



# A Donsker-Type Theorem for Log-Likelihood Processes

Zhonggen Su<sup>1</sup> · Hanchao Wang<sup>2</sup>

Received: 29 September 2018 / Revised: 24 May 2019 / Published online: 13 June 2019  
© Springer Science+Business Media, LLC, part of Springer Nature 2019

## Abstract

Let  $(\Omega, \mathcal{F}, (\mathcal{F})_{t \geq 0}, P)$  be a complete stochastic basis, and  $X$  be a semimartingale with predictable compensator  $(B, C, \nu)$ . Consider a family of probability measures  $\mathbf{P} = (P^{n, \psi}, \psi \in \Psi, n \geq 1)$ , where  $\Psi$  is an index set,  $P^{n, \psi} \stackrel{\text{loc}}{\ll} P$ , and denote the likelihood ratio process by  $Z_t^{n, \psi} = \frac{dP^{n, \psi}|_{\mathcal{F}_t}}{dP|_{\mathcal{F}_t}}$ . Under some regularity conditions in terms of logarithm entropy and Hellinger processes, we prove that  $\log Z_t^n$  converges weakly to a Gaussian process in  $\ell^\infty(\Psi)$  as  $n \rightarrow \infty$  for each fixed  $t > 0$ .

**Keywords** Hellinger process of order zero · Log-likelihood process · Semimartingale · Weak convergence

**Mathematics Subject Classification (2010)** 60F05 · 60F17

## 1 Introduction and Preliminaries

The celebrated Donsker theorem is a functional extension of the central limit theorem in probability theory. Plenty of research on this topic has come out in the past decades. The reader is referred to classic books, and papers like Dudley [3], Gine and Zinn [4], Ossiander [10], Andersen et. al [1], Liptser and Shiryaev [7], van der Geer [11], Billingsley [2], Jacod and Shiryaev [5] for both theoretical framework and wide applications. A primary purpose of the present paper is to establish a certain Donsker theorem for log-likelihood processes indexed by an arbitrary set. In this section, we first introduce some basic notions about log-likelihood processes and martingale representation property.

---

✉ Hanchao Wang  
wanghanchao@sdu.edu.cn  
Zhonggen Su  
suzhonggen@zju.edu.cn

<sup>1</sup> School of Mathematical Science, Zhejiang University, Hangzhou, China

<sup>2</sup> Zhongtai Securities Institute for Financial Studies, Shandong University, Jinan 250100, China

Throughout this paper, we follow the standard definitions and notations of martingale theory, which can be found in the book by Jacod and Shiryaev [5]. Let  $(\Omega, \mathcal{F}, (\mathcal{F})_{t \geq 0}, P)$  be a complete stochastic basis. Fix a semimartingale  $X$  on it, and assume that all  $P$ -martingales have a representation property relative to  $X$ . Denote by the triplet  $(B, C, \nu)$  the predictable characteristic of  $X$  (associated with some bounded truncation function). More precisely, if  $\Delta X_t = X_t - X_{t-}$  denotes the jump of  $X$  at time  $t$ , then  $X_t - \sum_{s \leq t} (\Delta X_s - h_\tau(\Delta X_s))$ , where  $h_\tau(x) = x1_{(|x| \leq \tau)}$ , is a special semimartingale, which can be uniquely divided into a bounded variation process and a local martingale process. The  $B$  is a bounded variation process of  $X - \sum_{s \leq \cdot} (\Delta X_s - h_\tau(\Delta X_s))$ . Let  $X^c$  be the continuous local martingale part of  $X$ , then

$$C_t = \langle X^c, X^c \rangle_t.$$

Let  $\mu$  be the jump measure of  $X$  defined by

$$\mu(\omega, dt, dx) = \sum_s \mathbf{1}_{(\Delta X_s(\omega) \neq 0)} \varepsilon_{(s, \Delta X_s(\omega))}(dt, dx) \tag{1.1}$$

where  $\varepsilon_{(s, \Delta X_s(\omega))}$  denotes the Dirac measure at point  $(s, \Delta X_s(\omega))$ . The  $\nu$  is the unique predictable compensator of  $\mu$  (up to a  $P$ -null set). Namely,  $\nu$  is a predictable random measure such that for any predictable function<sup>1</sup>  $W$ ,  $W * (\mu - \nu) = W * \mu - W * \nu$  is a local martingale, where the  $W * \mu$  is defined by

$$W * \mu_t(\omega) = \begin{cases} \int_0^t \int_{\mathbb{R}} W(\omega, s, x) \mu(\omega; ds, dx) & \text{if } \int_0^t \int_{\mathbb{R}} |W(\omega, s, x)| \mu(\omega; ds, dx) < \infty, \\ +\infty & \text{otherwise.} \end{cases}$$

(see Section 2.1 of Chapter 2 in Jacod and Shiryaev [5] for more details). Note the predictable quadratic variation is given by

$$\langle W * (\mu - \nu) \rangle_t = (W - \hat{W})^2 * \nu_t + \sum_{s \leq t} (1 - a_s) \hat{W}_s^2, \tag{1.2}$$

where

$$\hat{W} = \hat{W}(\omega, t) = \int_{\mathbb{R}} W(\omega, t, x) \nu(\omega, \{t\} \times dx) \tag{1.3}$$

and

$$a_t(\omega) = \nu(\omega, \{t\} \times \mathbb{R}). \tag{1.4}$$

<sup>1</sup> Let  $\tilde{\Omega} = \Omega \times \mathbb{R}_+ \times \mathbb{R}$ ,  $\tilde{\mathcal{P}} = \mathcal{P} \otimes \mathcal{B}$ , where  $\mathcal{B}$  is a Borel  $\sigma$ -field on  $\mathbb{R}$  and  $\mathcal{P}$  a  $\sigma$ -field generated by all left continuous adapted processes on  $\Omega \times \mathbb{R}_+$ . The predictable function is a  $\tilde{\mathcal{P}}$ -measurable function on  $\tilde{\Omega}$ .

It follows from Corollary 1.19 of Chapter 2 in Jacod and Shiryaev [5] that  $a_t = 0$  is equivalent to the fact  $X$  is a quasi-left continuous process.<sup>2</sup> Specially, for a process with independent increments,  $a_t = 0$  means this process has no fixed time of discontinuity.<sup>3</sup> Thus, we may and do choose a good version of both  $\hat{W}$  and  $a$  such that  $\hat{W}$  is the predictable projection of  $W(\omega, t, \Delta X_t)\mathbf{1}_{(\Delta X_t \neq 0)}$  and  $a_t \leq 1$ . In particular,

$$E(\mathbf{1}_{(\Delta X_t=0)} | \mathcal{F}_{t-}) = 1 - a_t. \tag{1.5}$$

Now consider another probability measure  $P'$  such that

$$P' \lll^{loc} P, \tag{1.6}$$

which means that for any  $t \geq 0$ ,  $P' |_{\mathcal{F}_t} \lll P |_{\mathcal{F}_t}$ . Define the likelihood ratio process

$$Z_t = \frac{dP' |_{\mathcal{F}_t}}{dP |_{\mathcal{F}_t}}. \tag{1.7}$$

It follows from Chapter III in Jacod and Shiryaev [5] that  $Z_t$  is a local martingale.

Since by assumption all  $P$ -martingales have a representation property relative to  $X$ , according to Theorem 5.19 of Chapter III in Jacod and Shiryaev [5],  $Z_t$  has the following representation: there is a predictable process  $\beta$  and a nonnegative predictable function  $Y$  on  $\tilde{\Omega}$  such that

$$Z_t = \begin{cases} Z_0 e^{N_t - \frac{1}{2}\beta^2 \cdot C_t} \prod_{s \leq t} (1 + \Delta N_s) e^{-\Delta N_s}, & (\omega, t) \in \Delta, \\ 0, & \text{otherwise.} \end{cases} \tag{1.8}$$

Here

$$N_t = \beta \cdot X_t^c + \left( Y - 1 + \frac{\hat{Y} - a}{1 - a} \mathbf{1}_{(a < 1)} \right) * (\mu - \nu)_t$$

and  $\Delta$  is a random set defined as follows

$$\begin{aligned} \sigma &= \inf\{t : \hat{Y}_t > 1 \text{ or } \hat{Y}_t < 1 \text{ and } a_t = 1\}, \\ H_t &= \frac{1}{2}\beta^2 \cdot C_t + (1 - \sqrt{Y})^2 \mathbf{1}_{[0, \sigma)} * \nu_t + \sum_{s \leq t} \left( \sqrt{1 - a_s} - \sqrt{1 - \hat{Y}_s} \right)^2 \mathbf{1}_{(s < \sigma)}, \\ T_n &= \inf\{t : H_t \geq n\}, \\ \Delta &= [0, \sigma) \cap (\cup_n [0, T_n]). \end{aligned}$$

Note that  $\beta$ ,  $Y$  and  $\Delta$  depend on  $P'$ . In fact,  $Y$  can be explicitly represented as follows. Let  $M_\mu^P$  be a measure on  $(\tilde{\Omega}, \tilde{\mathcal{P}})$  where  $\tilde{\Omega} := \Omega \times \mathbb{R}_+ \times \mathbb{R}$ ,  $\tilde{\mathcal{P}} := \mathcal{P} \otimes \mathcal{B}$ , such that

<sup>2</sup> A quasi-left continuous process  $X$  is a càdlàg adapted process such that for any increasing stopping times  $T_n$  with limit  $T$ ,  $\lim_{n \rightarrow \infty} X_{T_n} = X_T$ .

<sup>3</sup>  $t$  is called as the fixed time of discontinuity if  $P(\Delta X_t \neq 0) > 0$ .

$M_\mu^P(W) = E(W * \mu)_\infty$  for all measurable nonnegative functions  $W$ . Then  $Y$  is the conditional expectation of  $\frac{Z}{Z_-}$  with respect to  $\tilde{\mathcal{P}}$  under  $M_\mu^P$ , namely

$$Y = M_\mu^P \left( \frac{Z}{Z_-} \mid \tilde{\mathcal{P}} \right).$$

Define the log-likelihood process  $L_t$  by

$$L_t = \log Z_t. \tag{1.9}$$

This process has been a well-studied object in the context of both stochastic processes and statistical inferences. Obviously,

$$\begin{aligned} L_t = & \log Z_0 + \beta \cdot X_t^c - \frac{1}{2}\beta^2 \cdot C_t + \left( Y - 1 + \frac{\hat{Y} - a}{1 - a} \mathbf{1}_{(a < 1)} \right) * (\mu - \nu)_t \\ & + \sum_{s \leq t} (\log(1 + \Delta N_s) - \Delta N_s). \end{aligned}$$

Assume we are given a family of probability measures  $\mathbf{P}^n = \{P^{n,\psi} : \psi \in \Psi\}$  on  $(\Omega, \mathcal{F})$ , indexed by an arbitrary non-empty set  $\Psi$ , and assume

$$P^{n,\psi} \overset{\text{loc}}{\ll} P$$

for every  $n > 0$  and  $\psi \in \Psi$ . We shall be mainly interested in the sequence of likelihood ratio processes  $Z_t^{n,\psi}$ . The main purpose of the paper is to establish a certain Donsker theorem for log-likelihood processes  $\log Z_t^n$  in  $\ell^\infty(\Psi)$  as  $n \rightarrow \infty$ , where we denote by  $\ell^\infty(\Psi)$  the space of bounded real-valued functions defined on  $\Psi$ .

It seems hard to develop directly an invariance principle for  $\log Z^n$  due to complicated structure. To the best of our knowledge, there are only few works in this area, such as Le Cam [6], Vostrikova [13] and so on. The reader can find some interesting results in Nishiyama [8] and [9], where  $\log Z^n$  is assumed to be very special continuous semimartingales and discrete time semimartingales, respectively. It is a challenging problem to extend Nishiyama’s work to general setting. To attack such a problem, we shall combine stochastic calculus techniques and chaining arguments with the Kakutani–Hellinger distance for probability measures. In particular, we shall characterize the regularity of  $\ell^\infty(\Psi)$ -valued log-likelihood processes in terms of the Kakutani–Hellinger distance and the Hellinger processes.

The rest of the paper is organized as follows. We will first make some necessary assumptions and then state our main result in Sect. 2. The proof of main result is given in Sect. 3, which consists of several lemmas and two propositions.

## 2 Main Result

To state our main results, we need some more notations and make some technical assumptions. Start with the Kakutani–Hellinger distance between two probability measures  $P$  and  $P'$ . Assume that  $Q$  is a third probability measure on  $(\Omega, \mathcal{F})$  such that

$$P \lll Q, \quad P' \lll Q \tag{2.1}$$

Let

$$Z = \frac{dP}{dQ}, \quad Z' = \frac{dP'}{dQ} \tag{2.2}$$

and define the Kakutani–Hellinger distance by

$$\rho^2(P, P') = \frac{1}{2} \int_{\Omega} (\sqrt{Z} - \sqrt{Z'})^2 dQ. \tag{2.3}$$

It is easy to check that  $\rho(P, P')$  is a metric in the space of probability measures and does not depend on the probability measure  $Q$ . Note

$$\rho^2(P, P') = 1 - E_Q \sqrt{ZZ'} \tag{2.4}$$

For  $0 < \alpha < 1$ , call  $\check{H}(\alpha; P, P') = E_Q(Z^\alpha (Z')^{1-\alpha})$  the Hellinger integral of order  $\alpha$ . We remark that  $\check{H}(\alpha; P, P') \rightarrow 1$  as  $\alpha \rightarrow 0$  if  $P' \lll P$ .

Proceed to introduce the Hellinger processes. Assume that

$$P \lll^{loc} Q, \quad P' \lll^{loc} Q \tag{2.5}$$

and define

$$Z_t = \frac{dP|_{\mathcal{F}_t}}{dQ|_{\mathcal{F}_t}}, \quad Z'_t = \frac{dP'|_{\mathcal{F}_t}}{dQ|_{\mathcal{F}_t}}. \tag{2.6}$$

Then for each  $0 < \alpha < 1$ , there is a unique predictable increasing process  $h(\alpha; P, P')$ , called the Hellinger process of order  $\alpha$ , such that

(i)

$$h(\alpha; P, P')_0 = 0$$

(ii)

$$h(\alpha; P, P')_t = \mathbf{1}_{\cup[0, S_t]} \cdot h(\alpha; P, P')_t$$

(iii)

$Y(\alpha)_t + Y(\alpha)_- * \nu_t$  is local martingale

where

$$S_n = \inf\{t : Z_t > n \text{ or } Z'_t > n\}$$

and

$$Y(\alpha)_t = Z_t^\alpha (Z'_t)^{1-\alpha}.$$

One can extend the above Hellinger process to order zero and even to a general function. Given a function  $\psi : \mathbb{R} \mapsto \mathbb{R}$  such that

$$\frac{\psi(x)}{|x-1|^2 \wedge |x-1|} \quad (2.7)$$

is bounded with convention  $\frac{0}{0} = 0$  and  $\psi(1) = 0$ . Denote

$$J(\psi; P, P')_t = \sum_{s \leq t} \frac{Z'_s}{Z'_{s-}} \psi\left(\frac{Z_s/Z_{s-}}{Z'_s/Z'_{s-}}\right), \quad (2.8)$$

then there is a predictable increasing process, denoted by  $\iota(\psi; P, P')$ , such that

(i')

$$\iota(\psi; P, P')_0 = 0$$

(ii')

$$\iota(\psi; P, P')_t = \mathbf{1}_{\cup[0, T_n]} \cdot \iota(\psi; P, P')_t$$

(iii')

$J(\psi; P, P')_{T_n \wedge t} - \iota(\psi; P, P')_{T_n \wedge t}$  is local martingale .

Call  $\iota(\psi; P, P')$  the Hellinger process of order 0 associated with  $\psi$ . In particular, if

$$\psi(x) = \begin{cases} 1, & x = 0, \\ 0, & x > 0, \end{cases}$$

then we simply call  $\iota(\psi; P, P')$  the Hellinger process of order 0.

In general, it is rather complicated to compute  $h(\alpha; P, P')$ . However, we fortunately have the following explicit formula in the special case  $P' \ll P$ :

$$h(\alpha; P, P') = \frac{\alpha(1-\alpha)}{2} \beta^2 \cdot C + \varphi_\alpha(1, Y) * \nu + \sum_{s \leq t} \varphi_\alpha(1 - a_s, 1 - \hat{Y}_s)$$

In particular,

$$h\left(\frac{1}{2}; P, P'\right) = \frac{1}{8} \beta^2 \cdot C + \frac{1}{2} (1 - \sqrt{Y})^2 * \nu + \frac{1}{2} \sum_{s \leq t} \left( \sqrt{1 - a_s} - \sqrt{1 - \hat{Y}_s} \right)^2,$$

$$l(\psi; P, P') = Y \psi\left(\frac{1}{Y}\right) * \nu + \sum_{s \leq t} (1 - a'_s) \psi\left(\frac{1 - a_s}{1 - a'_s}\right).$$

Our technical assumptions mainly involve three aspects: the predictable envelope of  $\{Y^{n,\psi}, \psi \in \Psi\}$ , the Kakutani–Hellinger distance between probability measures  $P^{n,\psi}$  and the size of index set  $\Psi$ .

For every  $n > 0$ , denote the essence supremum  $\bar{Y}^n(\Psi) = [\sup_{\psi \in \Psi} Y^{n,\psi}]_{\bar{\mathcal{P}}, M_v^P}$ . This is the predictable envelope of  $\{Y^{n,\psi}, \psi \in \Psi\}$  used in Definitions 2.1 and 2.3 of Nishiyama [9].

**Assumption 1** For any  $n > 0, \psi \in \Psi, \Delta^{n,\psi} \equiv \Omega \times [0, 1]$  and  $0 \leq a < 1$ . Moreover,  $\{Y^{n,\psi}, \psi \in \Psi\}$  attains their predictable envelope for every  $n > 0$ , namely, there is a  $\psi_0 \in \Psi$  such that  $Y^{n,\psi_0} = [\sup_{\psi \in \Psi} Y^{n,\psi}]_{\bar{\mathcal{P}}, M_v^P}$ .

**Assumption 2** For every  $\varepsilon > 0$ , as  $n \rightarrow \infty$

$$\sup_{\psi \in \Psi} l(f_{1+\varepsilon}, P, P^{n,\psi})_t \xrightarrow{P} 0 \tag{2.9}$$

where  $f_{1+\varepsilon}(x) = |x - 1| 1_{\{1/(1+\varepsilon) < x < 1+\varepsilon\}}$ .

There is a nonnegative definite continuous function  $V_t$  on  $\Psi \times \Psi$ , such that as  $n \rightarrow \infty$ ,

$$\sup_{\psi \in \Psi} \left| h\left(\frac{1}{2}; P, P^{n,\psi}\right)_t - \frac{1}{8} V_t^{\psi,\psi} \right| \xrightarrow{P} 0 \tag{2.10}$$

and for every  $\psi, \phi \in \Psi$ ,

$$h\left(\frac{1}{2}; P^{n,\psi}, P^{n,\phi}\right)_t \xrightarrow{P} \frac{1}{8} V_t^{\psi,\phi}. \tag{2.11}$$

Let  $\Psi$  be an arbitrary set,  $\Delta_\Pi$  a positive rational number.  $\Pi = \{\Pi(\varepsilon)\}_{\varepsilon \in (0, \Delta_\Pi]}$  is called a decreasing series of finite partitions (DFP) of  $\Psi$  if

(i) each  $\Pi(\varepsilon) = \{\Psi(\varepsilon; k) : 1 \leq k \leq N_\Pi(\varepsilon)\}$  is a finite partition of  $\Psi$ , namely

$$\Psi = \bigcup_{k=1}^{N_\Pi(\varepsilon)} \Psi(\varepsilon; k);$$

- (ii)  $N_\Pi(\Delta_\Pi) = 1$  and  $\lim_{\varepsilon \downarrow 0} N_\Pi(\varepsilon) = \infty$ ;
- (iii)  $N_\Pi(\varepsilon) \geq N_\Pi(\varepsilon')$  as  $\varepsilon \leq \varepsilon'$ .

Given a  $0 < \varepsilon \leq \Delta_\Pi$ , the  $\varepsilon$ -entropy  $H_\Pi(\varepsilon)$  is defined by

$$H_\Pi(\varepsilon) = \sqrt{\log(1 + N_\Pi(\varepsilon))}.$$

**Assumption 3** There exists a decreasing series of finite partitions,  $\Pi$ , of  $\Psi$  such that as  $n \rightarrow \infty$

$$\int_0^{\Delta_\Pi} H_\Pi(\varepsilon) d\varepsilon < \infty$$

and

$$\|\mathfrak{H}^n\|_\Pi = O_P(1)$$

where

$$\|\mathfrak{H}^n\|_\Pi^2 = \sup_{\varepsilon \in (0, \Delta_\Pi] \cap \mathbb{Q}} \max_{1 \leq k \leq N_\Pi(\varepsilon)} \max_{\psi, \phi \in \Psi(\varepsilon, k)} \frac{1}{\varepsilon^2} h\left(\frac{1}{2}; P^{n, \psi}, P^{n, \phi}\right)_1.$$

We are now ready to state our main result as follows.

**Theorem 2.1** Under Assumptions 1, 2 and 3, we have

$$L_1^n \Rightarrow G \quad \text{in } \ell^\infty(\Psi), \tag{2.12}$$

where  $G$  stands for a Gaussian element in  $\ell^\infty(\Psi)$ , each  $d$ -dimensional marginal  $(G^{\psi_1}, \dots, G^{\psi_d})$  is a normal random vector with mean

$$\vec{\mu} = -\frac{1}{2}(V_1^{\psi_i, \psi_i}, 1 \leq i \leq d) \tag{2.13}$$

and covariance structure

$$\Sigma = (V_1^{\psi_i, \psi_j})_{1 \leq i, j \leq d}. \tag{2.14}$$

The proof is given in Sect. 2. For the sake of comparison, we review an earlier result due to Nishiyama [9] in the discrete time case. Let  $(\mathcal{F}_i)_{i \geq 0}$  be a discrete time stochastic



basis, and  $\mathbf{P}^n = \{P^{n,\psi} : \psi \in \Psi\}$  a family of probability measures on  $(\Omega, \mathcal{F})$ , such that

$$P^{n,\psi} \stackrel{\text{loc}}{\ll} P. \tag{2.15}$$

Define

$$W_i^{n,\psi} = \frac{dP^{n,\psi}|_{\mathcal{F}_i}}{dP|_{\mathcal{F}_i}} \tag{2.16}$$

and

$$\xi_i^{n,\psi} = \sqrt{\frac{W_i^{n,\psi}}{W_{i-1}^{n,\psi}} - 1} \tag{2.17}$$

Nishiyama [9] studied weak convergence for log-likelihood processes  $\log W_n^n$  in  $\ell^\infty(\Psi)$  and obtained a similar result to (2.12) under some integrability assumptions involving  $\xi^n$ 's and entropy conditions. More specifically, assume

(i) for every  $\varepsilon > 0$

$$\sum_{i=1}^n E\left(\left(\sup_{\psi \in \Psi} \xi_i^{n,\psi}\right)^2 1_{\{\sup_{\psi \in \Psi} \xi_i^{n,\psi} > \varepsilon\}} \middle| \mathcal{F}_{i-1}\right) \xrightarrow{P} 0;$$

(ii) there exists a decreasing series of finite partitions,  $\Pi$ , of  $\Psi$  such that

$$\sup_{\varepsilon \in (0, \Delta_\Pi] \cap \mathbb{Q}} \max_{1 \leq k \leq N_\Pi(\varepsilon)} \frac{1}{\varepsilon^2} \sum_{i=1}^n \sup_{\psi, \phi \in \Psi(\varepsilon, k)} E(|\xi_i^{n,\psi} - \xi_i^{n,\phi}|^2 | \mathcal{F}_{i-1}) = O_P(1) \tag{2.18}$$

and

$$\int_0^{\Delta_\Pi} H_\Pi(\varepsilon) d\varepsilon < \infty; \tag{2.19}$$

(iii) there is a  $V : \Psi \times \Psi \rightarrow R$  such that

$$\sup_{\psi \in \Psi} \left| \sum_{i=1}^n 4E^*((\xi_i^{n,\psi})^2 | \mathcal{F}_{i-1}) - V^{\psi,\psi} \right| \xrightarrow{P} 0;$$

and for  $\psi, \phi \in \Psi$

$$\sum_{i=1}^n 4E\left(\xi_i^{n,\psi} \xi_i^{n,\phi} | \mathcal{F}_{i-1}\right) \xrightarrow{P} V^{\psi,\phi}.$$

Then

$$\log W_n^n \Rightarrow G \quad \text{in } \ell^\infty(\Psi), \tag{2.20}$$

where  $G$  stands for a Gaussian element in  $\ell^\infty(\Psi)$ , each  $d$ -dimensional marginal  $(G^{\psi_1}, \dots, G^{\psi_d})$  is a normal random vector with mean

$$\vec{\mu} = -\frac{1}{2}(V^{\psi_i, \psi_i}, 1 \leq i \leq d)$$

and covariance structure

$$\Sigma = (V^{\psi_i, \psi_j})_{1 \leq i, j \leq d}.$$

To conclude the Introduction, two more remarks are given .

**Remark 2.2** Observe in the discrete time case the Hellinger process can be computed as follows.

$$h\left(\frac{1}{2}; P, P^{n, \psi}\right) = 2 \sum_{i=1}^n E\left[(\xi_i^{n, \psi})^2 \middle| \mathcal{F}_{i-1}\right] \tag{2.21}$$

and

$${}_t(f_{1+\varepsilon}, P, P^{n, \psi}) = \sum_{i=1}^n E\left[(\xi_i^{n, \psi})^2 1_{\{\frac{1}{1+\varepsilon} \leq (\xi_i^{n, \psi} + 1)^2 < 1+\varepsilon\}} \middle| \mathcal{F}_{i-1}\right] \tag{2.22}$$

Thus, there is to some extent a similarity between our assumptions in Theorem 2.1 and Nishiyama’s assumptions. However, it seems neater to use the Hellinger processes in continuous time case.

The integrability condition (Assumption 3) of partitioning entropy plays an important role in the proof of Theorem 2.1. It is possible to use the metric entropy condition, but we need to introduce a suitable pseudo-metric in the index set  $\Psi$ . The Hellinger processes would also be very likely a good candidate.

**Remark 2.3** It is rather interesting to consider the limiting behavior of the process  $\log Z^n$  in  $\ell^\infty([0, 1] \times \Psi)$ . To this end, we need to establish a tightness criterion in the space  $[0, 1] \times \Psi$ . This is more complicate, and will be left to the future work.

### 3 Proofs

Let us start with a decomposition. Observe that

$$N_t^{n, \psi} = \beta^{n, \psi} \cdot X_t^c + \left(Y^{n, \psi} - 1 + \frac{\hat{Y}^{n, \psi} - a}{1 - a}\right) * (\mu - \nu)_t.$$

It is easy to see

$$\Delta N_t^{n,\psi} = (Y^{n,\psi}(t, \Delta X_t) - 1)\mathbf{1}_{(\Delta X_t \neq 0)} - \frac{\hat{Y}_t^{n,\psi} - a_t}{1 - a_t} \mathbf{1}_{(\Delta X_t = 0)},$$

and so we have

$$\begin{aligned} \Delta L_t^{n,\psi} &= \log(1 + \Delta N_t^{n,\psi}) \\ &= \log(Y^{n,\psi}(t, \Delta X_t)) \mathbf{1}_{(\Delta X_t \neq 0)} + \log\left(1 - \frac{\hat{Y}_t^{n,\psi} - a_t}{1 - a_t}\right) \mathbf{1}_{(\Delta X_t = 0)}. \end{aligned}$$

Let  $\mu^{n,\psi}$  be the jump measure of  $L^{n,\psi}$  defined by

$$\mu_t^{n,\psi} = \sum_{s \leq t} \mathbf{1}_{(\Delta L_s^{n,\psi} \neq 0)} \mathcal{E}_{(s, \Delta L_s^{n,\psi})}$$

and  $\nu^{n,\psi}$  the corresponding predictable compensator. Then for any predictable function  $W(\omega, t, x)$ ,

$$\begin{aligned} W * \mu^{n,\psi} &= \sum_{s \leq t} W(\log Y^{n,\psi}(t, \Delta X_t)) \mathbf{1}_{(\Delta X_t \neq 0)} \\ &\quad + \sum_{s \leq t} W\left(\log\left(1 - \frac{\hat{Y}_s^{n,\psi} - a_s}{1 - a_s}\right)\right) \mathbf{1}_{(\Delta X_t = 0)} \\ &= W(\log Y^{n,\psi}) * \mu + \sum_{s \leq t} W\left(\log\left(1 - \frac{\hat{Y}_s^{n,\psi} - a_s}{1 - a_s}\right)\right) \mathbf{1}_{(\Delta X_s = 0)}, \end{aligned}$$

and so by the fact that  $1 - a_s$  is the predictable projection of  $\mathbf{1}_{(\Delta X_s = 0)}$ ,

$$W * \nu^{n,\psi} = W(\log Y^{n,\psi}) * \nu + \sum_{s \leq t} W\left(\log\left(1 - \frac{\hat{Y}_s^{n,\psi} - a_s}{1 - a_s}\right)\right) (1 - a_s). \tag{3.1}$$

Given a positive number  $\tau$ , consider the truncation function

$$h_\tau(x) = x \mathbf{1}_{(|x| \leq \tau)} \tag{3.2}$$

and define

$$\check{L}_t^{n,\tau,\psi} = \sum_{s \leq t} (\Delta L_s^{n,\psi} - h_\tau(\Delta L_s^{n,\psi})) \tag{3.3}$$

Thus, combined together, we easily have a canonical decomposition

$$L_t^{n,\psi} = \check{L}_t^{n,\tau,\psi} + \beta^{n,\psi} \cdot X_t^c - \frac{1}{2}(\beta^{n,\psi})^2 \cdot C_t + h_\tau * (\mu^{n,\psi} - \nu^{n,\psi})_t$$

$$\begin{aligned}
 &+ \sum_{s \leq t} \frac{\hat{Y}_s^{n,\psi} - a_s}{1 - a_s} - \left( Y^{n,\psi} - 1 + \frac{\hat{Y}^{n,\psi} - a}{1 - a} \right) * \nu_t \\
 &+ h_\tau (\log Y^{n,\psi}) * \nu_t + \sum_{s \leq t} \log \left( 1 - \frac{\hat{Y}_s^{n,\psi} - a_s}{1 - a_s} \right) (1 - a_s).
 \end{aligned}$$

For simplicity of writing, let

$$\begin{aligned}
 A_{1,t}^{n,\psi} &= \beta^{n,\psi} \cdot X_t^c - \frac{1}{2} (\beta^{n,\psi})^2 \cdot C_t, \\
 A_{2,t}^{n,\psi} &= h_\tau * (\mu^{n,\psi} - \nu^{n,\psi})_t, \\
 A_{3,t}^{n,\psi} &= \sum_{s \leq t} \frac{\hat{Y}_s^{n,\psi} - a_s}{1 - a_s} - \left( \hat{Y}^{n,\psi} - 1 + \frac{\hat{Y}^{n,\psi} - a}{1 - a} \right) * \nu_t \\
 &+ h_\tau (\log Y^{n,\psi}) * \nu_t + \sum_{s \leq t} \log \left( 1 - \frac{\hat{Y}_s^{n,\psi} - a_s}{1 - a_s} \right) (1 - a_s), \tag{3.4}
 \end{aligned}$$

Thus, we have

$$L_t^{n,\psi} = \check{L}_t^{n,\tau,\psi} + A_{1,t}^{n,\psi} + A_{2,t}^{n,\psi} + A_{3,t}^{n,\psi}$$

The proof of Theorem 2.1 will consist of a series of lemmas and propositions.

**Lemma 3.1** *Under Assumptions 1, 2 and 3, we have for each  $\psi \in \Psi$  and  $\tau > 0$ , as  $n \rightarrow \infty$*

$$\check{L}_t^{n,\tau,\psi} \xrightarrow{P} 0 \tag{3.5}$$

**Proof** Set

$$\Upsilon_s^{n,\psi} = \frac{1 - \hat{Y}_s^{n,\psi}}{1 - a_s}. \tag{3.6}$$

Then

$$\check{L}_t^{n,\tau,\psi} = \log(Y^{n,\psi}) \mathbf{1}_{(|\log(Y^{n,\psi})| > \tau)} * \mu_t + \sum_{s \leq t} \log(\Upsilon_s^{n,\psi}) \mathbf{1}_{(|\log(\Upsilon_s^{n,\psi})| > \tau)} \mathbf{1}_{(\Delta X_s = 0)}.$$

Also, for any  $\delta > 0$

$$\begin{aligned}
 &|\log(Y^{n,\psi})| \mathbf{1}_{(|\log(Y^{n,\psi})| > \tau)} * \nu_t \\
 &\leq |\log(Y^{n,\psi})| \mathbf{1}_{(|\log(Y^{n,\psi})| > \tau)} \mathbf{1}_{(|Y^{n,\psi} - 1| \leq \delta)} * \mu_t \\
 &\quad + |\log(Y^{n,\psi})| \mathbf{1}_{(|\log(Y^{n,\psi})| > \tau)} \mathbf{1}_{(|Y^{n,\psi} - 1| > \delta)} * \mu_t.
 \end{aligned}$$

By the Lenglart domination property (see page 35 of Jacod and Shiryaev [5]),

$$\begin{aligned}
 &P(|\log(Y^{n,\psi})| \mathbf{1}_{(|\log(Y^{n,\psi})|>\tau)} \mathbf{1}_{(|Y^{n,\psi}-1|\leq\delta)} * \mu_t > \varepsilon) \\
 &\leq \frac{\eta}{\varepsilon} + P(|\log(Y^{n,\psi})| \mathbf{1}_{(|\log(Y^{n,\psi})|>\tau)} \mathbf{1}_{(|Y^{n,\psi}-1|\leq\delta)} * \nu_t > \eta).
 \end{aligned}$$

Note for  $\delta < 1$  there is a positive constant  $c_\delta$  such that for any  $x > 0$

$$|\log x| \mathbf{1}_{(|x-1|\leq\delta)} \leq c_\delta |x - 1| \mathbf{1}_{(|x-1|\leq\delta)}, \tag{3.7}$$

so

$$\begin{aligned}
 &|\log(Y^{n,\psi})| \mathbf{1}_{(|\log(Y^{n,\psi})|>\tau)} \mathbf{1}_{(|Y^{n,\psi}-1|\leq\delta)} * \nu_t \\
 &\leq c_\delta |Y^{n,\psi} - 1| \mathbf{1}_{(e^{-\tau} < Y^{n,\psi} < e^\tau)^c} \mathbf{1}_{(|Y^{n,\psi}-1|\leq\delta)} * \nu_t \xrightarrow{P} 0.
 \end{aligned}$$

On the other hand, for each  $\varepsilon \in (0, 1)$

$$P(|\log(Y^{n,\psi})| \mathbf{1}_{(|\log(Y^{n,\psi})|>\tau)} \mathbf{1}_{(|Y^{n,\psi}-1|>\delta)} * \mu_t > \varepsilon) \leq P(\mathbf{1}_{(|Y^{n,\psi}-1|>\delta)} * \mu_t > \varepsilon).$$

Again, by the Lenglart domination property, it follows for any  $\eta > 0$

$$\begin{aligned}
 P(\mathbf{1}_{(|Y^{n,\psi}-1|>\delta)} * \mu_t > \varepsilon) &\leq \frac{\eta}{\varepsilon} + P(\mathbf{1}_{(|Y^{n,\psi}-1|>\delta)} * \nu_t > \eta) \\
 &\leq \frac{\eta}{\varepsilon} + P(|Y^{n,\psi} - 1| \mathbf{1}_{(|Y^{n,\psi}-1|>\delta)} * \nu_t > \delta\eta).
 \end{aligned}$$

Letting  $n \rightarrow \infty$  and then  $\eta \rightarrow 0$ , we have

$$P(\mathbf{1}_{(|Y^{n,\psi}-1|>\delta)} * \mu_t > \varepsilon) \rightarrow 0.$$

In combination, we have proved the desired statement. □

**Lemma 3.2** *Under Assumptions 1, 2 and 3, we have for each  $\psi \in \Psi$ , as  $n \rightarrow \infty$*

$$(Y^{n,\psi} - 1) * \nu_t \xrightarrow{P} 0, \tag{3.8}$$

$$(1 - \sqrt{Y^{n,\psi}})^2 * \nu_t \xrightarrow{P} 0, \tag{3.9}$$

$$\sum_{s \leq t} \left( \sqrt{1 - a_s} - \sqrt{1 - \hat{Y}_s^{n,\psi}} \right)^2 \xrightarrow{P} 0. \tag{3.10}$$

Consequently,

$$\frac{1}{8} (\beta^{n,\psi})^2 \cdot C_t \xrightarrow{P} V_t^\psi. \tag{3.11}$$

**Proof** Obviously, for any  $\varepsilon > 0$

$$|Y^{n,\psi} - 1| \leq \varepsilon + |Y^{n,\psi} - 1| \mathbf{1}_{(|Y^{n,\psi} - 1| > \varepsilon)}.$$

Also, by Assumption 2

$$\begin{aligned} |Y^{n,\psi} - 1| \mathbf{1}_{(|Y^{n,\psi} - 1| > \varepsilon)} * \nu_t &\leq \iota(h_{1+\varepsilon}; P, P^{n,\psi}) \\ &\xrightarrow{P} 0, \quad n \rightarrow \infty. \end{aligned}$$

The desired (3.8) holds.

Observe an elementary inequality: for any  $0 < \varepsilon < 1$ , there is a positive constant  $c_\varepsilon$  such that

$$(\sqrt{1+x} - 1)^2 \leq \begin{cases} |x|^2, & |x| \leq \varepsilon < 1, \\ c_\varepsilon |x|, & |x| > \varepsilon. \end{cases}$$

Then it follows

$$\begin{aligned} (\sqrt{Y^{n,\psi}} - 1)^2 * \nu_t &\leq |Y^{n,\psi} - 1|^2 \mathbf{1}_{(|Y^{n,\psi} - 1| \leq \varepsilon)} * \nu_t \\ &\quad + c_\varepsilon |Y^{n,\psi} - 1| \mathbf{1}_{(|Y^{n,\psi} - 1| > \varepsilon)} * \nu_t, \\ |Y^{n,\psi} - 1|^2 \mathbf{1}_{(|Y^{n,\psi} - 1| \leq \varepsilon)} * \nu_t &\leq \varepsilon^2 \nu([0, t] \times \mathbb{R}), \\ |Y^{n,\psi} - 1| \mathbf{1}_{(|Y^{n,\psi} - 1| > \varepsilon)} * \nu_t &\leq \iota(h_{1+\varepsilon}; P, P^{n,\psi}). \end{aligned}$$

Thus, under Assumption 2, we have by letting  $n \rightarrow \infty$  and then  $\varepsilon \rightarrow 0$

$$(\sqrt{Y^{n,\psi}} - 1)^2 * \nu_t \xrightarrow{P} 0.$$

For (3.10), note

$$\begin{aligned} \sum_{s \leq t} \left( \sqrt{1 - a_s} - \sqrt{1 - \hat{Y}_s^{n,\psi}} \right)^2 &= \sum_{s \leq t} (1 - a_s) \left( 1 - \sqrt{\frac{1 - \hat{Y}_s^{n,\psi}}{1 - a_s}} \right)^2 \\ &= \sum_{s \leq t} (1 - a_s) \left( 1 - \sqrt{\Upsilon_s^{n,\psi}} \right)^2, \end{aligned}$$

where  $\Upsilon_s^{n,\psi}$  is as in (3.6). Then it easily follows

$$\begin{aligned} \left(1 - \sqrt{\Upsilon_s^{n,\psi}}\right)^2 &\leq |\Upsilon_s^\psi - 1|^2 \mathbf{1}_{(|\Upsilon_s^{n,\psi} - 1| \leq \varepsilon)} + c_\varepsilon |\Upsilon_s^{n,\psi} - 1| \mathbf{1}_{(|\Upsilon_s^{n,\psi} - 1| > \varepsilon)}, \\ \sum_{s \leq t} (1 - a_s) |\Upsilon_s^{n,\psi} - 1|^2 \mathbf{1}_{(|\Upsilon_s^{n,\psi} - 1| \leq \varepsilon)} &\leq \varepsilon^2 \sum_{s \leq t} (1 - a_s), \\ \sum_{s \leq t} (1 - a_s) |\Upsilon_s^{n,\psi} - 1| \mathbf{1}_{(|\Upsilon_s^{n,\psi} - 1| > \varepsilon)} &\leq \iota(h_{1+\varepsilon}; P, P^{n,\psi}). \end{aligned}$$

Thus, under Assumption 2, we have by first letting  $n \rightarrow \infty$  and then  $\varepsilon \rightarrow 0$

$$\sum_{s \leq t} \left(\sqrt{1 - a_s} - \sqrt{1 - \hat{Y}_s^{n,\psi}}\right)^2 \xrightarrow{P} 0.$$

The proof is now complete. □

**Lemma 3.3** *Under Assumptions 1, 2 and 3, we have for each  $\psi \in \Psi$  and  $\tau > 0$ , as  $n \rightarrow \infty$*

$$h_\tau * (\mu^{n,\psi} - \nu^{n,\psi})_t \xrightarrow{P} 0.$$

**Proof** First, observe the quadratic variation of  $h_\tau * (\mu^{n,\psi} - \nu^{n,\psi})_t$  is given by

$$\begin{aligned} \langle h_\tau * (\mu^{n,\psi} - \nu^{n,\psi}) \rangle_t &= (h_\tau(\log Y^{n,\psi}))^2 * \nu_t \\ &\quad + \sum_{s \leq t} \left( h_\tau \left( \log \left( 1 - \frac{\hat{Y}_s^{n,\psi} - a_s}{1 - a_s} \right) \right) \right)^2 (1 - a_s). \end{aligned}$$

We shall prove that  $\langle h_\tau * (\mu^{n,\psi} - \nu^{n,\psi}) \rangle_t$  converges in probability to zero below. Note there is a  $\delta > 0$  such that for any  $\varepsilon < \delta$

$$|(h_\tau(\log(1 + x)))^2 - 4(1 - \sqrt{1 + x})^2| \leq \begin{cases} |x|^3, & |x| \leq \varepsilon < \delta, \\ c_\varepsilon |x|, & |x| > \varepsilon \end{cases}$$

where  $0 < c_\varepsilon < \infty$ . Thus, it follows for any  $\varepsilon < \delta$

$$\begin{aligned} (h_\tau(\log Y^{n,\psi}))^2 * \nu_t &\leq |Y^\psi - 1|^3 \mathbf{1}_{(|Y^\psi - 1| \leq \varepsilon)} * \nu_t \\ &\quad + c_\varepsilon |Y^{n,\psi} - 1| \mathbf{1}_{(|Y^{n,\psi} - 1| > \varepsilon)} * \nu_t. \end{aligned}$$

Hence, letting  $n \rightarrow \infty$  and then  $\varepsilon \rightarrow 0$  immediately yields

$$(h_\tau(\log Y^{n,\psi}))^2 * \nu_t \xrightarrow{P} 0.$$

A similar argument shows

$$\sum_{s \leq t} \left( h_\tau \left( \log \left( 1 - \frac{\hat{Y}_s^{n,\psi} - a_s}{1 - a_s} \right) \right) \right)^2 (1 - a_s) \xrightarrow{P} 0.$$

Combined together, we have the desired statement. □

**Lemma 3.4** *Under Assumptions 1, 2 and 3, we have for each  $\psi \in \Psi$  and  $\tau > 0$ , as  $n \rightarrow \infty$*

$$A_{3,t}^{n,\psi} \xrightarrow{P} 0.$$

**Proof** Note there is a  $\delta > 0$  such that for any  $\varepsilon < \delta$

$$\left| h_\tau(\log(1+x)) + 2(1 - \sqrt{1+x})^2 - x \right| \leq \begin{cases} |x|^3, & |x| \leq \varepsilon < \delta, \\ c_\varepsilon|x|, & |x| > \varepsilon, \end{cases}$$

where  $0 < c_\varepsilon < \infty$ . Thus, it follows for any  $\varepsilon < \delta$

$$\begin{aligned} & \left| h_\tau(\log Y^{n,\psi}) * v_t + 2(1 - \sqrt{Y^{n,\psi}})^2 * v_t - \left( \hat{Y}^{n,\psi} - 1 + \frac{\hat{Y}^{n,\psi} - a}{1-a} \right) * v_t \right| \\ & \leq |Y^{n,\psi} - 1|^3 \mathbf{1}_{(|Y^{n,\psi} - 1| \leq \varepsilon)} * v_t + c_\varepsilon |Y^{n,\psi} - 1| \mathbf{1}_{(|Y^{n,\psi} - 1| > \varepsilon)} * v_t. \end{aligned}$$

This implies

$$\left| h_\tau(\log Y^{n,\psi}) * v_t + 2(1 - \sqrt{Y^{n,\psi}})^2 * v_t - \left( \hat{Y}^{n,\psi} - 1 + \frac{\hat{Y}^{n,\psi} - a}{1-a} \right) * v_t \right| \xrightarrow{P} 0.$$

Since it was proved  $2(1 - \sqrt{Y^{n,\psi}})^2 * v_t \xrightarrow{P} 0$ , then we have

$$\left| h_\tau(\log Y^{n,\psi}) * v_t - \left( \hat{Y}^{n,\psi} - 1 + \frac{\hat{Y}^{n,\psi} - a}{1-a} \right) * v_t \right| \xrightarrow{P} 0.$$

Similarly, we have

$$\left| \sum_{s \leq t} \frac{\hat{Y}_s^{n,\psi} - a_s}{1 - a_s} + \sum_{s \leq t} h_\tau \left( \log \left( 1 - \frac{\hat{Y}_s^{n,\psi} - a_s}{1 - a_s} \right) \right) (1 - a_s) \right| \xrightarrow{P} 0.$$

Combined together, the proof is complete. □

**Proposition 3.5** *Under Assumptions 1, 2 and 3, every finite-dimensional marginal of  $L^n$  converges weakly.*

**Proof** From Lemmas 3.1–3.4,

$$\beta_s^{n,\psi} \cdot X_t^c - \frac{1}{2}(\beta_s^{n,\psi})^2 \cdot C_t$$

has non-degenerate limiting finite-dimensional marginal laws, and the other part of  $L^n$  asymptotically vanishes.



For every  $\psi$ , the process

$$\beta_s^{n,\psi} \cdot X_t^c - \frac{1}{2}(\beta_s^{n,\psi})^2 \cdot C_t$$

is a continuous semimartingale. Its predictable characteristics  $(\hat{B}^\psi, \hat{C}^\psi, 0)$  are

$$\begin{aligned} \hat{B}_t^{n,\psi} &= -\frac{1}{2}(\beta_t^{n,\psi})^2 \cdot C_t, \\ \hat{C}_t^{n,\psi} &= (\beta_t^{n,\psi})^2 \cdot C_t. \end{aligned}$$

By Lemma 3.2 and Assumption 2, there is a non-decreasing continuous function  $V$ , such that  $V_0 = 0$ ,

$$\sup_{\psi \in \Psi} |(\beta^{n,\psi})^2 \cdot C_t - V_t^{\psi,\psi}| \xrightarrow{P^*} 0, \tag{3.12}$$

$$(\beta^{n,\psi} \beta^{n,\psi}) \cdot C_t \xrightarrow{P} V_t^{\psi,\phi} \tag{3.13}$$

for every  $\psi, \phi \in \Psi$ .

The proposition is now concluded by Theorem VIII.3.6 of Jacod and Shiryaev [5]. □

Next we turn to verifying uniform tightness.

**Lemma 3.6** *Under Assumptions 1, 2 and 3, we have for each  $\tau > 0$ , as  $n \rightarrow \infty$*

$$\sup_{\psi \in \Psi} |\check{L}_t^{n,\tau,\psi}| \xrightarrow{P} 0. \tag{3.14}$$

**Proof** Recall

$$\begin{aligned} \check{L}_t^{n,\tau,\psi} &= \sum_{s \leq t} \log(Y^{n,\psi}(s, \Delta X_s)) \mathbf{1}\left(\left|\log Y^{n,\psi}(s, \Delta X_s)\right| > \tau\right) \mathbf{1}_{(\Delta X_s \neq 0)} \\ &+ \sum_{s \leq t} \log\left(1 - \frac{\hat{Y}_s^{n,\psi} - a_s}{1 - a_s}\right) \mathbf{1}\left(\left|\log\left(1 - \frac{\hat{Y}_s^{n,\psi} - a_s}{1 - a_s}\right)\right| > \tau\right) \mathbf{1}_{(\Delta X_s = 0)}. \end{aligned} \tag{3.15}$$

Let us prove

$$\sup_{\psi \in \Psi} \sum_{s \leq t} \left| \log(Y^{n,\psi}(s, \Delta X_s)) \right| \mathbf{1}\left(\left|\log Y^{n,\psi}(s, \Delta X_s)\right| > \tau\right) \mathbf{1}_{(\Delta X_s \neq 0)} \xrightarrow{P} 0 \tag{3.16}$$

and

$$\sup_{\psi \in \Psi} \sum_{s \leq t} \log\left(1 - \frac{\hat{Y}_s^{n,\psi} - a_s}{1 - a_s}\right) \mathbf{1}\left(\left|\log\left(1 - \frac{\hat{Y}_s^{n,\psi} - a_s}{1 - a_s}\right)\right| > \tau\right) \mathbf{1}_{(\Delta X_s = 0)} \xrightarrow{P} 0. \tag{3.17}$$

For (3.16), note

$$|\log(Y^{n,\psi}(s, \Delta X_s))| \mathbf{1}\left(\left|\log Y^{n,\psi}(s, \Delta X_s)\right| > \tau\right) = \log(Y^{n,\psi}(s, \Delta X_s)) \mathbf{1}\left(\log Y^{n,\psi}(s, \Delta X_s) > \tau\right) - \log(Y^{n,\psi}(s, \Delta X_s)) \mathbf{1}\left(\log Y^{n,\psi}(s, \Delta X_s) < -\tau\right).$$

Thus, we need only to prove

$$\sum_{s \leq t} \sup_{\psi \in \Psi} \log(Y^{n,\psi}(s, \Delta X_s)) \mathbf{1}_{(\log Y^{n,\psi}(s, \Delta X_s) > \tau)} \mathbf{1}_{(\Delta X_s \neq 0)} \xrightarrow{P} 0 \tag{3.18}$$

and

$$\sum_{s \leq t} \inf_{\psi \in \Psi} \log(Y^{n,\psi}(s, \Delta X_s)) \mathbf{1}_{(\log Y^{n,\psi}(s, \Delta X_s) < -\tau)} \mathbf{1}_{(\Delta X_s \neq 0)} \xrightarrow{P} 0. \tag{3.19}$$

Let us first look at (3.18). Set

$$\bar{Y}^{n,\Psi}(\omega; s, x) = \left[ \sup_{\psi \in \Psi} Y^{n,\psi}(\omega; s, x) \right]_{\tilde{\mathcal{P}}, M_\mu^P}. \tag{3.20}$$

Then

$$\begin{aligned} & \sum_{s \leq t} \sup_{\psi \in \Psi} \log(Y^{n,\psi}(s, \Delta X_s)) \mathbf{1}_{(\log Y^{n,\psi}(s, \Delta X_s) > \tau)} \mathbf{1}_{(\Delta X_s \neq 0)} \\ & \leq \sum_{s \leq t} \log(\bar{Y}^{n,\Psi}(\omega; s, \Delta X_s)) \mathbf{1}_{(\log \bar{Y}^{n,\Psi}(\omega; s, \Delta X_s) > \tau)} \mathbf{1}_{(\Delta X_s \neq 0)} \\ & = \log(\bar{Y}^{n,\Psi}) \mathbf{1}_{(\log \bar{Y}^{n,\Psi} > \tau)} * \mu_t. \end{aligned} \tag{3.21}$$

For any  $\varepsilon, \eta > 0$ ,

$$P(\log(\bar{Y}^{n,\Psi}) \mathbf{1}_{(\log \bar{Y}^{n,\Psi} > \tau)} * \mu_t > \varepsilon) \leq \frac{\eta}{\varepsilon} + P(\log(\bar{Y}^{n,\Psi}) \mathbf{1}_{(\log \bar{Y}^{n,\Psi} > \tau)} * \nu_t > \eta).$$

Note  $x - 1 > \log x$  if  $\log x > 0$ . Then

$$\log(\bar{Y}^{n,\Psi}) \mathbf{1}_{(\log \bar{Y}^{n,\Psi} > \tau)} \leq (\bar{Y}^{n,\Psi} - 1) \mathbf{1}_{(\bar{Y}^{n,\Psi} - 1 > \tau)}. \tag{3.22}$$

Recalling the definition of  $\iota(f_{1+\varepsilon}, P, P^{n,\psi})$  and Assumptions 1 and 2, we can obtain (3.18). The proofs of (3.19) and (3.17) are similar. □

**Lemma 3.7** *Under Assumptions 1, 2 and 3, for any  $\varepsilon, \eta > 0$ , there is a  $\delta > 0$  and a partition  $\Pi(\delta) = \{\Psi(\delta), 1 \leq k \leq N(\delta)\}$  such that*

$$\limsup_{n \rightarrow \infty} P \left( \sup_{1 \leq k \leq N(\delta)} \sup_{\phi, \psi \in \Psi(\delta, k)} |A_t^{n,2,\psi} - A_t^{n,2,\phi}| > \varepsilon \right) \leq \eta. \tag{3.23}$$

**Proof** Let us fix  $\varepsilon, \eta > 0$ . First, note

$$\begin{aligned} A_t^{n,2,\psi} &= h_\tau(\log Y^{n,\psi}) * (\mu - \nu)_t \\ &+ \sum_{s \leq t} h_\tau \left( \log \left( 1 - \frac{\hat{Y}_s^{n,\psi} - a_s}{1 - a_s} \right) \right) \mathbf{1}_{(\Delta X_s=0)} \\ &- h_\tau \left( \log \left( 1 - \frac{\hat{Y}_s^{n,\psi} - a_s}{1 - a_s} \right) \right) (1 - a_s). \end{aligned}$$

Let

$$J_t^{n,1,\psi} = h_\tau(\log Y^{n,\psi}) * (\mu - \nu)_t$$

and

$$\begin{aligned} J_t^{n,2,\psi} &= \sum_{s \leq t} \left[ h_\tau \left( \log \left( 1 - \frac{\hat{Y}_s^{n,\psi} - a_s}{1 - a_s} \right) \right) \mathbf{1}_{(\Delta X_s=0)} \right. \\ &\quad \left. - h_\tau \left( \log \left( 1 - \frac{\hat{Y}_s^{n,\psi} - a_s}{1 - a_s} \right) \right) (1 - a_s) \right]. \end{aligned}$$

We shall treat  $J_t^{n,1,\psi}$  and  $J_t^{n,2,\psi}$  separately below. Let us only focus on the  $J_t^{n,2,\psi}$  since the  $J_t^{n,1,\psi}$  is similar and simpler.

According to Assumption 3, there is a sufficiently large positive finite constant  $K$  such that

$$\limsup_{n \rightarrow \infty} P \left( \|\mathfrak{H}^n\|_{\Pi,t} > K \right) \leq \frac{\eta}{4}. \tag{3.24}$$

Thus, we only need to condition on the event  $\{\mathfrak{H}^n\|_{\Pi,t} > K\}$ . In particular, we shall prove

$$E^* \max_{1 \leq k \leq N_\Pi(\delta)} \sup_{\psi, \phi \in \Psi(\delta,k)} \left| J_t^{2,\psi} - J_t^{2,\phi} \right| \mathbf{1}_{\{\|\mathfrak{H}\|_{\Pi,t} \leq K\}} \leq c_{11} \int_0^\delta H_\Pi(\varepsilon) d\varepsilon. \tag{3.25}$$

Assuming (3.25), we can take  $\delta$  so small that

$$c_{11} \int_0^\delta H_\Pi(\varepsilon) d\varepsilon < \frac{\varepsilon \eta}{4}, \tag{3.26}$$

from which it follows by the Markov inequality

$$\limsup_{n \rightarrow \infty} P \left( \sup_{1 \leq k \leq N_\Pi(\delta)} \sup_{\phi, \psi \in \Psi(\delta,k)} |J_t^{n,2,\psi} - J_t^{n,2,\phi}| > \varepsilon, \|\mathfrak{H}\|_{\Pi,t} \leq K \right) \leq \frac{\eta}{4}$$

This in turn together with (3.24) implies

$$\limsup_{n \rightarrow \infty} P \left( \sup_{1 \leq k \leq N_{\Pi}(\delta)} \sup_{\phi, \psi \in \Psi(\delta, k)} |J_t^{n,2,\psi} - J_t^{n,2,\phi}| > \varepsilon \right) \leq \frac{\eta}{2}.$$

It remains to prove (3.25). For every integer  $p \geq 0$ , construct a nested refinement partition  $\Pi(2^{-p}\delta) = \{\Psi(2^{-p}\delta; k), 1 \leq k \leq N_{\Pi}(2^{-p}\delta)\}$  of  $\Psi$ , and then choose an element  $\psi_{p,k}$  from each partitioning set  $\Psi(2^{-p}\delta; k)$  in such a way that

$$\{\psi_{p,k} : 1 \leq k \leq N_{\Pi}(2^{-p}\delta)\} \subset \{\psi_{p+1,k} : 1 \leq k \leq N_{\Pi}(2^{-p-1}\delta)\}. \tag{3.27}$$

For every  $\psi \in \Psi$  and each  $p \geq 0$ , define  $\pi_p \psi = \psi_{p,k}$  and  $\Pi_p \psi = \Psi(2^{-p}\delta; k)$  whenever  $\psi \in \Psi(2^{-p}\delta; k)$ . Obviously,  $\Pi_p \psi \subseteq \Pi_{p-1} \psi$ . Define

$$W(\Pi_p \psi)_t = \left[ \sup_{\varphi, \phi \in \Pi_p \psi} \left| h_{\tau} \left( \log \left( 1 - \left( \frac{\hat{Y}_s^{n,\varphi} - a_s}{1 - a_s} \right) \right) \right) - h_{\tau} \left( \log \left( 1 - \left( \frac{\hat{Y}_s^{n,\phi} - a_s}{1 - a_s} \right) \right) \right) \right] \Big|_{\tilde{\mathcal{F}}, M_{\mu}^p}.$$

Note  $W(\Pi_p \psi) \leq W(\Pi_{p-1} \psi)$ . Set

$$\alpha_p = \frac{2^{-p+1}\delta}{H_{\Pi}(2^{-p-1}\delta)} K, \quad p \geq 0 \tag{3.28}$$

and

$$A_0(\psi) = \mathbf{1}_{\{W(\Pi_0 \psi) \leq \alpha_0\}}, \quad B_0(\psi) = \mathbf{1}_{\{W(\Pi_0 \psi) > \alpha_0\}}$$

and for  $p \geq 1$

$$A_p(\psi) = \mathbf{1}_{\{W(\Pi_0 \psi) \leq \alpha_0, \dots, W(\Pi_{p-1} \psi) \leq \alpha_{p-1}, W(\Pi_p \psi) \leq \alpha_p\}},$$

$$B_p(\psi) = \mathbf{1}_{\{W(\Pi_0 \psi) \leq \alpha_0, \dots, W(\Pi_{p-1} \psi) \leq \alpha_{p-1}, W(\Pi_p \psi) > \alpha_p\}}.$$

It is easy to see

$$A_0 + B_0 = 1$$

and for each  $p \geq 1$

$$A_p + B_p = A_{p-1}.$$

Hence, it follows for any  $q \geq 1$

$$A_q + B_q + B_{q-1} + \dots + B_0 = 1.$$

Note  $\pi_0\psi = \pi_0\phi$  if  $\psi, \phi \in \Psi(\delta, k)$ , and so we have

$$\sup_{\phi, \psi \in \Psi(\delta, k)} \left| J_t^{n,2,\psi} - J_t^{n,2,\phi} \right| \leq 2 \sup_{\psi \in \Psi} \left| J_t^{n,2,\psi} - J_t^{n,2,\pi_0\psi} \right|.$$

It is now enough to show

$$E^* \sup_{\psi \in \Psi} \left| J_t^{n,2,\psi} - J_t^{n,2,\pi_0\psi} \right| 1_{\{\|\mathcal{F}\|_{\Pi,t} \leq K\}} \leq c_{12} \int_0^\delta H_\Pi(\varepsilon) d\varepsilon. \tag{3.29}$$

We have the following identity

$$\begin{aligned} J_t^{n,2,\psi} - J_t^{n,2,\pi_0\psi} &= \sum_{p=0}^q (J_t^{n,2,\psi} - J_t^{n,2,\pi_p\psi}) B_p(\psi) \\ &\quad + (J_t^{n,2,\psi} - J_t^{n,2,\pi_q\psi}) A_q(\psi) \\ &\quad + \sum_{p=1}^q (J_t^{n,2,\pi_p\psi} - J_t^{n,2,\pi_{p-1}\psi}) A_{p-1}(\psi). \end{aligned} \tag{3.30}$$

Denote

$$M_t^{n,p,\psi} = (J_t^{n,2,\psi} - J_t^{n,2,\pi_p\psi}) B_p(\psi).$$

We shall only establish (3.29) for  $M^{n,p,\psi}$  since the other three terms in RHS of (3.30) can be similarly treated.

Obviously,  $M^{n,p,\psi}$  is a local martingale, and

$$|\Delta M_t^{n,p,\psi}| \leq |W(\Pi_{p-1}(\psi))_t| B_p(\psi) \leq \alpha_{p-1}.$$

On the other hand, the predictable quadratic variation of  $M_t^{n,p,\psi}$  satisfies

$$\begin{aligned} \langle M^{n,p,\psi} \rangle_t &= \sum_{s \leq t} \left[ h_\tau \left( \log \left( 1 - \frac{\hat{Y}_s^{n,\psi} - a_s}{1 - a_s} \right) \right) \right. \\ &\quad \left. - h_\tau \left( \log \left( 1 - \frac{\hat{Y}_s^{n,\pi_p\psi} - a_s}{1 - a_s} \right) \right) \right]^2 (1 - a_s). \end{aligned}$$

By an elementary calculation, for any  $\gamma \in (0, 1)$  there is a constant  $c_\gamma$  such that

$$|\log x - \log y| \leq c_\gamma |\sqrt{x} - \sqrt{y}|$$

whenever  $x, y \in [1 - \gamma, 1 + \gamma]$ . Then it follows

$$\begin{aligned} \langle M^{n,p,\psi} \rangle_1 &\leq \sum_{s \leq t} \left[ \sqrt{1 - \frac{\hat{Y}_s^{n,\psi} - a_s}{1 - a_s}} - \sqrt{1 - \frac{\hat{Y}_s^{n,\pi_p\psi} - a_s}{1 - a_s}} \right]^2 (1 - a_s) \\ &\leq \sum_{s \leq t} \left[ \sqrt{1 - \hat{Y}_s^{n,\psi}} - \sqrt{1 - \hat{Y}_s^{n,\pi_p\psi}} \right]^2 \\ &\leq 2h \left( \frac{1}{2}; P^{n,\psi}, P^{n,\pi_p\psi} \right) \leq 2(2^{-p}\delta)^2 \|\mathfrak{H}\|_{\Pi,t}^2. \end{aligned}$$

By Bernstein–Freedman’s inequality (see Lemma 3.2 of Nishiyama [8]) for local martingale with bounded jumps, it follows for  $\varepsilon > 0$ ,

$$P(|M_t^{n,p,\psi}| > \varepsilon, \|\mathfrak{H}\|_{\Pi,t} \leq K) \leq 2 \exp \left( - \frac{\varepsilon^2}{2[\alpha_{p-1}\varepsilon + (2^{-p}\delta K)^2]} \right).$$

This, in turn together with Lemma 2.2.10 of van der Vaart and Wellner [12], yields

$$\begin{aligned} E^* \sup_{\psi \in \Psi} |M^{n,p,\psi}| \mathbf{1}_{\{\|\mathfrak{H}\|_{\Pi,1} \leq K\}} &\leq \alpha_{p-1} H_{\Pi}(2^{-p}\delta)^2 + 2^{-p} \delta c_{13} H_{\Pi}(2^{-p}\delta) \\ &\leq c_{13} 2^{-p} \delta H_{\Pi}(2^{-p}\delta) \end{aligned}$$

and

$$\begin{aligned} E^* \sup_{\psi \in \Psi} \left| \sum_{p=0}^q J_t^{n,2,\psi} - J_t^{n,2,\pi_p\psi} \right| B_p(\psi) \mathbf{1}_{\{\|\mathfrak{H}\|_{\Pi,1} \leq K\}} &\leq c_{13} \sum_{p=0}^q 2^{-p} \delta H_{\Pi}(2^{-p}\delta) \\ &\leq c_{14} \int_0^{\delta} H_{\Pi}(\varepsilon) d\varepsilon. \end{aligned}$$

Thus, (3.25) is obtained, and so complete the proof. □

**Lemma 3.8** *Under Assumptions 1, 2 and 3, for any  $\varepsilon, \eta > 0$ , there is a  $\delta > 0$  and a partition  $\Pi(\delta) = \{\Psi(\delta), 1 \leq k \leq N(\delta)\}$  such that*

$$\limsup_{n \rightarrow \infty} P \left( \sup_{1 \leq k \leq N_{\Pi}(\delta)} \sup_{\phi, \psi \in \Psi(\delta,k)} |A_t^{n,1,\psi} - A_t^{n,1,\phi}| > \varepsilon \right) \leq \eta.$$

**Proof** Recall

$$A_t^{n,1,\psi} = \beta^{n,\psi} \cdot X_t^c - \frac{1}{2}(\beta^{n,\psi})^2 \cdot C_t.$$

It is enough to prove the following two statements

$$\limsup_{n \rightarrow \infty} P \left( \sup_{1 \leq k \leq N_{\Pi}(\delta)} \sup_{\phi, \psi \in \Psi(\delta, k)} |(\beta^{n, \psi})^2 \cdot C_t - (\beta^{n, \phi})^2 \cdot C_t| > \varepsilon \right) \leq \frac{\eta}{2} \tag{3.31}$$

and

$$\limsup_{n \rightarrow \infty} P \left( \sup_{1 \leq k \leq N_{\Pi}(\delta)} \sup_{\phi, \psi \in \Psi(\delta, k)} |\beta^{n, \psi} \cdot X_t^c - \beta^{n, \phi} \cdot X_t^c| > \varepsilon \right) \leq \frac{\eta}{2}. \tag{3.32}$$

We shall concentrate on proving (3.31) below since (3.32) is similar. The proof is completely similar to that of Lemma 3.7 with some minor modifications. For every integer  $p \geq 0$ , choose an element  $\psi_{p,k}$  from each partitioning set  $\Psi(2^{-p}\delta; k)$  in such a way that

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

and define  $\pi_p \psi = \psi_{p,k}$  and  $\Pi_p \psi = \Psi(2^{-p}\delta; k)$  whenever  $\psi \in \Psi(2^{-p}\delta; k)$ . Note

$$\begin{aligned} & \sup_{1 \leq k \leq N_{\Pi}(\delta)} \sup_{\phi, \psi \in \Psi(\delta, k)} |(\beta^{n, \psi})^2 \cdot C_t - (\beta^{n, \phi})^2 \cdot C_t| \\ & \leq 2 \sup_{\psi \in \Psi} |(\beta_s^{n, \psi})^2 \cdot C_t - (\beta_s^{n, \pi_0 \psi})^2 \cdot C_t|, \end{aligned}$$

then a main step is to prove

$$E^* \sup_{\psi \in \Psi} |(\beta^{n, \psi})^2 \cdot C_t - (\beta^{n, \pi_0 \psi})^2 \cdot C_t| \mathbf{1}_{\{\|\Xi\|_{\Pi, 1} \leq K\}} \leq c_{14} \int_0^\delta H_{\Pi}(\varepsilon) d\varepsilon. \tag{3.33}$$

To this end, for  $p \geq 0$ , set

$$\Gamma(\Pi_p \psi)_t = \left[ \sup_{\varphi, \phi \in \Pi_p \psi} |(\beta_s^{n, \varphi})^2 - (\beta_s^{n, \phi})^2| \right]_{\tilde{\mathcal{P}}, M_{\mu}^p}.$$

Obviously,  $\Gamma(\Pi_p \psi) \leq \Gamma(\Pi_{p-1} \psi)$ . Define

$$\begin{aligned} D_0(\psi) &= \mathbf{1}_{\{\Gamma(\Pi_0 \psi) \leq \alpha_0\}}, & E_0(\psi) &= \mathbf{1}_{\{\Gamma(\Pi_0 \psi) > \alpha_0\}}, \\ D_p(\psi) &= \mathbf{1}_{\{\Gamma(\Pi_0 \psi) \leq \alpha_0, \dots, \Gamma(\Pi_{p-1} \psi) \leq \alpha_{p-1}, \Gamma(\Pi_p \psi) \leq \alpha_p\}}, \\ E_p(\psi) &= \mathbf{1}_{\{\Gamma(\Pi_0 \psi) \leq \alpha_0, \dots, \Gamma(\Pi_{p-1} \psi) \leq \alpha_{p-1}, \Gamma(\Pi_p \psi) > \alpha_p\}}, \end{aligned}$$

where  $\alpha_p$  is as in (3.28).

Note we have the following identity

$$\begin{aligned}
 (\beta^{n,\psi})^2 \cdot C_t - (\beta^{n,\pi_0\psi})^2 \cdot C_t &= \sum_{p=1}^q ((\beta^{n,\pi_p\psi})^2 \cdot C_t - (\beta^{n,\pi_{p-1}\psi})^2 \cdot C_t) D_{p-1}(\psi) \\
 &\quad + ((\beta^{n,\psi})^2 \cdot C_t - (\beta^{n,\pi_q\psi})^2 \cdot C_t) D_q(\psi) \\
 &\quad + \sum_{p=0}^q ((\beta^{n,\psi})^2 \cdot C_t - (\beta^{n,\pi_p\psi})^2 \cdot C_t) E_p(\psi).
 \end{aligned}$$

and

$$\begin{aligned}
 ((\beta^{n,\psi})^2 \cdot C_t - (\beta^{n,\pi_p\psi})^2 \cdot C_t) E_p(\psi) &\leq \Gamma(\Pi_p\psi)(C_t - C_0) E_p(\psi) \\
 &\leq \frac{\Gamma(\Pi_p\psi)^2}{\alpha_p} (C_t - C_0) E_p(\psi).
 \end{aligned}$$

In addition, it is easy to see

$$E^* \sup_{\psi \in \Psi} ((\beta^{n,\pi_p\psi})^2 \cdot C_t - (\beta^{n,\pi_{p-1}\psi})^2 \cdot C_t) D_p(\psi) \leq 2^{-p} \delta H_{\Pi}(2^{-p} \delta)$$

by Schwarz’s inequality. Thus,

$$\begin{aligned}
 E^* \sup_{\psi \in \Psi} |(\beta^{n,\psi})^2 \cdot C_t - (\beta^{n,\pi_0\psi})^2 \cdot C_t| \mathbf{1}_{\{\|\delta\|_{\Pi,1} \leq K\}} \\
 \leq \sum_{p=1}^q K_2 2^{-p-1} \delta H_{\Pi}(2^{-p-1} \delta) \\
 \leq K_2 \int_0^{\delta} H_{\Pi}(\varepsilon) d\varepsilon.
 \end{aligned}$$

Thus, (3.31) is proved. We complete the proof of this lemma. □

We can obtain the following proposition by Lemmas 3.6–3.8.

**Proposition 3.9** *Under Assumptions 1, 2 and 3, for any  $\varepsilon, \eta > 0$ , there is a  $\delta > 0$  and a partition  $\Pi(\delta) = \{\Psi(\delta), 1 \leq k \leq N(\delta)\}$  such that*

$$\limsup_{n \rightarrow \infty} P \left( \sup_{1 \leq k \leq N_{\Pi}(\delta)} \sup_{\phi, \psi \in \Psi(\delta,k)} |L^{n,\psi} - L^{n,\phi}| > \varepsilon \right) \leq \eta.$$

**The proof of Theorem 2.1** Proposition 3.9 implies the asymptotic equicontinuity of  $L^n$ , and the asymptotic marginal distribution of  $L^n$  is obtained by Proposition 3.5. Then we can obtain Theorem 2.1 by these two propositions and Theorem 1.1 in Nishiyama [9]. □



**Acknowledgements** The authors would like to thank the anonymous referees and the Associate Editor for careful reading and constructive comments. This work was supported by the National Natural Science Foundation of China (Nos. 11371317, 11701331, 11731012, 11871425), Fundamental Research Funds for Central Universities, Shandong Provincial Natural Science Foundation (No. ZR2017QA007) and Young Scholars Program of Shandong University.

## References

1. Andersen, N., Giné, E., Ossiander, M., Zinn, J.: The central limit theorem and the law of iterated logarithm for empirical processes under local conditions. *Probab. Theory Relat. Fields* **77**, 271–305 (1988)
2. Billingsley, P.: *Convergence of Probability Measures*, 2nd edn. Wiley, New York (1999)
3. Dudley, R.: Central limit theorems for empirical measures. *Ann. Probab.* **6**, 899–929 (1978)
4. Giné, E., Zinn, J.: Some limit theorems for empirical processes. *Ann. Probab.* **12**, 929–989 (1984)
5. Jacod, J., Shiryaev, A.: *Limit theorems for stochastic processes*. Grundlehren der Mathematischen Wissenschaften, vol. 288, 2nd edn. Springer, Berlin (2003)
6. Le Cam, L.: Locally asymptotically normal families of distributions. *Univ. Calif. Publ. Stat.* **3**, 27–98 (1960)
7. Liptser, R., Shiryaev, A.: *Theory of Martingale*. Kluwer, Dordrecht (1989)
8. Nishiyama, Y.: Some central limit theorems for  $l^\infty$ -valued semimartingales and their applications. *Probab. Theory Relat. Fields* **108**, 459–494 (1997)
9. Nishiyama, Y.: Weak convergence of some classes of martingales with jumps. *Ann. Probab.* **28**, 685–712 (2000)
10. Ossiander, M.: A central limit theorem under metric entropy with  $L_2$  bracketing. *Ann. Probab.* **15**, 897–919 (1987)
11. van de Geer, S.: Exponential inequalities for martingales, with application to maximum likelihood estimation for counting processes. *Ann. Stat.* **23**, 1779–1801 (1995)
12. van der Vaart, A., Wellner, J.: *Weak Convergence and Empirical Processes*. Springer, New York (1996)
13. Vostrikova, L.: Functional limit theorems for the likelihood ratio processes. *Ann. Univ. Sci. Budapest. Sect. Comput.* **6**, 145–182 (1985)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.