# A note on weak convergence of random step processes

Márton Ispány* and Gyula Pap*,◇

* University of Debrecen, Faculty of Informatics, Pf. 12, H–4010 Debrecen, Hungary

e–mails: Ispany.Marton@inf.unideb.hu (M. Ispány), Pap.Gyula@inf.unideb.hu (G. Pap).

◇ Corresponding author.

*April 14, 2009.*

**Abstract**

First, sufficient conditions are given for a triangular array of random vectors such that the sequence of related random step functions converges towards a (not necessarily time homogeneous) diffusion process. These conditions are weaker and easier to check than the existing ones in the literature, and they are derived from a very general semimartingale convergence theorem due to Jacod and Shiryaev, which is hard to use directly.

Next, sufficient conditions are given for convergence of stochastic integrals of random step functions, where the integrands are functionals of the integrators. This result covers situations which can not be handled by existing ones.

## 1 Introduction

The aim of the present paper is to obtain a useful theorem concerning convergence of step processes towards a diffusion process. We derive sufficient conditions (see Theorem 2.1 and Corollary 2.2) from a very general semimartingale convergence theorem due to Jacod and Shiryaev [5, Theorem IX.3.39]. (This theorem of Jacod and Shiryaev is hard to use directly, since one has to check the local strong majoration hypothesis, the local condition on big jumps, local uniqueness for the associated martingale problem, and the continuity condition.) Theorem 2.1 can also be considered as a generalization of the sufficient part of the functional martingale central limit theorem (see, e.g., Jacod and Shiryaev [5, Theorem VII.3.4]), but Theorem 2.1 allows not necessarily time homogeneous diffusion limit processes as well. Similarly, Corollary 2.2 can be considered as a generalization of the sufficient part of the Lindeberg-Feller functional central limit theorem (see, e.g., Jacod and Shiryaev [5, Theorem VII.5.4]).

There are several diffusion approximations in the literature, but they contain assumptions which are stronger and more complicated to check. For example, Ethier and Kurtz [2, Theorem 7.4.1] deals only with the time homogeneous case, and their conditions (4.3)—(4.7) are hard to check. The result of Joffe and Métivier [7, Theorem 3.3.1] is not easy to use, since their conditions $(H_1)$ and $(H_4)$ are rather complicated to check. Gikhman and Skorokhod [3, Theorem 9.4.1] covers only convergence of Markov chains, and it contains Lipschitz conditions on the drift and diffusion coefficient of the limiting diffusion process, and assumes finite $2+\delta$ moments for some $\delta > 0$. Our Theorem 2.1 and Corollary 2.2 are valid not only for martingales or Markov chains, since we do not suppose any dependence structure. The conditions are natural, since uniform convergence on compacts in probability (ucp) is involved. (The role of the topology of the ucp is nicely explained by Kurtz and Protter [10].)

We also develope sufficient conditions (see Theorem 3.2 and Corollary 3.3) for convergence of stochastic integrals of random step functions, where the integrand is a functional of the integrator. We mention that our result covers situations which can not be handled by the convergence theorems of Jacod and Shiryaev [5, Theorem IX.5.12, Theorem IX.5.16, Corollary IX.5.18, Remark IX.5.19]. There is a nice theory of convergence of stochastic integrals due to Jakubowski, Mémin and Pagès [6] and to Kurtz and Protter [8], [9], [10]. The key result of this theory says that if $(\mathcal{U}^n)_{n\in\mathbb{N}}$ is a uniformly tight sequence of semimartingales

*2000 Mathematics Subject Classifications*: 60F17, 60J60.

*Key words and phrases*: Weak convergence of semimartingales, diffusion process.

(or, equivalently, it has uniformly controlled variations) then it is good in the sense that $(\mathcal{U}^n, \mathcal{V}^n, \mathcal{Y}^n) \xrightarrow{\mathcal{L}} (\mathcal{U}, \mathcal{V}, \mathcal{Y})$ whenever $(\mathcal{U}^n, \mathcal{V}^n) \xrightarrow{\mathcal{L}} (\mathcal{U}, \mathcal{V})$, where $\mathcal{Y}_t^n := \int_0^t \mathcal{V}_{s-}^n \, \mathrm{d}\mathcal{U}_s^n$ and $\mathcal{Y}_t := \int_0^t \mathcal{V}_{s-} \, \mathrm{d}\mathcal{U}_s$. In our Theorem 3.2 and Corollary 3.3, the sequence $(\mathcal{U}^n)_{n\in\mathbb{N}}$ of semimartingales is not necessarily good (see Example 2.3).

In the proofs the simple structure of the approximating step processes and the almost sure continuity of the limiting diffusion process play a crucial role.

As an application of these results, a Feller type diffusion approximation can be derived for critical multitype branching processes with immigration if the offspring mean matrix is primitive, and the asymptotic behavior of the conditional least squares estimator of the offspring mean matrix may be established, see Ispány and Pap [4], which will be the content of a forthcoming paper.

## 2 Convergence of step processes to diffusion processes

A process $(\mathcal{U}_t)_{t\in\mathbb{R}_+}$ with values in $\mathbb{R}^d$ is called a diffusion process if it is a weak solution of a stochastic differential equation

$$\mathrm{d}\mathcal{U}_t = \beta(t, \mathcal{U}_t)\,\mathrm{d}t + \gamma(t, \mathcal{U}_t)\,\mathrm{d}\mathcal{W}_t, \qquad t \in \mathbb{R}_+, \tag{2.1}$$

where $\mathbb{R}_+$ denotes the set of nonnegative real numbers, $\beta : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ and $\gamma : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d\times r}$ are Borel functions and $(\mathcal{W}_t)_{t\in\mathbb{R}_+}$ is an $r$-dimensional standard Wiener process.

If $(\Omega, \mathcal{A}, \mathsf{P})$ is a probability space, $\mathcal{F} \subset \mathcal{A}$ is a $\sigma$-algebra, and $\xi : \Omega \to \mathbb{R}^d$ is a random variable with $\mathsf{E}(\|\xi\|^2 \mid \mathcal{F}) < \infty$ then $\mathrm{Var}(\xi \mid \mathcal{F})$ will denote the conditional variance matrix defined by

$$\mathrm{Var}(\xi \mid \mathcal{F}) := \mathsf{E}\Big(\big(\xi - \mathsf{E}(\xi \mid \mathcal{F})\big)\big(\xi - \mathsf{E}(\xi \mid \mathcal{F})\big)^\top \,\Big|\, \mathcal{F}\Big).$$

(Here and in the sequel, $\|x\|$ denotes the Euclidean norm of a (column) vector $x \in \mathbb{R}^d$, $A^\top$ and $\operatorname{tr} A$ denote the transpose and the trace of a matrix $A$, respectively.) The set of all nonnegative integers and the set of all positive integers will be denoted by $\mathbb{Z}_+$ and $\mathbb{N}$, respectively. The lower integer part and the positive part of $x \in \mathbb{R}$ will be denoted by $\lfloor x \rfloor$ and $x_+$, respectively.

**Theorem 2.1** *Let $\beta : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ and $\gamma : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d\times r}$ be continuous functions. Assume that the SDE (2.1) has a unique weak solution with $\mathcal{U}_0 = u_0$ for all $u_0 \in \mathbb{R}^d$. Let $\eta$ be a probability measure on $\mathbb{R}^d$, and let $(\mathcal{U}_t)_{t\in\mathbb{R}_+}$ be a solution of (2.1) with initial distribution $\eta$. For each $n \in \mathbb{N}$, let $(U_k^n)_{k\in\mathbb{Z}_+}$ be a sequence of random variables with values in $\mathbb{R}^d$ adapted to a filtration $(\mathcal{F}_k^n)_{k\in\mathbb{Z}_+}$. Let*

$$\mathcal{U}_t^n := \sum_{k=0}^{\lfloor nt \rfloor} U_k^n\,, \qquad t \in \mathbb{R}_+, \quad n \in \mathbb{N}.$$

*Let $h : \mathbb{R}^d \to \mathbb{R}^d$ be a continuous function with compact support satisfying $h(x) = x$ in a neighborhood of 0. Suppose $U_0^n \xrightarrow{\mathcal{L}} \eta$, and for each $T > 0$,*

(i) $\displaystyle\sup_{t\in[0,T]} \left\| \sum_{k=1}^{\lfloor nt \rfloor} \mathsf{E}(h(U_k^n) \mid \mathcal{F}_{k-1}^n) - \int_0^t \beta(s, \mathcal{U}_s^n)\,\mathrm{d}s \right\| \xrightarrow{\mathsf{P}} 0,$

(ii) $\displaystyle\sup_{t\in[0,T]} \left\| \sum_{k=1}^{\lfloor nt \rfloor} \mathrm{Var}(h(U_k^n) \mid \mathcal{F}_{k-1}^n) - \int_0^t \gamma(s, \mathcal{U}_s^n)\gamma(s, \mathcal{U}_s^n)^\top \mathrm{d}s \right\| \xrightarrow{\mathsf{P}} 0,$

(iii) $\displaystyle\sum_{k=1}^{\lfloor nT \rfloor} \mathsf{P}(\|U_k^n\| > \theta \mid \mathcal{F}_{k-1}^n) \xrightarrow{\mathsf{P}} 0$ *for all* $\theta > 0$.

*Then $\mathcal{U}^n \xrightarrow{\mathcal{L}} \mathcal{U}$ as $n \to \infty$, i.e., the distributions of $\mathcal{U}^n$ on the Skorokhod space $\mathbb{D}(\mathbb{R}^d)$ converge weakly to the distribution of $\mathcal{U}$ on $\mathbb{D}(\mathbb{R}^d)$.*

**Proof.** The process $(\mathcal{U}_t)_{t\in\mathbb{R}_+}$ is a semimartingale with characteristics $(\mathcal{B},\mathcal{C},0)$, where $\mathcal{B}_t := \int_0^t \beta(s,\mathcal{U}_s)\,\mathrm{d}s$, $\mathcal{C}_t := \int_0^t \gamma(s,\mathcal{U}_s)\gamma(s,\mathcal{U}_s)^\top \mathrm{d}s$ (see Jacod and Shiryaev [5, III. § 2c]). In general, $\operatorname{var}\mathcal{B}$ and $\operatorname{tr}\mathcal{C}$ do not necessarily satisfy majoration hypothesis, where $\operatorname{var}\alpha$ denotes the total variation of a function $\alpha\in\mathbb{D}(\mathbb{R}^d)$. So we fix $T>0$, and stop the characteristics at $T$, that is, we consider the processes $\big(\mathcal{B}_t^T\big)_{t\in\mathbb{R}_+}$ and $\big(\mathcal{C}_t^T\big)_{t\in\mathbb{R}_+}$ defined by

$$\mathcal{B}_t^T := \int_0^{t\wedge T} \beta(s,\mathcal{U}_s)\,\mathrm{d}s, \quad \mathcal{C}_t^T := \int_0^{t\wedge T} \gamma(s,\mathcal{U}_s)\gamma(s,\mathcal{U}_s)^\top \mathrm{d}s,$$

where $t\wedge T := \inf\{t,T\}$. Clearly, the stopped process $\big(\mathcal{U}_t^T\big)_{t\in\mathbb{R}_+}$ defined by $\mathcal{U}_t^T := \mathcal{U}_{t\wedge T}$ is a semimartingale with characteristics $\big(\mathcal{B}^T,\mathcal{C}^T,0\big)$.

We will also consider the stopped processes $\big(\mathcal{U}_t^{n,T}\big)_{t\in\mathbb{R}_+}$, $n\in\mathbb{N}$, defined by $\mathcal{U}_t^{n,T} := \mathcal{U}_{t\wedge T}^n$. We will check that all hypotheses of Theorem IX.3.39 of Jacod and Shiryaev [5] are fulfilled.

Firstly, we check the local strong majoration hypothesis. For each $a>0$, consider the mapping $\tau_a : \mathbb{D}(\mathbb{R}^d)\to[0,\infty]$ defined by $\tau_a(\alpha) := \inf\{t\in\mathbb{R}_+ : |\alpha(t)|\geqslant a \text{ or } |\alpha(t-)|\geqslant a\}$ for $\alpha\in\mathbb{D}(\mathbb{R}^d)$, where $\inf\emptyset := \infty$. Then the stopped processes $\big(\operatorname{var}\mathcal{B}^T_{t\wedge\tau_a(\mathcal{U}^T)}\big)_{t\in\mathbb{R}_+}$ and $\big(\operatorname{tr}\mathcal{C}^T_{t\wedge\tau_a(\mathcal{U}^T)}\big)_{t\in\mathbb{R}_+}$ are strongly majorized by the functions $t\mapsto b_{a,T}t$ and $t\mapsto c_{a,T}t$ respectively, where

$$b_{a,T} := \sup_{t\in[0,T]}\sup_{\|x\|\leqslant a}\|\beta(t,x)\|, \qquad c_{a,T} := \sup_{t\in[0,T]}\sup_{\|x\|\leqslant a}\|\gamma(t,x)\|^2.$$

Indeed, for all $s,t\in\mathbb{R}_+$ with $s<t$, we have

$$\begin{aligned}
\operatorname{var}\mathcal{B}^T_{t\wedge\tau_a(\mathcal{U}^T)} - \operatorname{var}\mathcal{B}^T_{s\wedge\tau_a(\mathcal{U}^T)} &= \int_{s\wedge T\wedge\tau_a(\mathcal{U}^T)}^{t\wedge T\wedge\tau_a(\mathcal{U}^T)} \|\beta(u,\mathcal{U}_u^T)\|\,\mathrm{d}u,\\
\operatorname{tr}\mathcal{C}^T_{t\wedge\tau_a(\mathcal{U}^T)} - \operatorname{tr}\mathcal{C}^T_{s\wedge\tau_a(\mathcal{U}^T)} &= \int_{s\wedge T\wedge\tau_a(\mathcal{U}^T)}^{t\wedge T\wedge\tau_a(\mathcal{U}^T)} \|\gamma(u,\mathcal{U}_u^T)\|^2\,\mathrm{d}u.
\end{aligned}$$

The process $\big(\mathcal{U}_t^T\big)_{t\in\mathbb{R}_+}$ is a.s. continuous, hence $u\leqslant t\wedge T\wedge\tau_a(\mathcal{U}^T)$ implies $\|\mathcal{U}_u^T\|\leqslant a$ a.s, thus

$$\|\beta(u,\mathcal{U}_u^T)\|\leqslant b_{a,T} \quad \text{a.s.}, \qquad \|\gamma(u,\mathcal{U}_u^T)\|^2\leqslant c_{a,T} \quad \text{a.s.}$$

Consequently

$$\begin{aligned}
\int_{s\wedge T\wedge\tau_a(\mathcal{U}^T)}^{t\wedge T\wedge\tau_a(\mathcal{U}^T)} \|\beta(u,\mathcal{U}_u^T)\|\,\mathrm{d}u &\leqslant b_{a,T}t - b_{a,T}s \qquad \text{a.s.},\\
\int_{s\wedge T\wedge\tau_a(\mathcal{U}^T)}^{t\wedge T\wedge\tau_a(\mathcal{U}^T)} \|\gamma(u,\mathcal{U}_u^T)\|^2\,\mathrm{d}u &\leqslant c_{a,T}t - c_{a,T}s \qquad \text{a.s.},
\end{aligned}$$

hence the local strong majoration hypothesis holds.

The local condition on big jumps is obviously satisfied, since the third characteristic of the semimartingale $\big(\mathcal{U}_t^T\big)_{t\in\mathbb{R}_+}$ is $0$. By the assumption, the martingale problem associated to the characteristics $(\mathcal{B}^T,\mathcal{C}^T,0)$ admits a unique solution for each initial value $u_0\in\mathbb{R}^d$, thus Theorem III.2.40 of Jacod and Shiryaev [5] yields local uniqueness for the corresponding martingale problem as in Corollary III.2.41. The continuity conditions are clearly implied by the continuity of the functions $\beta$ and $\gamma$. Convergence of the initial distributions holds trivially.

For each $n\in\mathbb{N}$, the stopped process $\big(\mathcal{U}_t^{n,T}\big)_{t\in\mathbb{R}_+}$ is also a semimartingale with characteristics

$$\begin{aligned}
\mathcal{B}_t^{n,T} &:= \sum_{k=1}^{\lfloor n(t\wedge T)\rfloor} \mathsf{E}\big(h(U_k^n)\mid \mathcal{F}_{k-1}^n\big),\\
\mathcal{C}_t^{n,T} &:= 0,\\
\nu^{n,T}([0,t]\times g) &:= \sum_{k=1}^{\lfloor n(t\wedge T)\rfloor} \mathsf{E}\big(g(U_k^n)\mathbb{1}_{\{U_k^n\neq 0\}}\,\big|\,\mathcal{F}_{k-1}^n\big)
\end{aligned}$$

for $g : \mathbb{R}^d \to \mathbb{R}_+$ Borel functions, and modified second characteristic

$$\widetilde{\mathcal{C}}_t^{n,T} := \sum_{k=1}^{\lfloor n(t\wedge T)\rfloor} \operatorname{Var}(h(U_k^n) \mid \mathcal{F}_{k-1}^n)$$

(see Jacod and Shiryaev [5, II.3.14, II.3.18]). For each $a > 0$, assumptions (i)—(iii) imply

$$\sup_{t\in[0,T]} \left\| \mathcal{B}^{n,T}_{t\wedge\tau_a(\mathcal{U}^T)} - \int_0^{t\wedge\tau_a(\mathcal{U}^T)} \beta(s,\mathcal{U}_s^n)\,\mathrm{d}s \right\| \xrightarrow{\mathsf{P}} 0,$$
$$\sup_{t\in[0,T]} \left\| \widetilde{\mathcal{C}}^{n,T}_{t\wedge\tau_a(\mathcal{U}^T)} - \int_0^{t\wedge\tau_a(\mathcal{U}^T)} \gamma(s,\mathcal{U}_s^n)\gamma(s,\mathcal{U}_s^n)^\top \mathrm{d}s \right\| \xrightarrow{\mathsf{P}} 0,$$
$$\nu^{n,T}\big([0,\tau_a(\mathcal{U}^T)]\times g_c\big) \xrightarrow{\mathsf{P}} 0 \qquad \text{for all } c > 0,$$

where $g_c : \mathbb{R}^d \to \mathbb{R}_+$ is defined by

$$g_c(x) := (c\|x\| - 1)_+ \wedge 1. \tag{2.2}$$

(Indeed, $g_c(x) \leqslant \mathbb{1}_{\{\|x\|>1/c\}}$ for all $x \in \mathbb{R}^d$). Therefore all hypotheses of Theorem IX.3.39 of Jacod and Shiryaev [5] are fulfilled, hence for all $T > 0$, $\mathcal{U}^{n,T} \xrightarrow{\mathcal{L}} \mathcal{U}^T$. This implies that the finite dimensional distributions of the processes $\mathcal{U}^n$ converge to the corresponding finite dimensional distributions of the process $\mathcal{U}$ (see Jacod and Shiryaev [5, VI.3.14]).

The aim of the following discussion is to show the tightness of $\{\mathcal{U}^n : n \in \mathbb{N}\}$, which will imply $\mathcal{U}^n \xrightarrow{\mathcal{L}} \mathcal{U}$. For each $T > 0$, by Prokhorov's Theorem, convergence $\mathcal{U}^{n,T} \xrightarrow{\mathcal{L}} \mathcal{U}^T$ implies tightness of $\{\mathcal{U}^{n,T} : n \in \mathbb{N}\}$. By Theorem VI.3.21 of Jacod and Shiryaev [5], this implies

$$\mathsf{P}\Big(\sup_{t\in[0,T]} \|\mathcal{U}_t^{n,T}\| > K\Big) \to 0 \qquad \text{as } n\to\infty \text{ and } K\to\infty,$$
$$\mathsf{P}\left(w_T'(\mathcal{U}^{n,T},\theta) > \delta\right) \to 0 \qquad \text{as } n\to\infty \text{ and } \theta\downarrow 0 \text{ for all } \delta > 0,$$

where $w_T'(\alpha,\cdot)$ denotes the "modulus of continuity" on $[0,T]$ for a function $\alpha \in \mathbb{D}(\mathbb{R}^d)$ (see Jacod and Shiryaev [5, VI.1.8]). Since the above convergences hold for all $T > 0$, we conclude for all $T > 0$ that

$$\mathsf{P}\Big(\sup_{t\in[0,T]} \|\mathcal{U}_t^{n}\| > K\Big) \to 0 \qquad \text{as } n\to\infty \text{ and } K\to\infty,$$
$$\mathsf{P}\left(w_T'(\mathcal{U}^{n},\theta) > \delta\right) \to 0 \qquad \text{as } n\to\infty \text{ and } \theta\downarrow 0 \text{ for all } \delta > 0.$$

Again by Theorem VI.3.21 of Jacod and Shiryaev [5], this implies tightness of $\{\mathcal{U}^n : n \in \mathbb{N}\}$, and we obtain $\mathcal{U}^n \xrightarrow{\mathcal{L}} \mathcal{U}$. □

**Corollary 2.2** *Let* $\beta$, $\gamma$, $\eta$, $(U_k^n)_{k\in\mathbb{Z}_+}$, $(\mathcal{F}_k^n)_{k\in\mathbb{Z}_+}$ *and* $\mathcal{U}^n$ *for* $n \in \mathbb{N}$ *be as in Theorem 2.1. Suppose that* $\mathsf{E}\big(\|U_k^n\|^2 \,\big|\, \mathcal{F}_{k-1}^n\big) < \infty$ *for all* $n, k \in \mathbb{N}$. *Assume that the SDE* (2.1) *has a unique weak solution with* $\mathcal{U}_0 = u_0$ *for all* $u_0 \in \mathbb{R}^d$. *Let* $(\mathcal{U}_t)_{t\in\mathbb{R}_+}$ *be a solution of* (2.1) *with initial distribution* $\eta$. *Suppose* $U_0^n \xrightarrow{\mathcal{L}} \eta$, *and for each* $T > 0$,

(i) $\displaystyle\sup_{t\in[0,T]} \left\| \sum_{k=1}^{\lfloor nt\rfloor} \mathsf{E}(U_k^n \mid \mathcal{F}_{k-1}^n) - \int_0^t \beta(s,\mathcal{U}_s^n)\,\mathrm{d}s \right\| \xrightarrow{\mathsf{P}} 0,$

(ii) $\displaystyle\sup_{t\in[0,T]} \left\| \sum_{k=1}^{\lfloor nt\rfloor} \operatorname{Var}(U_k^n \mid \mathcal{F}_{k-1}^n) - \int_0^t \gamma(s,\mathcal{U}_s^n)\gamma(s,\mathcal{U}_s^n)^\top \mathrm{d}s \right\| \xrightarrow{\mathsf{P}} 0,$

(iii) $\displaystyle\sum_{k=1}^{\lfloor nT\rfloor} \mathsf{E}\big(\|U_k^n\|^2 \mathbb{1}_{\{\|U_k^n\|>\theta\}} \,\big|\, \mathcal{F}_{k-1}^n\big) \xrightarrow{\mathsf{P}} 0$ *for all* $\theta > 0$.

*Then* $\mathcal{U}^n \xrightarrow{\mathcal{L}} \mathcal{U}$ *as* $n \to \infty$.

**Proof.** Clearly, there exists $K \geqslant 1$ such that $h(x) = x$ for $\|x\| \leqslant 1/K$, $h(x) = 0$ for $\|x\| \geqslant K$, and $\|h(x)\| \leqslant K$ for all $x \in \mathbb{R}^d$. Hence $h(x) - x = 0$ for $\|x\| \leqslant 1/K$ and $\|h(x) - x\| \leqslant \|h(x)\| + \|x\| \leqslant K + \|x\| \leqslant (K^2+1)\|x\|$ for $\|x\| \geqslant 1/K$. Thus, we conclude

$$\|h(x) - x\| \leqslant (K^2+1)\|x\| \mathbb{1}_{\{\|x\|\geqslant 1/K\}} \leqslant (K^2+1)K\|x\|^2 \mathbb{1}_{\{\|x\|\geqslant 1/K\}} \tag{2.3}$$

for all $x \in \mathbb{R}^d$. For all $T > 0$ and all $t \in [0,T]$, applying (2.3), we get

$$\begin{aligned} \left\| \sum_{k=1}^{\lfloor nt \rfloor} \mathsf{E}(h(U_k^n) \mid \mathcal{F}_{k-1}^n) - \sum_{k=1}^{\lfloor nt \rfloor} \mathsf{E}(U_k^n \mid \mathcal{F}_{k-1}^n) \right\| &\leqslant \sum_{k=1}^{\lfloor nT \rfloor} \mathsf{E}\big(\|h(U_k^n) - U_k^n\| \,\big|\, \mathcal{F}_{k-1}^n\big) \\ &\leqslant (K^2+1)K \sum_{k=1}^{\lfloor nT \rfloor} \mathsf{E}\Big(\|U_k^n\|^2 \mathbb{1}_{\{\|U_k^n\|\geqslant 1/K\}} \mid \mathcal{F}_{k-1}^n\Big), \end{aligned}$$

which together with assumptions (i) and (iii) of this corollary imply condition (i) of Theorem 2.1. We have

$$\begin{aligned} \operatorname{Var}\big(h(U_k^n) \,\big|\, \mathcal{F}_{k-1}^n\big) - \operatorname{Var}\big(U_k^n \,\big|\, \mathcal{F}_{k-1}^n\big) &= \mathsf{E}\big(h(U_k^n)h(U_k^n)^\top - U_k^n(U_k^n)^\top \,\big|\, \mathcal{F}_{k-1}^n\big) \\ &\quad + \big(\mathsf{E}\big(h(U_k^n) \,\big|\, \mathcal{F}_{k-1}^n\big)\mathsf{E}\big(h(U_k^n)^\top \,\big|\, \mathcal{F}_{k-1}^n\big) - \mathsf{E}\big(U_k^n \,\big|\, \mathcal{F}_{k-1}^n\big)\mathsf{E}\big((U_k^n)^\top \,\big|\, \mathcal{F}_{k-1}^n\big)\big). \end{aligned}$$

For arbitrary matrices $A, B, C, D \in \mathbb{R}^{d\times r}$, we have

$$\|AB^\top - CD^\top\| \leqslant \|A - C\| \cdot \|B\| + \|A\| \cdot \|B - D\| + \|A - C\| \cdot \|B - D\|,$$

hence applying (2.3) and $\|h(x)\| \leqslant K$ valid for all $x \in \mathbb{R}^d$, we obtain

$$\begin{aligned} \sum_{k=1}^{\lfloor nt \rfloor} \big\|\mathsf{E}\big(h(U_k^n)h(U_k^n)^\top - U_k^n(U_k^n)^\top \,\big|\, \mathcal{F}_{k-1}^n\big)\big\| &\leqslant \sum_{k=1}^{\lfloor nt \rfloor} \mathsf{E}\big(2\|h(U_k^n) - U_k^n\|\|h(U_k^n)\| + \|h(U_k^n) - U_k^n\|^2 \,\big|\, \mathcal{F}_{k-1}^n\big) \\ &\leqslant (K^2+1)(3K^2+1) \sum_{k=1}^{\lfloor nt \rfloor} \mathsf{E}\big(\|U_k^n\|^2 \mathbb{1}_{\{\|U_k^n\|\geqslant 1/K\}} \,\big|\, \mathcal{F}_{k-1}^n\big). \end{aligned}$$

In a similar way, we obtain

$$\begin{aligned} &\sum_{k=1}^{\lfloor nt \rfloor} \big\|\mathsf{E}\big(h(U_k^n) \,\big|\, \mathcal{F}_{k-1}^n\big)\mathsf{E}\big(h(U_k^n)^\top \,\big|\, \mathcal{F}_{k-1}^n\big) - \mathsf{E}\big(U_k^n \,\big|\, \mathcal{F}_{k-1}^n\big)\mathsf{E}\big((U_k^n)^\top \,\big|\, \mathcal{F}_{k-1}^n\big)\big\| \\ &\leqslant 2K^2(K^2+1) \sum_{k=1}^{\lfloor nt \rfloor} \mathsf{E}\big(\|U_k^n\|^2 \mathbb{1}_{\{\|U_k^n\|\geqslant 1/K\}} \,\big|\, \mathcal{F}_{k-1}^n\big) + K^2(K^2+1)^2 \left( \sum_{k=1}^{\lfloor nt \rfloor} \mathsf{E}\big(\|U_k^n\|^2 \mathbb{1}_{\{\|U_k^n\|\geqslant 1/K\}} \,\big|\, \mathcal{F}_{k-1}^n\big) \right)^2. \end{aligned}$$

These inequalities together with assumptions (ii) and (iii) of this corollary imply condition (ii) of Theorem 2.1. We have

$$\mathsf{P}\big(\|U_k^n\| > \theta \,\big|\, \mathcal{F}_{k-1}^n\big) \leqslant \theta^{-2}\mathsf{E}\Big(\|U_k^n\|^2 \mathbb{1}_{\{\|U_k^n\|\geqslant \theta\}} \mid \mathcal{F}_{k-1}^n\Big),$$

thus assumption (iii) of this corollary implies (iii) of Theorem 2.1. □

**Example 2.3** We give an example for a system $(U_k^n)_{n\in\mathbb{N},\,k\in\mathbb{Z}_+}$ of random variables satisfying conditions (i)—(iii) of Corollary 2.2, such that the sequence $(\mathcal{U}^n)_{n\in\mathbb{N}}$ of semimartingales is not good (see the Introduction).

Let $(\eta_k)_{k\in\mathbb{N}}$ be independent standard normal random variables. Let $U_0^n := 0$, $U_{3j}^n := -\eta_j/\sqrt{n}$, $U_{3j-1}^n := U_{3j-2}^n := \eta_j/\sqrt{n}$ and $\mathcal{F}_{j-1}^n := \sigma(U_0^n, \dots, U_{j-1}^n)$ for $j, n \in \mathbb{N}$. Then conditions (i)—(iii) of Corollary 2.2 are satisfied with $\beta = 0$ and $\gamma = 1/\sqrt{3}$. For each $n \in \mathbb{N}$, let

$$\int_0^t \mathcal{U}_{s-}^n \, \mathrm{d}\mathcal{U}_s^n = \sum_{k=1}^{\lfloor nt \rfloor} U_k^n \sum_{j=1}^{k-1} U_j^n = \frac{1}{2}(\mathcal{U}_t^n)^2 - \frac{1}{2}\sum_{k=1}^{\lfloor nt \rfloor} (U_k^n)^2.$$

Then, by Corollary 2.2, $\mathcal{U}^n \xrightarrow{\mathcal{L}} \mathcal{U} := \mathcal{W}/\sqrt{3}$, where $(\mathcal{W}_t)_{t\in\mathbb{R}_+}$ is a standard Wiener process. Moreover, $\sum_{k=1}^{\lfloor nt \rfloor} (U_k^n)^2 \xrightarrow{\mathsf{P}} t$, hence $\int_0^t \mathcal{U}_{s-}^n \, \mathrm{d}\mathcal{U}_s^n \xrightarrow{\mathcal{L}} \frac{1}{6}(\mathcal{W}_t)^2 - \frac{1}{2}t$. But, by Itô's formula, $\int_0^t \mathcal{U}_{s-} \, \mathrm{d}\mathcal{U}_s = \frac{1}{6}((\mathcal{W}_t)^2 - t)$, thus the sequence $\left(\int_0^t \mathcal{U}_{s-}^n \, \mathrm{d}\mathcal{U}_s^n\right)_{n\in\mathbb{N}}$ does not converge to $\int_0^t \mathcal{U}_{s-} \, \mathrm{d}\mathcal{U}_s$. Consequently, the sequence $(\mathcal{U}^n)_{n\in\mathbb{N}}$ of semimartingales is not good.

## 3 Convergence of integrals of step processes

For a function $\alpha \in \mathbb{D}(\mathbb{R}^d)$ and for a sequence $(\alpha_n)_{n\in\mathbb{N}}$ in $\mathbb{D}(\mathbb{R}^d)$, we write $\alpha_n \xrightarrow{\mathrm{lu}} \alpha$ if $(\alpha_n)_{n\in\mathbb{N}}$ converges to $\alpha$ locally uniformly, i.e., if $\sup_{t\in[0,T]} \|\alpha_n(t) - \alpha(t)\| \to 0$ as $n \to \infty$ for all $T > 0$. The space of all continuous functions $\alpha : \mathbb{R}_+ \to \mathbb{R}^d$ will be denoted by $\mathbb{C}(\mathbb{R}^d)$. For measurable mappings $\Phi : \mathbb{D}(\mathbb{R}^d) \to \mathbb{D}(\mathbb{R}^p)$ and $\Phi_n : \mathbb{D}(\mathbb{R}^d) \to \mathbb{D}(\mathbb{R}^p)$, $n \in \mathbb{N}$, we will denote by $C_{\Phi,(\Phi_n)_{n\in\mathbb{N}}}$ the set of all functions $\alpha \in \mathbb{C}(\mathbb{R}^d)$ such that $\Phi(\alpha) \in \mathbb{C}(\mathbb{R}^p)$ and $\Phi_n(\alpha_n) \xrightarrow{\mathrm{lu}} \Phi(\alpha)$ whenever $\alpha_n \xrightarrow{\mathrm{lu}} \alpha$ with $\alpha_n \in \mathbb{D}(\mathbb{R}^d)$, $n \in \mathbb{N}$. If $\Phi_n = \Phi$ for all $n \in \mathbb{N}$ then we write simply $C_\Phi$ instead of $C_{\Phi,(\Phi_n)_{n\in\mathbb{N}}}$. Further, $\widetilde{C}_{\Phi,(\Phi_n)_{n\in\mathbb{N}}}$ will denote the set of all functions $\alpha \in C_{\Phi,(\Phi_n)_{n\in\mathbb{N}}}$ such that $\Phi(\alpha_n) \xrightarrow{\mathrm{lu}} \Phi(\alpha)$ whenever $\alpha_n \xrightarrow{\mathrm{lu}} \alpha$ with $\alpha_n \in \mathbb{D}(\mathbb{R}^d)$, $n \in \mathbb{N}$. Finally, $D_{\Phi,(\Phi_n)_{n\in\mathbb{N}}}$ will denote the set of all functions $\alpha \in \mathbb{D}(\mathbb{R}^d)$ such that $\Phi_n(\alpha_n) \to \Phi(\alpha)$ in $\mathbb{D}(\mathbb{R}^p)$ whenever $\alpha_n \to \alpha$ in $\mathbb{D}(\mathbb{R}^d)$ with $\alpha_n \in \mathbb{D}(\mathbb{R}^d)$, $n \in \mathbb{N}$. We need the following version of the continuous mapping theorem several times.

**Lemma 3.1** *Let $(\mathcal{U}_t)_{t\in\mathbb{R}_+}$ and $(\mathcal{U}^n_t)_{t\in\mathbb{R}_+}$, $n \in \mathbb{N}$, be stochastic processes with values in $\mathbb{R}^d$ such that $\mathcal{U}^n \xrightarrow{\mathcal{L}} \mathcal{U}$. Let $\Phi : \mathbb{D}(\mathbb{R}^d) \to \mathbb{D}(\mathbb{R}^p)$ and $\Phi_n : \mathbb{D}(\mathbb{R}^d) \to \mathbb{D}(\mathbb{R}^p)$, $n \in \mathbb{N}$, be measurable mappings such that $\mathsf{P}(\mathcal{U} \in C_{\Phi,(\Phi_n)_{n\in\mathbb{N}}}) = 1$. Then $\Phi_n(\mathcal{U}^n) \xrightarrow{\mathcal{L}} \Phi(\mathcal{U})$.*

**Proof.** In view of the continuous mapping theorem (see, e.g., Billingsley [1, Theorem 5.5]), it suffices to check that $\mathsf{P}(\mathcal{U} \in D_{\Phi,(\Phi_n)_{n\in\mathbb{N}}}) = 1$. For a function $\alpha \in \mathbb{C}(\mathbb{R}^d)$, $\alpha_n \to \alpha$ in $\mathbb{D}(\mathbb{R}^d)$ if and only if $\alpha_n \xrightarrow{\mathrm{lu}} \alpha$ (see, e.g., Jacod and Shiryaev [5, VI.1.17]). Consequently, $\mathbb{C}(\mathbb{R}^d) \cap \Phi^{-1}(\mathbb{C}(\mathbb{R}^p)) \cap D_{\Phi,(\Phi_n)_{n\in\mathbb{N}}} = C_{\Phi,(\Phi_n)_{n\in\mathbb{N}}}$ implying $D_{\Phi,(\Phi_n)_{n\in\mathbb{N}}} \supset C_{\Phi,(\Phi_n)_{n\in\mathbb{N}}}$. □

**Theorem 3.2** *Let $\beta$, $\gamma$, $\eta$, $(U^n_k)_{k\in\mathbb{Z}_+}$, $(\mathcal{F}^n_k)_{k\in\mathbb{Z}_+}$ and $\mathcal{U}^n$ for $n \in \mathbb{N}$ be as in Theorem 2.1. Assume that the SDE (2.1) has a unique weak solution with $\mathcal{U}_0 = u_0$ for all $u_0 \in \mathbb{R}^d$. Let $(\mathcal{U}_t)_{t\in\mathbb{R}_+}$ be a solution of (2.1) with initial distribution $\eta$.*

*For each $n \in \mathbb{N}$ and $k \in \mathbb{Z}_+$, let $\psi_{n,k} : (\mathbb{R}^d)^{k+1} \to \mathbb{R}^p$ be a Borel function, and let $\Psi_n : \mathbb{D}(\mathbb{R}^d) \to \mathbb{D}(\mathbb{R}^p)$ be defined by*

$$\Psi_n(\alpha)(t) := \psi_{n,\lfloor nt\rfloor}\Big(\alpha\big(\tfrac{1}{n}\big) - \alpha(0), \dots, \alpha\big(\tfrac{\lfloor nt\rfloor}{n}\big) - \alpha\big(\tfrac{\lfloor nt\rfloor - 1}{n}\big)\Big)$$

*for $\alpha \in \mathbb{D}(\mathbb{R}^d)$. Let*

$$\begin{aligned}
V^n_k &:= \psi_{n,k}(U^n_0, \dots, U^n_k), \qquad k \in \mathbb{Z}_+, \quad n \in \mathbb{N},\\
\mathcal{V}^n_t &:= V^n_{\lfloor nt\rfloor} = \Psi_n(\mathcal{U}^n)_t, \qquad t \in \mathbb{R}_+, \quad n \in \mathbb{N},\\
\mathcal{Y}^n_t &:= \sum_{k=1}^{\lfloor nt\rfloor} V^n_{k-1} \otimes U^n_k = \int_0^t \mathcal{V}^n_{s-} \otimes \mathrm{d}\mathcal{U}^n_s, \qquad t \in \mathbb{R}_+, \quad n \in \mathbb{N}.
\end{aligned}$$

*Let $\Psi : \mathbb{D}(\mathbb{R}^d) \to \mathbb{D}(\mathbb{R}^p)$ be a measurable mapping such that $\mathsf{P}(\mathcal{U} \in \widetilde{C}_{\Psi,(\Psi_n)_{n\in\mathbb{N}}}) = 1$. Let*

$$\mathcal{V}_t := \Psi(\mathcal{U})_t, \qquad \mathcal{Y}_t := \int_0^t \mathcal{V}_{s-} \otimes \mathrm{d}\mathcal{U}_s, \qquad t \in \mathbb{R}_+.$$

*Let the mappings $\beta' : \mathbb{D}(\mathbb{R}^d) \to \mathbb{D}(\mathbb{R}^d \times \mathbb{R}^{pd})$ and $\gamma' : \mathbb{D}(\mathbb{R}^d) \to \mathbb{D}(\mathbb{R}^{d\times r} \times \mathbb{R}^{(pd)\times r})$ be defined by*

$$\beta'(\alpha)(s) := \begin{bmatrix} \beta(s, \alpha(s)) \\ \Psi(\alpha)(s) \otimes \beta(s, \alpha(s)) \end{bmatrix}, \quad \gamma'(\alpha)(s) := \begin{bmatrix} \gamma(s, \alpha(s)) \\ \Psi(\alpha)(s) \otimes \gamma(s, \alpha(s)) \end{bmatrix}.$$

*Let $h' : \mathbb{R}^d \times \mathbb{R}^{pd} \to \mathbb{R}^d \times \mathbb{R}^{pd}$ be a continuous function with compact support satisfying $h'(x) = x$ in a neighborhood of 0. Suppose $U^n_0 \xrightarrow{\mathcal{L}} \eta$, and for each $T > 0$,*

(i) $\displaystyle\sup_{t\in[0,T]} \left\| \sum_{k=1}^{\lfloor nt\rfloor} \mathsf{E}(h'(U^n_k, V^n_{k-1} \otimes U^n_k) \mid \mathcal{F}^n_{k-1}) - \int_0^t \beta'(\mathcal{U}^n)_s \,\mathrm{d}s \right\| \xrightarrow{\mathsf{P}} 0,$

(ii) $\sup_{t\in[0,T]}\left\|\sum_{k=1}^{\lfloor nt\rfloor}\mathrm{Var}(h'(U_k^n, V_{k-1}^n\otimes U_k^n)\,|\,\mathcal{F}_{k-1}^n)-\int_0^t\gamma'(\mathcal{U}^n)_s\gamma'(\mathcal{U}^n)_s^\top\mathrm{d}s\right\| \stackrel{\mathsf{P}}{\longrightarrow} 0,$

(iii) $\sum_{k=1}^{\lfloor nT\rfloor}\mathsf{P}\big(\|U_k^n\|(1+\|V_{k-1}^n\|)>\theta\,\big|\,\mathcal{F}_{k-1}^n\big) \stackrel{\mathsf{P}}{\longrightarrow} 0$ *for all* $\theta>0$.

*Then* $(\mathcal{U}^n,\mathcal{V}^n,\mathcal{Y}^n) \stackrel{\mathcal{L}}{\longrightarrow} (\mathcal{U},\mathcal{V},\mathcal{Y})$ *as* $n\to\infty$.

**Proof.** Our first aim is to prove $(\mathcal{U}^n,\mathcal{Y}^n) \stackrel{\mathcal{L}}{\longrightarrow} (\mathcal{U},\mathcal{Y})$. We start by showing that the sequence $(\mathcal{U}^n,\mathcal{Y}^n)_{n\in\mathbb{N}}$ is tight in $\mathbb{D}(\mathbb{R}^d\times\mathbb{R}^{pd})$, and for this we will use Theorem VI.4.18 of Jacod and Shiryaev [5]. By the assumptions, the sequence $(\mathcal{U}_0^n,\mathcal{Y}_0^n)=(U_0^n,0)$, $n\in\mathbb{N}$, is weakly convergent, thus obviously tight in $\mathbb{R}^d\times\mathbb{R}^{pd}$, hence condition (i) of Theorem VI.4.18 of Jacod and Shiryaev [5] holds. For each $n\in\mathbb{N}$, the process $(\mathcal{U}_t^n,\mathcal{Y}_t^n)_{t\in\mathbb{R}_+}$ is a semimartingale with characteristics $(\mathcal{B}'^n,\mathcal{C}'^n,\nu'^n)$ relative to the truncation function $h'$ given by

$$\begin{aligned}
\mathcal{B}_t'^n &:= \sum_{k=1}^{\lfloor nt\rfloor}\mathsf{E}(h'(U_k^n, V_{k-1}^n\otimes U_k^n)\mid\mathcal{F}_{k-1}^n),\\
\mathcal{C}_t'^n &:= 0,\\
\nu'^n([0,t]\times g) &:= \sum_{k=1}^{\lfloor nt\rfloor}\mathsf{E}\Big(g(U_k^n, V_{k-1}^n\otimes U_k^n)\mathbb{1}_{\{(U_k^n,V_{k-1}^n\otimes U_k^n)\neq 0\}}\mid\mathcal{F}_{k-1}^n\Big)
\end{aligned}$$

for $g:\mathbb{R}^d\times\mathbb{R}^{pd}\to\mathbb{R}_+$ Borel functions, and modified second characteristic

$$\widetilde{\mathcal{C}}_t'^n := \sum_{k=1}^{\lfloor nt\rfloor}\mathrm{Var}(h'(U_k^n, V_{k-1}^n\otimes U_k^n)\mid\mathcal{F}_{k-1}^n)$$

(see Jacod and Shiryaev [5, II.3.14, II.3.18]). For all $T>0$, $\theta>0$, $\varepsilon>0$,

$$\begin{aligned}
&\mathsf{P}\Big(\nu'^n\big([0,T]\times\mathbb{1}_{\{\|x\|>\theta\}}\big)>\varepsilon\Big)\\
&=\mathsf{P}\Bigg(\sum_{k=1}^{\lfloor nT\rfloor}\mathsf{P}(\|(U_k^n, V_{k-1}^n\otimes U_k^n)\|>\theta\mid\mathcal{F}_{k-1}^n)>\varepsilon\Bigg)\to 0
\end{aligned}$$

by assumption (iii), hence condition (ii) of Theorem VI.4.18 of Jacod and Shiryaev [5] holds.

In order to check condition (iii) of Theorem VI.4.18 of Jacod and Shiryaev [5], first we will show

$$\int_0^t\beta'(\mathcal{U}^n)_s\,\mathrm{d}s \stackrel{\mathcal{L}}{\longrightarrow} \int_0^t\beta'(\mathcal{U})_s\,\mathrm{d}s \qquad \text{in } \mathbb{D}(\mathbb{R}^d\times\mathbb{R}^{pd}). \tag{3.1}$$

We will apply Lemma 3.1. We have $\int_0^t\beta'(\mathcal{U})_s\,\mathrm{d}s=\Phi_{\beta'}(\mathcal{U})_t$, and for each $n\in\mathbb{N}$, $\int_0^t\beta'(\mathcal{U}^n)_s\,\mathrm{d}s=\Phi_{\beta'}(\mathcal{U}^n)_t$ with the measurable mapping $\Phi_{\beta'}:\mathbb{D}(\mathbb{R}^d)\to\mathbb{D}(\mathbb{R}^d\times\mathbb{R}^{pd})$ given by

$$\Phi_{\beta'}(\alpha)(t) := \int_0^t\beta'(\alpha)(s)\,\mathrm{d}s, \qquad \alpha\in\mathbb{D}(\mathbb{R}^d), \quad t\in\mathbb{R}_+.$$

Observe that assumptions (i)–(iii) imply that conditions (i)–(iii) of Theorem 2.1 hold, thus we conclude $\mathcal{U}^n \stackrel{\mathcal{L}}{\longrightarrow} \mathcal{U}$ as $n\to\infty$. In order to show $\mathsf{P}\big(\mathcal{U}\in C_{\Phi_{\beta'}}\big)=1$, it is enough to check $C_{\Phi_{\beta'}}\supset\widetilde{C}_{\Psi,(\Psi_n)_{n\in\mathbb{N}}}$. Clearly $\Phi_{\beta'}(\mathbb{C}(\mathbb{R}^d))\subset\mathbb{C}(\mathbb{R}^d\times\mathbb{R}^{pd})$. Now we fix $T>0$, a function $\alpha\in\widetilde{C}_{\Psi,(\Psi_n)_{n\in\mathbb{N}}}$ and a sequence $(\alpha_n)_{n\in\mathbb{N}}$ in $\mathbb{D}(\mathbb{R}^d)$ with $\alpha_n \stackrel{\mathrm{lu}}{\longrightarrow} \alpha$. Obviously

$$\sup_{t\in[0,T]}\|\Phi_{\beta'}(\alpha_n)-\Phi_{\beta'}(\alpha)\| \leqslant T\sup_{t\in[0,T]}\|\beta'(\alpha_n)(t)-\beta'(\alpha)(t)\|,$$

hence it suffices to show

$$\sup_{t\in[0,T]}\|\beta(t,\alpha_n(t))-\beta(t,\alpha(t))\|\to 0, \tag{3.2}$$

$$\sup_{t\in[0,T]}\|\Psi(\alpha_n)(t)\otimes\beta(t,\alpha_n(t))-\Psi(\alpha)(t)\otimes\beta(t,\alpha(t))\|\to 0. \tag{3.3}$$

For sufficiently large $n \in \mathbb{N}$, we have $\sup_{t\in[0,T]} \|\alpha_n(t) - \alpha(t)\| \leqslant 1$, thus $\sup_{t\in[0,T]} \|\alpha_n(t)\| \leqslant 1 + \sup_{t\in[0,T]} \|\alpha(t)\| < \infty$. The function $\beta$ is uniformly continuous on the compact set $[0,T] \times \{x \in \mathbb{R}^d : \|x\| \leqslant 1 + \sup_{t\in[0,T]} \|\alpha(t)\|\}$, hence (3.2) holds. Moreover,

$$\begin{aligned}&\|\Psi(\alpha_n)(t) \otimes \beta(t,\alpha_n(t)) - \Psi(\alpha)(t) \otimes \beta(t,\alpha(t))\| \\ &\leqslant \|\Psi(\alpha_n)(t) - \Psi(\alpha)(t)\|\|\beta(t,\alpha_n(t))\| + \|\beta(t,\alpha_n(t)) - \beta(t,\alpha(t))\|\|\Psi(\alpha)(t)\|.\end{aligned}$$

Continuity of $\Psi(\alpha)$ implies $\sup_{t\in[0,T]} \|\Psi(\alpha)(t)\| < \infty$. For sufficiently large $n \in \mathbb{N}$, $\sup_{t\in[0,T]} \|\beta(t,\alpha_n(t))\| \leqslant 1 + \sup_{t\in[0,T]} \|\beta(t,\alpha(t))\| < \infty$ (by convergence (3.2) and by continuity of $\alpha$ and $\beta$). By $\Psi(\alpha_n) \xrightarrow{\mathrm{lu}} \Psi(\alpha)$, (3.3) is also satisfied, and we conclude $C_{\Phi_{\beta'}} \supset \widetilde{C}_{\Psi,(\Psi_n)_{n\in\mathbb{N}}}$. Consequently, $\mathsf{P}(\mathcal{U} \in C_{\Phi_{\beta'}}) = 1$, and by Lemma 3.1, we obtain (3.1). If $\alpha \in \mathbb{C}(\mathbb{R}^d)$ and $(\alpha_n)_{n\in\mathbb{N}}$ is a sequence in $\mathbb{D}(\mathbb{R}^d)$ with $\alpha_n \xrightarrow{\mathrm{lu}} \alpha$ then for all $T > 0$, $\sup_{t\in[0,T]} \|\alpha_n(t)\| \to \sup_{t\in[0,T]} \|\alpha(t)\|$ as $n \to \infty$. (See, e.g., Proposition VI.2.4 of Jacod and Shiryaev [5].) Hence, by the continuous mapping theorem, we obtain

$$\sup_{t\in[0,T]} \left\| \int_0^t \beta'(\mathcal{U}^n)_s \,\mathrm{d}s - \int_0^t \beta'(\mathcal{U})_s \,\mathrm{d}s \right\| \xrightarrow{\mathcal{L}} 0 \qquad \text{as } n \to \infty.$$

This together with assumption (i) implies

$$\sup_{t\in[0,T]} \left\| \mathcal{B}'^n_t - \int_0^t \beta'(\mathcal{U})_s \,\mathrm{d}s \right\| \xrightarrow{\mathsf{P}} 0 \qquad \text{as } n \to \infty \text{ for all } T > 0. \tag{3.4}$$

Particularly, the sequence $(\mathcal{B}'^n)_{n\in\mathbb{N}}$ is $C$-tight in $\mathbb{D}(\mathbb{R}^d \times \mathbb{R}^{pd})$. Indeed, the Skorokhod topology is coarser than the local uniform topology, hence (3.4) implies $\varrho(\mathcal{B}'^n, \Psi_{\beta'}(\mathcal{U})) \xrightarrow{\mathsf{P}} 0$, where $\varrho$ denotes a distance on $\mathbb{D}(\mathbb{R}^d)$ compatible with the Skorokhod topology. Consequently, $\mathcal{B}'^n \xrightarrow{\mathcal{L}} \Psi_{\beta'}(\mathcal{U})$ with $\mathsf{P}(\Psi_{\beta'}(\mathcal{U}) \in \mathbb{C}(\mathbb{R}^d \times \mathbb{R}^{pd})) = 1$. In a similar way, the sequence $(\widetilde{\mathcal{C}}'^n)_{n\in\mathbb{N}}$ is $C$-tight in $\mathbb{D}(\mathbb{R}^{d\times r} \times \mathbb{R}^{(pd)\times r})$. Moreover, assumption (iii) yields

$$\nu'^n([0,T] \times g_c) \xrightarrow{\mathsf{P}} 0 \qquad \text{as } n \to \infty \tag{3.5}$$

for all $T > 0$ and all $c > 0$, where $g_c : \mathbb{R}^d \times \mathbb{R}^{pd} \to \mathbb{R}_+$ is defined by (2.2). Therefore all hypotheses of Theorem VI.4.18 of Jacod and Shiryaev [5] are fulfilled, hence we conclude that the sequence $(\mathcal{U}^n, \mathcal{Y}^n)_{n\in\mathbb{N}}$ is tight in $\mathbb{D}(\mathbb{R}^d \times \mathbb{R}^{pd})$.

It remains to prove that if a sub-sequence, still denoted by $(\mathcal{U}^n, \mathcal{Y}^n)_{n\in\mathbb{N}}$, weakly converges to a limit distribution then the limit is the distribution of $(\mathcal{U}, \mathcal{Y})$. For this we will apply Theorem IX.2.22 of Jacod and Shiryaev [5]. The process $(\mathcal{U}_t, \mathcal{Y}_t)_{t\in\mathbb{R}_+}$ is a semimartingale with characteristics $(\mathcal{B}', \mathcal{C}', 0)$, where

$$\mathcal{B}'_t := \int_0^t \beta'(\mathcal{U})_s \,\mathrm{d}s, \qquad \mathcal{C}'_t := \int_0^t \gamma'(\mathcal{U})_s \gamma'(\mathcal{U})_s^\top \mathrm{d}s$$

(see Jacod and Shiryaev [5, IX.5.3]). By Remark IX.2.23 of Jacod and Shiryaev [5], assumptions (i)–(iii) of Theorem 3.2 imply that condition (i) of Theorem IX.2.22 in [5] is met. To prove the continuity condition (ii) of Theorem IX.2.22 in [5], consider the measurable mapping $\Phi : \mathbb{D}(\mathbb{R}^d) \to \mathbb{D}\big(\mathbb{R}^d \times (\mathbb{R}^d \times \mathbb{R}^{pd}) \times (\mathbb{R}^{d\times r} \times \mathbb{R}^{(pd)\times r})\big)$ given by

$$\Phi(\alpha)(t) := \big(\alpha(t), \Phi_{\beta'}(\alpha)(t), \Phi_{\gamma'}(\alpha)(t)\big), \qquad \alpha \in \mathbb{D}(\mathbb{R}^d), \quad t \in \mathbb{R}_+.$$

As we have already proved, $\mathsf{P}\big(\mathcal{U} \in C_{\Phi_{\beta'}} \cap C_{\Phi_{\gamma'}}\big) = 1$. The local uniform topology on $\mathbb{D}(\mathbb{R}^m)$ is the $m$-fold product of the local uniform topology on $\mathbb{D}(\mathbb{R})$, hence we obtain $C_\Phi \supset C_{\Phi_{\beta'}} \cap C_{\Phi_{\gamma'}}$. Using again that the Skorokhod topology is coarser than the local uniform topology, we conclude $D_\Phi \supset C_\Phi$. Consequently, the continuity condition $\mathsf{P}(\mathcal{U} \in D_\Phi) = 1$ holds. Hence all hypotheses of Theorem IX.2.22 of Jacod and Shiryaev [5] are met, therefore $(\mathcal{U}^n, \mathcal{Y}^n) \xrightarrow{\mathcal{L}} (\mathcal{U}, \mathcal{Y})$. Again by Lemma (3.1), we obtain $(\mathcal{U}^n, \mathcal{V}^n, \mathcal{Y}^n) \xrightarrow{\mathcal{L}} (\mathcal{U}, \mathcal{V}, \mathcal{Y})$. □

**Corollary 3.3** *Let $\beta$, $\gamma$, $\eta$, $(U_k^n)_{k\in\mathbb{Z}_+}$, $(\mathcal{F}_k^n)_{k\in\mathbb{Z}_+}$ and $\mathcal{U}^n$ for $n \in \mathbb{N}$ be as in Theorem 2.1. Suppose that $\mathsf{E}\big(\|U_k^n\|^2 \,\big|\, \mathcal{F}_{k-1}^n\big) < \infty$ for all $n, k \in \mathbb{N}$. Assume that the SDE (2.1) has a unique weak solution with $\mathcal{U}_0 = u_0$ for all $u_0 \in \mathbb{R}^d$. Let $(\mathcal{U}_t)_{t\in\mathbb{R}_+}$ be a solution with initial distribution $\eta$. Let $\Psi$, $\mathcal{V}$, $\mathcal{Y}$, $\beta'$, $\gamma'$, $(\psi_{n,k})_{k\in\mathbb{N}}$, $\Psi_n$, $(V_k^n)_{k\in\mathbb{Z}_+}$, $\mathcal{V}^n$ and $\mathcal{Y}^n$ for $n \in \mathbb{N}$ be as in Theorem 3.2. Suppose that $\mathsf{P}\big(\mathcal{U} \in \widetilde{C}_{\Psi,(\Psi_n)_{n\in\mathbb{N}}}\big) = 1$. Suppose $U_0^n \xrightarrow{\mathcal{L}} \eta$, and for each $T > 0$,*

(i) $\sup_{t\in[0,T]} \left\| \sum_{k=1}^{\lfloor nt \rfloor} \mathsf{E}\left( \begin{bmatrix} U_k^n \\ V_{k-1}^n \otimes U_k^n) \end{bmatrix} \,\Big|\, \mathcal{F}_{k-1}^n \right) - \int_0^t \beta'(\mathcal{U}^n)_s \, \mathrm{d}s \right\| \xrightarrow{\mathsf{P}} 0,$

(ii) $\sup_{t\in[0,T]} \left\| \sum_{k=1}^{\lfloor nt \rfloor} \mathrm{Var}\left( \begin{bmatrix} U_k^n \\ V_{k-1}^n \otimes U_k^n) \end{bmatrix} \,\Big|\, \mathcal{F}_{k-1}^n \right) - \int_0^t \gamma'(\mathcal{U}^n)_s \gamma'(\mathcal{U}^n)_s^\top \mathrm{d}s \right\| \xrightarrow{\mathsf{P}} 0,$

(iii) $\sum_{k=1}^{\lfloor nT \rfloor} \mathsf{E}\Big( \|U_k^n\|^2 (1+\|V_{k-1}^n\|^2) \mathbb{1}_{\{\|U_k^n\|(1+\|V_{k-1}^n\|)>\theta\}} \Big| \mathcal{F}_{k-1}^n \Big) \xrightarrow{\mathsf{P}} 0$ *for all* $\theta > 0$.

*Then* $(\mathcal{U}^n, \mathcal{V}^n, \mathcal{Y}^n) \xrightarrow{\mathcal{L}} (\mathcal{U}, \mathcal{V}, \mathcal{Y})$.

**Proof.** This follows from Theorem 3.2 in the same way as Corollary 2.2 from Theorem 2.1. □

**Acknowledgements.** The authors have been supported by the Hungarian Scientific Research Fund under Grant No. OTKA T-048544 and OTKA T-079128.

# References

[1] P. Billingsley, *Convergence of Probability Measures* (John Wiley & Sons, Inc., New York-London-Sydney, 1968). MR0233396

[2] S. N. Ethier and T. G. Kurtz, *Markov processes. Characterization and convergence* (Wiley, New York, 1986). MR0838085

[3] I. I. Gikhman and A. V. Skorokhod, *Introduction to the theory of random processes* (W. B. Saunders Co., Philadelphia, Pa.-London-Toronto, Ont., 1969). MR0247660

[4] M. Ispány and G. Pap, Weak convergence of step processes and an application for critical multitype branching processes with immigration (2007). http://arxiv.org/abs/math.PR/0701803

[5] J. Jacod and A. N. Shiryaev, *Limit Theorems for Stochastic Processes*, 2nd ed. (Springer-Verlag, Berlin, 2003). MR1943877

[6] A. Jakubowski, J. Mémin and G. Pagès, Convergence en loi des suites d'intégrales stochastiques sur l'espace $D^1$ de Skorokhod, *Probab. Theory Related Fields*, **81** (1989), 111–137. MR0981569

[7] A. Joffe and M. Métivier, Weak convergence of sequences of semimartingales with applications to multitype branching processes, *Adv. in Appl. Probab.*, **18** (1986), 20–65. MR0827331

[8] T. G. Kurtz and Ph. E. Protter, Weak limit theorems for stochastic integrals and stochastic differential equations, *Ann. Probab.*, **19** (1991), 1035–1070. MR1112406

[9] T. G. Kurtz and Ph. E. Protter, Weak convergence of stochastic integrals and differential equations, *Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995)*, 1–41, (Lecture Notes in Math., 1627, Springer, Berlin, 1996). MR1431298

[10] T. G. Kurtz and Ph. E. Protter, Weak convergence of stochastic integrals and differential equations. *Working paper*, 2004,
http://www.orie.cornell.edu/~protter/WebPapers/KPwkConvStochIntI.pdf