) \\ &\qquad \le 5K 2^{-p}\delta \tilde{H}_\Pi(2^{-p}\delta),\end{aligned}$$

where it should be noted that "$\sup_{\psi\in\Psi}$" of the left-hand side is actually "$\max_{1\le k\le N_\Pi(2^{-p})}$." We therefore obtain

$$\begin{aligned}E|(I_1')| 1_{\{\|\mathscr{W}\|_{\Pi,\tau}\le K\}} &\lesssim 5K \sum_{p=1}^{q} 2^{-p}\delta \tilde{H}_\Pi(2^{-p}\delta) \\ &\le 5K \int_0^\delta \tilde{H}_\Pi(\varepsilon)\, d\varepsilon.\end{aligned}$$

Exactly the same calculation as for (I_1') yields some bounds for (II_1') and (III), which lead to inequality (2.4). □

PROOF OF THEOREM 2.5 (II). Due to Lemma 2.4, it suffices to show the assertion in the case of Π being a NFP. We extend given NFP $\Pi = \{\Pi(\varepsilon)\}_{\varepsilon\in(0,\Delta_\Pi]}$ to $\Pi = \{\Pi(\varepsilon)\}_{\varepsilon\in(0,2\Delta_\Pi]}$ where $N_\Pi(\varepsilon) = 1$ for all $\varepsilon \in [\Delta_\Pi, 2\Delta_\Pi]$. In order to apply the assertion (i) with $\delta = 2\Delta_\Pi$, we consider the truncated processes $X_t^{a,\psi}$ with $a = a(2\Delta_\Pi, K) = 2\Delta_\Pi K/\sqrt{\log 2}$; notice that

$$\begin{aligned}\sup_{t\in[0,\tau]}\sup_{\psi,\phi\in\Psi}|X_t^\psi - X_t^\phi| \le{} & \sup_{t\in[0,\tau]}\sup_{\psi,\phi\in\Psi}|X_t^{a,\psi} - X_t^{a,\phi}| \\ & +2\overline{W}1_{\{\overline{W}>a\}} * \mu_\tau + 2\overline{W}1_{\{\overline{W}>a\}} * \nu_\tau.\end{aligned}$$

First we have

$$\begin{aligned}\overline{W}1_{\{\overline{W}>a\}} * \nu_\tau &\le \frac{|\overline{W}|^2 * \nu_\tau}{a} \\ &\le \frac{L}{a} \quad \text{on the set} \quad \{|\overline{W}|^2 * \nu_\tau \le L\}.\end{aligned}$$

Next, let us introduce the predictable time

$$S = \inf\{t \in \mathbb{R}_+ : |\overline{W}|^2 * \nu_t > L\}.$$

Take an announcing sequence $\{S_n\}$ for S [see I.2.16 of Jacod and Shiryaev (1987)]. Since $0 \le S_n < S$ almost surely on the set $\{S > 0\}$, it holds that $|\overline{W}|^2 * \nu_{S_n} \le L$ almost surely. Thus it follows also from Doob's stopping theorem that

$$\begin{aligned}E\overline{W}1_{\{\overline{W}>a\}} * \mu_{S_n\wedge T_m} &= E\overline{W}1_{\{\overline{W}>a\}} * \nu_{S_n\wedge T_m} \\ &\le \frac{E|\overline{W}|^2 * \nu_{S_n\wedge T_m}}{a} \le \frac{L}{a},\end{aligned}$$

where $\{T_m\}$ is a localizing sequence for the local martingale $t \rightsquigarrow \overline{W}1_{\{\overline{W}>a\}} * (\mu - \nu)_t$. By letting $n, m \to \infty$, we obtain $E\overline{W}1_{\{\overline{W}>a\}} * \mu_S \le L/a$. The predictable time S appeared in this inequality can be replaced by τ on the set $\{|\overline{W}|^2 * \nu_\tau \le L\}$.

Hence it follows from the assertion (i) with $\delta = 2\Delta_\Pi$ that

$$\begin{aligned}&E^* \sup_{t\in[0,\tau]}\sup_{\psi,\phi\in\Psi}|X_t^\psi - X_t^\phi| 1_{\{\|\mathscr{W}\|_{\Pi,\tau}\le K,\ |\overline{W}|^2*\nu_\tau \le L\}} \\ &\quad\lesssim K\int_0^{2\Delta_\Pi} \tilde{H}_\Pi(\varepsilon)\,d\varepsilon + 4\frac{L}{2\Delta_\Pi K/\sqrt{\log 2}} \\ &\quad\le 2\left\{K\int_0^{\Delta_\Pi} \tilde{H}_\Pi(\varepsilon)\,d\varepsilon + \frac{L}{\Delta_\Pi K}\right\}.\end{aligned}$$

□

3. Weak convergence theorems. Let $(E, \mathscr{E})$ be a Blackwell space and Ψ an arbitrary set. For every $n \in \mathbb{N}$, let μ^n be an E-valued multivariate point process defined on a stochastic basis $\mathbf{B}^n = (\Omega^n, \mathscr{F}^n, \mathbf{F}^n = (\mathscr{F}_t^n)_{t\in\mathbb{R}_+}, P^n)$, and ν^n a "good' version of the predictable compensator of μ^n. Let $\mathscr{W}^n = \{W^{n,\psi} : \psi \in \Psi\}$ be a family of predictable functions on $\tilde{\Omega}^n = \Omega^n \times \mathbb{R}_+ \times E$ indexed by Ψ. Let a DFP Π of Ψ be given. Notice that $(E, \mathscr{E})$, Ψ and Π do not depend on n, while all other objects are indexed by $n \in \mathbb{N}$. In the same way as Section 2, we introduce the following notations:

1. the predictable envelope $\overline{W}^n$ of $\mathscr{W}^n$;
2. the quadratic Π-modulus $\|\mathscr{W}^n\|_\Pi$ of $\mathscr{W}^n$.

Further, let a stopping time τ^n on $\mathbf{B}^n$ be given. Throughout this section, we shall assume:

$$\text{the process} \quad t \rightsquigarrow \overline{W}^n * \nu_t^n \quad \text{is locally integrable;} \tag{3.1}$$

$$\nu^n([0, \tau^n] \times E) < \infty \quad P^n\text{-almost surely.} \tag{3.2}$$

As in Section 2, we define the local martingales $t \rightsquigarrow X_t^{n,\psi}$ and the locally square-integrable martingales $t \rightsquigarrow X^{n,a,\psi}$ on $\mathbf{B}^n$ by

$$X_t^{n,\psi} = W^{n,\psi} * (\mu^n - \nu^n)_t \quad \forall\, t \in \mathbb{R}_+ \ \forall\, \psi \in \Psi$$

and

$$X_t^{n,a,\psi} = W^{n,\psi} 1_{\{\overline{W} \le a\}} * (\mu^n - \nu^n)_t \qquad \forall\, t \in \mathbb{R}_+ \ \forall\, \psi \in \Psi \ \forall\, a > 0,$$

respectively. We will derive the asymptotic behavior of the processes $\psi \rightsquigarrow X_{\tau^n}^{n,\psi}$ and $(t, \psi) \rightsquigarrow X_t^{n,\psi}$, as $n \to \infty$.

Let us now introduce several conditions. The first one is the *partitioning entropy condition*, which is a natural generalization of the metric entropy condition for L^2-bracketing in the i.i.d. case.

[PE] There exists a DFP Π of Ψ such that

$$\|\mathscr{W}^n\|_{\Pi, \tau^n} = O_{P^n}(1) \quad \text{and} \quad \int_0^{\Delta_\Pi} H_\Pi(\varepsilon)\, d\varepsilon < \infty.$$

Notice that, due to Lemma 2.4, under [PE] we can always construct a new NFP Π which satisfies the displayed conditions. Next, we shall also consider two kinds of *Lindeberg conditions*:

[L1] $\overline{W}^n 1_{\{\overline{W}^n > \varepsilon\}} * \nu_{\tau^n}^n \to_{P^n} 0$ for every $\varepsilon > 0$;

[L2] $|\overline{W}^n|^2 1_{\{\overline{W}^n > \varepsilon\}} * \nu_{\tau^n}^n \to_{P^n} 0$ for every $\varepsilon > 0$.

When we mention [L2], the assumption that

$$\text{the process} \quad t \rightsquigarrow |\overline{W}^n|^2 * \nu_t^n \quad \text{is locally integrable} \tag{3.3}$$

is also implicitly imposed in addition to (3.2), and in this case the process $t \rightsquigarrow X_t^{n,\psi}$ is a locally square-integrable martingale on $\mathbf{B}^n$. It is trivial that [L2] implies [L1].

Here, we introduce some conditions prescribing the asymptotic behavior of the quadratic covariations. Let S be a subset of $\mathbb{R}_+$, and suppose that the family $\{C_t^{(\psi,\phi)}: t \in \mathbb{R}_+, (\psi, \phi) \in \Psi^2\}$ of constants in the following satisfies that

$$t \rightsquigarrow C_t^{(\psi,\phi)} \text{ is continuous for every } (\psi, \phi) \in \Psi^2. \tag{3.4}$$

[C1] $[X^{n,\psi}, X^{n,\phi}]_t \to_{P^n} C_t^{(\psi,\phi)}$ for every $t \in S$ and $(\psi, \phi) \in \Psi^2$;

[C2] $\langle X^{n,\psi}, X^{n,\phi}\rangle_t \to_{P^n} C_t^{(\psi,\phi)}$ for every $t \in S$ and $(\psi, \phi) \in \Psi^2$;

[C1$_\mathrm{a}$] $[X^{n,a,\psi}, X^{n,a,\phi}]_t \to_{P^n} C_t^{(\psi,\phi)}$ for every $t \in S$ and $(\psi, \phi) \in \Psi^2$, for every $a > 0$;

[C2$_\mathrm{a}$] $\langle X^{n,a,\psi}, X^{n,a,\phi}\rangle_t \to_{P^n} C_t^{(\psi,\phi)}$ for every $t \in S$ and $(\psi, \phi) \in \Psi^2$, for every $a > 0$.

Similarly to the remark following [L2], the assumption (3.3) is implicitly imposed when we mention [C2]. It is well known that the quadratic covariations are given by

$$[X^{n,\psi}, X^{n,\phi}]_t = \sum_{s\le t} \Delta X_s^{n,\psi} \Delta X_s^{n,\phi}$$

and

$$\langle X^{n,\psi}, X^{n,\phi}\rangle_t = (W^{n,\psi} W^{n,\phi}) * \nu_t^n - \sum_{s\le t} \widehat{W}_s^{n,\psi} \widehat{W}_s^{n,\phi},$$

where $\widehat{W}_t^{n,\psi}(\omega) = \int_E W^{n,\psi}(\omega, t, x)\nu(\omega; \{t\} \times dx)$, respectively.

Using the constants $\{C_t^{\psi,\phi}\}$ appearing in the conditions above, we set (formally)

$$\rho((t,\psi),(s,\phi)) = \sqrt{C_t^{(\psi,\psi)} + C_s^{(\phi,\phi)} - 2C_{t\wedge s}^{(\psi,\phi)}} \tag{3.5}$$

for every $(t,\psi),(s,\phi) \in \mathbb{R}_+ \times \Psi$. Either of [C1], [C2], [C1$_\mathrm{a}$] or [C2$_\mathrm{a}$] implies that the value in the inside of the square-root is nonnegative for every $(t,\psi),(s,\phi) \in S \times \Psi$, hence the $\mathbb{R}_+$-valued function ρ is well defined by the formula (3.5) at least on $(S \times \Psi)^2$. Further, by virtue of (3.4), this is true also on $([0, \tau] \times \Psi)^2$ if S is a dense subset of the finite interval $[0, \tau]$ with τ being a constant.

LEMMA 3.1.

(i) *The condition* [L1] *implies the following*:

(i$_1$) $\overline{W}^n 1_{\{\overline{W}^n > \varepsilon\}} * \mu_{\tau^n}^n \to_{P^n} 0$ *for every* $\varepsilon > 0$;

(i$_2$) $\sup_{t\in[0,\tau^n]} \sup_{\psi\in\Psi} |X_t^{n,\psi} - X_t^{n,a,\psi}| \to_{P^{n*}} 0$ *for every* $a > 0$;

(i$_3$) $\sup_{t\in[0,\tau^n]} \Delta(\overline{W}^n * \mu_t^n) \to_{P^n} 0$ *and* $\sup_{t\in[0,\tau^n]} \Delta(\overline{W}^n * \nu_t^n) \to_{P^n} 0$;

(i$_4$) $\sup_{t\in[0,\tau^n]} \sup_{\psi\in\Psi} |\Delta X_t^{n,a,\psi}| \to_{P^{n*}} 0$ *for every* $a > 0$.

(ii) *Let* $\tau^n \equiv \tau$ *be a fixed constant, and suppose that* S *is a subset of the finite interval* $[0, \tau]$. *Then, under* [L1] *it holds that* [C1] $\Leftrightarrow$ [C1$_\mathrm{a}$] $\Leftrightarrow$ [C2$_\mathrm{a}$]. *Under* [L2], *the condition* [C2] *is also equivalent to either of them.*

PROOF. It follows from Lenglart's inequality that

$$P^n\left(\overline{W}^n 1_{\{\overline{W}^n > \varepsilon\}} * \mu_{\tau^n}^n \ge \eta\right) \le \eta + P^n\left(\overline{W}^n 1_{\{\overline{W}^n > \varepsilon\}} * \nu_{\tau^n}^n \ge \eta^2\right) \qquad \forall\, \eta > 0,$$

hence the condition [L1] implies (i_1). The assertions (i_2), (i_3) and (i_4) are immediate from (i_1).

Next we show the part (ii) of the lemma. By polarization it is enough to consider the case $\phi = \psi$. Observe that

$$\begin{aligned}
&\big|[X^{n,\psi}, X^{n,\psi}]_t - [X^{n,a,\psi}, X^{n,a,\psi}]_t\big| \\
&\quad = \left|\sum_{s\le t}(\Delta X_s^{n,\psi} + \Delta X_s^{n,a,\psi})(\Delta X_s^{n,\psi} - \Delta X_s^{n,a,\psi})\right| \\
&\quad \le 2\sum_{s\le t}\left|\Delta W^{n,\psi} 1_{\{\overline{W}>a\}} * (\mu^n - \nu^n)_s\right| \qquad \text{on the set } \Omega_1^n \\
&\quad \le 2\left|\overline{W}^n 1_{\{\overline{W}>a\}} * \mu_\tau^n + \overline{W}^n 1_{\{\overline{W}>a\}} * \nu_\tau^n\right|,
\end{aligned}$$

where $\Omega_1^n = \{\sup_{t\in[0,\tau]} |\Delta X_t^{n,\psi}| \le 1\} \cup \{\sup_{t\in[0,\tau]} |\Delta X_t^{n,a,\psi}| \le 1\}$. The assertion that [C1] $\Leftrightarrow$ [$C1_a$] under [L1] is now derived from (i_1), (i_3) and (i_4). Theorem VIII.3.6 of Jacod and Shiryaev (1987), with minor changes, says that [$C1_a$] $\Leftrightarrow$ [$C2_a$] under [L1] (see the Appendix for the details). The equivalence that [C2] $\Leftrightarrow$ [$C2_a$] under [L2] follows from the inequality

$$\begin{aligned}
&\big|\langle X^{n,\psi}, X^{n,\psi}\rangle_t - \langle X^{n,a,\psi}, X^{n,a,\psi}\rangle_t\big| \\
&\quad \le |\overline{W}^n|^2 1_{\{\overline{W}^n>a\}} * \nu_\tau^n \\
&\qquad + \sum_{t\le\tau}\int_E 2\overline{W}^n(t,x)\nu(\{t\}\times dx)\int_E \overline{W}^n(t,x)1_{\{\overline{W}^n(t,x)>a\}}\nu(\{t\}\times dx) \\
&\quad \le |\overline{W}^n|^2 1_{\{\overline{W}^n>a\}} * \nu_\tau^n + 2\overline{W}^n 1_{\{\overline{W}^n>a\}}\nu_\tau^n \quad \text{on the set } \Omega_2^n,
\end{aligned}$$

where $\Omega_2^n = \{\sup_{t\in[0,\tau]} \Delta(\overline{W}^n * \nu_t^n) \le 1\}$. $\square$

The first result of this section is concerned with the processes $\psi \rightsquigarrow X_{\tau^n}^{n,\psi}$.

THEOREM 3.2. *Consider the above situation with* (3.1) *and* (3.2). *Suppose that every finite-dimensional marginal of* $X_{\tau^n}^n = (X_{\tau^n}^{n,\psi} | \psi \in \Psi)$ *converges weakly to a* (*tight*,) *Borel law, and also that the conditions* [PE] *and* [L1] *are satisfied. Then* $X_{\tau^n}^n$ *converges weakly in* $\ell^\infty(\Psi)$ *to a tight, Borel law.*

The result above is a direct consequence of the next lemma, applying Theorem 1.1.

LEMMA 3.3. *The conditions* [PE] *and* [L1] *imply that for every* $\varepsilon, \eta > 0$ *there exists a finite partition* $\{\Psi_k : 1 \le k \le N\}$ *of* Ψ *such that*

$$\limsup_{n\to\infty} P^{n*}\left(\sup_{t\in[0,\tau^n]} \sup_{\substack{1\le k\le N \\ \psi,\phi\in\Psi_k}} |X_t^{n,\psi} - X_t^{n,\phi}| > \varepsilon\right) \le \eta.$$

PROOF. Take a NFP Π which satisfies the requirements of [PE]. Fix any $\varepsilon, \eta > 0$. First notice that for any $\delta \in (0, \Delta_\Pi]$ and $K > 0$,

$$(3.6)\quad P^{n*}\left(\sup_{t\in[0,\tau^n]}\sup_{\substack{1\le k\le N_\Pi(\delta)\\ \psi,\phi\in\Psi(\delta;k)}}|X_t^{n,a(\delta,K),\psi} - X_t^{n,a(\delta,K),\phi}| > \varepsilon\right) \le (I) + (II),$$

where the terms of the right-hand side are given by

$$(I) = P^n(\|\mathscr{W}^n\|_{\Pi,\tau^n} > K),$$

$$(II) = \frac{1}{\varepsilon}E^{n*}\sup_{t\in[0,\tau^n]}\sup_{\substack{1\le k\le N_\Pi(\delta)\\ \psi,\phi\in\Psi(\delta;k)}}|X_t^{n,a(\delta,K),\psi} - X_t^{n,a(\delta,K),\phi}|1_{\{\|\mathscr{W}^n\|_{\Pi,\tau^n}\le K\}},$$

where $a(\delta, K) = \delta K/\tilde{H}_\Pi(\delta/2)$. It follows from (i) of Theorem 2.5 that there exists a universal constant $C > 0$ such that

$$(3.7)\qquad (II) \le C\frac{K}{\varepsilon}\int_0^\delta \tilde{H}_\Pi(\epsilon)\,d\epsilon.$$

Now, the first condition of [PE] yields that there exists a constant $K = K_\eta > 0$ such that $\limsup_{n\to\infty}(I) \le \eta/2$. Next, since $\tilde{H}_\Pi(\epsilon) \le 1 + H_\Pi(\epsilon)$, the second condition of [PE] implies that we can choose a sufficiently small constant $\delta = \delta_{\varepsilon,\eta} > 0$ such that the right-hand side of (3.7) is not bigger than $\eta/2$. Consequently, (i_2) of Lemma 3.1 with $a = a(\delta_{\varepsilon,\eta}, K_\eta)$ yields the assertion. $\square$

The next result deals with the processes $(t,\psi)\rightsquigarrow X_t^{n,\psi}$.

THEOREM 3.4. *Consider the above situation with* (3.1) *and* (3.2) *where* $\tau^n \equiv \tau$ *is a fixed positive constant, and let* S *be a dense subset of the finite interval* $[0,\tau]$ *containing* τ. *Suppose that either* [PE] + [L1] + [C1] *or* [PE] + [L2] + [C2] *is satisfied. Then, it holds that* $X^n \Rightarrow_{P^n} X$ *in* $\ell^\infty([0,\tau]\times\Psi)$, *where each marginal* $(X_{t_1}^{\psi_1},\dots,X_{t_d}^{\psi_d})$ *has the normal distribution* $N(0,\Sigma)$ *with* $\Sigma = \{C_{t_i\wedge t_j}^{(\psi_i,\psi_j)}\}_{ij}$. *Furthermore, the formula* (3.5) *defines a pseudo-metric* ρ *on* $[0,\tau]\times\Psi$ *such that* $[0,\tau]\times\Psi$ *is totally bounded with respect to* ρ *and that almost all paths of* X *are uniformly* ρ-*continuous*.

The following lemma, which is rather well known, is used to show the result above.

LEMMA 3.5. *Under* [L1] + [C1], *for every* $\psi \in \Psi$ *and every* $\varepsilon, \eta > 0$ *there exists* $\delta > 0$ *such that*

$$\limsup_{n\to\infty} P^n\left(\sup_{\substack{t,s\in[0,\tau]\\ |t-s|\le\delta}}|X_t^{n,\psi} - X_s^{n,\psi}| > \varepsilon\right) \le \eta.$$

PROOF. Fix any $N \in \mathbb{N}$ for a while, and put $a = N^{-1}$. By (ii) of Lemma 3.1 we may assume [L1] + [C2$_a$]. It always holds that $C_0^{(\psi,\psi)} = 0$ and that $t \rightsquigarrow C_t^{(\psi,\psi)}$ is nondecreasing, because so does $t \rightsquigarrow \langle X^{n,a,\psi}, X^{n,a,\psi}\rangle_t$. We may assume $C_\tau^{(\psi,\psi)} > 0$ without loss of generality. Since $t \rightsquigarrow C_t^{(\psi,\psi)}$ is continuous and S is dense in $[0, \tau]$, we can choose some points $\tau_i \in S$ ($i = 1, \dots, N$) such that $C_{\tau_i}^{(\psi,\psi)} - C_{\tau_{i-1}}^{(\psi,\psi)} \le 2C_\tau^{(\psi,\psi)} N^{-1}$, where $\tau_0 = 0$. It follows from Bernstein–Freedman's inequality that for every $\varepsilon > 0$,

$$P^n\left(\sup_{t\in[\tau_{i-1},\tau_i]} |X_t^{n,a,\psi} - X_{\tau_{i-1}}^{n,a,\psi}| > \varepsilon,\ \Omega_N^n\right) \le 2\exp\left(-\frac{\varepsilon^2}{2[\varepsilon a + 3C_\tau^{(\psi,\psi)} N^{-1}]}\right),$$

where

$$\Omega_N^n = \bigcap_{i=1}^N \left\{\langle X^{n,a,\psi}, X^{n,a,\psi}\rangle_{\tau_i} - \langle X^{n,a,\psi}, X^{n,a,\psi}\rangle_{\tau_{i-1}} \le 3C_\tau^{(\psi,\psi)} N^{-1}\right\}.$$

Hence we have

$$P^n\left(\max_{1\le i\le N}\sup_{t\in[\tau_{i-1},\tau_i]} |X_t^{n,a,\psi} - X_{\tau_{i-1}}^{n,a,\psi}| > \varepsilon,\ \Omega_N^n\right) \le 2N\exp\left(-\frac{\varepsilon^2 N}{2[\varepsilon + 3C_\tau^{(\psi,\psi)}]}\right).$$

Here notice that $\lim_{n\to\infty} P^n(\Omega_N^n) = 1$. Choosing a large number N, and then letting $n \to \infty$, we can easily deduce the assertion from (i$_2$) of Lemma 3.1. □

PROOF OF THEOREM 3.4.. Let us check the conditions of Theorem 1.1. First, Theorem VIII.3.11 of Jacod and Shiryaev (1987) says that either of [L1] + [C1$_a$] or [L2] + [C2$_a$] implies the finite-dimensional convergence of $X^{n,a}$ for any $a > 0$ [recall also (i$_4$) of Lemma 3.1]. Thus the finite-dimensional convergence of X^n follows from (i$_2$) and (ii) of Lemma 3.1. The condition (iii) of Theorem 1.1 can be shown by means of Lemmas 3.3 and 3.5.

The last assertion of the theorem, concerning the pseudo-metric ρ defined by (3.5), follows from the Gaussian property of the limits by virtue of Example 1.5.10 of van der Vaart and Wellner (1996). □

Let us close this section with a brief explanation of the verification of [PE]. The partitioning entropy condition is essentially the same as the bracketing entropy condition. When a class of random weight functions $\mathscr{W}^n = \{W^{n,\psi}: \psi \in \Psi\}$ is given, the quadratic Π-modulus is the minimum value of the random coefficient K_n that appears in constructing the brackets, that is,

$$\sqrt{|W^{n,\varepsilon,l} - W^{n,\varepsilon,u}|^2 * \nu_{\tau^n}^n} \le K_n \varepsilon,$$

where $(W^{n,\varepsilon,l}, W^{n,\varepsilon,u})$ is a pair of the brackets of random weight functions. Notice that, in the i.i.d. case, the standard L^2-bracketing is $K_n \equiv 1$. In the present case, [PE] is checked by choosing a large K_n that is bounded in probability, and by choosing the brackets as above. Then, an appropriate DFP $\Pi = \{\Pi(\varepsilon)\}$ is constructed from the brackets. The reason why we introduced [PE] rather than the bracketing entropy is for simplicity, especially in the case where the weight functions are random; in practice, it is often more convenient to construct a DFP directly than to do it by way of the bracketing. When Ψ is a compact set of a Euclidean space, the partition $\Pi(\varepsilon)$ should be generated from some ε^q-balls. This $q > 0$ may be arbitrary because of the logarithm of the entropy condition.

4. Discrete-time case. Let $\{\xi_i\}_{i\in\mathbb{N}}$ be an $\ell^\infty(\Psi)$-valued martingale difference array on a discrete-time stochastic basis $\mathbf{B} = (\Omega, \mathscr{F}, \mathbf{F} = \{\mathscr{F}_i\}_{i\in\mathbb{N}_0}, P)$ (recall Definition 1.2). Based on the notation of Definition 2.1, we make the following definition.

DEFINITION 4.1. The adapted envelope $\{\overline{\xi}_i\}_{i\in\mathbb{N}}$ of $\{\xi_i\}_{i\in\mathbb{N}}$ is defined by

$$\overline{\xi}_i = \left[\sup_{\psi\in\Psi} |\xi_i^\psi|\right]_{\mathscr{F}_i,\,P} \qquad \forall\, i \in \mathbb{N}.$$

For a given DFP Π of Ψ, the quadratic Π-modulus $\|\xi\|_\Pi$ of $\{\xi_i\}_{i\in\mathbb{N}}$ is defined as the $\mathbb{R}_+ \cup \{\infty\}$-valued adapted process $\{\|\xi\|_{\Pi,\,i}\}_{i\in\mathbb{N}}$ given by

$$\|\xi\|_{\Pi,\,i} = \sup_{\varepsilon\in(0,\,\Delta_\Pi]\cap\mathbb{Q}} \max_{1\le k\le N_\Pi(\varepsilon)} \frac{\sqrt{\sum_{j=1}^i E_{j-1}|\xi_j(\Psi(\varepsilon;k))|^2}}{\varepsilon} \qquad \forall\, i \in \mathbb{N},$$

where

$$\xi_i(\Psi') = \left[\sup_{\psi,\,\phi\in\Psi'} |\xi_i^\psi - \xi_i^\phi|\right]_{\mathscr{F}_i,\,P} \qquad \forall\, i \in \mathbb{N} \quad \forall\, \Psi' \subset \Psi.$$

THEOREM 4.2. *Let $\{\xi_i\}_{i\in\mathbb{N}}$ be an $\ell^\infty(\Psi)$-valued martingale difference array defined on a discrete-time stochastic basis $\mathbf{B}$. Then, the following* (i) *and* (ii) *hold for any finite stopping time σ.*

(i) *For any NFP Π of Ψ and any constants $\delta \in (0, \Delta_\Pi]$ and $K > 0$,*

$$E^* \max_{1\le m\le\sigma} \sup_{\substack{1\le k\le N_\Pi(\delta)\\ \psi,\,\phi\in\Psi(\delta;k)}} \left|\sum_{i=1}^m (\xi_i^{a,\,\psi} - \xi_i^{a,\,\phi})\right| 1_{\{\|\xi\|_{\Pi,\,\sigma}\le K\}} \lesssim K \int_0^\delta \tilde{H}_\Pi(\varepsilon)\, d\varepsilon,$$

where $\xi_i^{a,\,\psi} = \xi_i^\psi 1_{\{\overline{\xi}_i\le a\}}$ with $a = a(\delta, K) = \delta K/\tilde{H}_\Pi(\delta/2)$.

(ii) *For any DFP Π of Ψ and any constants $K, L > 0$,*

$$E^* \max_{1\le m\le\sigma} \sup_{\psi,\,\phi\in\Psi} \left|\sum_{i=1}^m (\xi_i^\psi - \xi_i^\phi)\right| 1_{\{\|\xi\|_{\Pi,\,\sigma}\le K,\ \sum_1^\sigma E_{i-1}|\overline{\xi}_i|^2\le L\}} \lesssim K \int_0^{\Delta_\Pi} \tilde{H}_\Pi(\varepsilon)\, d\varepsilon + \frac{L}{\Delta_\Pi K}.$$

PROOF. Fix any $\delta, K > 0$, and define a_p, π_p and Π_p for every integer $p \ge 0$ in the same way as the first paragraph of the proof of Theorem 2.5(i). For every integer $q \geq 1$ we introduce the finite stopping time

$$\sigma_q = \inf\left\{ i \in \mathbb{N} : i > \frac{|\tilde{H}_\Pi(2^{-q-2}\delta)|^2}{16} - 1 \right\} \wedge \sigma.$$

Then, we have $\sigma_q \uparrow \sigma$ as $q \to \infty$ almost surely. Hence it is enough to show the assertion (i) with σ replaced by σ_q for every $q \geq 1$.

Now, choose some "good" versions of $\overline{\xi}_i$ and $\xi_i(\Pi_p \psi)$, $p = 0, 1, \ldots, q$ [recall the argument about (2.5)]. The rest of the proof is also quite similar to that of Theorem 2.5(i), although a careful discussion about the choice of versions of conditional expectations is needed here.

We first define

$$A_{i,\,p}(\psi) = 1_{\{\xi_i(\Pi_0\psi) \leq a_0, \ldots, \xi_i(\Pi_{p-1}\psi) \leq a_{p-1},\, \xi_i(\Pi_p\psi) \leq a_p\}}, \qquad p = 0, 1, \ldots, q,$$

$$B_{i,\,p}(\psi) = 1_{\{\xi_i(\Pi_0\psi) \leq a_0, \ldots, \xi_i(\Pi_{p-1}\psi) \leq a_{p-1},\, \xi_i(\Pi_p\psi) > a_p\}}, \qquad p = 1, \ldots, q,$$

$$B_{i,\,0}(\psi) = 1_{\{\xi_i(\Pi_0\psi) > a_0\}}.$$

Since $a_0 = 2\delta K \tilde{H}_\Pi(\delta/2)$ we have that $B_{i,\,0}(\psi) 1_{\{\overline{\xi}_i \leq a(\delta,\,K)\}} = 0$, and that

$$\begin{aligned} &E^*\left(\sup_{\psi\in\Psi} \left| E_{i-1}(\xi_i^{\psi} - \xi_i^{\pi_0\psi}) B_{i,\,0}(\psi) \right| 1_{\{\overline{\xi}_i \leq a(\delta,\,K)\}} \right) \\ &\qquad \leq E\left(\sup_{\psi\in\Psi} \left| E_{i-1}\xi_i(\Pi_0\psi) B_{i,\,0}(\psi) \right| 1_{\{\overline{\xi}_i \leq a(\delta,\,K)\}} \right) \\ &\qquad\qquad \text{since (4.3) holds identically} \\ &\qquad \leq 2E\left(\overline{\xi}_i B_{i,\,0}(\psi) 1_{\{\overline{\xi}_i \leq a(\delta,\,K)\}} \right) \;=\; 0, \end{aligned}$$

and thus

$$\sup_{\psi\in\Psi} \left| E_{i-1}(\xi_i^{\psi} - \xi_i^{\pi_0\psi}) B_{i,\,0}(\psi) \right| 1_{\{\overline{\xi}_i \leq a(\delta,\,K)\}} = 0 \tag{4.1}$$

almost surely. Noting also that $E_{i-1}\xi_i^{\psi} = E_{i-1}\xi_i^{\pi_0\psi} = 0$ almost surely (we may take them to be zero *identically* without loss of generality), we therefore can write in the same way as before that

$$\max_{1\leq m\leq\sigma_q} \sup_{\psi\in\Psi} \left| \sum_{i=1}^{m} (\xi_i^{a,\,\psi} - \xi_i^{a,\,\pi_0\psi}) \right| \leq (I_1') + 2(I_2) + (II_1') + 2(II_2) + (III), \tag{4.2}$$

where

$$(I_1') = \sup_{\psi\in\Psi} \sum_{p=1}^{q} \left| \sum_{i=1}^{\sigma_q} \left\{ \xi_i(\Pi_p\psi) B_{i,\,p}(\psi) - E_{i-1}\xi_i(\Pi_p\psi) B_{i,\,p}(\psi) \right\} \right|,$$

$$(I_2) = \sup_{\psi\in\Psi} \sum_{p=1}^{q} \sum_{i=1}^{\sigma_q} E_{i-1}\xi_i(\Pi_p\psi) B_{i,\,p}(\psi),$$

$$(II'_1) = \sup_{\psi\in\Psi}\left|\sum_{i=1}^{\sigma_q}\left\{\xi_i(\Pi_q\psi)A_{i,q}(\psi) - E_{i-1}\xi_i(\Pi_q\psi)A_{i,q}(\psi)\right\}\right|,$$

$$(II_2) = \sup_{\psi\in\Psi}\sum_{i=1}^{\sigma_p} E_{i-1}\xi_i(\Pi_q\psi)A_{i,q}(\psi),$$

$$(III) = \max_{1\le m\le\sigma_q}\sup_{\psi\in\Psi}\sum_{p=1}^{q}\left|\sum_{i=1}^{m}\left\{(\xi_i^{\pi_p\psi} - \xi_i^{\pi_{p-1}\psi})A_{i,p-1}(\psi)\right.\right.$$
$$\left.\left. - \sum_{i=1}^{m} E_{i-1}(\xi_i^{\pi_p\psi} - \xi_i^{\pi_{p-1}\psi})A_{i,p-1}(\psi)\right\}\right|.$$

In order to make this inequality hold *identically*, and to apply Bernstein–Freedman's inequality for local martingales with bounded jumps, we have to choose certain versions of conditional expectations for which the following inequalities hold *identically*:

$$|E_{i-1}(\xi_i^{\psi}-\xi_i^{\pi_0\psi})B_{i,0}(\psi)| \le E_{i-1}\xi_i(\Pi_0\psi)B_{i,0}(\psi), \tag{4.3}$$

$$\begin{aligned}|E_{i-1}(\xi_i^{\psi} - \xi_i^{\pi_p\psi})B_{i,p}(\psi)| &\le E_{i-1}\xi_i(\Pi_p\psi)B_{i,p}(\psi)\\ &\le a_{p-1}, \qquad p = 1,\dots,q,\end{aligned} \tag{4.4}$$

$$|E_{i-1}(\xi_i^{\pi_p\psi}-\xi_i^{\pi_{p-1}\psi})A_{i,p-1}(\psi)| \le a_{p-1}. \tag{4.5}$$

To do so, first choose some versions of the terms $E_{i-1}\xi_i(\Pi_p(\psi))B_{i,p}(\psi)$ of (4.3) and (4.4), which are nonnegative and the second inequalities of (4.4) are fulfilled, identically; next, on the exceptional sets of (4.1), (4.2), (4.3), (4.4) and (4.5), we define the values of all other conditional expectations as zero. Then, the values of $E_{i-1}\xi_i(\Pi_p\psi)B_{i,p}(\psi)$ and $E_{i-1}(\xi_i^{\pi_p\psi} - \xi_i^{\pi_{p-1}\psi})A_{i,p-1}(\psi)$ depend on ψ only through $\Pi_0\psi,\dots,\Pi_p\psi$ and $\pi_{p-1}\psi$, $\pi_p\psi$, respectively, while (4.1), (4.2), (4.3), (4.4) and (4.5) hold *identically* for *all* $\psi\in\Psi$. [See Nishiyama (1998) for more details.]

To get assertion (i), we can perform the estimations for terms (I'_1), (I_2), (II'_1), (II_2) and (III) on the right-hand side of (4.2) exactly in the same way as those of the proof of Theorem 2.5(i).

The assertion (ii) can be proved in the same way as that of Theorem 2.5(ii), paying attention to the choice of conditional expectations; introduce a continuous-time stochastic basis and repeat the argument with an announcing sequence [see page 14 and I.2.43 of Jacod and Shiryaev (1987)]. □

Let us turn to weak convergence results. We give some analogies of Theorems 3.2 and 3.4; those can be shown using Theorem 4.2(i) instead of Theorem 2.5(i); thus the proofs are omitted. Let Ψ be an arbitrary set and Π a DFP of Ψ. For every $n\in\mathbb{N}$, let $\{\xi_i^n\}_{i\in\mathbb{N}}$ be an $\ell^\infty(\Psi)$-valued martingale difference array on a discrete-time stochastic basis $\mathbf{B}^n = (\Omega^n, \mathscr{F}^n, \mathbf{F}^n = \{\mathscr{F}_i^n\}_{i\in\mathbb{N}_0}, P^n)$. In the same way as Section 3, we introduce the following notation:

1. the adapted envelope $\{\overline{\xi}_i^n\}_{i\in\mathbb{N}}$ of $\{\xi_i^n\}_{i\in\mathbb{N}}$;
2. the quadratic Π-modulus $\{\|\xi^n\|_{\Pi,i}\}_{i\in\mathbb{N}}$ of $\{\xi_i^n\}_{i\in\mathbb{N}}$.

We shall always assume

$$E^n\overline{\xi}_i^n < \infty \qquad \forall\, i \in \mathbb{N}. \tag{4.6}$$

For a given finite stopping time σ^n, we make the following conditions:

[PE′] There exists a DFP Π of Ψ such that

$$\|\xi^n\|_{\Pi,\sigma^n} = O_{P^n}(1) \quad \text{and} \quad \int_0^{\Delta_\Pi} H_\Pi(\varepsilon)\,d\varepsilon < \infty;$$

[L1′] $\sum_{i=1}^{\sigma^n} E_{i-1}^n \overline{\xi}_i^n 1_{\{\overline{\xi}_i^n>\varepsilon\}} \to_{P^n} 0$ for every $\varepsilon > 0$;

[L2′] $\sum_{i=1}^{\sigma^n} E_{i-1}^n |\overline{\xi}_i^n|^2 1_{\{\overline{\xi}_i^n>\varepsilon\}} \to_{P^n} 0$ for every $\varepsilon > 0$.

When we mention [L2′], the assumption that

$$E^n|\overline{\xi}_i^n|^2 < \infty \qquad \forall\, i \in \mathbb{N}, \tag{4.7}$$

which is stronger than (4.6), is implicitly imposed. It is clear that [L2′] implies [L1′].

COROLLARY 4.3. *Consider the above situation with* (4.6). *Suppose that every finite-dimensional marginal of* $X^n = (X^{n,\psi}|\psi \in \Psi)$ *given by* $X^{n,\psi} = \sum_{i=1}^{\sigma^n} \xi_i^{n,\psi}$ *converges weakly to a* (*tight*,) *Borel law*, *and also that the conditions* [PE′] *and* [L1′] *are satisfied. Then* X^n *converges weakly in* $\ell^\infty(\Psi)$ *to a tight*, *Borel law.*

Next, let us consider the process $(t,\psi) \rightsquigarrow X_t^{n,\psi}$ given by

$$X_t^{n,\psi} = \sum_{i=1}^{\sigma_t^n} \xi_i^{n,\psi} \qquad \forall\, t \in [0,\tau]\ \forall\, \psi \in \Psi, \tag{4.8}$$

where $\tau > 0$ is a constant, and $(\sigma_t^n)_{t\in[0,\tau]}$ is a family of finite stopping times on $\mathbf{B}^n$ such that $\sigma_0^n = 0$ and that each path $t \rightsquigarrow \sigma_t^n$ is increasing, cad, with jumps equal to 1. We introduce two kinds of conditions, in which the family $\{C_t^{(\psi,\phi)}: t \in \mathbb{R}_+, (\psi,\phi) \in \Psi^2\}$ of constants should satisfy (3.4):

[C1′] $\sum_{i=1}^{\sigma_t^n} \xi_i^{n,\psi}\xi_i^{n,\phi} \to_{P^n} C_t^{(\psi,\phi)}$ for every $t \in S$ and $(\psi,\phi) \in \Psi^2$;

[C2′] $\sum_{i=1}^{\sigma_t^n} E_{i-1}^n \xi_i^{n,\psi}\xi_i^{n,\phi} \to_{P^n} C_t^{(\psi,\phi)}$ for every $t \in S$ and $(\psi,\phi) \in \Psi^2$.

Similarly to the remark following [L2′], the assumption (4.7) is implicitly imposed when we mention [C2′].

COROLLARY 4.4. *Let* S *be a dense subset of the finite interval* $[0,\tau]$ *containing* τ. *Consider the above situation with* (4.6), *and assume* [PE′] *with* $\sigma^n = \sigma_\tau^n$. *Suppose also that either* [L1′] + [C1′] *or* [L2′] + [C2′] *is satisfied. Then*, *the same conclusion as Theorem* 3.4 *holds for the sequence of processes* $X^n = (X_t^{n,\psi}|(t,\psi) \in [0,\tau]\times\Psi)$ *defined by* (4.8).

Let us close this section with stating a generalization of Jain–Marcus's (1975) central limit theorem to the case of martingale difference arrays. We denote by $N(\Psi, \rho; \varepsilon)$ the ε-covering number of a pseudo-metric space (Ψ, ρ).

PROPOSITION 4.5. *Let (Ψ, ρ) be a totally bounded pseudo-metric space. For every $n \in \mathbb{N}$, let $\{\xi_i^n\}_{i\in\mathbb{N}}$ be an $\ell^\infty(\Psi)$-valued martingale difference array on a discrete-time stochastic basis $\mathbf{B}^n$ such that*

$$|\xi_i^{n,\psi} - \xi_i^{n,\phi}| \le K_i^n \rho(\psi, \phi) \qquad \forall\, \psi, \phi \in \Psi,$$

where $\{K_i^n\}_{i\in\mathbb{N}}$ is an $\mathbb{R}_+$-valued adapted process. For given finite stopping time σ^n, a sufficient condition for [PE′] *is*

$$\sum_{i=1}^{\sigma^n} E_{i-1}^n |K_i^n|^2 = O_{P^n}(1) \quad \textit{and} \quad \int_0^1 \sqrt{\log N(\Psi, \rho; \varepsilon)}\, d\varepsilon < \infty.$$

5. Log-likelihood ratio random fields.

5.1. *Results.* For every $n \in \mathbb{N}$, let $\mathbf{B}^n = (\Omega^n, \mathscr{F}^n, \mathbf{F}^n = \{\mathscr{F}_i^n\}_{i\in\mathbb{N}_0}, P^n)$ be a discrete-time stochastic basis. Let $\mathbf{P}^n = \{P^{n,\psi} : \psi \in \Psi\}$ be a family of probability measures on $(\Omega^n, \mathscr{F}^n)$, indexed by an arbitrary set Ψ, such that $P^{n,\psi} \ll P^n$ for every $\psi \in \Psi$. We denote

$$Z_i^{n,\psi} = \frac{dP_i^{n,\psi}}{dP_i^n},$$

where $P_i^{n,\psi}$ [resp. P_i^n] is the restriction of $P^{n,\psi}$ [resp. P^n] on the σ-field $\mathscr{F}_i^n$. We assume $P_0^{n,\psi} = P_0^n$ for every $\psi \in \Psi$, hence we can set $Z_0^{n,\psi} = 1$. For a given finite stopping time σ^n on $\mathbf{B}^n$, we suppose also that the random element $\log Z_{\sigma^n}^n = (\log Z_{\sigma^n}^{n,\psi} | \psi \in \Psi)$ takes values in $\ell^\infty(\Psi)$. Here we set

$$\zeta_i^{n,\psi} = \sqrt{\frac{Z_{i\wedge\sigma^n}^{n,\psi}}{Z_{(i-1)\wedge\sigma^n}^{n,\psi}}} - 1 \qquad \forall\, i \in \mathbb{N}\ \forall\, \psi \in \Psi.$$

THEOREM 5.1. *In the above situation, suppose that the following conditions hold*:

(a$_1$) $\sum_{i=1}^{\sigma^n} 4E_{i-1}^n \zeta_i^{n,\psi} \zeta_i^{n,\phi} \to_{P^n} C(\psi, \phi)$ (*some constant*) *for every* $\psi, \phi \in \Psi$;

(a$_2$) $\sup_{\psi\in\Psi} \left| \sum_{i=1}^{\sigma^n} 4E_{i-1}^n |\zeta_i^{n,\psi}|^2 - C(\psi, \psi) \right| \to_{P^{n*}} 0$;

(b) $\sum_{i=1}^{\sigma^n} E_{i-1}^n |\overline{\zeta_i^n}|^2 1_{\{|\overline{\zeta_i^n}| > \varepsilon\}} \to_{P^n} 0$ *for every* $\varepsilon > 0$;

(c) *there exists a DFP* Π *of* Ψ *such that*

$$\|\zeta^n\|_{\Pi,\sigma^n} = O_{P^n}(1) \quad \textit{and} \quad \int_0^{\Delta_\Pi} H_\Pi(\varepsilon)\, d\varepsilon < \infty.$$

Then, it holds that $\log Z^n_{\sigma^n} \Longrightarrow_{P^n} X$, *where each marginal* $(X(\psi_1), \dots, X(\psi_d))$ *has the normal distribution* $N(-\frac{1}{2}\mathrm{diag}\,\Sigma, \Sigma)$ *with* $\Sigma = \{C(\psi_i, \psi_j)\}_{ij}$. *Furthermore, the formula*

$$\rho(\psi, \phi) = \sqrt{C(\psi, \psi) + C(\phi, \phi) - 2C(\psi, \phi)} \qquad \forall\, \psi, \phi \in \Psi$$

defines a pseudo-metric on Ψ *such that* (Ψ, ρ) *is totally bounded and that almost all paths of* X *are uniformly* ρ*-continuous.*

REMARK. If a version of the conditional expectation $E^n_{i-1}\zeta_i^{n,\psi}\zeta_i^{n,\phi}$ satisfies the assumption (a_1), then so does any version. However, this is not true in (a_2); the assumption means that there exist *some* versions of $E^n_{i-1}|\zeta_i^{n,\psi}|^2$'s which satisfy the requirement.

EXAMPLE (Ergodic Markov chain). Let $\{X_i\}_{i\in\mathbb{N}_0}$ be an ergodic Markov chain, defined on a probability space $(\Omega, \mathscr{F}, P)$, with values in an arbitrary state space $(E, \mathscr{E})$. Let $\mu(dx)$ denote the initial distribution, $p(x, dy)$ the transition distribution, and $\pi(dx)$ the invariant distribution. Let us equip the space $\mathscr{L}^2 = \mathscr{L}^2(E \times E, \pi(dx)p(x, dy))$ with the "inner product" given by

$$\langle h_1, h_2\rangle_{\mathscr{L}^2} = \int_{E\times E} h_1(x, y)h_2(x, y)\pi(dx)p(x, dy) \qquad \forall\, h_1, h_2 \in \mathscr{L}^2.$$

The meaning of the quotation marks is that $\|h\|_{\mathscr{L}^2} = \sqrt{\langle h, h\rangle_{\mathscr{L}^2}}$ is merely a "semi-"norm. Next we define the subset $\mathscr{L}^2_0$ of $\mathscr{L}^2$ by

$$\mathscr{L}^2_0 = \left\{h \in \mathscr{L}^2\colon \quad \int_E h(x, y)p(x, dy) = 0 \quad \forall\, x \in E \quad \text{and} \quad h > -1\right\}.$$

Fix a subset $\mathscr{H} \subset \mathscr{L}^2_0$. For every $n \in \mathbb{N}$, let us consider a family of probability measures $\mathbf{P}^n = \{P^{n,h}\colon h \in \mathscr{H}\}$ on $(\Omega, \mathscr{F})$ such that, under $P^{n,h}$, the process $\{X_i\}_{i\in\mathbb{N}_0}$ is the Markov chain with initial distribution μ and transition distribution $p^{n,h}$ given by

$$p^{n,h}(x, dy) = \left(1 + \frac{h(x, y)}{\sqrt{n}}\right) p(x, dy).$$

Here we set $\mathscr{F}_i = \sigma\{X_0, \dots, X_i\}$. Then it holds that

$$Z_i^{n,h} = \frac{dP_i^{n,h}}{dP_i} = \prod_{j=1}^{i}\left(1 + \frac{h(X_{j-1}, X_j)}{\sqrt{n}}\right).$$

We need some more notation to state the following result, which concerns the asymptotic behavior of the process $\log Z^n_n = (\log Z^{n,h}_n | h \in \mathscr{H})$. For given $K \in \mathscr{L}^2(E, \pi(dx))$ we define the pseudo-metric ρ_K on $\mathscr{L}^2$ by

$$\rho_K(h_1, h_2) = \sup_{x\in E}\frac{\rho_x(h_1, h_2)}{|K(x)| \vee 1} \qquad \forall\, h_1, h_2 \in \mathscr{L}^2,$$

where

$$\rho_x(h_1, h_2) = \sqrt{\int_E |h_1(x, y) - h_2(x, y)|^2 p(x, dy)} \qquad \forall\, x \in E.$$

For every $\varepsilon > 0$ the ε-bracketing number $N_{[\ ]}(\mathscr{H}, \rho_K; \varepsilon)$ is defined as the smallest N such that: there exist N-pairs $l_k, u_k \in \mathscr{L}^2$ such that $\rho_K(l_k, u_k) < \varepsilon$ $(k = 1, \dots, N)$ and that for every $h \in \mathscr{H}$ the relation $l_k \leq h \leq u_k$ holds for some k.

PROPOSITION 5.2. *Let $\{X_i\}_{i \in \mathbb{N}_0}$, $(\Omega, \mathscr{F}, \mathbf{F} = \{\mathscr{F}_i\}_{i \in \mathbb{N}_0}, P)$ and $\mathbf{P}^n = \{P^{n,h}: h \in \mathscr{H}\}$ as above be given. Suppose that there exists $h^* \in \mathscr{L}^4(E \times E, \pi(dx) p(x, dy))$ such that $\sup_{h \in H} |h| \leq h^*$, and also that there exists $K \in \mathscr{L}^2(E, \pi(dx))$ such that*

$$\int_0^1 \sqrt{\log N_{[\ ]}(\mathscr{H}, \rho_K; \varepsilon)}\, d\varepsilon < \infty.$$

Then, it holds that $\log Z_n^n \Rightarrow_P X$ in $\ell^\infty(\mathscr{H})$, where $X(h) = -\frac{1}{2}\|h\|_{\mathscr{L}^2}^2 + G(h)$ and $h \rightsquigarrow G(h)$ is a zero-mean Gaussian process such that $EG(h_1)G(h_2) = \langle h_1, h_2 \rangle_{\mathscr{L}^2}$. Furthermore, almost all paths of X are uniformly $\|\cdot\|_{\mathscr{L}^2}$-continuous.

This result is easily derived from the ergodic theorem and Theorem 5.1, hence the proof is omitted.

Here we give a statistical application. Fix a subset $\mathscr{H} \subset \mathscr{L}_0^2$ such that $\|h\|_{\mathscr{L}^2} > 0$ for every $h \in \mathscr{H}$, and let us consider the testing problem:

$$\begin{aligned} &\text{hypothesis } H_0\colon p \\ &\quad\text{against } \quad H_1^n\colon p^{n,h} \quad \text{for some } h \in \mathscr{H}. \end{aligned}$$

We propose the test statistics

$$S^n = \sup_{h \in \mathscr{H}} \left| \tfrac{1}{2}\|h\|_{\mathscr{L}^2}^2 + \log Z_n^{n,h} \right|.$$

Assume the same conditions as in Proposition 5.2. Then, it holds that

$$S^n \Rightarrow P^{n,u} \sup_{h \in \mathscr{H}} |\langle h, u \rangle_{\mathscr{L}^2} + G(h)| \qquad \text{in } \mathbb{R} \qquad \forall\, u \in \{0\} \cup \mathscr{H},$$

where the process $h \rightsquigarrow G(h)$ is as above. This fact follows easily from Proposition 5.2, which implies local asymptotic normality and contiguity, together with Le Cam's third lemma and the continuous mapping theorem. In view of Anderson's lemma [e.g., Lemma 3.11.4 of van der Vaart and Wellner (1996)], the statistics S^n seems reasonable.

5.2. *Proof of Theorem* 5.1. Let us denote

$$\tilde{Z}_i^{n,\psi} = \frac{Z_{i \wedge \sigma^n}^{n,\psi}}{Z_{(i-1)\wedge \sigma^n}^{n,\psi}} \qquad \forall\, i \in \mathbb{N}\ \forall\, \psi \in \Psi,$$

$$\lambda_i^{n,a,\psi} = \log \tilde{Z}_i^{n,\psi} 1_{\{\bar{\zeta}_i^n \leq a\}} \qquad \forall\, i \in \mathbb{N}\ \forall\, \psi \in \Psi\ \forall\, a > 0.$$

The process $\psi \rightsquigarrow \log Z^{n,\psi}_{\sigma^n} = \sum_{i=1}^{\sigma^n} \log \tilde{Z}^{n,\psi}_i$ can be well approximated by the process $\psi \rightsquigarrow \Lambda^{n,a,\psi} = \sum_{i=1}^{\sigma^n} \lambda^{n,a,\psi}_i$. As a matter of fact, it holds that

$$\begin{aligned}\sup_{\psi\in\Psi} 1_{\{\log Z^{n,\psi}_{\sigma^n} \neq \Lambda^{n,a,\psi}\}} &\le \sum_{i=1}^{\sigma^n} 1_{\{\overline{\zeta}^n_i > a\}} \\ &\le \frac{1}{a^2} \sum_{i=1}^{\sigma^n} |\overline{\zeta}^n_i|^2 1_{\{\overline{\zeta}^n_i > a\}};\end{aligned}$$

hence using also Lenglart's inequality we obtain

$$\sup_{\psi\in\Psi} |\log Z^{n,\psi}_{\sigma^n} - \Lambda^{n,a,\psi}| \xrightarrow{P^{n*}} 0.$$

We consider the decomposition

$$\Lambda^{n,a,\psi} = \sum_{i=1}^{\sigma^n} E^n_{i-1}\lambda^{n,a,\psi}_i + \sum_{i=1}^{\sigma^n} \left\{\lambda^{n,a,\psi}_i - E^n_{i-1}\lambda^{n,a,\psi}_i\right\}. \tag{5.1}$$

We will derive the uniform convergence of the first term in (outer) probability and apply Corollary 4.4 to the martingale difference array $\{\xi^n_i\}_{i\in\mathbb{N}_0}$ of the second term, that is, $\xi^{n,\psi}_i = \lambda^{n,a,\psi}_i - E^n_{i-1}\lambda^{n,a,\psi}_i$. We use the following lemma which will be proved later.

LEMMA 5.3. *For every $a \in (0,1)$, there exist some versions of the conditional expectations $E^n_{i-1}\lambda^{n,a,\psi}_i$ such that*:

(i) *If* $\sup_{\psi\in\Psi} C(\psi,\psi) < \infty$ *then* $\sup_{\psi\in\Psi}\left|\sum_{i=1}^{\sigma^n} E^n_{i-1}\lambda^{n,a,\psi}_i + \frac{1}{2}C(\psi,\psi)\right| \to_{P^{n*}} 0$.

(ii) $\sum_{i=1}^{\sigma^n} E^n_{i-1}\lambda^{n,a,\psi}_i\lambda^{n,a,\phi}_i \to_{P^n} C(\psi,\phi)$ *for every* $\psi,\phi\in\Psi$.

(iii) $\sum_{i=1}^{\sigma^n} |E^n_{i-1}\lambda^{n,a,\psi}_i|^2 \to_{P^n} 0$ *for every* $\psi\in\Psi$.

REMARK (i) We will see later that the conditions of the theorem actually imply that $\sup_{\psi\in\Psi} C(\psi,\psi) < \infty$. (ii) The choice of versions of the conditional expectations $E^n_{i-1}\lambda^{n,a,\psi}_i\lambda^{n,a,\phi}_i$ is not important.

Let us proceed with the main part of the proof. It is clear that there exists a constant $\delta \in (0,1)$ such that $|\log x - \log y| \le 2|\sqrt{x} - \sqrt{y}|$ whenever $x, y \in [1-\delta, 1+\delta]$. We consider the decomposition (5.1) for $a = \sqrt{1+\delta} - 1$; then it holds that $\{x\colon |\sqrt{x} - 1| \le a\} \subset \{x\colon |x-1| \le \delta\}$.

First we show the weak convergence of the second term of the decomposition (5.1). The condition [C2′] is direct from (ii) and (iii) of Lemma 5.3. It is also easy to see that the assumption (b) implies the Lindeberg condition [L2′]. Finally, recalling the choice of δ and the relationship between a and δ, we have for any subset $\Psi' \subset \Psi$,

$$
\begin{aligned}
E_{i-1}^n & \left[\sup_{\psi,\phi\in\Psi'} |\lambda_i^{n,a,\psi} - \lambda_i^{n,a,\phi}|\right]^2_{\mathscr{F}_i^n,\,P^n} \\
&= E_{i-1}^n \left[\sup_{\psi,\,\phi\in\Psi} |\log \tilde{Z}_i^{n,\psi} - \log \tilde{Z}_i^{n,\phi}| 1_{\{\overline{\zeta}_i^n \le a\}}\right]^2_{\mathscr{F}_i^n,\,P^n} \\
&\le E_{i-1}^n \left[\sup_{\psi,\,\phi\in\Psi'} 2|\zeta_i^{n,\psi} - \zeta_i^{n,\phi}|\right]^2_{\mathscr{F}_i^n,\,P^n}.
\end{aligned}
$$

Thus the assumption (c) implies the condition [PE$'$]. Consequently, Corollary 4.4 yields $\sum_i^n \xi_i^n \Rightarrow_{P^n} Y$, where $(Y(\psi_1), \dots, Y(\psi_d))$ has the normal distribution $N(0, \Sigma)$ with $\Sigma = \{C(\psi_i, \psi_j)\}_{ij}$.

Next we consider the first term of the decomposition. Observe that

$$
\begin{aligned}
\sqrt{C(\psi,\psi)} &= \sqrt{E|Y(\psi)|^2} \\
&\le \sqrt{E|Y(\psi)-Y(\phi)|^2} + \sqrt{E|Y(\phi)|^2} = \rho(\psi,\phi) + \sqrt{C(\psi,\phi)}.
\end{aligned}
$$

The inequality above and the total boundedness of (Ψ, ρ), a consequence of Corollary 4.4, imply that $\sup_{\psi\in\Psi} C(\psi,\psi) < \infty$. Hence (i) of Lemma 5.3 works to show the uniform convergence of the first term of (5.1). Also, it is trivial from the above inequality that $\psi \rightsquigarrow \sqrt{C(\psi,\psi)}$ is uniformly ρ-continuous; thus so is $\psi \rightsquigarrow C(\psi,\psi)$.

PROOF OF LEMMA 5.3. For every $\varepsilon > 0$ we denote

$$
B^{n,\varepsilon}(\psi) = \sum_{i=1}^{\sigma^n} E_{i-1}^n \lambda_i^{n,\varepsilon,\psi},
$$

$$
C^{n,\varepsilon}(\psi,\,\phi) = \sum_{i=1}^{\sigma^n} E_{i-1}^n \lambda_i^{n,\varepsilon,\psi} \lambda_i^{n,\varepsilon,\phi}.
$$

Step 1. First we prove the following facts: for given $a \in (0,1)$ there exist constants $K_1, K_2, K_3 > 0$ such that for every $\varepsilon \in (0, a]$,

$$
\sup_{\psi\in\Psi} \left|B^{n,\varepsilon}(\psi) + \tfrac{1}{2} C(\psi,\,\psi)\right| \le \varepsilon K_1 + o_{P^n}(1), \tag{5.2}
$$

$$
\left|C^{n,\varepsilon}(\psi,\phi) - C(\psi,\phi)\right| \le \varepsilon K_2 + o_{P^n}(1) \qquad \forall\, \psi,\ \phi \in \Psi, \tag{5.3}
$$

$$
\sum_{i=1}^{\sigma^n} |E_{i-1}^n \lambda_i^{n,\varepsilon,\psi}| \le \tfrac{1}{2} C(\psi,\psi) + \varepsilon K_3 + o_{P^n}(1) \qquad \forall\, \psi \in \Psi. \tag{5.4}
$$

In order to show (5.2), first notice that there exists a constant $K > 0$ such that $|\log x - (x-1) + 2(\sqrt{x}-1)^2| \le K|\sqrt{x}-1|^3$ whenever $|\sqrt{x}-1| \le a$. Hence, for fixed $\varepsilon \in (0, a]$ we obtain

$$\left| B^{n,\varepsilon}(\psi) + 2\sum_{i=1}^{\sigma^n} E_{i-1}^n |\zeta_i^{n,\psi}|^2 1_{\{\overline{\zeta}_i^n \le \varepsilon\}} \right| \le \varepsilon K \sum_{i=1}^{\sigma^n} E_{i-1}^n |\zeta_i^{n,\psi}|^2 1_{\{\overline{\zeta}_i^n \le \varepsilon\}} + \left| \sum_{i=1}^{\sigma^n} E_{i-1}^n (\tilde{Z}_i^{n,\psi} - 1) 1_{\{\overline{\zeta}_i^n \le \varepsilon\}} \right|$$

almost surely. Since $E_{i-1}^n(\tilde{Z}_i^{n,\psi} - 1) = 0$ almost surely, the last term of the right-hand side equals

$$\begin{aligned} \left| \sum_{i=1}^{\sigma^n} E_{i-1}^n (\tilde{Z}_i^{n,\psi} - 1) 1_{\{\overline{\zeta}_i^n > \varepsilon\}} \right| &\le \sum_{i=1}^{\sigma^n} E_{i-1}^n |\tilde{Z}_i^{n,\psi} - 1| 1_{\{\overline{\zeta}_i^n > \varepsilon\}} \\ &\le \frac{\varepsilon + 2}{\varepsilon} \sum_{i=1}^{\sigma^n} E_{i-1}^n |\zeta_i^{n,\psi}|^2 1_{\{\overline{\zeta}_i^n > \varepsilon\}} \end{aligned}$$

almost surely. Thus we obtain

$$\begin{aligned} &\left| B^{n,\varepsilon}(\psi) + \frac{1}{2} C(\psi,\psi) \right| \\ &\quad \le (2 + \varepsilon K) \left| \sum_{i=1}^{\sigma^n} E_{i-1}^n |\zeta_i^{n,\psi}|^2 - \frac{1}{4} C(\psi,\psi) \right| + \varepsilon K \frac{1}{4} C(\psi,\psi) \\ &\qquad + \left(2 + \frac{\varepsilon + 2}{\varepsilon} \right) \sum_{i=1}^{\sigma^n} E_{i-1}^n |\overline{\zeta}_i^n|^2 1_{\{\overline{\zeta}_i^n > \varepsilon\}} \end{aligned} \tag{5.5}$$

almost surely. In order to get the estimate for all $\omega \in \Omega^n$, we can choose the versions of conditional expectations as follows: first, we may without loss of generality choose a version of $E_{i-1}^n |\overline{\zeta}_i^n|^2 1_{\{\overline{\zeta}_i^n > \varepsilon\}}$ which is nonnegative identically; next, on the union of all exceptional sets for the estimates appearing above, we define the values of all other conditional expectations as zero. Then, the inequality (5.5) holds *identically* for *all* $\psi \in \Psi$. By taking the supremum of (5.5) with respect to $\psi \in \Psi$ and letting $n \to \infty$, we obtain the assertion (5.2).

A similar argument yields (5.4). In fact, it is much easier than (5.2), because the assertion of (5.4) is ψ-wise, for which we do not need any argument about versions of conditional expectations. Also, it is easy to show (5.3) if we notice the following fact: for given $a \in (0, 1)$ there exists a constant $K > 0$ such that $|\log x \cdot \log y - 4(\sqrt{x} - 1)(\sqrt{y} - 1)| \le K \max\{|\sqrt{x} - 1|^3, |\sqrt{y} - 1|^3\}$ whenever $\max\{|\sqrt{x} - 1|, |\sqrt{y} - 1|\} \le a$.

Step 2. Next we prove the following facts:

$$\sup_{\psi \in \Psi} |B^{n,a}(\psi) - B^{n,\varepsilon}(\psi)| \xrightarrow{P^{n*}} 0 \qquad \forall \varepsilon \in (0, a), \tag{5.6}$$

$$|C^{n,a}(\psi,\phi) - C^{n,\varepsilon}(\psi,\phi)| \xrightarrow{P^n} 0 \qquad \forall \psi, \phi \in \Psi\ \forall \varepsilon \in (0, a). \tag{5.7}$$

In order to show (5.6), notice that for given $a \in (0,1)$ there exists a constant $K > 0$ such that $|\log x| \le K|\sqrt{x} - 1|^2$ whenever $|\sqrt{x} - 1| \le a$. For every $\varepsilon \in (0, a)$ it holds that

$$\begin{aligned}
&\left| \sum_{i=1}^{\sigma^n} E^n_{i-1} \lambda_i^{n,a,\psi} - \sum_{i=1}^{\sigma^n} E^n_{i-1} \lambda_i^{n,\varepsilon,\psi} \right| \\
&\qquad = \left| \sum_{i=1}^{\sigma^n} E^n_{i-1} \log \tilde{Z}_i^{n,\psi} 1_{\{\varepsilon < \overline{\zeta}_i^n \le a\}} \right| \\
&\qquad \le K \sum_{i=1}^{\sigma^n} E^n_{i-1} |\zeta_i^{n,\psi}|^2 1_{\{\varepsilon < |\zeta_i^n| \le a\}} \\
&\qquad \le K \sum_{i=1}^{\sigma^n} E^n_{i-1} |\overline{\zeta}_i^n|^2 1_{\{\overline{\zeta}_i^n > \varepsilon\}},
\end{aligned} \tag{5.8}$$

almost surely. We can choose some versions of conditional expectations such that the estimate above holds identically for all $\psi \in \Psi$, in the same way as in the proof of (5.2). Take the supremum of (5.8) with respect to $\psi \in \Psi$ and let $n \to \infty$; then we get (5.6). A similar computation yields (5.7).

Step 3. Now it is easy to see that (5.2) and (5.6) imply the assertion (i) and that (5.3) and (5.7) do so for assertion (ii); first choose $\varepsilon > 0$ small enough, and then let $n \to \infty$. In order to show assertion (iii), notice that for any $\varepsilon \in (0, a)$,

$$\begin{aligned}
&\left| \sum_{i=1}^{\sigma^n} |E^n_{i-1} \lambda_i^{n,a,\psi}|^2 - \sum_{i=1}^{\sigma^n} |E^n_{i-1} \lambda_i^{n,\varepsilon,\psi}|^2 \right| \\
&\qquad = \left| \sum_{i=1}^{\sigma^n} E^n_{i-1} (\lambda_i^{n,a,\psi} + \lambda_i^{n,\varepsilon,\psi}) E^n_{i-1} (\lambda_i^{n,a,\psi} - \lambda_i^{n,\varepsilon,\psi}) \right| \\
&\qquad \le 2|\log(1-a^2)| \sum_{i=1}^{\sigma^n} \left| E^n_{i-1} \log \tilde{Z}_i^{n,\psi} 1_{\{\varepsilon < \overline{\zeta}_i^n \le a\}} \right| \;=\; o_{P^n}(1);
\end{aligned}$$

hence

$$\begin{aligned}
\sum_{i=1}^{\sigma^n} |E^n_{i-1} \lambda_i^{n,a,\psi}|^2 &= \sum_{i=1}^{\sigma^n} |E^n_{i-1} \lambda_i^{n,\varepsilon,\psi}|^2 + o_{P^n}(1) \\
&\le |\log(1-\varepsilon^2)| \sum_{i=1}^{\sigma^n} |E^n_{i-1} \lambda_i^{n,\varepsilon,\psi}| + o_{P^n}(1).
\end{aligned}$$

We therefore obtain (iii) by virtue of (5.4); first choose $\varepsilon > 0$ small enough and then let $n \to \infty$. □

APPENDIX

Supplement to the Proof of Lemma 3.1. We state here a proof of the equivalence $[\mathrm{C1_a}] \Leftrightarrow [\mathrm{C2_a}]$ under [L1] in (ii) of Lemma 3.1, following exactly the same line as that of Theorem VIII.3.6 of Jacod and Shiryaev (1987).

Fix any $\psi \in \Psi$ and $a > 0$, and we set $Y^n = [X^{n,a,\psi}, X^{n,a,\psi}] - \langle X^{n,a,\psi}, X^{n,a,\psi}\rangle$. We will prove that $\sup_{s\in[0,t]} |\Delta Y^n_s| \xrightarrow{P^n} 0$ for every $t \in S$ under either [L1] + [C1$_\mathrm{a}$] or [L1] + [C2$_\mathrm{a}$]. Since $X^{n,a,\psi}$ is a locally square-integrable martingale, we have that Y^n is a local martingale and that so is $|Y^n|^2 - [Y^n, Y^n]$ [see Proposition I.4.50 of Jacod and Shiryaev (1987)]. Hence Lenglart's inequality yields that for every $\varepsilon, \eta > 0$,

$$P^n\left(\sup_{s\in[0,t]} |Y^n_s|^2 \geq \varepsilon\right) \leq \frac{1}{\varepsilon}\left(\eta + E^n \sup_{s\in[0,t]} \Delta[Y^n, Y^n]_s\right) + P^n([Y^n, Y^n]_t \geq \eta)$$
$$\leq \frac{2\eta}{\varepsilon} + \left(\frac{16a^4}{\varepsilon} + 1\right) P^n([Y^n, Y^n]_t \geq \eta),$$

because $\Delta[Y^n, Y^n] = |\Delta Y^n|^2 \leq (|\Delta[X^{n,a,\psi}, X^{n,a,\psi}]|^2 \vee |\Delta\langle X^{n,a,\psi}, X^{n,a,\psi}\rangle|^2) \leq 16a^4$. Thus it suffices to show that

$$[Y^n, Y^n]_t \xrightarrow{P^n} 0 \qquad \forall\, t \in S \tag{A.1}$$

under either [L1] + [C1$_\mathrm{a}$] or [L1] + [C2$_\mathrm{a}$].

Since the local martingale Y^n has finite variation, we have

$$\begin{aligned}[Y^n, Y^n]_t &= \sum_{s\leq t} |\Delta Y^n_s|^2 \\ &\leq \sum_{s\leq t} |\Delta[X^{n,a,\psi}, X^{n,a,\psi}]_s|^2 + \sum_{s\leq t} |\Delta\langle X^{n,a,\psi}, X^{n,a,\psi}\rangle_s|^2 \\ &\leq \alpha^n_t A^n_t + \beta^n_t B^n_t,\end{aligned}$$

where

$$\alpha^n_t = \sup_{s\in[0,t]} \Delta[X^{n,a,\psi}, X^{n,a,\psi}]_s, \qquad A^n_t = [X^{n,a,\psi}, X^{n,a,\psi}]_t,$$
$$\beta^n_t = \sup_{s\in[0,t]} \Delta\langle X^{n,a,\psi}, X^{n,a,\psi}\rangle_s, \qquad B^n_t = \langle X^{n,a,\psi}, X^{n,a,\psi}\rangle_t.$$

Using (ii$_3$) of Lemma 3.1, we obtain that $\alpha^n_t \to_{P^n} 0$ and $\beta^n_t \to_{P^n} 0$ for every $t \in S$, under [L1]. On the other hand, Lenglart's inequality yields that

$$P^n(A^n_t \geq \varepsilon) \leq \frac{\eta + 2a^2}{\varepsilon} + P^n(B^n_t \geq \eta) \qquad \forall\, \varepsilon, \eta > 0$$

and that

$$P^n(B^n_t \geq \varepsilon) \leq \frac{\eta + 4a^2}{\varepsilon} + P^n(A^n_t \geq \eta) \qquad \forall\, \varepsilon, \eta > 0.$$

Hence [C2$_\mathrm{a}$] implies that $A^n_t = O_{P^n}(1)$, and [C1$_\mathrm{a}$] does that $B^n_t = O_{P^n}(1)$. The assertion (A.1) has been established. $\square$

Acknowledgments. I thank Richard D. Gill for his constant encouragement. A major part of this research was performed in the Department of Mathematics at the University of Utrecht, while the author held a fellowship for research abroad from the Japan Society for the Promotion of Science.

REFERENCES

ANDERSEN, N. T., GINÉ, E., OSSIANDER, M. and ZINN, J. (1988). The central limit theorem and the law of iterated logarithm for empirical processes under local conditions. *Probab. Theory Related Fields* **77** 271–305.

ARCONES, M. A. and YU, B. (1994). Central limit theorems for empirical and U-processes of stationary mixing sequences. *J. Theoret. Probab.* **7** 47–71.

BAE, J. and LEVENTAL, S. (1995). Uniform CLT for Markov chains and its invariance principle: a martingale approach. *J. Theoret. Probab.* **8** 549–570.

DOUKHAN, P., MASSART, P. and RIO, E. (1995). Invariance principles for absolutely regular empirical processes. *Ann. Inst. H. Poincaré Probab. Statist.* **31** 393–427.

DUDLEY, R. M. (1978). Central limit theorems for empirical measures. *Ann. Probab.* **6** 899–929. [Correction (1979) **7** 909–911.]

FREEDMAN, D. A. (1975). On tail probabilities for martingales. *Ann. Probab.* **3** 100–118.

JACOD, J. and SHIRYAEV, A. N. (1987). *Limit Theorems for Stochastic Processes.* Springer, Berlin.

JAIN, N. and MARCUS, M. (1975). Central limit theorems for $C(S)$-valued random variables. *J. Funct. Anal.* **19** 216–231.

LIPTSER, R. S. and SHIRYAEV A. N. (1989). *Theory of Martingales.* Kluwer, Dordrecht.

NISHIYAMA, Y. (1996). A central limit theorem for ℓ^∞-valued martingale difference arrays and its application. Preprint 971, Dept. Mathematics, Utrecht Univ.

NISHIYAMA, Y. (1997). Some central limit theorems for ℓ^∞-valued semimartingales and their applications. *Probab. Theory Related Fields* **108** 459–494.

NISHIYAMA, Y. (1999). A maximal inequality for continuous martingales and M-estimation in a Gaussian white noise model. *Ann. Statist.* **27** 675–696.

NISHIYAMA, Y. (1998). *Entropy methods for martingales.* Ph.D. thesis, Univ. Utrecht.

OSSIANDER, M. (1987). A central limit theorem under metric entropy with L_2 bracketing. *Ann. Probab.* **15** 897–919.

VAN DE GEER, S. (1995). Exponential inequalities for martingales, with application to maximum likelihood estimation for counting processes. *Ann. Statist.* **23** 1779–1801.

VAN DER VAART, A. W. and WELLNER, J. A. (1996). *Weak Convergence and Empirical Processes: With Applications to Statistics.* Springer, New York.

VOSTRIKOVA, L. (1987). On the weak convergence of likelihood ratio processes of general statistical parametric models. *Stochastics* **23** 277–298.

INSTITUTE OF STATISTICAL MATHEMATICS
4-6-7 MINAMI-AZABU
MINATO-KU
TOKYO 106-8569
JAPAN
E-MAIL: nisiyama@ism.ac.jp