The Annals of Probability
1998, Vol. 26, No. 1, 267–307

ASYMPTOTIC ERROR DISTRIBUTIONS FOR THE EULER METHOD FOR STOCHASTIC DIFFERENTIAL EQUATIONS

By Jean Jacod and Philip Protter[1]

Université Pierre et Marie Curie and Purdue University

We are interested in the rate of convergence of the Euler scheme approximation of the solution to a stochastic differential equation driven by a general (possibly discontinuous) semimartingale, and by the asymptotic behavior of the associated normalized error. It is well known that for Itô's equations the rate is $1/\sqrt{n}$; we provide a necessary and sufficient condition for this rate to be $1/\sqrt{n}$ when the driving semimartingale is a continuous martingale, or a continuous semimartingale under a mild additional assumption; we also prove that in these cases the normalized error processes converge in law.

The rate can also differ from $1/\sqrt{n}$: this is the case for instance if the driving process is deterministic, or if it is a Lévy process without a Brownian component. It is again $1/\sqrt{n}$ when the driving process is Lévy with a nonvanishing Brownian component, but then the normalized error processes converge in law in the finite-dimensional sense only, while the discretized normalized error processes converge in law in the Skorohod sense, and the limit is given an explicit form.

1. Introduction. The classical Itô-type stochastic differential equation (SDE) is of the form

$$X_t = x_0 + \int_0^t a(X_s)\,dW_s + \int_0^t b(X_s)\,ds \tag{1.1}$$

with a, b matrices of functions and W a multidimensional Brownian motion. By replacing dW_t and dt with a vector of semimartingales dY_t we consider the more general equation

$$X_t = x_0 + \int_0^t f(X_{s-})\,dY_s, \tag{1.2}$$

where f denotes a matrix $f = (f^{ij})$ of functions. In applications one often wants to solve (1.2) numerically, when possible. Because of simulation difficulties, and because one often combines a numerical solution of (1.2) with a (slow) Monte Carlo technique, it is usually advisable to solve (1.2) numerically with an Euler scheme, rather than a more complicated, faster one. (See the survey paper of Talay [16] for a discussion of this issue.)

Without loss we will take the time interval to be $[0, 1]$ rather than $[0, T]$ for some (nonrandom) $T > 0$. We will assume $[0, 1]$ is partitioned by $\Pi^n =$

Received May 1996; revised April 1997.

[1]Supported in part by ONR Grant N00014-96-1-0262, NSF Grant INT-94-01109 and a Purdue University Global Initiative Faculty Grant.

AMS 1991 *subject classifications*. Primary 60H10, 65U05; secondary 60G44, 60F17.

Key words and phrases. Stochastic differential equations, Euler scheme, error distributions, Lévy processes, numerical approximation.

$\{0 = t_0 < t_1 < \cdots < t_n = 1\}$ with $t_i = i/n$, $0 \le i \le n$. Rates of convergence will thus be given relative to this partition scheme. For equation (1.1) if $a \equiv 0$, then the rate of convergence of the Euler scheme is classically $1/n$; if a does not vanish, then it is also classical that the rate is $1/\sqrt{n}$. The distribution of the (normalized) asymptotic error, however, is not at all classical and was established only recently for (1.1) (see [8]).

In this paper we mainly aim to give a class of equations of type (1.2) that converge at the rate $1/\sqrt{n}$ and determine their asymptotic error, although we also examine some equations providing the rate $1/n$. To give a flavor of our results in a very simple setting, consider the one-dimensional case (for Y and X as well) when Y is continuous and is either (1) nondecreasing or (2) a local martingale. Denote by X^n the "continuous" Euler approximation for (1.2) and by $\overline{X}^n$ the "discretized" one (see Section 3 for the definitions), so the error processes are, respectively, $U^n = X^n - X$ and $\overline{U}^n = \overline{X}^n - X$.

The first situation corresponds to a purely deterministic problem:

THEOREM 1.1. *If Y is a nondecreasing continuous function, there is equivalence between the following:*

(a) *For $x_0 = 1$ and $f(x) = x$ [i.e., $X = e^Y$ in (1.2)], the sequence of numbers $nU_1^n = n\overline{U}_1^n$ is bounded.*

(b) *For all starting points x_0 and all C^1 functions f with at most linear growth, the functions nU^n and $n\overline{U}^n$ converge uniformly to a limit U.*

(c) *The function Y has the form*

$$Y_t = \int_0^t y_s\,ds \quad \textit{with} \int_0^1 y_s^2\,ds < \infty.$$

In this case, the limiting process U is the solution of the following linear equation:

$$U_t = \int_0^t f'(X_s)U_s y_s\,ds - \tfrac{1}{2}\int_0^t f(X_s)f'(X_s)y_s^2\,ds. \tag{1.3}$$

This covers in particular the case of an ordinary differential equation of the form

$$dX_t = f(X_t)y_t\,dt,$$

where the coefficient f is C^1 and $s \to y_s$ is a given nonnegative function. Then the Euler approximation converges at the rate $1/n$ on the interval $[0,1]$ *if and only if* we have $\int_0^1 y_s^2\,ds < \infty$, which seems to be a new result.

In the second situation, we denote by C the quadratic variation process of Y. We then have the following:

THEOREM 1.2. *If Y is a continuous local martingale, there is equivalence between the following:*

(a) *For all $x_0 = 1$ and $f(x) = x$ [i.e., $X = \mathscr{E}(Y)$, the Doléans exponential of X], the sequence of random variables* $\sup_t |\sqrt{n}\,U_t^n|$ *is tight.*

(b) *For all starting points* x_0 *and all* C^1*-functions* f *with at most linear growth, the processes* $(Y, \sqrt{n}\,U^n)$ *and* $(Y, \sqrt{n}\,\overline{U}^n)$ *converge in law to a limit* (Y, U).

(c) *The quadratic variation has the form*

$$C_t = \int_0^t c_s\,ds \quad \text{with } \int_0^1 c_s^2\,ds < \infty.$$

In this case, the limiting process U *is the solution of the following linear equation:*

$$U_t = \int_0^t f'(X_s)U_s\,dY_s - \frac{1}{\sqrt{2}}\int_0^t f(X_s)f'(X_s)c_s\,dW_s, \tag{1.4}$$

where W *is a standard Brownian motion, independent of* Y.

Note that in (1.4) we have some "additional" randomness provided by the extra Brownian motion W: this is a typical feature of the limiting error process, when the driving term Y itself is random.

Surprisingly, the situation is very different when the driving term Y is discontinuous. Consider, for example, the case where Y is a one-dimensional discontinuous Lévy process. Then two situations occur. First, if there is no Brownian part, then $\sqrt{n}\,U^n$ and $\sqrt{n}\,\overline{U}^n$ converge in law to 0, which means that the rate is faster than $1/\sqrt{n}$ (but we do not know the correct rate, or even if there is a rate at all). Second, if there is a Brownian part in Y, then $\sqrt{n}\,\overline{U}^n$ converges to a limit U, but $\sqrt{n}\,U^n$ does *not* converge in the usual sense (i.e., for the Skorohod topology on the set of càdlàg functions). It does converge to U, however, for weaker topologies: the one induced by convergence in (Lebesgue) measure, which is known as the Meyer–Zheng topology [11], and also the new S-topology introduced by Jakubowski [5].

The paper is organized as follows: In Section 2 some preliminaries are given, and this section may be skipped at first reading [except for the definitions of the so-called stable convergence and of the property ($\star$)]. Section 3 is devoted to general results (extending [8]) on rates of convergence. We have given in Section 4 some results in the case Y is of finite variation, because this is simpler than the general case while it shows already all the pathologies of this problem; this section may also be skipped, although it contains the proof of Theorem 1.1. Section 5 is devoted to continuous semimartingales, and this is the most useful part of this paper as far as applications are concerned, and it contains the proof of Theorem 1.2. Finally, the case of Lévy processes is considered in Section 6.

2. Preliminaries.

In this paper we will mainly be dealing with weak convergence in the Skorohod topology: weak convergence for this topology is denoted by "$\Rightarrow$". We need to give a review of and some complements to weak convergence.

First we recall some facts about *stable convergence*. Let X_n be a sequence of random variables with values in a Polish space E, all defined on the same probability space $(\Omega, \mathscr{F}, P)$. We say that X_n *converges stably in law* to X, written "$X_n \Rightarrow^{\text{stably}} X$", if X is an E-valued random variable defined on an extension $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P})$ of the original space and if

$$\lim_n E(Uf(X_n)) = \tilde{E}(Uf(X)) \tag{2.1}$$

for every bounded continuous $f: E \to \mathbb{R}$ and all bounded measurable random variables U. This convergence was introduced by Rényi [13] and studied by Aldous and Eagleson [1]; see also [4]. It is obviously stronger than convergence in law.

If Y is another variable with values in another Polish space F, we have the following equivalence:

LEMMA 2.1. *If $X_n \Rightarrow^{\text{stably}} X$, then we have $(Y, X_n) \Rightarrow^{\text{stably}} (Y, X)$ for the product topology on $E \times F$.*

Conversely, if (Y, X_n) weakly converges to a limit, we can realize this limit as (Y, X) with X defined on an extension of the space on which Y is defined, and $X_n \Rightarrow^{\text{stably}} X$ as soon as Y generates the σ-field $\mathscr{F}$.

PROOF. The first claim is trivial. Conversely, assume that (Y, X_n) weakly converges to a limit (Y', X'). Call $Q(y, dx)$ a version of the regular conditional distribution of X' given Y'. Set $\tilde{\Omega} = \Omega \times E$, and $\tilde{\mathscr{F}} = \mathscr{F} \otimes \mathscr{E}$, where $\mathscr{E}$ is the Borel σ-field of E, and $\tilde{P}(d\omega, dx) = P(d\omega)Q(Y(\omega), dx)$. We thus define an extension of the original space, with the "canonical" variable $X(\omega, x) = x$, and the pairs (Y, X) and (Y', X') clearly have the same law.

Observe that (2.1) holds for all $U = g(Y)$, where g is continuous and bounded on F, and what we need to prove is that it holds when g is measurable and bounded. However, we then can find a sequence g_q of bounded continuous functions such that $g_q(Y) \to g(Y)$ in $\mathbb{L}^1(P)$, and the result readily follows. □

Note that all this applies when X_n, X are $\mathbb{R}^d$-valued processes with càdlàg paths, as well as Y: we can then view them as random variables with values in the Skorohod space $\mathbb{D}$. However, in this situation we should be careful: the stable convergence of X_n implies the weak convergence of the pair (Y, X_n) for the product topology on $\mathbb{D}(\mathbb{R}^q) \times \mathbb{D}(\mathbb{R}^d)$, which is *not* the Skorohod topology on $\mathbb{D}(\mathbb{R}^{q+d})$, and we do not have in general weak convergence of (Y, X_n) in the usual sense.

Next, we prove a result on weak convergence and discretization which might be well known, but we could not find it in the literature. First, a standard result asserts that if x is a function belonging to $\mathbb{D}$ and if η_n is a sequence of increasing piecewise constant and right-continuous functions from $[0, 1]$ into $[0, 1]$ which converges to the identity, then the sequence of "discretized" functions $x \circ \eta_n$ converges to x. More generally, we have the following lemma.

LEMMA 2.2. *If a sequence X^n of (possibly multidimensional) processes weakly converges to X, then the processes $X^n_{\eta_n}$ weakly converge to the same limit.*

PROOF. By the Skorohod representation theorem, we can replace weak convergence by a.s. convergence, so that we only need to prove that if $x_n \to x$ in $\mathbb{D}$, then the sequence $y_n = x_n \circ \eta_n$ also converges to x. There are time-changes λ_n converging to the identity and such that $x_n - x \circ \lambda_n$ goes uniformly to 0. Then $y_n - x \circ \lambda'_n$ also goes uniformly to 0, where $\lambda'_n = \lambda_n \circ \eta_n$. Now we have recalled before stating the lemma that $x \circ \lambda'_n \to x$ in $\mathbb{D}$; since for the Skorohod topology we have $z_n + z'_n \to z + z'$ as soon as $z_n \to z$ and $z'_n \to z'$ *and* z is a continuous function, we are clearly finished. □

Next, we recall some facts about convergence of stochastic integrals, coming from the work of [6] and [7]. See [9] for an expository account. First recall that, for every $\delta > 0$, any semimartingale can be written as

$$X_t = X_0 + A(\delta)_t + M(\delta)_t + \sum_{s \le t} \Delta X_s 1_{\{|\Delta X_s| > \delta\}}, \tag{2.2}$$

where $A(\delta)$ is a predictable process with finite variation, null at 0, $M(\delta)$ is a local martingale null at 0, and ΔX_s denotes the jump size of X at time s. As usual $\langle M, M \rangle$ denotes the predictable bracket of two local martingales M and N, if it exists. All these notions are relative to some filtration $(\mathscr{F}_t)$ on our probability space. We also write, for any (possibly multidimensional) process V:

$$V^\star = \sup_{t \in [0,1]} \|V_t\|. \tag{2.3}$$

DEFINITION. Let $X^n = (X^{n,i})_{1 \le i \le d}$ be a sequence of $\mathbb{R}^d$-valued semimartingales, with $A(\delta)^{n,i}$ and $M(\delta)^{n,i}$ associated with $X^{n,i}$ as in (2.2). We say that the sequence (X^n) satisfies $(\star)$ if for some $\delta > 0$ and for each i the sequence

$$\langle M(\delta)^{n,i}, M(\delta)^{n,i} \rangle_1 + \int_0^1 |dA(\delta)^{n,i}_s| + \sum_{0 < s \le 1} |\Delta X^{n,i}_s| 1_{\{|\Delta X^{n,i}_s| > \delta\}}$$

is tight. This notion does not depend on the particular choice of $\delta > 0$ [recall that our time interval here is $[0, 1]$; it is important to emphazise that this notion does depend on the underlying filtrations $(\mathscr{F}^n_t)$].

It turns out that this property is equivalent to the notion of *uniform tightness* (UT) as introduced by Jakubowski, Mémin and Pagès [6] (see, e.g., [10] for the equivalence). Since the time interval here is $[0, 1]$, it is also equivalent to the condition of *uniformly controlled variation* (UCV) in [9]. Its usefulness derives from the following fundamental set of results (see, e.g., [9]). Below, we denote by $H \cdot X$ the stochastic integral process of H w.r.t. X, and it is understood that these two processes have matching dimensions.

THEOREM 2.3. *Let X^n and Y^n be two sequences of $\mathbb{R}^d$-valued semimartingales, relative to the filtrations $(\mathscr{F}_t^n)$.*

(a) *If both sequences X^n and Y^n have ($\star$), then so has the sequence $X^n + Y^n$.*

(b) *If each X^n is of finite variation and if the sequence $\int_0^1 |dX_s^n|$ is tight, then the sequence X^n has ($\star$).*

(c) *Let H^n be a sequence of $(\mathscr{F}_t^n)$-predictable processes such that the sequence $H^{n\star}$ is tight. If the sequence X^n has ($\star$), so has the sequence $H^n \cdot X^n$.*

(d) *Let H^n and H'^n be two sequences of $(\mathscr{F}_t^n)$-predictable processes such that the sequence $H^{n\star}$ is tight and that $(H^n - H'^n)^\star \to^P 0$. If the sequence X^n has ($\star$), then $(H^n \cdot X^n - H'^n \cdot X^n)^\star \to^P 0$.*

(e) *Suppose that X^n weakly converges. Then ($\star$) is necessary and sufficient for the following property (called goodness):*

> *For any sequence H^n of $(\mathscr{F}_t^n)$-adapted, right-continuous and left-hand limited processes such that the sequence (H^n, X^n) weakly converges to a limit (H, X), then X is a semimartingale w.r.t. the filtration generated by the process (H, X), and we have $(H^n, X^n, H_-^n \cdot X^n) \Rightarrow (H, X, H_- \cdot X)$.*

We finally turn our attention to stochastic differential equations. General results are available (see, e.g., [9], [14] and [15]), but we confine ourselves to linear equations of the type

$$X_t = J_t + \int_0^t X_{s-} H_s \, dY_s, \tag{2.4}$$

where Y is a given semimartingale, J is an adapted càdlàg process and H is a predictable process. All these terms can be multidimensional, with matching dimensions.

Let us begin with a comparison lemma, where X' is the solution of another equation (2.4) associated with J' and H', and with *the same* semimartingale Y.

LEMMA 2.4. *For all $\varepsilon > 0$, $A > 0$, there is a constant K depending on ε and A and on the semimartingale Y such that for all $\eta > 0$, $u > 0$, $v \in (0, A]$, $w \in (0, u]$ we have*

$$\begin{aligned} P((X - X')^\star > \eta) \le \varepsilon &+ P(H^\star > A) + P(J^\star > u) \\ &+ P((H - H')^\star > v) + P((J - J')^\star > w) \\ &+ \frac{uv + w}{\eta} K. \end{aligned} \tag{2.5}$$

PROOF. Let us first introduce notation: if Z is a càdlàg process and T a stopping time, we write Z^{T-} for the process $Z_t^{T-} = Z_t 1_{[0,T)}(t) + Z_{T-} 1_{[T,1]}(t)$.

We will use the "slicing technique" of Doléans-Dade (see [12]), which says three things. First, for any semimartingale Y and any $\alpha > 0$, $\varepsilon > 0$, there is a stopping time T such that the semimartingale Y^{T-} is α-sliceable and that $P(T \le 1) \le \varepsilon$. Second, if Y is α-sliceable for some α, then $E(\sup_t |\int_0^t H_s \, dY_s|) \le$

$K_Y E(H^\star)$, where K_Y only depends on Y. Third, if Y is α-sliceable for some α and if we consider (2.4) with $|H| \le A$, then $E(X^\star) \le K_{A,Y} E(J^\star)$ for a constant $K_{A,Y}$ depending on A, Y, provided $\alpha \le C_A$ for some $C_A > 0$ depending on A only.

Now we fix $A > 0$ and $\varepsilon > 0$, and we take $\alpha = C_A$. Then we choose a stopping time T such that $P(T \le 1) \le \varepsilon$ and that $\overline{Y} = Y^{T-}$ is α-sliceable. Then we set $S = \inf(t\colon |H_t| > A$ or $|J_t| > u$ or $|H_t - H'_t| > v$ or $|J_t - J'_t| > w) \wedge T$ and $\overline{J} = J^{S-}$, $\overline{J}' = J'^{S-}$, and we define the ith component of $\overline{H}$ as $\overline{H}^i = \overline{H}^i \wedge A \vee -A$, and similarly for $\overline{H}'$. These last two processes are predictable, and we can consider the solutions $\overline{X}$ and $\overline{X}'$ of (2.4), associated with $(\overline{J}, \overline{H}, \overline{Y})$ and $(\overline{J}', \overline{H}', \overline{Y})$, respectively. Note that

$$X = \overline{X}, \quad X' = \overline{X}' \quad \text{on the set } \{S > 1\}. \tag{2.6}$$

Note also that $\overline{X}'' = \overline{X}' - \overline{X}$ is the solution of (2.4) associated with $(\overline{J}'', \overline{H}, \overline{Y})$, where $\overline{J}''_t = \overline{J}'_t - \overline{J}_t + \int_0^t (\overline{H}'_s - \overline{H}_s)\overline{X}'_{s-}\, d\overline{Y}_s$. Using the properties of sliceable semimartingales recalled above, we get the following if $v \le A$ and $w \le u$:

$$E(\overline{X}''^\star) \le K E(\overline{J}''^\star), \qquad E(\overline{J}''^\star) \le w + K v E(\overline{X}'^\star),$$
$$E(\overline{X}'^\star) \le (u + w)K \le (u + A)K,$$

where K only depends on A and $\overline{Y}$, so indeed on A and ε and Y. Relation (2.5) readily follows from these estimates and from (2.6), once we observe that $P(S \le 1)$ is smaller that the sum of the first four terms on the right side of (2.5). □

Now we consider a sequence of SDE's like (2.4):

$$X^n_t = J^n_t + \int_0^t X^n_{s-} H^n_s \, dY_s, \tag{2.7}$$

all defined on the same filtered probability space and with the same dimensions. Also let ρ_n be an auxiliary sequence of random variables with values in some Polish space E, all defined on the same space again.

THEOREM 2.5. (a) *Tightness of both sequences $J^{n\star}$ and $H^{n,\star}$ implies tightness of the sequence $X^{n\star}$.*

(b) *Suppose that we have another equation (2.7) with solution X'^n and coefficients J'^n and H'^n. If the sequences $J^{n\star}$ and $H^{n,\star}$ are tight and if $(J^n - J'^n)^\star \to^P 0$ and $(H^n - H'^n)^\star \to^P 0$, then $(X^n - X'^n)^\star \to^P 0$.*

(c) *Let $V^n_t = \int_0^t H^n_s \, dY_s$. Suppose that the sequence $H^{n\star}$ is tight and that the sequence (J^n, V^n, ρ^n) stably converges to a limit (J, V, ρ) defined on some extension of the space. Then V is a semimartingale on the extension [w.r.t. the filtration generated by the pair (J, V)], and $(J^n, V^n, X^n, \rho^n) \Rightarrow^{\text{stably}} (J, V, X, \rho)$, where X is the unique solution of*

$$X_t = J_t + \int_0^t X_{s-} \, dV_s. \tag{2.8}$$

Statement (a) has been proved by Słomiński [15], while (b) and (c) are variations on the so-called stability results for SDE's [(c) is due again to Słomiński [14], while (b) has a slightly new formulation], and we give the proof for the reader's convenience. We have stated this theorem in a simple form, which is enough for our purposes, but it still holds for nonlinear equations with Lipschitz-continuous coefficients. Also Y might be replaced by a sequence Y^n: in this case it is necessary to add the assumption that the sequence Y^n has ($\star$), which implies that in fact it is "uniformly" sliceable in some sense.

PROOF OF THEOREM 2.5. (a) Relation (2.5) applied with $J'=0$ and $H'=0$ yields

$$P(X^{n\star}>\eta)\le\varepsilon+2P(H^{n\star}\ge A)+2P(J^{n\star}\ge u)+\frac{u}{\eta}K_{\varepsilon,A}, \tag{2.9}$$

where $K_{\varepsilon,A}$ is a constant depending on ε, A and Y. If we choose first ε arbitrarily, then A,u big, then η big, we obtain that the left side of (2.8) is smaller than 2ε, hence (a) holds.

(b) Similarly,

$$\begin{aligned}P((X^n-X'^n)^\star>\eta)\le\varepsilon&+P(H^{n\star}>A)+P(J^{n\star}>u)\\&+P((H^n-H'^n)^\star>v)+P((J^n-J'^n)^\star>w)\\&+\frac{uv+w}{\eta}K_{\varepsilon,A}.\end{aligned}$$

So we obtain the result by choosing first ε, η arbitrarily, then A,u big, then v,w small, then n big.

(c) The assumptions ensure that the sequence V^n has ($\star$). Thus if we do not introduce the variables ρ^n and if we replace stable convergence by ordinary (weak) convergence, this result is well known (see, e.g., [9]). Since stable convergence is just weak convergence of $(U.J^n,V^n,\rho^n)$ to (J,V,ρ) for any random variable U on the original probability space, our statement is proved. □

3. The fundamental result on the error distribution.

We let $Y=(Y^i)_{1\le i\le d}$ be a semimartingale on a stochastic basis $(\Omega,\mathscr{F},(\mathscr{F}_t),P)$. We always assume that $Y_0=0$ (this is of course not a restriction here). The time interval is $[0,1]$. We consider the q-dimensional SDE:

$$dX_t=f(X_{t-})\,dY_t,\qquad X_0=x_0. \tag{3.1}$$

Here $x_0\in\mathbb{R}^q$ and f is a continuously differentiable function from $\mathbb{R}^q$ into $\mathbb{R}^q\otimes\mathbb{R}^d$ with linear growth [i.e., $\|f(x)\|\le K(1+\|x\|)$ for some constant K].

One knows that (3.1) has a unique (strong) solution. We consider the *Euler continuous approximation* X^n given by

$$dX^n_t=f(X^n_{\varphi_n(t)})\,dY_t,\qquad X^n_0=x_0, \tag{3.2}$$

where $\varphi_n(t) = [nt]/n$ if $nt \notin \mathbb{N}$ and $\varphi_n(t) = t - 1/n$ if $nt \in \mathbb{N}$, and the *Euler discontinuous approximation* $\overline{X}^n$ given by

$$\overline{X}^n_t = X^n_{[nt]/n}. \tag{3.3}$$

The corresponding error processes are denoted by

$$U^n_t = X^n_t - X_t, \qquad \overline{U}^n_t = \overline{X}^n_t - X_{[nt]/n} = U^n_{[nt]/n}. \tag{3.4}$$

THEOREM 3.1. *If f is locally Lipschitz continuous and with linear growth, then $U^{n\star}$ and $\overline{U}^{n\star}$ tend to 0 in probability.*

This result is known when f is globally Lipschitz [14] or when f is bounded [8], but we need this general form below.

PROOF OF THEOREM 3.1. The second statement follows clearly from the first one. For the first claim, we consider functions $h_m \in C^1_b(\mathbb{R}^q)$ with $1_{\{\|x\| \le m\}} \le h_m(x) \le 1_{\{\|x\| \le m+1\}}$, and set $f_m(x) = f(x)h_m(x)$. Let $X(m)$ be the solution of (3.1) with the coefficient f_m and $X^n(m)$ be the corresponding Euler approximations.

Observe that $X = X(m)$ is $X^\star < m$ and that $X^n = X^n(m)$ if $X^n(m)^\star < m$. Hence $U^n = X^n(m) - X(m)$ on the set $\{X^\star < m-1, (X^n(m) - X(m))^\star \le 1\}$, and thus for $\varepsilon \in (0,1]$ we get

$$P(U^{n\star} > \varepsilon) \le P(X^\star \ge m-1) + P((X^n(m) - X(m))^\star > \varepsilon).$$

Since $X^n(m) \to X(m)$ in probability uniformly on $[0,1]$ as $n \to \infty$ for each $m \ge 1$ (see, e.g., [8]) and since $\lim_m P(X^\star \ge m-1) = 0$, the result follows. □

Next, let us examine rates of convergence. By this, we mean a sequence α_n of constants going to $+\infty$, such that the processes $\alpha_n U^n$ or $\alpha_n \overline{U}^n$ are tight, with nontrivial limiting processes. If this is the case, we also are interested in the "error processes" which are the limits of either one of these two sequences. Indeed, as far as applications are concerned, the usual Euler scheme gives us $\overline{X}^n$ and thus we would prefer to have results on $\overline{U}^n$, but mathematically speaking the processes U^n are easier to handle.

Here we give an improvement on a result by Kurtz and Protter [8], who essentially proved the implication (a)⇒(b) below. For this, we need to introduce some notation. For any process V we write

$$\Delta^n_i V = V_{i/n} - V_{(i-1)/n}, \qquad V^{(n)}_t = V_t - V_{[nt]/n}. \tag{3.5}$$

For any two semimartingales U, V, we write

$$Z^n_t(U, V) = \int_0^t U^{(n)}_{s-}\, dV_s. \tag{3.6}$$

THEOREM 3.2. *Let $Z^n = (Z^{n,ij} := Z^n(Y^i, Y^j))_{1\le i,\, j\le d}$, and let (α_n) be a deterministic sequence of positive numbers. There is equivalence between the following:*

(a) *The sequence $\alpha_n Z^n$ has ($\star$) and $(Y, \alpha_n Z^n) \Rightarrow (Y, Z)$.*

(b) *For any starting point x_0 and any C^1 function f with linear growth, the sequence $\alpha_n U^n$ has ($\star$) and $(Y, \alpha_n U^n) \Rightarrow (Y, U)$.*

In this case, we can realize the limits Z and U above on the same extension of the space on which Y is defined, and they are connected by

$$dU_t^i = \sum_{j=1}^{d} \sum_{k=1}^{q} f_k'^{ij}(X_{t-})\Big[U_{t-}^k dY_t^j - \sum_{\ell=1}^{d} f^{k\ell}(X_{t-})\, dZ_t^{\ell j}\Big], \qquad U_0^i = 0, \tag{3.7}$$

($f_k'^{ij}$ is the kth partial derivative of f^{ij}), and $(Y, \alpha_n Z^n, \alpha_n U^n) \Rightarrow (Y, Z, U)$ and also the sequence $\overline{U}^n$ stably converges in law to U.

REMARK 3.1. In view of Lemma 2.1, in (a) and (b) above we also have stable convergence in law of $\alpha_n Z^n$ and $\alpha_n U^n$, but these stable convergences are not enough to imply the convergences in (a) and (b).

Note also that we do *not* have $(Y, \alpha_n \overline{U}^n) \Rightarrow (Y, U)$ (except when Y or Z is continuous) in the last claim.

REMARK 3.2. We will see later that when Y is continuous, then assumption (a) is satisfied under mild hypotheses on Y. It is also satisfied when Y has jumps *and* each jump time is contained in an interval of constancy of Y, provided the "continuous part" of Y satisfies again the mild assumptions referred to above. In all other cases, we conjecture that indeed either the limits in this theorem are all 0 (i.e., the rate α_n is not the correct rate) or the sequence $\alpha_n Z^n$ is not even tight for the Skorohod topology. This conjecture is supported by the results of Sections 4.2 and 6.

REMARK 3.3. As we shall see in the proof, (a) is in fact equivalent to the property (b) stated for a single (judiciously chosen) equation; namely, let $\lambda \in \mathbb{R} \setminus \{0\}$ be such that

$$P(\Delta Y_t^i = -\lambda \text{ for some } i \le d \text{ and } t \in (0,1]) = 0. \tag{3.8}$$

Then it is enough to have (b) for the d^2-dimensional equation

$$X_t^{ij} = \delta_{ij} + (1 - \delta_{ij})Y_t^j + \lambda \int_0^t X_{s-}^{ij}\, dY_s^i, \qquad 1 \le i, j \le d. \tag{3.9}$$

PROOF OF THEOREM 3.2. (i) The implication (a)$\Rightarrow$(b), as well as (3.7), is proved in [8] in the case when ∇f is bounded. The fact that $\alpha_n U^n$ satisfies ($\star$) immediately follows from (3.21) in [8] and from the fact that $\alpha_n Z^n$ satisfies ($\star$).

When f is C^1 with linear growth, let f_m and $X(m)$ be defined as in Theorem 1.1, and let $U^n(m)$ be associated with $X(m)$ by (3.4). Then

∇f_m is bounded. We know that $(Y, \alpha_n Z^n, \alpha_n U^n(m))$ weakly converges to $(Y, Z, U(m))$ for all m, where $U(m)$ is the solution of (3.7) written for f_m and $X(m)$.

Denote by U the solution of (3.7) for f and X. As in Theorem 3.1 we have $U = U(m)$ and $U^n(m) = U^n$ on the set $\{X^\star < m-1,\ U^n(m)^\star \le 1\}$, while $U^n(m)^\star \to^P 0$ as $n \to \infty$; hence

$$\lim_m \limsup_n P(U \neq U^m \text{ or } U^n \neq U^n(m)) = 0.$$

Then $(Y, \alpha_n Z^n, \alpha^n U^n(m)) \Rightarrow (Y, Z, U(m))$ for all m implies $(Y, \alpha_n Z^n, \alpha_n U^n) \Rightarrow (Y, Z, U)$, and we have (b) and (3.7). The last claim is obvious by Lemma 2.1, since from what precedes we have stable convergence of U^n to U.

(ii) Suppose now that we have (b) for (3.9). We have $U^n = (U^{n,ij})_{1\le i,\, j\le d}$ and $X^n = (X^{n,ij})_{1\le i,\, j\le d}$. A simple calculation shows that

$$U_t^{n,ij} = \begin{cases} \displaystyle\int_0^t \big[\lambda U_{s-}^{n,ii} dY_s^i - \lambda^2 X_{\varphi_n(s)}^{n,ii} dZ_s^{n,ii}\big], & \text{if } j = i, \\ \displaystyle\int_0^t \big[\lambda U_{s-}^{n,ij} dY_s^i - \lambda^2 X_{\varphi_n(s)}^{n,ij} dZ_s^{n,ii}\big] - \lambda Z_t^{n,ji}, & \text{if } j \neq i. \end{cases} \tag{3.10}$$

Observe that, for each i, $X^{ii} = \mathscr{E}(\lambda Y^i)$ (the Doléans–Dade or stochastic exponential). Thus (3.8) implies that a.s. X^{ii} does not vanish, and $X^{n,ii}$ does not vanish either on $[0,1]$ for n large enough (because of Theorem 3.1). Hence (3.10) can be "inverted" to yield

$$Z_t^{n,ji} = \begin{cases} \displaystyle\int_0^t \left[\left(\frac{U_{s-}^{n,ii}}{\lambda X_{\varphi_n(s)}^{n,ii}}\right) dY_s^i - \left(\frac{1}{\lambda^2 X_{\varphi_n(s)}^{n,ii}}\right) dU_s^{n,ii}\right], & \text{if } i = j, \\ \displaystyle\int_0^t \left[U_{s-}^{n,ij} dY_s^i - \lambda X_{\varphi_n(s)}^{n,ij}\, dZ_s^{n,ii}\right] - \frac{U_t^{n,ij}}{\lambda}, & \text{if } i \neq j. \end{cases} \tag{3.11}$$

One deduces from the hypothesis that $(Y, \alpha_n U^n, X^n_{[n\cdot]/n}) \Rightarrow (Y, U, X)$, and the pair $(Y, \alpha_n U^n)$ has ($\star$), so (a) readily follows from Theorems 2.3 and 3.1. $\square$

In view of Remark 3.2, the following result has some interest: although not providing the limit of the error process, it actually gives the convergence rate (recall that if a sequence V^n has ($\star$), then a fortiori the sequence $V^{n\star}$ is tight). Note that Słomiński has proved the implication (a)⇒(b) when f is globally Lipschitz (see [15]).

THEOREM 3.3. *With the notation of Theorem 3.2, there is equivalence between the following:*

(a) *The sequence $\alpha_n Z^n$ has ($\star$).*

(b) *For any starting point x_0 and any C^1-function f with linear growth, the sequence $\alpha_n U^n$ has ($\star$).*

PROOF. Assume first (b) and consider (3.9) for a λ having (3.8), so (3.11) holds. We have seen that $1/X^{n,ii}_{\varphi_n(s)}$ goes to $1/X^{ii}_s$ uniformly in s, in probability, while X^{ii} does not vanish, hence the sequence $(1/X^{n,ii}_{\varphi_n(\cdot)})^\star$ is tight. Then ($\star$) for (U^n) yields that $(U^{n,ii}/X^{n,ii}_{\varphi_n(\cdot)})^\star$ also is tight, and Theorem 2.3(a) applied to the first part of (3.11) yields that $(Z^{n,ii})$ has ($\star$). Then apply the same result to the second part of (3.11) to get that $(Z^{n,ij})$ has ($\star$) for $i \neq j$, and (a) holds.

For the converse, we need to introduce the equation satisfied by U^n, which is (with matrix notation)

$$dU^n_t = (f(X^n_t) - f(X_t))\,dY_t - (f(X^n_{t-}) - f(X^n_{\varphi_n(t)}))\,dY_t.$$

Now, with any càdlàg process V we set

$$k(V)^n_t = f(V_{\varphi_n(t)}) \int_0^1 \nabla f(V_{\varphi_n(t)} + u(V_{t-} - V_{\varphi_n(t)}))\,du, \tag{3.12}$$

which is left-continuous. Apply Taylor's expansion and the fact that $X^n_{t-} - X^n_{\varphi_n(t)} = f(X^n_{\varphi_n(t)})(Y_{t-} - Y_{\varphi_n(t)})$ to get the following, where $\tilde{X}^n_{t-}$ is in between X_{t-} and X^n_{t-}:

$$d(\alpha_n U^n)_t = (\alpha_n U^n)_{t-} \nabla f(\tilde{X}^n_{t-})\,dY_t - k(X^n)^n_t\,d(\alpha_n Z^n)_t. \tag{3.13}$$

The sequence $k(X^n)^{n\star}$ is tight by Theorem 3.1, so the sequence

$$\int_0^{\cdot} k(X^n)^n_t\,d(\alpha_n Z^n)_t$$

has ($\star$) as soon as (a) holds by Theorem 2.3. Since the sequence $\nabla f(\tilde{X}^n)^\star$ is also tight, Theorem 2.5 gives the tightness of the sequence $(\alpha_n U^{n,\star})$, and another application of Theorem 2.3 yields (b). □

4. Processes with finite variation.

We treat here the case where the driving process Y in (3.1) is of finite variation, with the rate $1/n$ in view. In this case (3.1) is truly an "ω-wise" (or deterministic) equation, and the reason for looking at this case is not practical importance but rather methodological implications. When Y is continuous, we find a necessary and sufficient condition for getting this rate $1/n$, and this seems to be new even though it concerns only "ordinary differential equations."

However, when Y has jumps together with a nontrivial continuous part, the picture changes radically: the rate is still $1/n$, in the sense that $nU^{n\star}$ and $n\overline{U}^{n\star}$ remain tight, but in the deterministic case these processes have no limit. In the random case, rather mild conditions imply the convergence of these processes to a limit involving "additional randomness."

For simplicity we only consider the one-dimensional case. Extensions to several dimensions are straightforward and left to the reader.

4.1. *The continuous case.*

Here we assume that Y is of finite variation and continuous. Remember that $Y_0 = 0$. We write $Z^n = Z^n(Y, Y)$ [see (3.6)]. An integration by parts shows that

$$Z_t^n = \tfrac{1}{2}\biggl(\sum_{i=1}^{[nt]} (\Delta_i^n Y)^2 + (Y(n)_t)^2\biggr) \geq 0. \tag{4.1}$$

THEOREM 4.1. *Assume that Y is continuous with finite variation. There is equivalence between the following:*

(a) *We have*

$$Y_t = \int_0^t y_s\,ds, \qquad \int_0^1 y_s^2\,ds < \infty \quad \text{a.s.} \tag{4.2}$$

(b) *The sequence of random variables $(2^n Z_1^{2^n})_{n\geq 1}$ is tight.*
(c) *The sequence of random variables $(nZ_1^n)_{n\geq 1}$ is tight.*
(d) $\sup_n nZ_1^n < \infty$ *a.s.*
(e) *The processes nZ^n converge a.s. uniformly in time to a process Z.*

Moreover in this case we have $Z_t = \frac{1}{2}\int_0^t y_s^2\,ds$ and $\sup_n \int_0^1 |dZ_s^n| < \infty$ a.s.

Note that (e) and the last claim imply condition (a) of Theorem 3.2, with $\alpha_n = n$, while the latter implies (c); hence the above result completely solves the question of whether the rate for U_n is $1/n$ or not, for processes Y as above. Note also that all statements above are "ω-wise" (a.s.), that is, this result is deterministic in nature, and indeed we begin with two lemmas which are concerned with the deterministic case.

LEMMA 4.2. *Assume that Y is continuous with finite variation and deterministic. Only two cases are possible:*

(i) *Equation (4.2) holds and $\sup_n nZ_1^n < \infty$.*
(ii) *Equation (4.2) does not hold and $2^n Z_1^{2^n} \to \infty$.*

PROOF. Consider the following measures on $(0,1]$: $\lambda(dt) = dt$, $\mu(dt) = dY_t$, $\mu'(dt) = |dY_t|$; and set $\rho = 1 + \mu'((0,1])$. Then $\nu = (1/\rho)(\mu' + \lambda)$ is a probability measure, and we introduce the Radon–Nikodym derivatives:

$$V = \frac{d\lambda}{d\nu}, \qquad U = \frac{d\mu}{d\nu},$$

which satisfy $\{V = 0\} \subset \{|U| = \rho\}$. With the convention $a/0 = +\infty$ (resp., $-\infty$) if $a > 0$ (resp., $a < 0$), we also set $L = U/V$.

Let $\mathscr{G}_n$ be the σ-field of $(0,1]$ generated by the intervals $(j/n, (j+1)/n]$, and set

$$V_n = \frac{d\lambda}{d\nu}\bigg|_{\mathscr{G}_n}, \qquad U_n = \frac{d\mu}{d\nu}\bigg|_{\mathscr{G}_n}, \qquad L_n = \frac{U_n}{V_n}.$$

These are finite-valued functions, with $V_n = \nu(V|\mathscr{G}_n)$ and $U_n = \nu(U|\mathscr{G}_n)$. A simple computation yields $L_n(s) = n\Delta_i^n Y$ for $s \in ((i-1)/n, i/n]$, so by (4.1):

$$nZ_1^n = \int L_n^2\,d\lambda = \int L_n^2 V_n\,d\mu'. \tag{4.3}$$

The sequences U_{2^n} and V_{2^n} are uniformly integrable ν-martingales w.r.t. the filtration $(\mathscr{G}_{2^n})$, converging ν-a.s. to U and V, hence L_{2^n} ν-a.s. converges to L (because $\{V=0\}\subset\{|U|=\rho>0\}$). On the set $\{V>0\}$ we have $L^2_{2^n}V_{2^n}\to L^2V$ ν-a.s., while on the complement we have $L^2_{2^n}V_{2^n}=L_{2^n}U_{2^n}\to+\infty$ ν-a.s. Then (4.3) and Fatou's lemma yield

$$\liminf_n 2^n Z_1^{2^n} \ge \int L^2\,d\lambda + \infty\mu'(V=0). \tag{4.4}$$

If $\liminf_n 2^n Z_1^{2^n}<\infty$, we deduce $\mu'(V=0)=0$ from (4.4); hence the first half of (4.2) holds true with $y=L$; another application of (4.4) yields the second half of (4.2). Conversely, assume (4.2): we have $L=y$, which belongs to $L^2(d\lambda)$, and $L_n=\lambda(L|\mathscr{G}_n)$; hence $\sup_n nZ_1^n=\sup_n\int L_n^2 d\lambda<\infty$. This completes the proof of the lemma. □

LEMMA 4.3. *Assume that Y is deterministic and satisfies (4.2) with y piecewise constant. Then nZ_t^n converges uniformly in t to $Z_t=\frac12\int_0^t y_s^2\,ds$.*

PROOF. We have $y_s=u_i$ for $t_{i-1}<s<t_i$, where $0=t_0<\cdots<t_k=1$. Let $C=\sup_i|u_i|$ and $\tau_n(s)=s-[ns]/n$. We have $|Y_s^n|\le C/n$, and also $Y_s^n=u_i\tau_n(s)$ if $t_{i-1}<[ns]/n\le s<t_i$. Hence

$$nZ_t^n = n\sum_{i=1}^k u_i^2\int_{t_{i-1}\wedge t}^{t_i\wedge t}\tau_n(s)\,ds + R_n(t),$$

where $|R_n(t)|\le 2knC^2/n^2\to 0$. A simple calculation shows that $\int_u^v n\tau_n(s)\,ds\to(v-u)/2$; hence

$$Z_t'^n := n\sum_{i=1}^k u_i^2\int_{t_{i-1}\wedge t}^{t_i\wedge t}\tau_n(s)\,ds \to \frac12\sum_{i=1}^k u_i^2(t_i\wedge t - t_{i-1}\wedge t) = Z_t.$$

Since Z'^n and Z are continuous nondecreasing, this convergence is uniform in t over $[0,1]$. □

PROOF OF THEOREM 4.1. That (e)⇒(d)⇒(c)⇒(b) is obvious. Let B be the set of all ω such that the function $t\to Y_t(\omega)$ does not satisfy (4.2). By Lemma 4.2, $2^nZ_1^{2^n}(\omega)\to\infty$ if $\omega\in B$, then (b) implies $P(B)=0$ and (a).

It remains to prove that (a) implies (e) and the last claim, so below we assume (4.2). A density argument shows that there is a sequence $y(p)$ of bounded processes, piecewise constant in time, such that

$$\int_0^1 y(p)_s^2\,ds \le \int_0^1 y_s^2\,ds, \qquad \eta_p := \int_0^1 |y_s - y(p)_s|^2\,ds \to 0 \quad \text{a.s.} \tag{4.5}$$

Let $Y(p)_t=\int_0^t y(p)_s\,ds$, with the associated processes $Z^n(p)=Z^n(Y(p),Y(p))$, and set $Z_t=\frac12\int_0^t y_s^2\,ds$ and $Z(p)_t=\frac12\int_0^t y(p)_s^2\,ds$. By Lemma 4.3, $Z^n(p)$ converges for all ω and p to $Z(p)$, uniformly in time, as $n\to\infty$. By (4.5) it is obvious that $Z(p)\to Z$ uniformly in time as $p\to\infty$, a.s. So in order to prove (e) it is enough to show that $u_p^n = \sup_t |nZ_t^n(p)_t - nZ_t^n|$ satisfies $\sup_n u_p^n \to 0$ a.s. as $p \to \infty$.

By the Cauchy–Schwarz inequality, and with $\eta = \int_0^1 y_s^2\, ds$, we get

$$\begin{aligned} u_p^n &\le n \int_0^1 |Y(p)_s^n y(p)_s - Y_s^n y_s|\, ds \\ &\le n \int_0^1 |Y(p)_s^n||y(p)_s - y_s|\, ds + n \int_0^1 |Y(p)_s^n - Y_s^n||y_s|\, ds \\ &\le n\sqrt{\eta_p} \left(\int_0^1 |Y(p)_s^n|^2\, ds\right)^{1/2} + m\sqrt{\eta} \left(\int_0^1 |Y(p)_s^n - Y_s^n|^2\, ds\right)^{1/2}. \end{aligned}$$

Now $n|Y(p)_s^n| \le (n \int_{[ns]/n}^s y(p)_s^2\, ds)^{1/2}$; hence

$$\int_0^1 |Y(p)_s^n|^2\, ds \le \frac{1}{n} \int_0^1 ds \int_{[ns]/n}^s y(p)_u^2\, du \le \frac{1}{n^2} \int_0^1 y(p)_u^2\, du \le \frac{\eta}{n^2},$$

and similarly $\int_0^1 |Y(p)_s^n - Y_s^n|^2\, ds \le \eta_p/n^2$. Therefore $u_p^n \le 2\sqrt{\eta\eta_p} \to 0$ as $p \to \infty$; hence we have (e) with Z as above.

Finally, $\int_0^1 |dZ_s^n| \le Z^n(Y', Y')_1$, where $Y'_t = \int_0^t |y_s|\, ds$. Since Y' also satisfies (4.2) we deduce the last claim. □

PROOF OF THEOREM 1.1. That (b)⇒(a) is obvious, and (c) implies (b) and the last claim by Theorems 3.2 and 4.1.

Now suppose (a). A simple computation shows that $X_1^n = X_1 \prod_{i=1}^n ((1 + \Delta_i^n Y) \exp -\Delta_i^n Y)$. By hypothesis $0 \le -nU_1^n \le K$ for some constant K (remember that Y is increasing); then if $A^n = \sum_{i=1}^n (\Delta_i^n Y - \log(1 + \Delta_i^n Y))$, it follows that $nA^n \le K'$ for another constant K'; hence $n \sum_{i=1}^{[nt]} (\Delta_i^n Y)^2 \le K''$ for yet another constant K''. So (4.1) and Lemma 4.2 give (c). □

4.2. *The discontinuous case.* In the remainder of this section, we study the case where $Y = A + B$, with A a continuous process with finite variation, and with B as follows:

$$B_t = \sum_{j \ge 1} b_j 1_{\{T_j \le t\}}, \tag{4.6}$$

where $T_0 = 0$, T_j is $[0, 1] \cup \{\infty\}$-valued, nondecreasing in j and with $T_j < T_{j+1}$ if $T_j < 1$, and also $K = \inf(j\colon T_{j+1} = \infty)$ is a.s. finite and $b_j \ne 0$ if $j \le K$.

When $A = 0$, the situation is particularly simple:

THEOREM 4.4. *If $A = 0$, then $Z^n(\omega) = 0$ for all n large enough (depending on ω).*

PROOF. Let ω be fixed. For n large enough, each interval $((i-1)/n, i/n]$ contains at most one jump time of B, in which case the property $Z_t^n(\omega) = 0$ is readily verified. □

In the above situation, Theorem 3.2 readily applies, but we have even more, namely, that $U^n = \overline{U}^n = 0$ as soon as n is large enough.

The situation when $A \neq 0$ is more surprising, and we will find out that the sequences nZ^n and nU^n do not converge in law for the Skorohod topology, unless the process A is constant on a neighborhood of each T_j. However, under mild assumptions the processes $n\overline{U}^n$ and $n\overline{Z}^n$ indeed converge, where

$$\overline{Z}^n_t = Z^n_{[nt]/n}. \tag{4.7}$$

We need some additional notation below:

$$T_+(n, j) = \inf\left(\frac{i}{n}: i \geq 1,\ \frac{i}{n} \geq T_j\right), \qquad T_-(n, j) = T_+(n, j) - \frac{1}{n}; \tag{4.8}$$

$$\alpha^n_j = n(A_{T_j} - A_{T_-(n,j)}), \quad \beta^n_j = n(A_{T_+(n,j)} - A_{T_j}), \quad \gamma^n_j = \alpha^n_j + \beta^n_j; \tag{4.9}$$

$$\overline{\beta}^n_j = \int_{T_j}^{T_+(n,j)} |dA_s|; \tag{4.10}$$

with $\alpha^n_j = \beta^n_j = \overline{\beta}^n_j = 0$ if $j > K$. We are not especially interested in the processes $n\overline{Z}^n$, but they are simpler than $n\overline{U}^n$ and so we start with them.

THEOREM 4.5. *We have equivalence between the following:*

(a) $n\overline{Z}^n \Rightarrow^{\text{stably}} Z$.

(b) *The process A has (4.2) (with a density a, say), and $(\gamma^n_j)_{j\geq 1} \Rightarrow^{\text{stably}} (\gamma_j)_{j\geq 1}$ for the product topology on $\mathbb{R}^{\mathbb{N}}$.*

In this case, the limits in (a) *and* (b) *above are connected by*

$$Z_t = \tfrac{1}{2}\int_0^t a_s^2\, ds + \sum_{j\geq 1} b_j \gamma_j 1_{\{T_j \leq t\}}, \tag{4.11}$$

and furthermore the sequence nZ^n converges stably in finite-dimensional laws along the (dense) set $J = \{t: P(\Delta Y_t \neq 0) = 0\}$.

PROOF. Let us write $C^n = nZ^n(A, A)$ and $D^n = nZ^n(A, B) + nZ^n(B, A)$, and associate with these $\overline{C}^n$ and $\overline{D}^n$ as in (4.7). On the set Ω_n where each interval $((i-1)/n, i/n]$ contains at most one T_j, we have $nZ^n = C^n + D^n$. Observe also that

$$\overline{D}^n_t = \sum_{j\geq 1} b_j \gamma^n_j 1_{\{T_+(n,j) \leq t\}} \quad \text{on } \Omega_n. \tag{4.12}$$

On the other hand, (4.1) shows that $\Delta\overline{C}^n_{T_+(j,n)} = (1/2n)(\gamma^n_j)^2$ on $\{T_j \leq 1\}$. Hence

$$\Delta(n\overline{Z}^n)_{T_+(n,j)} = \frac{1}{2n}(\gamma^n_j)^2 + b_j\gamma^n_j \quad \text{on } \Omega_n \cap \{T_+(n, j) \leq 1\}. \tag{4.13}$$

(i) Assume (a). By (4.9) and the continuity of A, we have $\gamma_j^n/n \to 0$ for each j. We also have $\Omega_n \to \Omega$. Now, (a) implies tightness for each sequence (4.13) (j fixed), and since $b_j \neq 0$ we deduce that each sequence γ_j^n is tight. Then (4.12) yields that the sequence $\overline{D}^n$ is tight, which in turn yields together with (a) again that the sequence $\overline{C}_1^n = C_1^n$ is tight. At this point, Theorem 4.1 gives that A satisfies (4.2) and that C^n converges a.s. uniformly to $C_t = \frac{1}{2}\int_0^t a_s^2\,ds$.

By well-known properties of stable convergence, we deduce from this, from the fact that $\Omega_n \to \Omega$ and from (a) that $\overline{D}^n \Rightarrow^{\text{stably}} Z - C$. In view of (4.12), this gives the second half of (b) and the relation (4.11).

(ii) Assume (b). By Theorem 4.1, C^n converges a.s. uniformly to C, as given above, so it is enough (using $\Omega_n \to \Omega$ again) to prove that the right-hand side of (4.12) stably converges: in view of (4.8) this readily follows from the second part of (b).

(iii) It remains to prove the last claim. Exactly as before, it is enough to prove that the processes D^n stably converge in finite-dimensional laws to the process defined by the last sum in (4.11). Since D^n is constant over each interval $[T_+(n, j), T_-(n, j+1))$, this is clearly equivalent to the stable convergence of the sequence $\overline{D}^n$ finite-dimensionally in law along J, and this property readily follows from the stable convergence (for Skorohod topology) of $\overline{D}^n$ to D, as seen before, because J is exactly the complements of the fixed times of discontinuity of D. □

Now we turn to U^n and $\overline{U}^n$, for which we give only a sufficient condition.

THEOREM 4.6. *Assume that the process A has (4.2) with a density a and that each sequence $(\overline{\beta}_j^n)_{n\geq 1}$ is tight. Then the sequences (nZ^n) and (nU^n) have ($\star$). If further $(\alpha_j^n, \beta_j^n)_{j\geq 1} \Rightarrow^{\text{stably}} (\alpha_j, \beta_j)_{j\geq 1}$, then for any starting point x_0 and any C^1-function f with linear growth we have the following:*

(a) *The sequence $(n\overline{U}^n) \Rightarrow^{\text{stably}} U$, where U is the unique solution of the following linear SDE:*

$$U_t = \int_0^t f'(X_{s-})U_{s-}\,dY_s - \tfrac{1}{2}\int_0^t f(X_s)f'(X_s)a_s^2\,ds$$
$$- \sum_{j:\,T_j\leq t} b_j f(X_{T_j-})\biggl(\alpha_j f'(X_{T_j-}) + \beta_j \int_0^1 f'(X_{T_j-} + u\Delta X_{T_j})\,du\biggr). \tag{4.14}$$

(b) *The sequence (nU^n) converges stably in finite-dimensional laws along J (see Theorem 4.5) to U.*

Observe that another way of writing (4.14) is as follows:

$$U_t = \int_0^t f'(X_{s-})U_{s-}\,dY_s - \tfrac{1}{2}\int_0^t f(X_s)f'(X_s)a_s^2\,ds$$
$$- \sum_{j:\,T_j\leq t} \bigl(\Delta X_{T_j} f'(X_{T_j-})\alpha_j + (f(X_{T_j}) - f(X_{T_j-}))\beta_j\bigr). \tag{4.15}$$

Before giving the proof, let us provide some comments and examples.

REMARK 4.1. Equations (3.7) and (4.14) are *different*, unless $b_j\beta_j = 0$ for all j. This means in particular that the convergence of nU^n described above is *not* in the Skorohod sense, otherwise one could apply Theorem 3.2, and similarly for nZ^n in Theorem 4.5. In fact, what prevents Skorohod convergence is that in each interval $[T_j, T_+(n, j)]$ we have for these processes a "big" jump at T_j and also a continuous part with a nonvanishing increment.

In fact, the results obtained in (b) above show that the sequence nU^n converges to U in the topology of convergence in measure for functions on $[0, 1]$ (also called Meyer–Zheng topology), and indeed this sequence also converges in the S-topology introduced by Jakubowski [5]. The same holds for nZ^n. For these topologies, a result like Theorem 2.3 does not hold, explaining why the limit U satisfies another equation than (3.7).

REMARK 4.2. In a sense, the most interesting aspect of this subsection consists in "negative" results (no Skorohod convergence for nU^n, no convergence at all if the conditions above are not met). However, if one is interested in "positive" results, one can check as a by-product of the following proof that the sequence $n\overline{U}^n$ is (Skorohod)-tight and the sequence $nU^{n\star}$ is tight as soon as A satisfies (4.2) and all sequences α_j^n and $\overline{\beta}_j^n$ are tight, even if we do not have convergence.

Observe that we may have (4.2) for A and yet the tightness above may fail: take, for example, $T_1 = \frac{1}{2}$, $T_2 = \infty$, $b_1 = 1$, and $a_s = |s - \frac{1}{2}|^{-1/3}$. In this case $nZ_1^n \to \infty$, and the rate of convergence (if it exists) is not $1/n$.

REMARK 4.3. As said before, (3.1) is in principle not random. However, if Y is really not random we have no chance of getting a limit in the previous theorem.

Take, for example, $A_t = t$ and $B_t = 1_{\{T\le t\}}$ for some $T = p/q$ with p and q relative primes. The conditions of Theorem 4.5 hold, and $n\overline{Z}^n$ converges for the Skorohod topology to $Z = \frac{1}{2}A + B$. However, the conditions of Theorem 4.6 do not hold because α_1^n takes successively all values $0, 1/q, \dots, (q-1)/q$ as n varies. In fact the functions nU^n and $n\overline{U}^n$ have q distinct limit functions, solutions of the (nonrandom) differential equations for $i = 0, 1, \dots, q-1$:

$$\begin{aligned} U(i)_t = \int_0^t f'(X_{s-})U(i)_{s-}\,ds - \frac{1}{2}\int_0^t f(X_s)f'(X_s)\,ds \\ + \biggl(f'(X_{T-})U(i)_{T-} - \frac{i}{q}f'(X_{T-})\Delta X_T \\ - \frac{q-i}{q}(f(X_T) - f(X_{X_{T-}}))\biggr)1_{\{T\le t\}}. \end{aligned}$$

REMARK 4.4. In fact the existence of a limit for the sequences α_j^n and β_j^n is connected with the asymptotic behavior of the fractional part of the variables nT_j. This fractional part is known to converge (even stably) if T_j admits a density regular enough (see, e.g., [2]), while of course it does not converge if

T_j is deterministic. Another factor which might ensure convergence is enough randomness in the density a_s. In all these cases, the limit U features “more randomness” than Y, as seen from the fact that U (or sometimes even Z) is defined on genuine extensions of the original space.

Here is an example where convergence comes from the randomness of T_j: suppose that Y is as in Remark 4.3, but T is uniform on $(0,1]$. Then $n\overline{Z}^n$ tends ω-wise to Z as above, while $n\overline{U}^n$ stably converges to the solution of the following equation, where α denotes a random variable, uniform and independent of T:

$$U_t = \int_0^t f'(X_{s-})U_{s-}\,ds - \frac{1}{2}\int_0^t f(X_s)f'(X_s)\,ds \\ + \Big(f'(X_{T-})U_{T-} - \alpha f'(X_{T-})\Delta X_T - (1-\alpha)(f(X_T) - f(X_{T-}))\Big)1_{\{T\le t\}}.$$

Here is another example, where convergence comes from the randomness of a_s: let W be a standard Brownian motion and set

$$B_t = 1_{\{t\ge 1\}} \quad \text{and} \quad a_s = (1-s)^{-1/2}W_{1-s}1_{\{|W_{1-s}|\le\sqrt{1-s}\}}.$$

Observe that $|a_s| \le 1$. It is easy, by a scaling argument, to check that here $n\overline{Z}^n$ converges stably to $Z = \frac{1}{2}A + UB$, where U is a standard normal variable, independent of Y.

REMARK 4.5. We have left out the case when Y has infinitely many jumps on $(0,1]$: nothing is known in general for this case; see, however, Section 6 when Y has in addition independent increments.

PROOF OF THEOREM 4.6. (i) First we prove that the sequence nZ^n is $(\star)$. Since Z^n is of finite variation, it suffices to show that $n\int_0^1 |dZ_t^n|$ forms a tight sequence.

With the notation of the proof of Theorem 4.5, $nZ^n = C^n + D^n$ on Ω_n, so a simple computation shows that

$$n\int_0^1 |dZ_s^n| \le 2\int_0^1 |dC_s^n| + \sum_{i=1}^K |b_j|(|\alpha_j^n| + \overline{\beta}_j^n) \quad \text{on the set } \Omega_n. \tag{4.16}$$

Now $\Omega_n \to \Omega$, while $\int_0^1 |dC_s^n|$ is tight by Theorem 4.1, so the result immediately follows from the fact that K is a.s. finite and from the tightness of all sequences $(\alpha_j^n)_n$ and $(\overline{\beta}_j^n)_n$ [recall K is defined just after (4.6)]. In view of Theorem 3.3 it then follows that nU^n is $(\star)$, and in particular the sequence $nU^{n\star}$ is tight.

(ii) Recall that nU^n is the solution of (3.13), with $\alpha_n = n$, and introduce the solution V_n of the following linear equation:

$$dV_t^n = V_{t-}^n f'(X_{t-})\,dY_t - k(X)_t^n\,d(nZ^n)_t, \qquad V_0^n = 0. \tag{4.17}$$

Theorems 2.2(d), 3.1 and 2.5(b) give that $(nU^n - V^n)^\star \to^P 0$, so in order to prove (a) and (b) we can replace the processes nU^n and $n\overline{U}^n$ by V^n and $\overline{V}_t^n = V_{[nt]/n}^n$, respectively.

Also let U be the solution of (4.14), on an extension of the original probability space which supports the limits (α_j, β_j). Let us introduce the processes

$$W(j)^n_t = \overline{V}^n_{t\wedge T_+(n,j)}, \tag{4.18}$$

$$W(j)'^n_t = \begin{cases} W(j)^n_t, & \text{if } t \le T_+(n,j), \\ W(j)^n_{T_+(n,j)} + V^n_t - V^n_{T_+(n,j)}, & \text{if } T_+(n,j) < t < T_{j+1}, \\ W(j)^n_{T_+(n,j)} + V^n_{T_{j+1}-} - V^n_{T_+(n,j)}, & \text{if } T_{j+1} \le t, \end{cases} \tag{4.19}$$

$$W(j)_t = U_{t\wedge T_j}, \qquad W(j)'_t = \begin{cases} U_t, & \text{if } t < T_{j+1}, \\ U_{T_{j+1}-}, & \text{if } t \ge T_{j+1}. \end{cases} \tag{4.20}$$

Let us also write ρ^n for the double sequence $(\alpha^n_j, \beta^n_j)_{j\ge1}$ and $\rho = (\alpha_j, \beta_j)_{j\ge1}$. Consider the following property:

$$(\rho^n, W(j)^n) \Rightarrow^{\text{stably}} (\rho, W(j)). \tag{H$_j$}$$

By hypothesis (H_0) holds. If (H_j) holds for all j, then (a) follows, because $K < \infty$ a.s.

(iii) Suppose that (H_j) holds. First if $H(n,j)$ is the interval $(T_+(n,j), T_{j+1})$ we deduce from (4.17) that

$$\begin{aligned} W(j)'^n_t = W(j)^n_t &+ \int_0^t W(j)'^n_{s-} a_s f'(X_{s-}) 1_{H(n,j)}(s)\, ds \\ &- \int_0^t k(X)^n_s 1_{H(n,j)}(s)\, dC^n_s. \end{aligned} \tag{4.21}$$

Recall that $C^n = nZ^n(A,A)$ converges a.s. uniformly to the process $C_t = \frac{1}{2}\int_0^t a_s^2\, ds$. Further, set

$$J^n_t = W(j)^n_t - \int_0^t k(X)^n_s 1_{H(n,j)}(s)\, dC^n_s,$$

$$J_t = W(j)_t - \int_0^t f(X_s) f'(X_s) 1_{(T_j, T_{j+1})}(s)\, dC_s,$$

$$L^n_t = \int_0^t a_s f'(X_{s-}) 1_{H(n,j)}(s)\, ds, \qquad L_t = \int_0^t a_s f'(X_s) 1_{(T_j, T_{j+1})}(s)\, ds,$$

so (4.21) becomes

$$W(j)'^n_t = J^n_t + \int_0^t W(j)'^n_{s-}\, dL^n_s. \tag{4.22}$$

Now, X is continuous on the interval (T_j, T_{j+1}), so $(L^n - L)^\star \to^P 0$ and the last term in (4.21) converges uniformly in probability to

$$\int_0^t f(X_s) f'(X_s) 1_{(T_j, T_{j+1})}(s)\, dC_s,$$

and since this process is continuous, as well as L, it follows from (H_j) that $(J^n, L^n, \rho^n) \Rightarrow^{\text{stably}} (J, L, \rho)$. Hence Theorem 2.5(c) yields that

$$(W(j)'^n, \rho^n) \Rightarrow^{\text{stably}} (W(j)', \rho).$$

Using Lemma 2.2 and the fact that $W(j)'^n$ and $W(j)'$ are constant on $[T_{j+1}, 1]$, we finally deduce that if $\overline{W}(j)^n_t = W(j)^n_{[nt]/n}$, then

$$(\overline{W}(j)'^n_{\cdot\wedge T_-(n,j+1)}, \rho^n, W(j)'^n_{T_{j+1}-}) \ \Rightarrow^{\text{stably}} \ (W(j)', \rho, W(j)'_{T_{j+1}-}). \tag{4.23}$$

(iv) Now, set $\delta_n := V^n_{T_+(n,j+1)} - V^n_{T_-(n,j+1)}$. By (3.13) we have $\delta_n = u_n + v_n$ on the set $\Omega_n \cap \{T_{j+1} \le 1\}$, where

$$u_n = \int_{T_-(n,\,j+1)}^{T_+(n,\,j+1)} V^n_{s-} f'(X_{s-}) a_s\, ds - \int_{T_-(n,\,j+1)}^{T_+(n,\,j+1)} k(X)^n_s \, dC^n_s,$$

$$v_n = b_{j+1}\biggl(V^n_{T_{j+1}-} f'(X_{T_{j+1}-}) - k(X)^n_{T_{j+1}} \alpha^n_{j+1} - \int_{T_{j+1}}^{T_+(n,j+1)} k(X)^n_s a_s \, ds\biggr).$$

First, the sequences $V^{n\star}$ and $k(X)^{n\star}$ are tight, so one deduces that $u_n \to^P 0$. Next the sequences

$$k(X)^n_{T_{j+1}} - (ff')(X_{T_{j+1}-})$$

and

$$\int_{T_{j+1}}^{T_+(n,\,j+1)} k(X)^n_s a_s \, ds - \beta^n_{j+1} f(X_{T_{j+1}-}) \int_0^1 f'(X_{T_{j+1}-} + u\Delta X_{T_{j+1}})\, du$$

converge to 0 in probability. Furthermore, $V^n_{T_{j+1}-} = W(j)'^n_{T_{j+1}-}$. Therefore if

$$\begin{aligned}\delta = b_{j+1}\biggl(&W(j)'_{T_{j+1}-} f'(X_{T_{j+1}-}) - \alpha_{j+1}(ff')(X_{T_{j+1}-}) \\ &- \beta_{j+1} \int_0^1 f'(X_{T_{j+1}-} + u\Delta X_{T_{j+1}})\, du\biggr),\end{aligned}$$

one deduces from (4.23) that

$$\bigl(\overline{W}(j)'^n_{\cdot\wedge T_-(n,\,j+1)}, \rho^n, \delta_n\bigr) \ \Rightarrow^{\text{stably}} \ (W(j)', \rho, \delta). \tag{4.24}$$

However,

$$W(j+1)^n = \overline{W}(j)'^n_{\cdot\wedge T_-(n,j+1)} + \delta_n 1_{[T_+(n,\,j+1),\,1]},$$

while $W(j+1) = W(j)'_{\cdot\wedge T_{j+1}-} + \delta 1_{[T_{j+1},1]}$. Thus (4.24) yields ($\mathrm{H}_{j+1}$), and the proof of (a) follows by induction on j.

(v) Finally, on the set $\{T_+(n, j) < t < T_-(n, j+1)\}$ we have $V^n_t = W(j)'^n_t$ and $U_t = W(j)'_t$. Since $W(j)'^n \Rightarrow^{\text{stably}} W(j)'$ and since U is continuous outside all T_j's, the first claim of (b) is obvious, and we are finished. □

5. Continuous semimartingales.

5.1. *The martingale case.* Here we consider the case where the driving process $Y=(Y^i)_{1\le i\le d}$ in (3.1) is a *continuous* d-dimensional local martingale, null at 0. We denote by $C=(C^{ij})_{1\le i,\,j\le d}$ the quadratic variation process, that is, $C^{ij}=\langle Y^i,Y^j\rangle$, and we write $Z^n=(Z^{n,ij})_{1\le i,j\le d}$, where $Z^{n,\,ij}=Z^n(Y^i,Y^j)$. We introduce the d^4-dimensional processes D^n whose components are

$$D^{n,\,ijk\ell}=n\langle Z^{n,\,ik},Z^{n,\,j\ell}\rangle. \tag{5.1}$$

The main result is as follows:

THEOREM 5.1. *Assume that Y is a continuous local martingale. There is equivalence between the following:*

(a) *We have (with c being a $d\times d$ symmetric nonnegative matrix-valued predictable process)*

$$C_t=\int_0^t c_s\,ds,\qquad \int_0^1 \|c_s\|^2\,ds<\infty. \tag{5.2}$$

(b) *For each i the sequence of random variables $(\sqrt{n}\,Z^{n,\,ii\star})_{n\ge1}$ is tight.*

(c) *For each i the sequence of random variables $(D^{n,\,iiii})$ is tight.*

In this case, and if σ is a $d\times q$ matrix valued process such that $c=\sigma\sigma^\dagger$ ($\sigma^\dagger$ stands for the transpose; such processes always exist, for $q\ge d$ at least), the sequence $\sqrt{n}\,Z^n$ stably converges in law to a process Z given by

$$Z_t^{ij}=\frac{1}{\sqrt{2}}\sum_{1\le k,\,\ell\le q}\int_0^t \sigma_s^{ik}\sigma_s^{j\ell}\,dW_s^{k\ell}, \tag{5.3}$$

where $(W^{ij})_{1\le i,\,j\le q}$ is a standard q^2-dimensional Brownian motion defined on an extension of the space on which Y is defined and independent of Y. Moreover, we also have $(Y,\sqrt{n}\,Z^n)\Rightarrow(Y,Z)$ and the sequence $\sqrt{n}\,Z^n$ has ($\star$), and the following convergence holds a.s. uniformly in time:

$$D_t^{n,\,ijk\ell}\to D_t^{ijk\ell}:=\tfrac12\sum_{1\le u,\,v\le q}\int_0^t c_s^{ik}c_s^{j\ell}\,ds. \tag{5.4}$$

Note that (5.2) is the minimal condition under which the process Z of (5.3) is well defined. We divide the proof into several steps.

LEMMA 5.2. *Conditions* (b) *and* (c) *in Theorem 5.1 are equivalent, and they imply* (a).

PROOF. Observe that $dC_t=c_t\,dA_t$ for some increasing continuous process A and some $d\times d$ symmetric nonnegative matrix-valued process c. This last property readily implies that (a) is equivalent to the fact that each C^{ii} satisfies

(4.2). So indeed to prove our lemma we can and will suppose that Y, hence Z^n and D^n as well, are one-dimensional.

(i) Set $S(n, p) = \inf(t\colon \sqrt{n}|Z_t^n| \geq p)$ and $T(n, p) = \inf(t\colon D_t \geq p)$, and also $\tilde{Z}_t^n = \sqrt{n} \sup_{s \leq t} |Z_s^n|$.

First assume (b). Then $\sup_n P(S(n, p) < 1) \leq \sup_n P(Z^{n\star} \geq p) \to 0$ as $p \to \infty$, while $E(D^n_{S(n,p)}) \leq p^2$; hence

$$P(D_1^n > q) \geq \frac{p^2}{q} + P(S(n, p) < 1)$$

goes to 0 uniformly in n as $q \to \infty$, and (c) holds. Conversely (c) yields that $\sup_n P(T(n, p) < 1) \to 0$ as $p \to \infty$, while Doob's inequality yields $E((\tilde{Z}^n_{T(n, p)})^2) \leq 4p$, so we deduce (b) exactly as above.

(ii) From now on we assume (b) and (c). Note that $T_p = \inf(t\colon C_t \geq p)$ has $P(T_p < 1) \to 0$ as $p \to \infty$. If we stop the process Y at time T_p, the corresponding processes C, Z^n, D^n are also stopped at T_p. So clearly it suffices to prove (a) for the stopped processes: in other words we can and will assume that C_1 is bounded by a constant.

For every stopping time T set

$$\xi(n, i, T) = n \int_{T\wedge (i-1)/n}^{T\wedge i/n} (Y_s^{(n)})^2 \, dC_s$$

[recall notation (3.5)]. Observe that if N is a continuous martingale null at 0, then

$$N_t^4 = 4\int_0^t N_s^3 \, dN_s + 6\int_0^t N_s^2 \, d\langle N, N\rangle_s.$$

If furthermore the variable $\langle N, N\rangle_1$ is bounded, the process $\int_0^t N_s^3 \, dN_s$ is a martingale; hence $E(N_1^4) = 6E(\int_0^1 N_s^2 d\langle N, N\rangle_s)$. Apply this to $N_t = Y_{t\wedge T\wedge (i/n)} - Y_{t\wedge T\wedge (i-1/n)}$ to get $E(\xi(n, i, T)) = (n/6)E((Y_{t\wedge T\wedge (i/n)} - Y_{t\wedge T\wedge (i-1/n)})^4)$. On the other hand a Burkholder–Gundy inequality yields a universal constant K such that

$$E\big((C_{t\wedge T\wedge (i/n)} - C_{t\wedge T\wedge (i-1/n)})^2\big) \leq KE\big((Y_{t\wedge T\wedge (i/n)} - Y_{t\wedge T\wedge (i-1/n)})^4\big).$$

Therefore

$$E\big((C_{t\wedge T\wedge (i/n)} - C_{t\wedge T\wedge (i-1/n)})^2\big) \leq 6KE(\xi(n, i, T)). \tag{5.5}$$

Observe that $D_t^n = n\int_0^t (Y_s^{(n)})^2 dC_s$; hence $\sum_{i=1}^n \xi(n, i, T(n, p)) = D^n_{1\wedge T(n, p)} \geq p$. Letting

$$\Gamma_t^n = n\left(\sum_{i=1}^{[nt]} (\Delta_i^n C)^2 + (C_t^{(n)})^2\right) = n\sum_{i=1}^n (C_{t\wedge (i/n)} - C_{t\wedge (i-1/n)})^2,$$

we deduce from (5.5) that $E(\Gamma^n_{T(n, p)}) \leq 6Kp$.

Finally, let $\gamma_n = n\sum_{i=1}^n (\Delta_i^n C)^2$. Since $\gamma_n = \Gamma^n_{T(n,\,p)}$ on the set $\{T(n,\,p) \geq 1\}$, we have

$$P(\gamma_n > q) \leq P(T(n,\,p) < 1) + P(\Gamma^n_{T(n,\,p)} > q) \tag{5.6}$$
$$\leq P(T(n,\,p) < 1) + \frac{6Kp}{q}.$$

We have $\lim_p \sup_n P(T(n,\,p) < 1) = 0$ by (c); hence (3.6) yields that the sequence γ_n is tight, and (a) follows from Theorem 4.1. □

Now we assume (5.2), and we let σ be a $d \times q$ matrix-valued process such that $c = \sigma\sigma^\dagger$, for some q. Up to enlarging the space, we can assume that there is a Wiener process $W' = (W'^i)_{1\leq i \leq q}$ such that

$$Y_t = \int_0^t \sigma_s \, dW'_s. \tag{5.7}$$

By virtue of [3], if we prove that for all $t \in (0, 1]$ and all $i, j \leq d$ and $k \leq q$ we have, with notation (5.1) and (5.4),

$$D_t^n \to^P D_t, \qquad \sqrt{n}\,\langle Z^{n,\,ij}, W'^k\rangle_t \to^P 0, \tag{5.8}$$

then the processes $\sqrt{n}\,Z^n$ will converge stably in law to the process Z of (5.3): we deduce that the pair $(Y, \sqrt{n}\,Z^n)$ converges in law to (Y, Z) for the product topology on $\mathbb{D}(\mathbb{R}^d) \times \mathbb{D}(\mathbb{R}^{d^2})$, and since all these processes are continuous we also have convergence for the Skorohod topology on $\mathbb{D}(\mathbb{R}^{d+d^2})$. Further, (5.4) follows from (5.8), and it implies ($\star$) for the sequence $\sqrt{n}\,Z^n$. In other words, to prove Theorem 5.1 it remains only to show (5.8).

We begin with a lemma.

LEMMA 5.3. *We have (5.8) as soon as the process σ has the form*

$$\sigma_s = \sum_{i=1}^m A_{i-1} 1_{(t_{i-1},\,t_i]}(s), \tag{5.9}$$

where $0 = t_0 < t_1 < \cdots < t_m = 1$ and where each A_i is a bounded $\mathscr{F}_{t_i}$-measurable random variable with values in $d \times q$ matrices.

PROOF. (i) Set $\tau_n(u) = u - [nu]/n$. By a Burkholder–Gundy inequality we have, for some constant K,

$$E\big((Y_t^{(n)})^4\big) \leq K/n^2. \tag{5.10}$$

Recall that $Y^{(n)}$ is defined in (3.5). Since $Y_u^{(n)} = A_r W_u'^{(n)}$ for $t_r < [nu]/n \leq u \leq t_{r+1}$, simple computations show that, for $t_r < [nu]/n \leq u \leq v \leq t_{r+1}$, and with $B_r = A_r A_r^\dagger$,

$$E(Y_u^{i,\,(n)} Y_v^{j,\,(n)} | \mathscr{F}_{s_r}) = B_r^{ij} \tau_n(u) 1_{\{[nu]=[nv]\}}, \tag{5.11}$$

(5.12)
$$E(Y_u^{i,(n)} Y_v^{i,(n)} Y_u^{j,(n)} Y_v^{j,(n)} | \mathscr{F}_{s_r}) = \tau_n(u)\tau_n(v)(B_r^{ij})^2 + \tau_n(u)^2\big((B_r^{ij})^2 + B_r^{ii} B_r^{jj}\big) 1_{\{[nu]=[nv]\}}.$$

(ii) Let us fix r and t such that $0 < t \le t_{r+1} - t_r$. We have

$$D_{t_r+t}^{n,ijk\ell} - D_{t_r}^{n,ijk\ell} = n B_r^{j\ell} \int_{t_r}^{t_r+t} Y_u^{i,(n)} Y_u^{k,(n)}\, du,$$
$$D_{t_r+t}^{ijk\ell} - D_{t_r}^{ijk\ell} = \tfrac{1}{2} B_r^{j\ell} B_r^{ik} t,$$
$$\sqrt{n}\, \langle Z^{n,ij}, W'^k\rangle_{t_r+t} - \sqrt{n}\, \langle Z^{n,ij}, W'^k\rangle_{t_r} = \sqrt{n}\, B_r^{jk} \int_{t_r}^{t_r+t} Y_u^{i,(n)}\, du.$$

So it is enough to prove that

$$n \int_{t_r}^{t_r+t} Y_u^{i,(n)} Y_u^{k,(n)} du \to^{\mathbb{L}^2} \frac{t}{2} B_r^{ik}, \qquad \sqrt{n} \int_{t_r}^{t_r+t} Y_u^{i,(n)} du \to^{\mathbb{L}^2} 0.$$

(iii) Setting $s(n) = ([nt_r]+1)/n$, we have $s(n) \to t_r$ and

$$n \int_{t_r \vee s(n)}^{t_r+t} Y_u^{i,(n)} Y_u^{k,(n)} du \to^{\mathbb{L}^2} 0 \quad \text{and} \quad \sqrt{n} \int_{t_r \vee s(n)}^{t_r+t} Y_u^{i,(n)} du \to^{\mathbb{L}^2} 0$$

by (5.10). So it remains to prove that

$$\alpha_n := n \int_{s(n)}^{t_r+t} Y_u^{i,(n)} Y_u^{k,(n)}\, du - \frac{t}{2} B_r^{ik} \ \to^{\mathbb{L}^2}\ 0,$$
$$\beta_n := \sqrt{n} \int_{s(n)}^{t_r+t} Y_u^{i,(n)} du \ \to^{\mathbb{L}^2}\ 0.$$

Using (5.10) and $0 \le \tau_n(u) \le 1/n$ and the boundedness of B_r, we get

$$E(\beta_n^2) \le n \int_{[s(n),\, t_r+t]^2} \frac{K}{n} 1_{\{[nu]=[nv]\}}\, du\, dv \le \frac{2tK}{n} \ \to\ 0$$

for some constant K. Similarly,

$$E(\alpha_n^2) = n^2 \int_{[s(n),\, t_r+t]^2} E(Y_u^{i,(n)} Y_u^{k,(n)} Y_v^{i,(n)} Y_v^{k,(n)})\, du\, dv + \frac{t^2}{4} E\big((B_r^{ik})^2\big) - nt \int_{s(n)}^{t_r+t} E(B_r^{ik} Y_u^{i,(n)} Y_u^{k,(n)})\, du.$$

On the one hand, we have, by (5.11),

$$nt \int_{s(n)}^{t_r+t} E(B_r^{ik} Y_u^{i,(n)} Y_u^{k,(n)})\, du \ \to\ \frac{t^2}{2} E((B_r^{ik})^2).$$

On the other hand, (5.12) yields

$$\big| E\big(Y_u^{i,(n)} Y_u^{k,(n)} Y_v^{i,(n)} Y_v^{k,(n)}\big) - \tau_n(u)\tau_n(v) E\big((B_r^{ik})^2\big)\big| \le \frac{K}{n^2} 1_{\{[nu]=[nv]\}},$$

and thus

$$n^2 \int_{[s(n),\, t_r+t]^2} E(Y_u^{i,(n)} Y_u^{k,(n)} Y_v^{i,(n)} Y_v^{k,(n)})\, du\, dv \;\to\; \frac{t^2}{4} E((B_r^{ik})^2).$$

Putting these results together gives $E(\alpha_n^2) \to 0$, and we are finished. □

PROOF OF THEOREM 5.1. It remains to prove that (a) implies (5.8) in the general case. Let $T_p = \inf(t\colon \int_0^t \|c_s\|^2 ds \ge p)$. Since (a) yields $P(T_p < 1) \to 0$, by localization it is clearly enough to prove the result for the processes stopped at time T_p, which amounts to assuming that $\int_0^1 \|c_s\|^2\, ds$ is bounded by a constant p.

The rest of the argument parallels Theorem 3.1. By a density result, and if c is such that $c = \sigma\sigma^\dagger$, there exists a sequence $\sigma(q)$ of processes of the form (5.9), such that

$$\eta_q := \int_0^1 \|\sigma_s - \sigma(q)_s\|^4\, ds \;\to\; 0,$$
(5.13)
$$\int_0^1 \|\sigma(q)_s\|^4\, ds \;\le\; \int_0^1 \|\sigma_s\|^4\, ds \le p.$$

Let $Y(q)_t = \int_0^t \sigma(q)_s dW'_s$, with the associated processes $Z(q)^n$ and $D(q)^n$ [see (5.1)] and $D(q)$ [see (5.4)]. By Lemma 5.3, for each q we have that the following converge for all t:

(5.14) $$D(q)_t^n \to^P D(q)_t, \qquad \sqrt{n}\,\langle Z(q)^{n,ij}, W'^k\rangle_t \to^P 0.$$

We have, with $c(q) = \sigma(q)\sigma(q)^\dagger$,

$$\begin{aligned}\big|D(q)_t^{n,ijk\ell} - D_t^{n,ijk\ell}\big| &= n\left|\int_0^t (Y(q)_s^{i,(n)} Y(q)_s^{k,(n)} c(q)_s j\ell - Y_s^{i,(n)} Y_s^{k,(n)} c_s^{j\ell})\, ds\right| \\ &\le n \int_0^t \|Y(q)_s^{(n)}\|^2 \|\sigma(q)_s - \sigma_s\|(\|\sigma(q)_s\| + \|\sigma_s\|)\, ds \\ &\quad + \int_0^t \|Y(q)_s^{(n)} - Y_s^{(n)}\|(\|Y(q)_s^{(n)}\| + \|Y_s^{(n)}\|)\|\sigma_s\|^2\, ds.\end{aligned}$$

By combining the Burkholder–Gundy and Cauchy–Schwarz inequalities, we get that

$$E\big(\|Y_s^{(n)}\|^4\big) \le \frac{K}{n} E\bigg(\int_{[ns]/n}^s \|\sigma(q)_u\|^4\, du\bigg),$$

thus $\int_0^t E(\|Y_s^{(n)}\|^4)\, ds \le K/n^2$ by (5.13) for some constant K changing from line to line, and also $\int_0^t E(\|Y(q)_s^{(n)}\|^4)\, ds \le K/n^2$. The same argument shows that

$$\int_0^t E(\|Y(q)_s^{(n)} - Y_s^{(n)}\|^4)\, ds \le \frac{K\eta_q}{n^2}.$$

Thus (5.13) and a repeated use of the Cauchy–Schwarz inequality gives $E(|D_t^{n,ijk\ell} - D_t^{n,ijk\ell}|) \le K\eta_q^{1/4}$, so by (5.14) we get the first part of (5.8). The second part of (5.8) is proved similarly (it is in fact a bit simpler). □

Let us now state a corollary of the previous result, which contains Theorem 1.2 as a particular case.

COROLLARY 5.4. *Assume that Y is a local martingale. There is equivalence between the following:*

(a) *We have (5.2).*
(b) *For $q = 1$ and $x_0 = 1$ and $f^i(x) = x\delta_{ij}$ [i.e., $X = \mathscr{E}(Y^j)$] the sequence $\sqrt{n}\,U^{n\star}$ is tight, for each $j = 1, 2, \dots, d$.*
(c) *For all starting points x_0 and all C^1 functions f with linear growth, the sequences $(Y, \sqrt{n}\,U^n)$ and $(Y, \sqrt{n}\,\overline{U}^n)$ weakly converge to a limit (Y, U).*

In this case U is the solution of the linear equation (3.2), with Z given in (5.3).

PROOF. That (c)⇒(b) is obvious, and (a) implies (c) and the last claim, due to Theorems 3.2 and 5.1 [the ($\star$) property for $\sqrt{n}\,Z^n$ follows from (5.4)].

Now assume (b), and fix j. Then with f, x_0 corresponding to $X = \mathscr{E}(Y^j)$, we have [as in (3.11)]

$$\sqrt{n}\,Z_t^{n,jj} = \int_0^t \bigl[(\sqrt{n}\,U_s^n / X_{\varphi_n(s)}^n)\,dY_s^j - (1/X_{\varphi_n(s)}^n)\,d(\sqrt{n}\,U^n)_s\bigr].$$

A sequence of continuous local martingales has ($\star$) as soon as the sequence of their suprema is tight, so here the sequence $\sqrt{n}\,U^n$ has ($\star$). By Theorem 3.1, the fact that X does not vanish and Theorem 2.3(c), we deduce that the sequence $\sqrt{n}\,Z^{n,jj}$ has ($\star$), and (a) follows from Theorem 5.1. □

5.2. *The semimartingale case.* Now we suppose that $Y = M + A$, where A is a continuous adapted process of finite variation and M is a continuous local martingale, both being d-dimensional and null at 0. Again $C = (C^{ij})_{1\le i,\, j\le d}$ is the quadratic variation process, that is, $C^{ij} = \langle Y^i, Y^j\rangle = \langle M^i, M^j\rangle$, and we write $Z^n = (Z^{n,ij})_{1\le i,\, j\le d}$, where $Z^{n,ij} = Z^n(Y^i, Y^j)$.

We do not know what happens in the general case, and only partial results are available, when A has the form

$$A_t^i = \int_0^t a_s^i\,ds \quad \text{with} \quad \int_0^1 (a_s^i)^2\,ds < \infty \quad \text{a.s.} \tag{5.15}$$

THEOREM 5.5. *Assume that Y is a continuous semimartingale and that (5.15) holds. Then there is equivalence between the following:*

(a) *We have (5.2).*
(b) *The sequence of processes $(\sqrt{n}\,Z^n)$ has ($\star$) and is tight.*

In this case, the sequence $(\sqrt{n}\, Z^n)$ *stably converges in law to a process* Z *of the form (5.3) (where* $c = \sigma\sigma^\dagger$*) and has* $(\star)$ *and we also have* $(Y, \sqrt{n}\, Z^n) \Rightarrow (Y, Z)$.

PROOF. (i) We set $F^{n,ij} = \sqrt{n}\, Z^n(M^i, M^j)$, $G^{n,ij} = \sqrt{n}\, Z^n(M^i, A^j)$, and $H^{n,ij} = \sqrt{n}\, Z^n(A^i, M^j)$ and $K^{n,ij} = \sqrt{n}\, Z^n(A^i, A^j)$, so $\sqrt{n}\, Z^n = F^n + G^n + H^n + K^n$.

By Theorem 4.1 the sequence K^n tends in variation to 0, a.s., and a fortiori has $(\star)$ (more precisely, the variations of the processes $\sqrt{n}\, K^{n,ii}$ are bounded a.s. uniformly in n, for each i, and a simple extension of Theorem 4.1 shows that this is also true for $\sqrt{n}\, K^{n,ij}$).

(ii) Suppose that (b) holds. Each $H^{n,ij}$ is a continuous local martingale, and

$$\langle H^{n,ij}, H^{n,ij}\rangle_t = n \int_0^t (A_s^{i,(n)})^2\, dC_s^{jj}. \tag{5.16}$$

By (5.15) and the Cauchy–Schwarz inequality, we obtain $n(A_s^{i,(n)})^2 \leq \int_0^1 (a_u^i)^2\, du$ for all n, s. Thus the sequence of processes $(\langle H^{n,ij}, H^{n,ij}\rangle)$ is tight, and it follows (see, e.g., [4]) that the sequence $(H^{n,ij})$ is tight and has $(\star)$, and of course all its limiting processes are continuous.

Then, using part (i) above and Theorem 2.3(a), we see by difference that the sequence $F^n + G^n$ is also tight and has $(\star)$. Then the quadratic variation processes, which are the same as the quadratic variations of the processes F^n, are also tight; in other words, the local martingale M satisfies condition (c) of Theorem 5.1, and (a) follows.

(iii) Now we assume that (5.2), as well as (5.15), holds. In view of part (i) and of Theorem 5.1, it remains to prove that both sequences (G^n) and (H^n) have $(\star)$ and weakly converge to 0. Since this is a componentwise property, we can and will assume that $d = 1$, and exactly as in Theorem 5.1 we also can and will assume that both random variables $\int_0^1 a_s^2\, ds$ and $\int_0^1 c_s^2\, ds$ are bounded by a constant α.

Let us first consider H^n. This is a continuous local martingale satisfying, by (5.16),

$$\langle H^n, H^n\rangle_1 \leq \sum_{i=1}^n \int_{(i-1)/n}^{i/n} a_s^2\, ds \int_{(i-1)/n}^{i/n} c_s\, ds \leq \alpha\gamma_n,$$

where $\gamma_n = \sup_{1\leq i\leq n} \int_{(i-1)/n}^{i/n} c_s\, ds$. We have $|\gamma_n| \leq \sqrt{\alpha}$ and $\gamma_n(\omega) \to 0$ for all ω. Thus $E(\langle H^n, H^n\rangle_1) \to 0$, and thus the sequence (H^n) has $(\star)$ and weakly converges to 0.

Next consider G^n. Let $\varepsilon > 0$ and let S and T be two stopping times such that $T \leq S \leq T + \varepsilon$ and $S \leq 1$. We have by the Cauchy–Schwarz inequality

$$E\biggl(\int_T^S |dG_s^n|\biggr) = E\biggl(\int_T^S \sqrt{n}\, |a_s M_s^{(n)}|\, ds\biggr)$$
$$\leq \biggl(E\biggl(\int_T^S a_s^2\, ds\biggr)\; E\biggl(n \int_0^1 |M_s^{(n)}|^2\, ds\biggr)\biggr)^{1/2}.$$

We also have

$$
\begin{aligned}
E\left(n\int_0^1 |M_s^{(n)}|^2\,ds\right) &= E\left(n\int_0^1 ds\int_{[ns]/n}^s c_u\,du\right)\\
&= E\left(\int_0^1 c_u([nu]+1-nu)\,du\right)\\
&\le E(C_1)\le\sqrt{\alpha}.
\end{aligned}
\tag{5.17}
$$

Therefore $E(\int_T^S |dG_s^n|)\le \alpha^{1/4}\sqrt{E(w(\varepsilon))}$, where $w(\varepsilon)=\sup_t\int_t^{t+\varepsilon}a_s^2\,ds$. Since we have $\lim_{\varepsilon\to 0} w(\varepsilon)=0$ and $w(\varepsilon)\le\alpha$, then $E(w(\varepsilon))\to 0$ and we can apply Aldous's criterion (see, e.g., [4]) and deduce that the sequence of variation processes of the processes G^n is tight, which implies in particular that the sequence (G^n) has $(\star)$.

It remains to prove that $G^n\Rightarrow 0$. In a first step we set $\overline{G}_t^n=\int_0^t\sqrt{n}\,M_s^{(n)}ds$, which is the process G^n above when $a_t=1$ (i.e., $A_t=t$). So the sequence $(\overline{G}^n)$ is tight, and we also have

$$E((\overline{G}_t^n)^2)=2n\int_0^t du\int_0^u E(M_u^{(n)}M_v^{(n)})\,dv.$$

Since M is a martingale, $E(M_u^{(n)}M_v^{(n)})$ equals 0 if $[nv]\ne[nu]$ and $E(\int_{[nv]/n}^{v\wedge u}c_s\,ds)$ if $[nv]=[nu]$. Hence

$$E\big((\overline{G}_t^n)^2\big)=2E\left(\int_0^t du\int_{[nu]/n}^u dv\int_{[nu]/n}^v c_s\,ds\right)\le\frac{2}{n}E(C_1)\to 0.$$

It follows from all these that $\overline{G}^n\Rightarrow 0$.

In a second step, we assume that a is of the form (5.9). Then

$$G_t^n=\sum_{k=1}^m A_{k-1}(\overline{G}_{t\wedge t_k}^n-\overline{G}_{t\wedge t_{k-1}}^n),$$

so $G^n\Rightarrow 0$ by the first step. Finally, in the general case there is a sequence $a(p)$ of processes of the form (5.9) such that

$$\delta_p:=E\left(\int_0^1 |a(p)_s-a_s|^2\,ds\right)\to 0.$$

Setting $A(p)_t=\int_0^t a(p)_s\,ds$ and $G(p)^n=\sqrt{n}\,Z^n(M,A(p))$, we have $G(p)^n\Rightarrow 0$ for every p. On the other hand,

$$
\begin{aligned}
E\Big(\sup_t|G_t^n-G(p)_t^n|\Big)&\le E\left(\int_0^1|\sqrt{n}\,M_s^{(n)}|a_s-a(p)_s|ds\right)\\
&\le E\left(\int_0^1 n(M_s^{(n)})^2\,ds\right)^{1/2}\sqrt{\delta_p}\le\sqrt{\alpha\delta_p}
\end{aligned}
$$

by (5.17). Thus $G^n\Rightarrow 0$, and we are finished. □

REMARK. Theorem 5.5 puts us in the situation where Theorem 3.2 applies: if Y is a continuous semimartingale with (5.15) and (5.2), for any starting point x_0 and any C^1-function f with linear growth the processes $\sqrt{n}\,U^n$ and $\sqrt{n}\,\overline{U}^n$ weakly converge to the solution of (3.7), with Z given by (5.3). □

6. Lévy processes.

In this last section we suppose that the driving process Y is a Lévy process, and to simplify we assume that it is one-dimensional (an extension to the multidimensional situation is rather straightforward). The characteristics of Y are (b, c, F), where $b \in \mathbb{R}$, $c \geq 0$ and F is a positive measure on $\mathbb{R}$ with $F(\{0\}) = 0$ and $\int x^2 \wedge 1\, F(dx) < \infty$. We denote by μ the jump random measure of Y, and we set $\nu(dt, dx) = dt \otimes F(dx)$, so Y has the form (see [4])

$$Y_t = bt + Y_t^c + x1_{\{|x|\leq 1\}} \star (\mu - \nu) + x1_{\{|x|>1\}} \star \mu, \tag{6.1}$$

where Y^c is the continuous martingale part of Y: it is 0 if $c = 0$ and $Y/\sqrt{c}$ is a standard Brownian motion otherwise, and its quadratic variation process is ct. Further, the "$\star$" in (6.1) indicates the stochastic integral of a predictable function w.r.t. a random measure (see [4]). Set

$$Z^n = Z^n(Y, Y), \qquad \overline{Z}_t^n = Z_{[nt]/n}^n. \tag{6.2}$$

Here, if $F = 0$ the process Y is a continuous semimartingale, to which the results of the previous section readily apply. On the contrary, when $F \neq 0$ the situation resembles that of Section 4 for discontinuous processes: we do not have convergence of $(\sqrt{n}\,Z^n)$ and $(\sqrt{n}\,U^n)$ in the Skorohod sense (unless $Y^c = 0$), but only finite-dimensional convergence in law, while the sequences $(\sqrt{n}\,\overline{Z}^n)$ and $(\sqrt{n}\,\overline{U}^n)$ weakly converge.

Let us first describe the limiting processes Z and U. We take (possibly on an extension of the space on which Y is defined) the following:

1. a standard Brownian motion W;
2. two sequences $(V'_n)_{n\geq 1}$ and $(V''_n)_{n\geq 1}$ of standard normal variables;
3. a sequence $(\chi_n)_{n\geq 1}$ of uniform variables on $(0, 1)$;

in such a way that all these terms are *mutually independent*, and are *independent of* Y as well. We also set

$$V_n = \sqrt{\chi_n}\, V'_n + \sqrt{1 - \chi_n}\, V''_n, \tag{6.3}$$

which gives another sequence of independent standard normal variables.

Let us also denote by $(S_n)_{n\geq 1}$ an arbitrary ordering of all jump times of Y, consisting of stopping times taking values in $(0, 1] \cup \{\infty\}$: if $F(\mathbb{R}) < \infty$, we may choose the sequence S_n to be increasing and the variable $K = \inf(n\colon S_n > 1)$ is a.s. finite; otherwise the S_n's cannot be ordered as an increasing sequence.

Now we are ready to describe the limiting processes for $\sqrt{n}\,Z^n$ and $\sqrt{n}\,U^n$. First, the limit of $\sqrt{n}\,Z^n$ will be

$$Z_t = \frac{c}{\sqrt{2}} W_t + \sqrt{c} \sum_{n\geq 1} V_n \,\Delta Y_{S_n} 1_{[S_n,1]}(t). \tag{6.4}$$

Note that $Z = 0$ if $c = 0$. Since $\sum_{n\geq 1} (\Delta Y_{S_n})^2 < \infty$ a.s., it is not difficult to check that the last sum in (6.4) converges in $\mathbb{L}^2$, conditionally on the σ-field $\mathscr{F}$, and so converges in probability. There is another (more abstract) way of describing Z: it is, conditionally on $\mathscr{F}$, a Gaussian martingale null at 0 and with angle bracket [this bracket is not an $(\mathscr{F}_t)$-predictable process, but conditionally on $\mathscr{F}$ it becomes deterministic]

$$\langle Z, Z\rangle_t = \frac{c^2 t}{2} + c \sum_{s\leq t} (\Delta Y_s)^2 \tag{6.5}$$

or, equivalently, it is a Gaussian centered process with covariance function $(s,t) \rightsquigarrow \langle Z, Z\rangle_{s\wedge t}$ as given in (6.5). That these two descriptions characterize the same process (conditionally on $\mathscr{F}$) is easy, and it shows in particular that (6.4) does not depend on the particular choice of the sequence (S_n).

Next the limit of $\sqrt{n}\,U^n$ will be the unique solution of the following linear equation:

$$U_t = \int_0^t f'(X_{s-}) U_{s-}\, dY_s - \overline{Z}(f)_t, \tag{6.6}$$

where

$$\begin{aligned} \overline{Z}(f)_t = {} & \frac{c}{\sqrt{2}} \int_0^t (ff')(X_{s-})\, dW_s \\ & + \sqrt{c} \sum_{n:\, S_n\leq t} \biggl[\sqrt{\chi_n} V'_n (ff')(X_{S_n-}) + \sqrt{1-\chi_n} V''_n f(X_{S_n-}) \\ & \qquad \times \int_0^1 f'(X_{S_n-} + u\Delta X_{S_n})\, du \biggr] \Delta Y_{S_n}. \end{aligned} \tag{6.7}$$

Exactly as in (4.14), we may also write $\overline{Z}(f)$ as

$$\begin{aligned} \overline{Z}(f)_t = {} & \frac{c}{\sqrt{2}} \int_0^t (ff')(X_{s-})\, dW_s \\ & + \sqrt{c} \sum_{n:\, S_n\leq t} \Bigl[\sqrt{\chi_n} V'_n f'(X_{S_n-}) \,\Delta X_{S_n} \\ & \qquad + \sqrt{1-\chi_n} V''_n (f(X_{S_n}) - f(X_{S_n-})) \Bigr]. \end{aligned} \tag{6.8}$$

As in (6.4), the series on the right side of (6.7) and (6.8) are converging in measure. As for Z, another more abstract way of describing $\overline{Z}(f)$ is that, conditionally on $\mathscr{F}$, it is a Gaussian martingale null at 0 and with bracket

$$(6.9)\qquad \begin{aligned}\langle \overline{Z}(f), \overline{Z}(f)\rangle_t &= \frac{c^2}{2}\int_0^t (ff')(X_s)\,dW_s \\ &\quad + \frac{c}{2}\sum_{s\le t}\big(f'(X_{s-})^2(\Delta X_s)^2 + (f(X_s)-f(X_{s-}))^2\big).\end{aligned}$$

Here is the main result of this section. We exclude the following two simple cases:

1. $F=0$ and $c=0$ (i.e., $Y_t = bt$); then, by Theorem 4.1, the sequence (nZ^n) has ($\star$) and converges uniformly to $Z_t = b^2t/2$, and both sequences nU^n and $n\overline{U}^n$ converge uniformly to the unique solution of $dU_t = (U_t f'(X_t)b - f(X_t)f'(X_t)b^2/2)dt$ starting at 0 (here we are in the case of an ordinary, nonrandom, differential equation);
2. $F=0$ and $c>0$; then, by Theorem 5.5, the sequence $(\sqrt{n}\,Z^n)$ has ($\star$) and converges stably in law to $Z = cW/\sqrt{2}$ [as in (6.4) with $S_n=\infty$ for all n, i.e., the last sum in (6.4) disappears], and the sequences $\sqrt{n}\,U^n$ and $\sqrt{n}\,\overline{U}^n$ converge stably in law to the unique solution of $dU_t = U_t f'(X_t)\,dY_t - f(X_t)f'(X_t)\,dZ_t$ starting at 0, by Theorem 3.2.

THEOREM 6.1. *Assume that Y is a Lévy process such that $F\neq 0$. Let x_0 be any starting point, let f be any C^1-function with linear growth, and consider (3.1). The sequences $(\sqrt{n}\,Z^n)$ and $(\sqrt{n}\,U^n)$ have ($\star$), and we have the following:*

(a) *If $c=0$, the sequences $(\sqrt{n}\,Z^n)$, $(\sqrt{n}\,U^n)$ and $(\sqrt{n}\,\overline{U}^n)$ weakly converge to 0.*

(b) *If $c>0$, the sequence $(\sqrt{n}\,\overline{Z}^n, \sqrt{n}\,\overline{U}^n)$ stably converges in law to (Z,U), as given by (6.4) and (6.6)–(6.7), and the sequence $(\sqrt{n}\,Z^n, \sqrt{n}\,U^n)$ stably converges in finite-dimensional laws to the same limit.*

In case (b) above, the sequence $\sqrt{n}\,Z^n$ is *not* tight: if it were, by taking a convergent subsequence we could apply Theorem 3.2, but then the limit of $(\sqrt{n}\,U^n)$ would be given by (3.7), which is *not* the same equation as (6.6).

PROOF. The proof of Theorem 6.1 will go through several steps.

STEP 1 (Suppressing big jumps). Here we assume that Theorem 6.1 holds for Lévy processes having bounded jumps. Let Y be an arbitrary Lévy process, and set, for any $p>1$,

$$Y(p)_t = bt + Y_t^c + x1_{\{|x|\le 1\}}\star(\mu-\nu) + x1_{\{1<|x|\le p\}}\star\mu.$$

Let $X(p)$ be the solution of (3.1) relative to $Y(p)$, and let $Z(p)^n$, $U(p)^n$, $Z(p)$ and $U(p)$ be the processes associated with $Y(p)$ and $X(p)$. Observe that in the definition of $Z(p)$ and $U(p)$ we can use the same sequence of stopping times (S_n) and the same terms (W, V_n', V_n'', χ_n) as for Z and U.

Also let $\Omega_p = \{\omega: |\Delta Y_s|\le p\ \forall\ s\in(0,1]\}$. Then Ω_p increases to Ω, while on Ω_p we have the following equalities between processes: $X(p)=X$, $Z(p)^n =$

Z^n, $U(p)^n = U^n$, $Z(p) = Z$, $U(p) = U$. Since our theorem holds for each $Y(p)$, it follows that it also holds for Y.

Therefore, from now on we assume that *the jumps of Y are bounded by a constant*, that is, the measure F has compact support.

STEP 2 [The ($\star$) property]. For a moment, let M and N be two martingales with angle brackets $\langle M, M\rangle_t = \alpha t$ and $\langle N, N\rangle_t = \beta t$, and set $A_t = at$. Then

$$\begin{aligned} E(\langle Z^n(M,N), Z^n(M,N)\rangle_1) &= \beta \int_0^1 E((M_s^{(n)})^2)\,ds \\ &= \alpha\beta \int_0^1 \left(s - \frac{[ns]}{n}\right) ds = \frac{\alpha\beta}{2n}, \end{aligned} \tag{6.10}$$

$$E(\langle Z^n(A,N), Z^n(A,N)\rangle_1) = \beta a^2 \int_0^1 \left(s - \frac{[ns]}{n}\right)^2 ds = \frac{\beta a^2}{3n^2}, \tag{6.11}$$

$$\begin{aligned} E\left(\int_0^1 |dZ^n(M,A)_s|\right) &= |a| \int_0^1 E(|M_s^{(n)}|)\,ds \\ &\le |a| \int_0^1 \left(\alpha\left(s - \frac{[ns]}{n}\right)\right)^{1/2} ds \le \frac{2|a|\sqrt{\alpha}}{3\sqrt{n}}, \end{aligned} \tag{6.12}$$

$$E\left(\int_0^1 |dZ^n(A,A)_s|\right) = a^2 \int_0^1 \left(s - \frac{[ns]}{n}\right) ds = \frac{a^2}{2n}. \tag{6.13}$$

Let us come back to Y. Since F has compact support, we can set

$$b' = b + \int_{|x|>1} xF(dx), \qquad \alpha = c + \int x^2 F(dx),$$
$$B_t = b't, \qquad M = Y^c + x \star (\mu - \nu),$$

so that $Y = B + M$, while $\langle M, M\rangle_t = \alpha t$. Then $Z^n = Z^n(M,M) + Z^n(M,B) + Z^n(B,M) + Z^n(B,B)$. The two sequences of local martingales $(\sqrt{n}\, Z^n(M,M))$ and $(\sqrt{n}\, Z^n(B,M))$ have ($\star$) by (6.10) and (6.11), and the two sequences of processes with finite variation $(\sqrt{n}\, Z^n(M,B))$ and $(\sqrt{n}\, Z^n(B,B))$ have ($\star$) by (6.12) and (6.13). Hence the sequence $(\sqrt{n}\, Z^n)$ has ($\star$), as well as $(\sqrt{n}\, U^n)$ by Theorem 3.3.

STEP 3 (Suppressing small jumps). (i) For $\varepsilon > 0$ we set

$$M^\varepsilon = x1_{\{|x|\le\varepsilon\}} \star (\mu - \nu), \qquad N^\varepsilon = x1_{\{|x|>\varepsilon\}} \star (\mu - \nu), \qquad A^\varepsilon = x1_{\{|x|>\varepsilon\}} \star \mu,$$

$$b_\varepsilon = b' - \int_{|x|>\varepsilon} xF(dx), \qquad \rho_\varepsilon = \int_{|x|>\varepsilon} x^2 F(dx), \qquad B_t^\varepsilon = b_\varepsilon t.$$

Then (6.1) readily yields

$$Y = B + Y^c + M^\varepsilon + N^\varepsilon = B^\varepsilon + Y^c + M^\varepsilon + A^\varepsilon.$$

A simple computation, using the bilinearity of $(U,V) \rightsquigarrow Z^n(U,V)$, gives

$$\sqrt{n}\, Z^n = F^{n,\varepsilon} + G^{n,\varepsilon}, \tag{6.14}$$

where

$$F^{n,\varepsilon} = H^{n,\varepsilon} + I^{n,\varepsilon}, \qquad G^{n,\varepsilon} = J^{n,\varepsilon} + K^{n,\varepsilon} + L^{n,\varepsilon}$$

and

$$H^{n,\varepsilon} = \sqrt{n}\, Z^n(B^\varepsilon + Y^c, B^\varepsilon + Y^c),$$

$$I^{n,\varepsilon} = \sqrt{n}\,(Z^n(A^\varepsilon, Y^c) + Z^n(Y^c, A^\varepsilon)),$$

$$J^{n,\varepsilon} = \sqrt{n}\,(Z^n(M^\varepsilon, M^\varepsilon + N^\varepsilon + Y^c) + Z^n(N^\varepsilon + Y^c + B, M^\varepsilon)),$$

$$K^{n,\varepsilon} = \sqrt{n}\, Z^n(N^\varepsilon + B, N^\varepsilon + B), \qquad L^{n,\varepsilon} = \sqrt{n}\,(Z^n(M^\varepsilon, B) - Z^n(B^\varepsilon, B^\varepsilon)).$$

(ii) Observe that $\langle N^\varepsilon, N^\varepsilon\rangle_t = \rho_\varepsilon t$ and $\langle Y^c, Y^c\rangle_t = ct$ and $\langle M^\varepsilon, M^\varepsilon\rangle_t = (\rho_0 - \rho_\varepsilon)t$. We deduce from (6.12) and (6.13) that $E(\int_0^1 |dL_s^{n,\varepsilon}|) \le 2|b'|\sqrt{\rho_0 - \rho_\varepsilon}/3 + b_\varepsilon^2/\sqrt{n}$, so

$$\lim_{\varepsilon\to 0} \limsup_n E\left(\int_0^1 |dL_s^{n,\varepsilon}|\right) = 0. \tag{6.15}$$

Next, use (6.10) and (6.11) to obtain that the local martingale $J^{n,\varepsilon}$ has $\langle J^{n,\varepsilon}, J^{n,\varepsilon}\rangle_1 \le 6(\rho_0 - \rho_\varepsilon)(\rho_0 + c + b'^2)t$. Therefore, using Doob's inequality,

$$\lim_{\varepsilon\to 0} \sup_n E(\langle J^{n,\varepsilon}, J^{n,\varepsilon}\rangle_1) = 0. \tag{6.16}$$

Next, the process $B + N^\varepsilon$ is the sum of a continuous process with finite variation having (4.2) and a process of the form (4.6), and the associated variables α_i^n and $\overline{\beta}_i^n$ [see (4.9) and (4.10)] are bounded by a constant independent of n, and the number of jumps of this process is a Poisson random variable. Then $\sup_n E(\sqrt{n}\int_0^1 |dK_t^{n,\varepsilon}|) < \infty$, and

$$\lim_n E\left(\int_0^1 |dK_s^{n,\varepsilon}|\right) = 0 \qquad \forall\ \varepsilon > 0. \tag{6.17}$$

Putting together (6.15), (6.16) and (6.17), we readily deduce that if δ^n is a sequence of predictable processes such that $\delta^{n\star}$ is tight, then

$$\lim_{\varepsilon\to 0} \limsup_n P((\delta^n \cdot G^{n,\varepsilon})^\star > \eta) = 0 \qquad \forall\ \eta > 0. \tag{6.18}$$

(iii) We can now prove (a). Suppose that $c = 0$. Then $Y^c = 0$, so $I^{n,\varepsilon} = 0$ and $H^{n,\varepsilon} = \sqrt{n}\, Z^n(B^\varepsilon, B^\varepsilon)$ converges weakly to 0 by Theorem 4.1 for each $\varepsilon > 0$. Then $F^{n,\varepsilon} \Rightarrow 0$ as $n \to \infty$, and combining this with (6.14) and (6.18) with $\delta^n = 1$ yields that $\sqrt{n}\, Z^n \Rightarrow 0$. Then we can apply Theorem 3.2 to obtain $\sqrt{n}\, U^n \Rightarrow 0$, and thus (a) holds.

(iv) From now on we suppose that $c > 0$. Recall that $\sqrt{n}\, U^n$ is the solution of (3.13), with $\alpha_n = \sqrt{n}$, and introduce the solution V_n of the following linear equation:

$$dV_t^n = V_{t-}^n f'(X_{t-})\, dY_t - k(X)_t^n\, d(\sqrt{n}\, Z^n)_t, \qquad V_0^n = 0. \tag{6.19}$$

As in the proof of Theorem 4.6, Theorems 2.2(d), 3.1 and 2.5(b) and the ($\star$) property of $\sqrt{n}\,Z^n$ give that $(\sqrt{n}\,U^n - V^n)^\star \to^P 0$.

Next, we introduce the solution $V^{n,\varepsilon}$ of the following linear equation:

$$dV_t^{n,\varepsilon} = V_{t-}^{n,\varepsilon} f'(X_{t-})\,dY_t - k(X)_t^n\,dF_t^{n,\varepsilon}, \qquad V_0^{n,\varepsilon} = 0. \tag{6.20}$$

Also set $R_t^n = \int_0^t k(X)_t^n\,d(\sqrt{n}\,Z^n)_t$ and $R_t^{n,\varepsilon} = \int_0^t k(X)_t^n\,dF_t^{n,\varepsilon}$. Lemma 2.4 yields

$$P((V^n - V^{n,\varepsilon})^\star > \eta) \le \varepsilon' + P(f(X)^\star > A) + P(R^{n\star} > u) + P((R^n - R^{n,\varepsilon})^\star > w) + \frac{w}{\eta} K_{A,\varepsilon'}$$

for a constant $K_{A,\varepsilon'}$ depending on A, ε'. Since $\sqrt{n}\,Z^n$ has ($\star$), and $k(X^n)^\star$ is tight, Theorem 2.2(c) shows that $R^{n\star}$ is tight. On the other hand, $R_t^n - R_t^{n,\varepsilon} = \int_0^t k(X)_t^n\,dG_t^{n,\varepsilon}$, so (6.18) implies that $\lim_{\varepsilon\to 0}\limsup_n P((R^n - R^{n,\varepsilon})^\star > w) = 0$ for all $w > 0$. Thus one readily deduces, by taking ε' arbitrary, then A and u big, then w small, that

$$\lim_{\varepsilon\to 0}\limsup_n P((V^n - V^{n,\varepsilon})^\star > \eta) = 0$$

for all $\eta > 0$. In view of what precedes, we thus obtain

$$\lim_{\varepsilon\to 0}\limsup_n P((\sqrt{n}\,U^n - V^{n,\varepsilon})^\star > \eta) = 0 \qquad \forall\, \eta 0. \tag{6.21}$$

On the other hand, define $Z(\varepsilon)$ and $\overline{Z}(\varepsilon, f)$ by (6.4) and (6.7) [or (6.8)], except that in the sum of the right side we add the indicator function of the set $\{|\Delta Y_{S_n}| > \varepsilon\}$. It is easy to check that

$$\lim_{\varepsilon\to 0} P\big((Z - Z(\varepsilon))^\star + (\overline{Z}(f) - \overline{Z}(\varepsilon, f))^\star > \eta\big) = 0 \qquad \forall\, \eta > 0. \tag{6.22}$$

Then if $U(\varepsilon)$ is the solution of (6.6) with $\overline{Z}(\varepsilon, f)$ instead of $\overline{Z}(f)$, we also have, by (6.22) and Theorem 2.5,

$$\lim_{\varepsilon\to 0} P((U - U(\varepsilon))^\star > \eta) = 0 \qquad \forall\, \eta > 0. \tag{6.23}$$

Putting together (6.18), (6.21), (6.22) and (6.23), we see that in order to obtain (a), it is enough to prove that for each $\varepsilon > 0$ and if $\overline{F}_t^{n,\varepsilon} = F_{[nt]/n}^{n,\varepsilon}$ and $\overline{V}_t^{n,\varepsilon} = V_{[nt]/n}^{n,\varepsilon}$, then

$$(\overline{F}^{n,\varepsilon}, \overline{V}^{n,\varepsilon}) \Rightarrow^{\text{stably}} (Z(\varepsilon), U(\varepsilon)), \tag{6.24}$$

(6.25) $(F^{n,\varepsilon}, V^{n,\varepsilon})$ stably converges in finite-dimensional law to $(Z(\varepsilon), U(\varepsilon))$.

STEP 4. From now on we fix $\varepsilon > 0$. We denote by $0 < T_1 < \cdots < T_n < \cdots$ the successive jump times of Y with size bigger than ε. The number of T_n

with $T_n < \infty$ (or, equivalently, $T_n \le 1$) is a Poisson random variable K, with parameter $F(\mathbb{R} \setminus [-\varepsilon, \varepsilon])$. The processes $Z(\varepsilon)$ and $Z(\varepsilon, f)$ are given by

$$Z(\varepsilon)_t = \frac{c}{\sqrt{2}} W_t + \sqrt{c} \sum_{n \ge 1} \zeta_n \Delta Y_{T_n} 1_{[T_n, 1]}(t) \tag{6.26}$$

and

$$\begin{aligned} \overline{Z}(\varepsilon, f)_t = {} & \frac{c}{\sqrt{2}} \int_0^t (ff')(X_{s-}) dW_s \\ & + \sqrt{c} \sum_{n: T_n \le t} \bigg[\sqrt{\xi_n} \zeta_n' (ff')(X_{T_n -}) \\ & \qquad + \sqrt{1 - \xi_n} \zeta_n'' f(X_{T_n -}) \int_0^1 f'(X_{T_n -} + u \Delta X_{T_n}) du \bigg] \Delta Y_{T_n}, \end{aligned} \tag{6.27}$$

where the family $(W, \xi_n, \zeta_n', \zeta_n'')$ has the same properties as (W, χ_n, V_n', V_n'') and

$$\zeta_n = \sqrt{\xi_n} \zeta_n' + \sqrt{1 - \xi_n} \zeta_n'' \tag{6.28}$$

(this comes from relabelling the χ_n, V_n', V_n'''s).

Now, we associate with each T_j the times $T_+(n, j)$ and $T_-(n, j)$ by (4.8), and replace (4.9) by

$$\begin{aligned} & \alpha_j^n = \sqrt{n}\, \Delta Y_{T_j} (Y^c_{T_j} - Y^c_{T_-(n,j)}), \qquad \beta_j^n = \sqrt{n}\, \Delta Y_{T_j} (Y^c_{T_+(n,j)} - Y^c_{T_j}), \\ & \gamma_j^n = \alpha_j^n + \beta_j^n, \end{aligned} \tag{6.29}$$

these quantities being 0 if $T_j = \infty$. We also write $\rho_n = (\alpha_j^n, \beta_j^n)_{j \ge 1}$. Finally, set

$$\begin{aligned} & \alpha_j = \sqrt{c \xi_j} \zeta_j' \Delta Y_{T_j}, \qquad \beta_j = \sqrt{c(1 - \xi_j)} \zeta_j'' \Delta Y_{T_j}, \\ & \gamma_j = \alpha_j + \beta_j = \sqrt{c} \zeta_j \Delta Y_{T_j} \end{aligned} \tag{6.30}$$

and $\rho = (\alpha_j, \beta_j)_{j \ge 1}$.

LEMMA 6.2. *We have* $(H^{n, \varepsilon}, \rho_n) \Rightarrow^{\text{stably}} ((c/\sqrt{2}) W, \rho)$.

PROOF. For simplicity we write $H = (c/\sqrt{2}) W$. Let $Y' = Y - Y^c - A^\varepsilon$, and

$$R_t^n = Y_t^c - \sum_{j \le 1} (Y^c_{t \wedge T_+(n,j)} - Y^c_{t \wedge T_+(n,j)}),$$

$$S_t^n = H_t^{n, \varepsilon} - \sum_{j \le 1} (H^{n, \varepsilon}_{t \wedge T_+(n,j)} - H^{n, \varepsilon}_{t \wedge T_+(n,j)}).$$

In order to prove the result, we need to show that

$$E(h(Y) g(H^{n, \varepsilon}, \rho_n)) \to E(h(Y) g(H, \rho))$$

for all bounded functions h and uniformly continuous bounded functions g. By the same density argument as in Lemma 2.1, it is enough to prove this when $h(Y) = u(Y')v(Y^c)w(A^\varepsilon)$, where u, v, w are bounded functions, with in addition v continuous. Now, clearly $(R^n - Y^c)^\star \to 0$, and Theorem 5.5 yields $H^{n,\varepsilon} \Rightarrow^{\text{stably}} H$, so $(S^n - H^{n,\varepsilon})^\star \to^P 0$. Hence it suffices to prove that

$$E(u(Y')v(R^n)w(A^\varepsilon)g(S^n, \rho_n)) \to E(u(Y')v(Y^c)w(A^\varepsilon)g(H, \rho)). \tag{6.31}$$

In fact it is even enough to prove (6.31) when w depends only on the k first jump times and sizes of A^ε, and g depends only on S^n and on the k first variables α_j^n and β_j^n. Further, the set Ω_n (depending also on k) on which each interval $((i-1)/n, i/n]$ contains at most one T_j tends to Ω; hence we can put the indicator function in the left expectation of (6.31). So it remains to prove that

$$\begin{aligned} &E\big[u(Y')v(R^n)w((T_j, \Delta Y_{T_j})_{1\le j\le k})g(S^n, (\alpha_j^n, \beta_j^n)_{1\le j\le k})1_{\Omega_n}\big] \\ &\qquad \to E\big[u(Y')v(Y^c)w((T_j, \Delta Y_{T_j})_{1\le j\le k})g(H, (\alpha_j, \beta_j)_{1\le j\le k})\big]. \end{aligned} \tag{6.32}$$

Now, (6.29) and the independence of the increments of Y^c over all the intervals $(T_-(n, j), T_+(n, j)]$ from all the other random terms appearing in the left side of (6.32) yield that in this left side we can replace α_j^n and β_j^n by $\sqrt{c\xi_j^n}\zeta_j'\Delta Y_{T_j}$ and $\sqrt{c(1-\xi_j^n)}\zeta_j''\Delta Y_{T_j}$, where $\xi_i^n = n(T_j - T_-(n, j))$ is the fractional part of nT_j. Using once more $(R^n - Y^c)^\star \to 0$ and $(S^n - H^{n,\varepsilon})^\star \to^P 0$, we then see that this left-hand side has the same limit as

$$\begin{aligned} &E\Big[u(Y')v(Y^c)w((T_j, \Delta Y_{T_j})_{1\le j\le k}) \\ &\qquad \times g\Big(H^{n,\varepsilon}, \Big(\sqrt{c\xi_j^n}\zeta_j'\Delta Y_{T_j}, \sqrt{c(1-\xi_j^n)}\zeta_j''\Delta Y_{T_j}\Big)_{1\le j\le k}\Big)1_{\Omega_n}\Big]. \end{aligned}$$

Now if F_k and G_k denote the laws of the k-tuplets $(\Delta Y_{T_j})_{1\le j\le k}$ and $(T_j)_{1\le j\le k}$ (which are independent), the previous expression becomes (with $\{u\}$ denoting the fractional part of u)

$$\begin{aligned} &\int F_k(dx_1, \dots, dx_k)G_k(dt_1, \dots, dt_k)1_{\bigcap_{1<i<k}\{[nt_i]<[nt_{i+1}]\}} \\ &\quad \times E\Big[u(Y')v(Y^c)w((t_j, x_j)_{1\le j\le k}) \\ &\qquad \times g\Big(H^{n,\varepsilon}, \Big(\sqrt{c\{t_j/n\}}\zeta_j' x_j, \sqrt{c(1-\{t_j/n\})}\zeta_j'' x_j\Big)_{1\le j\le k}\Big)\Big]. \end{aligned}$$

Now we can use the property $H^{n,\varepsilon} \Rightarrow^{\text{stably}} H$ and the uniform continuity of g to get that the above has the same limit as

$$\begin{aligned} &\int F_k(dx_1, \dots, dx_k)G_k(dt_1, \dots, dt_k)1_{\bigcap_{1<i<k}\{[nt_i]<[nt_{i+1}]\}} \\ &\quad \times E\Big[u(Y')v(Y^c)w((t_j, x_j)_{1\le j\le k}) \\ &\qquad \times g\Big(H, \Big(\sqrt{c\{t_j/n\}}\zeta_j' x_j, \sqrt{c(1-\{t_j/n\})}\zeta_j'' x_j\Big)_{1\le j\le k}\Big)\Big] \end{aligned}$$

$$= E\Big[u(Y')v(Y^c)w((T_j, \Delta Y_{T_j})_{1\le j\le k}) \times g\Big(H, \Big(\sqrt{c\xi_j^n}\zeta_j' x_j, \sqrt{c(1-\xi_j^n)}\zeta_j'' x_j\Big)_{1\le j\le k}\Big)\Big].$$

At this point, we need to prove that the above converges to the right-hand side of (6.31). Because $T_1, \dots, T_k$ is independent of $Y^c, Y', H, \zeta_j', \zeta_j'', \Delta Y_{T_j}$, this amounts to proving that $(\xi_j^n)_{1\le j\le k} \Rightarrow^{\text{stably}} (\xi_j)_{1\le j\le k}$. However, since G_k has a regular density on its support, a trivial extension of one-dimensional results of Tukey [17] (see also [2]) gives this property. □

STEP 5. Now we turn to the proof of (6.24) and (6.25), which will reproduce the proof of Theorem 5.1 in a more complicated situation.

(i) The sets Ω_n on which each interval $((i-1)/n, i/n]$ contains at most one T_j tend to Ω. Then similarly to (4.18)–(4.20) we set

$$W(j)_t^n = \overline{V}^{n,\varepsilon}_{t\wedge T_+(n,j)},$$

$$W(j)_t'^n = \begin{cases} W(j)_t^n, & \text{if } t \le T_+(n, j), \\ W(j)^n_{T_+(n,j)} + V_t^{n,\varepsilon} - V^{n,\varepsilon}_{T_+(n,j)}, & \text{if } T_+(n,j) < t < T_{j+1}, \\ W(j)^n_{T_+(n,j)} + V^{n,\varepsilon}_{T_{j+1}-} - V^{n,\varepsilon}_{T_+(n,j)}, & \text{if } T_{j+1} \le t, \end{cases}$$

$$W(j)_t = U(\varepsilon)_{t\wedge T_j}, \qquad W(j)_t' = \begin{cases} U(\varepsilon)_t, & \text{if } t < T_{j+1}, \\ U(\varepsilon)_{T_{j+1}-}, & \text{if } t \ge T_{j+1}. \end{cases}$$

We also denote by $F(j)^n$, $F(j)'^n$, $F(j)$ and $F(j)'$ the processes obtained by replacing above $(\overline{V}^{n,\varepsilon}, V^{n,\varepsilon}, U(\varepsilon))$ by $(\overline{F}^{n,\varepsilon}, F^{n,\varepsilon}, Z(\varepsilon))$. We consider the property (recall that $H = (c/\sqrt{2})W$)

$$(\mathsf{H}_j) \qquad \big(\rho^n, H^{n,\varepsilon}, (F(j)^n, W(j)^n)\big) \Rightarrow^{\text{stably}} \big(\rho, H, (F(j), W(j))\big).$$

Observe that (H_0) holds by Lemma 6.2. If we have (H_j) for all j, then (6.24) holds, because $K < \infty$ a.s.

(ii) Suppose that (H_j) holds. Let $H(n, j)$ be the interval $(T_+(n, j), T_{j+1})$. Then [see (6.21)]

$$F(j)_t'^n = F(j)_t^n + \int_0^t 1_{H(n,j)}(s)\, dH_s^{n,\varepsilon} \tag{6.33}$$

and

$$W(j)_t'^n = W(j)_t^n + \int_0^t W(j)_{s-}'^n f'(X_{s-}) 1_{H(n,j)}(s)\, dY_s^c - \int_0^t k(X)_s^n 1_{H(n,j)}(s)\, dH_s^{n,\varepsilon}. \tag{6.34}$$

Then (4.22) holds with

$$J_t^n = W(j)_t^n - \int_0^t k(X)_s^n 1_{H(n,j)}(s)\, dH_s^{n,\varepsilon},$$

$$J_t = W(j)_t - \int_0^t f(X_{s-})f'(X_{s-})1_{(T_j,T_{j+1})}(s)\,dH_s,$$

$$L_t^n = \int_0^t f'(X_{s-})1_{H(n,j)}(s)\,dY_s^c, \qquad L_t = \int_0^t f'(X_{s-})1_{(T_j,T_{j+1})}(s)\,dY_s^c.$$

Clearly $(L^n - L)^\star \to^P 0$, while the sequence $k(X)_t^n$ is bounded (in n and t) by a finite random variable and converges to $(ff')(X_{s-})$ at each continuity point of Y, while the sequence $H^{n,\varepsilon}$ has ($\star$) and these processes are continuous. So $(J^n - J'^n)^\star \to^P 0$ if $J_t'^n = W(j)_t^n - \int_0^t (ff')(X_{s-})1_{H(n,j)}(s)dH_s^{n,\varepsilon}$. Then (H_j) yields $(J^n, L^n, \rho^n, H^{n,\varepsilon}) \Rightarrow^{\text{stably}} (J, L, \rho, H)$; hence Theorem 2.5(c) and (6.33) give that $((F(j)'^n, W(j)'^n), \rho^n, H^{n,\varepsilon}) \Rightarrow^{\text{stably}} ((F(j)', W(j)'), \rho, H)$. Therefore if $\overline{F}(j)_t^n = F(j)_{[nt]/n}^n$ and $\overline{W}(j)_t^n = W(j)_{[nt]/n}^n$, we get

$$\begin{aligned}(6.35)\qquad &\big((\overline{F}(j)^n_{\cdot\wedge T_-(n,\,j+1)}, \overline{W}(j)'^n_{\cdot\wedge T_-(n,j+1)}), F(j)'^n_{T_{j+1}-}, W(j)'^n_{T_{j+1}-}, \rho^n, H^{n,\varepsilon}\big)\\ &\quad\Rightarrow^{\text{stably}} \big((F(j)', W(j)'), F(j)'_{T_{j+1}-}, W(j)'_{T_{j+1}-}, \rho, H\big).\end{aligned}$$

(iii) Set $\mu_n := F^{n,\varepsilon}_{T_+(n,j+1)} - F^{n,\varepsilon}_{T_-(n,\,j+1)}$ and $\delta_n := V^{n,\varepsilon}_{T_+(n,j+1)} - V^{n,\varepsilon}_{T_-(n,\,j+1)}$. On the set $\Omega_n \cap \{T_j \le 1\}$ we have $\mu_n = a_n + \gamma_{j+1}^n$ and $\delta_n = u_n + v_n$, where (with $\overline{Y} = Y - A^\varepsilon$)

$$\begin{aligned}a_n &= H^{n,\varepsilon}_{T_+(n,\,j+1)} - H^{n,\varepsilon}_{T_-(n,\,j+1)},\\ u_n &= \int_{T_-(n,\,j+1)}^{T_+(n,\,j+1)} V^{n,\varepsilon}_{s-} f'(X_{s-})\,d\overline{Y}_s - \int_{T_-(n,\,j+1)}^{T_+(n,\,j+1)} k(X)_s^n\,dH_s^{n,\varepsilon},\\ v_n &= \Delta Y_{T_{j+1}}(V^{n,\varepsilon}_{T_{j+1}-}f'(X_{T_{j+1}-}) - k(X)^n_{T_{j+1}}\alpha^n_{j+1} - \int_{T_{j+1}}^{T_+(n,\,j+1)} k(X)_s^n\,dY_s^c).\end{aligned}$$

First, the sequences $V^{n\star}$ and $k(X)^{n\star}$ are tight, $\overline{Y}$ is continuous at time T_j, and $H^{n,\varepsilon} \Rightarrow H$ with H continuous, so one deduces that $a_n \to^P 0$ and $u_n \to^P 0$. Next the sequences

$$k(X)^n_{T_{j+1}} - (ff')(X_{T_{j+1}-})$$

and

$$\int_{T_{j+1}}^{T_+(n,\,j+1)} k(X)_s^n\,dY_s^c - \beta^n_{j+1} f(X_{T_{j+1}-})\int_0^1 f'(X_{T_{j+1}-} + u\Delta X_{T_{j+1}})\,du$$

converge to 0 in probability. Further, $F^{n,\varepsilon}_{T_{j+1}-} = F(j)'^n_{T_{j+1}-}$ and $V^{n,\varepsilon}_{T_{j+1}-} = W(j)'^n_{T_{j+1}-}$. Thus if

$$\begin{aligned}\delta = \Delta Y_{T_{j+1}}\bigg(&W(j)'_{T_{j+1}-}f'(X_{T_{j+1}-}) - \alpha_{j+1}(ff')(X_{T_{j+1}-})\\ &- \beta_{j+1}\int_0^1 f'(X_{T_{j+1}-} + u\Delta X_{T_{j+1}})\,du\bigg),\end{aligned}$$

one deduces from (6.35) that

$$((\overline{F}(j)'^n_{\cdot\wedge T_-(n,\,j+1)}, \overline{W}(j)'^n_{\cdot\wedge T_-(n,j+1)}), \mu_n, \delta_n, \rho^n, H^{n,\varepsilon}) \Rightarrow^{\text{stably}} ((F(j)', W(j)'), \gamma_{j+1}, \delta, \rho, H). \tag{6.36}$$

However,

$$F(j+1)^n = \overline{F}(j)'^n_{\cdot\wedge T_-(n,\,j+1)} + \mu_n 1_{[T_+(n,\,j+1),\,1]},$$
$$W(j+1)^n = \overline{W}(j)'^n_{\cdot\wedge T_-(n,\,j+1)} + \delta_n 1_{[T_+(n,\,j+1),\,1]},$$
$$F(j+1) = F(j)'_{\cdot\wedge T_{j+1}-} + \gamma_{j+1} 1_{[T_{j+1},\,1]},$$
$$W(j+1) = W(j)'_{\cdot\wedge T_{j+1}-} + \delta 1_{[T_{j+1},\,1]}.$$

Thus (6.36) yields (H_{j+1}), and the proof of (6.24) follows by induction on j.

(iv) Finally, on the set $\{T_+(n,j) < t < T_-(n,j+1)\}$ we have $F^{n,\varepsilon}_t = F(j)'^n_t$, $Z(\varepsilon)_t = F(j)'_t$, $V^{n,\varepsilon}_t = W(j)'^n_t$ and $U(\varepsilon)_t = W(j)'_t$. Since

$$(F(j)'^n, W(j)'^n) \Rightarrow^{\text{stably}} (F(j)', W(j)')$$

and since $Z(\varepsilon)$ and $U(\varepsilon)$ have no fixed times of discontinuity, we deduce (6.25), and the proof of Theorem 6.1 is complete. □

Acknowledgment. The authors want to thank an anonymous referee, whose remarks led to an improvement of the paper.

REFERENCES

[1] ALDOUS, D. J. and EAGLESON, G. K. (1978). On mixing and stability of limit theorems. *Ann. Probab.* 6 325–331.

[2] DELATTRE, S. and JACOD, J. (1996). A central limit theorem for normalized functions of the increments of a diffusion process, in presence of round-off errors. *Bernoulli*. To appear.

[3] JACOD, J. (1996). On continuous conditional Gaussian martingales and stable convergence in law. Publ. Lab. Probabilités 339, Univ. Paris-6.

[4] JACOD, J. and SHIRYAEV, A. N. (1987). *Limit Theorem for Stochastic Processes*. Springer, Berlin.

[5] JAKUBOWSKI, A. (1995). A non-Skorohod topology on the Skorohod space. Preprint, Torun Univ.

[6] JAKUBOWSKI, A., MÉMIN, J. and PAGÈS, G. (1989). Convergence en loi des suites d'intégrales stochastiques sur l'espace D^1 de Skorohod. *Probab. Theory Related Fields* 81 111–137.

[7] KURTZ, T. G. and PROTTER, P. (1991). Weak limit theorems for stochastic integrals and stochastic differential equations. *Ann. Probab.* 19 1035–1070.

[8] KURTZ, T. G. and PROTTER, P. (1991). Wong–Zakai corrections, random evolutions and numerical schemes for SDEs. In *Stochastic Analysis* 331–346. Academic Press, New York.

[9] KURTZ, T. G. and PROTTER, P. (1996). Weak convergence of stochastic integrals and differential equations. *1995 CIME School in Probability. Lecture Notes in Math.* 1627 1–41. Springer, Berlin.

[10] MÉMIN, J. and SŁOMIŃSKI, L. (1991). Condition UT et stabilité en loi des solutions d'équations différentielles stochastiques. *Séminaire de Probabilités XXV. Lecture Notes in Math.* 1485 162–177. Springer, Berlin.

[11] MEYER, P. A. and ZHENG, W. A. (1984). Tightness criteria for laws of semimartingales. *Ann. Inst. H. Poincaré Probab. Statist.* 20 353–372.

[12] PROTTER, P. (1990). *Stochastic Integration and Differential Equations*. Springer, Berlin.
[13] RÉNYI, A. (1963). On stable sequences of events. *Sankyā Ser. A* 25 293–302.
[14] SŁOMIŃSKI, L. (1989). Stability of strong solutions of stochastic differential equations. *Stochastic Process. Appl.* 31 173–202.
[15] SŁOMIŃSKI, L. (1996). Stability of stochastic differential equations driven by general semimartingales. *Dissertationes Math.* 349 1–113.
[16] TALAY, D. (1995). Simulation of stochastic differential systems. *Probabilistic Methods in Applied Physics. Lecture Notes in Physics* 451 63–106. Springer, New York.
[17] TUKEY, J. W. (1939). On the distribution of the fractional part of a statistical variable. *Mat Sb.* (*N.S.*) 4 561–562.

LABORATOIRE DE PROBABILITÉS
(CNRS URA 224)
UNIVERSITÉ PIERRE ET MARIE CURIE
TOUR 56, 4 PLACE JUSSIEU
75 252 PARIS CEDEX
FRANCE

DEPARTMENT OF MATHEMATICS
PURDUE UNIVERSITY
WEST LAFAYETTE, INDIANA 47907-1395
E-MAIL: protter@math.purdue.edu