A Wong-Zakai Approximation of Stochastic Differential Equations Driven by a General Semimartingale

Xianming Liu

School of Mathematics and Statistics,

Huazhong University of Science and Technology, Wuhan 430074, China

E-mail: xmliu@hust.edu.cn

Guangyue Han

Department of Mathematics,

The University of Hong Kong, Hong Kong, China

E-mail: ghan@hku.hk

February 23, 2019

Abstract

We examine a Wong-Zakai type approximation of a family of stochastic differential equations driven by a general càdlàg semimartingale. For such an approximation, compared with the pointwise convergence result by Kurtz, Pardoux and Protter [12, Theorem 6.5], we establish stronger convergence results under the Skorokhod M_1-topology, which, among other possible applications, implies the convergence of the first passage time of the solution to the approximating stochastic differential equation.

Key words: Wong-Zakai approximation, stochastic differential equation, semimartingale, the Skorokhod M_1-topology, random time change

1 Introduction

Let $L = \{L(t); 0 \le t < \infty\}$ be a stochastically continuous càdlàg semimartingale [19] defined on a probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$. For any $\epsilon > 0$, let L^ϵ be the smooth approximation [7, 12, 26] of L defined by

$$L^\epsilon(t) := \frac{1}{\epsilon}\int_{t-\epsilon}^{t} L(s)ds, \quad 0 \le t < \infty, \tag{1.1}$$

and let $X^\epsilon = \{X^\epsilon(t); 0 \le t < \infty\}$ be the solution to the following random differential equation

$$dX^\epsilon(t) = b(X^\epsilon(t))dt + f(X^\epsilon(t))dL^\epsilon(t), \quad X^\epsilon(0) = X_0, \tag{1.2}$$

where b, f are some functions from $\mathbb{R}$ to $\mathbb{R}$ satisfying certain regularity conditions, and X_0 is an $\mathcal{F}_0$-measurable random variable.

In this paper, we will establish some convergence results on X^ϵ as ϵ tends to 0. More precisely, we will show that, in some sense, X^ϵ converges to X, where X is the solution to the following stochastic differential equation:

$$X(t) = X_0 + \int_0^t b(X(s))ds + \int_0^t f(X(s-)) \diamond dL(s), \tag{1.3}$$

where $\diamond$ denotes Marcus integral. Here, let us add that (1.3) is in fact a Marcus canonical equation and can be equivalently rewritten as

$$\begin{aligned} X(t) &= X_0 + \textstyle\int_0^t b(X(s))ds + \int_0^t f(X(s)) \circ dL^c(s) + \int_0^t f(X(s-))dL^d(s) \\ &\quad + \textstyle\sum_{0<s\leq t}[\varphi(\Delta L(s)f; X(s-), 1) - X(s-) - f(X(s-))\Delta L(s)], \end{aligned} \tag{1.4}$$

where L^c and L^d are respectively the continuous and discontinuous parts of L, and $\circ$ denotes Stratonovich differential, and furthermore $\varphi(\sigma; u, t)$ is the flow generated by a vector field σ:

$$\frac{d\varphi(\sigma; u, t)}{dt} = f[\varphi(\sigma; u, t)], \quad \varphi(\sigma; u, 0) = u. \tag{1.5}$$

For more details about Marcus integral and canonical equations, we refer the reader to [1, 4, 9, 10, 12].

Note that the equation (1.2) is a perturbed version of (1.3) in the sense of Wong-Zakai [5, 12, 16, 22, 24, 25], the convergence result as above is somewhat expected. The real question, however, is in exactly what sense the result holds. For the special case $b = 0$, it has been shown by Kurtz, Pardoux and Protter [12] that for all but countably many t, $X^\epsilon(t)$ converges in probability to $X(t)$, as ϵ tends to 0. As detailed in the following theorem, we will show that for any $T > 0$, $\{X^\epsilon(t); 0 \leq t \leq T\}$ converges in probability to $\{X(t); 0 \leq t \leq T\}$ in $D([0, T], \mathbb{R})$ [1], the space of all the càdlàg functions over $[0, T]$, under the the Skorokhod M_1-topology, or simply put, X^ϵ converges in probability to X under the Skorokhod M_1-topology. Here we remark that the same convergence is not possible under the Skorokhod J_1-topology due to the simple fact that convergence under the J_1-topology retains continuity, whereas X^ϵ is continuous for any $\epsilon > 0$ and X may be discontinuous. For the precise definitions of the Skorokhod J_1 and M_1-topologies, see Appendix A.

Theorem 1.1. *Suppose that the functions b, f and f' are bounded and Lipschitz. Then, as ϵ tends to 0, X^ϵ converges in probability to X under the Skorokhod M_1-topology.*

For a real-valued stochastic process $W = \{W(t); 0 \leq t < \infty\}$ and a positive real number $a > 0$, we use $\tau_a(W)$ to denote the first passage time of W with respect to a, that is,

$$\tau_a(W) = \inf\{t \geq 0 : W(t) > a\}.$$

As an immediate corollary of Theorem 1.1, we have

[1]Following the usual practice in the theory of stochastic calculus, we will implicitly choose its unique càdlàg version for a given stochastic process.

Corollary 1.1. *For any positive real number $a > 0$, as ϵ tends to 0, $\tau_a(X^\epsilon)$ converges to $\tau_a(X)$ in distribution.*

Proof. The corollary follows from Theorem 1.1, the easily verifiable fact that for any $t, \epsilon > 0$,

$$\mathbb{P}(\tau_a(X^\epsilon) \le t) = \mathbb{P}(\sup_{s\in[0,t]} X^\epsilon(s) \ge a), \quad \mathbb{P}(\tau_a(X) \le t) = \mathbb{P}(\sup_{s\in[0,t]} X(s) \ge a),$$

and the fact that the supremum functional as above is continuous under the Skorokhod M_1-topology [20, Lemma 2.1]. □

The key tool that we used in this work is the so-called method of random time change (see, e.g., [11]), which is a well-known method that has also been used in [12]. On the other hand though, the power of this method somehow has not been fully utilized in [12]: Theorem 1.1 in this work, which is established through a short and simple argument, immediately implies that $X^\epsilon(t)$ converges in probability to $X(t)$ for all but countably many t, which further implies Theorem 6.5 in [12]. As a matter of fact, the power of this method can be further showcased in some special setting: For the case that L is a Lévy process, the method of Hintze and Pavlyukevich [6] can be adapted to show that as ϵ tends to 0, L^ϵ converges in probability to L under the Skorokhod M_1-topology, whereas our proof employing the method of random time change readily yields a stronger result stating that as ϵ tends to 0, L^ϵ converges *almost surely* to L under the Skorokhod M_1-topology (see Theorem 2.1 in Section 2). Here we remark that the proof of Theorem 3.1 in [6] is heavily dependent on the structure of the Lévy process and cannot carry over to the case when L is a general semimartingale, in which case our proof however aptly applies.

The remainder of this paper is organized as follows. In Section 2, we use a special case to illustrate the key methodology used in our proof. In Section 3, we prove Theorem 1.1, the main result of this paper. For self-containedness, we recall in Appendix A some basic notions and results on the Skorokhod J_1 and M_1-topologies.

2 A Special Case

Note that if we set $b \equiv 0$, $f \equiv 1$ and $X_0 = 0$, then X^ϵ is nothing but L^ϵ. This section, which is meant to be illustrative, is concerned with this special case, for which we will use the method of random time change to establish the following theorem:

Theorem 2.1. *As ϵ tends to 0, L^ϵ converges almost surely to to the semimartingale L under the Skorokhod M_1-topology.*

Proof. Let $[L] = [L, L]$ denote the quadratic variation of L, and let $[L]^c$ and $[L]^d$ denote its continuous and purely discontinuous parts, respectively. Define $\gamma^0(t) := [L]^d(t) + t$, and for any $\epsilon > 0$, define

$$\gamma^\epsilon(t) := \frac{1}{\epsilon}\int_{t-\epsilon}^{t} ([L]^d(s) + s)ds.$$

It can be shown that for any $t \ge 0$ and any $\epsilon > 0$, $\gamma^\epsilon(t) < \gamma^0(t) < \gamma^\epsilon(t+\epsilon)$. For any $\epsilon > 0$, let ς^ϵ be the generalized inverse of γ^ϵ, i.e., $\varsigma^\epsilon(t) := \inf\{s > 0 : \gamma^\epsilon(s) > t\}$. It can also be shown that for any $t \geq 0$ and any $\epsilon > 0$, $\varsigma^\epsilon(t) - \epsilon < \varsigma^0(t) < \varsigma^\epsilon(t)$, which implies that ς^ϵ converges to ς^0 uniformly over all t from any bounded time interval.

For any $\epsilon > 0$, define $Z^\epsilon(t) = L^\epsilon_{\varsigma^\epsilon(t)}$; in other words, the new process Z^ϵ is the original process L^ϵ reevaluated with respect to the new time scale $\varsigma^\epsilon(\cdot)$. It can be easily verified that $Z^\epsilon(t)$ is continuous in t.

The remainder of the proof consists of three steps as follows.

Step 1: In this step, we will show that as ϵ tends to 0, $\{Z^\epsilon(t)\}$ uniformly converges to a continuous process $\{Z(t)\}$.

Defining

$$\eta_-(t) = \sup\{s : \varsigma^0(s) < \varsigma^0(t)\}, \quad \eta^+(t) = \inf\{s : \varsigma^0(s) > \varsigma^0(t)\},$$

letting $\{\tau_i, i \in \mathbb{N}\}$ denote the sequence of all the jump times of L, we will deal with the following two cases.

Case 1: $t \in [0, \gamma^0(\tau_1-))$ or $t \in (\gamma^0(\tau_i), \gamma^0(\tau_{i+1}-))$ for some i.

In this case, we have $\eta_-(t) = \eta^+(t)$, and L is continuous at $\varsigma^0(t)$. Consequently,

$$\lim_{\epsilon \to 0+} Z^\epsilon(t) = \lim_{\epsilon \to 0+} L^\epsilon_{\varsigma^\epsilon(t)} = L_{\varsigma^0(t)}.$$

Case 2: $t \in [\gamma^0(\tau_i-), \gamma^0(\tau_i)]$.

In this case, $\varsigma^0(t) \equiv \tau_i$ and $\eta_-(t) \neq \eta^+(t)$, and L has a discontinuity at $\varsigma^0(t)$. It can be shown that

$$\lim_{\epsilon \to 0+} Z^\epsilon(\gamma^0(\tau_i-)) = \lim_{\epsilon \to 0+} L^\epsilon_{\varsigma^\epsilon \circ \gamma^0(\tau_i-)} = L_{\tau_i-} = L_{\varsigma^0(t)-}, \tag{2.1}$$

and

$$\lim_{\epsilon \to 0+} Z^\epsilon(\gamma^0(\tau_i)) = \lim_{\epsilon \to 0+} L^\epsilon_{\varsigma^\epsilon \circ \gamma^0(\tau_i)} = L_{\tau_i} = L_{\varsigma^0(t)}. \tag{2.2}$$

Moreover, it follows from the verifiable fact

$$\frac{dZ^\epsilon(t)}{dt} = \frac{L_{\varsigma^\epsilon(t)} - L_{\varsigma^\epsilon(t)-\epsilon}}{[L]^d_{\varsigma^\epsilon(t)} - [L]^d_{\varsigma^\epsilon(t)-\epsilon} + \epsilon}$$

that for any $t \in [\gamma^0(\tau_i-), \gamma^0(\tau_i)]$,

$$\lim_{\epsilon \to 0+} \frac{dZ^\epsilon(t)}{dt} = \frac{L_{\varsigma^0(t)} - L_{\varsigma^0(t)-}}{[L]^d_{\varsigma^0(t)} - [L]^d_{\varsigma^0(t)-}} = \frac{L_{\varsigma^0(t)} - L_{\varsigma^0(t)-}}{\eta^+(t) - \eta_-(t)}. \tag{2.3}$$

Consequently, it follows from (2.1),(2.2) and (2.3) that $\lim_{\epsilon \to 0+} Z^\epsilon(t) = Z(t)$ uniformly over all t from any bounded time interval, where Z is continuous and admits following expression:

$$Z(t) = \begin{cases} L_{\varsigma^0(t)}, & \text{if} \quad \eta_-(t) = \eta^+(t), \\ \frac{t-\eta_-(t)}{\eta^+(t)-\eta_-(t)} L_{\varsigma^0(t)} + \frac{\eta^+(t)-t}{\eta^+(t)-\eta_-(t)} L_{\varsigma^0(t)-}, & \text{if} \quad \eta_-(t) \neq \eta^+(t). \end{cases}$$

Step 2: This step will lead to the conclusion that as ϵ tends to 0, γ^ϵ converges almost surely to γ^0 under the Skorokhod M_1-topology. The proof of this step is postponed to next section (see Lemma 3.2).

Step 3: In this step, we will show that as ϵ tends to 0, L^ϵ converges almost surely to L under the Skorokhod M_1-topology, thereby completing the proof.

It follows from the facts that $\varsigma^0 \circ \gamma^0(t) = t$ and $\gamma^0(t) \notin [\gamma^0(\tau_i-), \gamma^0(\tau_i))$ for any t, τ_i that $Z_{\gamma^0(t)} \equiv L(t)$. Since Z^ϵ uniformly converges to Z, and γ^ϵ converges almost surely to γ^0 under the Skorokhod M_1-topology, we conclude that $L^\epsilon(\cdot) = Z^\epsilon_{\gamma^\epsilon(\cdot)}$ converges almost surely to $Z_{\gamma^0(\cdot)} = L(\cdot)$ under the Skorokhod M_1-topology.

□

Remark 2.1. *Compared to Theorem 1.1, Theorem 2.1 deals with a special setting and yields a stronger result. On the other hand, compared to Theorem* 3.1 *in [6], as mentioned in Section 1, Theorem 2.1 treats a more general setting yet still produces a stronger result.*

3 Proof of Theorem 1.1

The proof of Theorem 1.1 roughly follows the framework laid out in the proof of Theorem 2.1 and uses many notations defined therein.

For any $\epsilon > 0$, recall that Z^ϵ is defined as in the proof of Theorem 2.1, and define Y^ϵ as $Y^\epsilon(t) = X^\epsilon_{\varsigma^\epsilon(t)}$ for any t. It can be easily verified that $\{Z^\epsilon(t)\}$ and $\{Y^\epsilon(t)\}$ are continuous, and moreover $\{Y^\epsilon(t)\}$ is the unique solution to the following equation:

$$Y^\epsilon(t) = X_0 + \int_0^t b(Y^\epsilon(s))d\varsigma^\epsilon(s) + \int_0^t f(Y^\epsilon(s))dZ^\epsilon(s), \quad 0 \le t < \infty. \tag{3.1}$$

We will first prove the following lemma.

Lemma 3.1. *As ϵ tends to 0, Y^ϵ converges in probability to a process Y under the compact uniform topology. Moreover, Y is continuous and satisfies*

$$Y(t) = X_0 + \sum_i \left(\varphi\left(f\Delta L(\tau_i), Y_{\gamma^0(\tau_i-)}, \frac{t \wedge \gamma^0(\tau_i) - \gamma^0(\tau_i-)}{|\Delta L(\tau_i)|^2}\right) - Y_{\gamma^0(\tau_i-)} - f(Y_{\gamma^0(\tau_i-)})\Delta L(\tau_i)\right)$$
$$\times I_{[\gamma^0(\tau_i-),+\infty)} + \int_0^t b(Y(s))d\varsigma^0(s) + \int_0^t f(Y(s))dL_{\varsigma^0(s)} + \frac{1}{2}\int_0^t ff'(Y(s))d[L]^c_{\varsigma^0(s)}.$$

Proof. The proof largely follows from that of Theorem 6.5 in [12], so we only give a sketch emphasizing the key steps.

As shown in Section 2, ς^ϵ and Z^ϵ converge to ς^0 and Z, respectively, both uniformly over any bounded time interval, which immediately implies that U^ϵ converges almost surely to U under the Skorokhod J_1-topology, where the processes U^ϵ and U are defined as

$$U^\epsilon(t) := Z^\epsilon(t) - L_{\varsigma^\epsilon(t)}, \quad U(t) := Z(t) - L_{\varsigma^0(t)}.$$

And we note that (3.1) can be rewritten as

$$\begin{aligned}
Y^\epsilon(t) &= X_0 + \int_0^t b(Y^\epsilon(s))d\varsigma^\epsilon(s) + \int_0^t f(Y^\epsilon(s))dL_{\varsigma^\epsilon(s)} + \int_0^t f(Y^\epsilon(s))dU^\epsilon(s) \\
&= X_0 + \int_0^t b(Y^\epsilon(s))d\varsigma^\epsilon(s) + \int_0^t f(Y^\epsilon(s))dL_{\varsigma^\epsilon(s)} + f(Y^\epsilon(t))U^\epsilon(t) \\
&\quad - \int_0^t f'(Y^\epsilon(s))U^\epsilon(s)dY^\epsilon(s) - [f(Y^\epsilon), U^\epsilon](t) \\
&= X_0 + \int_0^t b(Y^\epsilon(s))d\varsigma^\epsilon(s) + \int_0^t f(Y^\epsilon(s))dL_{\varsigma^\epsilon(s)} + f(Y^\epsilon(t))U^\epsilon(t) \\
&\quad - \int_0^t f'(Y^\epsilon(s))f(Y^\epsilon(s))U^\epsilon(s)dZ^\epsilon(s) - \int_0^t f'(Y^\epsilon(s))b(Y^\epsilon(s))U^\epsilon(s)d\varsigma^\epsilon(s), \qquad (3.2)
\end{aligned}$$

where we have used the fact that $[f(Y^\epsilon), U^\epsilon] \equiv 0$.

By [12, Lemma 6.3], we infer that $\{\int_0^\cdot U^\epsilon(s)dZ^\epsilon(s)\}$ and $\{\varsigma^\epsilon(\cdot)\}$ are "good" (see Kurtz-Protter [13, 14]), and ς^ϵ converges to ς^0 uniformly over any bounded time interval, and moreover,

$$\int_0^t U^\epsilon(s)dZ^\epsilon(s) \to \frac{U(t)^2 - [L]_{\varsigma^0(t)}}{2} = \frac{(Z(t) - L_{\varsigma^0(t)})^2 - [L]_{\varsigma^0(t)}}{2} \qquad (3.3)$$

in probability under the Skorokhod J_1-topology. Then, parallel to the proof of Lemma 6.4 in [12], using the boundedness and Liptschitzness of b, f and f', we deduce that $f(Y^\epsilon(t))U^\epsilon(t)$ converges in distribution to $R(t)$ under the Skorokhod J_1-topology, where

$$R(t) = \sum_i I_{[\gamma^0(\tau_i-),\gamma^0(\tau_i))}(t) f\left(\varphi\left(f\Delta L(\tau_i), Y_{\gamma^0(\tau_i-)}, \frac{t-\gamma^0(\tau_i-)}{|\Delta L(\tau_i)|^2}\right)\right) U(t); \qquad (3.4)$$

here, $\{\tau_i, i \in \mathbb{N}\}$, as in the proof of Theorem 2.1, is the sequence of all the jump times of L. Moreover, by the definition of U^ϵ and ς^ϵ, we deduce that as ϵ tends to 0,

$$\int_0^t f'(Y^\epsilon(s))b(Y^\epsilon(s))U^\epsilon(s)d\varsigma^\epsilon(s) \to 0. \qquad (3.5)$$

Now, combining (3.2)-(3.5) as above, we deduce from [14] and [13, Theorem 5.4] that Y^ϵ converges in distribution to Y under the Skorokhod J_1-topology, where

$$Y(t) = X_0 + \int_0^t b(Y(s))d\varsigma^0(s) + \int_0^t f(Y(s))dL_{\varsigma^0(s)} + R(t) - \frac{1}{2}\int_0^t f'(Y(s))f(Y(s))d(U(s)^2 - [L]_{\varsigma^0(s)}).$$

Note that $U(t)$ can be further computed as

$$U(t) = Z(t) - L_{\varsigma^0(t)} = \begin{cases} 0, & \text{if } \eta_-(t) = \eta^+(t), \\ \frac{\eta^+(t)-t}{\eta^+(t)-\eta_-(t)}(L_{\varsigma^0(t)-} - L_{\varsigma^0(t)}), & \text{if } \eta_-(t) \neq \eta^+(t). \end{cases}$$

It then follows from the fact for any $t \in [\gamma^0(\tau_i-), \gamma^0(\tau_i))$,

$$\varsigma^0(t) \equiv \tau_i, \quad \eta^+(t) - \eta_-(t) = \Delta[L]^d_{\tau_i} = |\Delta L(\tau_i)|^2$$

that

$$
\begin{aligned}
Y(t) &= Y_{\gamma^0(\tau_i-)} + \int_{\gamma^0(\tau_i-)}^t b(Y(s))d\varsigma^0(s) + \int_{\gamma^0(\tau_i-)}^t f(Y(s))dL_{\varsigma^0(s)} \\
&\quad + f\left(\varphi\left(f\Delta L(\tau_i), Y_{\gamma^0(\tau_i-)}, \frac{t-\gamma^0(\tau_i-)}{|\Delta L(\tau_i)|^2}\right)\right)U(t) - \int_{\gamma^0(\tau_i-)}^t f'(Y(s))f(Y(s))\frac{s-\gamma^0(\tau_i)}{\eta^+(s)-\eta_-(s)}ds \\
&= Y_{\gamma^0(\tau_i-)} + f(Y_{\gamma^0(\tau_i-)})\Delta L(\tau_i) + f\left(\varphi\left(f\Delta L(\tau_i), Y_{\gamma^0(\tau_i-)}, \frac{t-\gamma^0(\tau_i-)}{|\Delta L(\tau_i)|^2}\right)\right)U(t) \\
&\quad - \int_{\gamma^0(\tau_i-)}^t f'(Y(s))f(Y(s))\frac{s-\gamma^0(\tau_i)}{\eta^+(s)-\eta_-(s)}ds.
\end{aligned}
$$

Consequently,

$$
\begin{aligned}
Y(t) &= X_0 + \sum_i \left(\varphi\left(f\Delta L(\tau_i), Y_{\gamma^0(\tau_i-)}, \frac{t\wedge\gamma^0(\tau_i)-\gamma^0(\tau_i-)}{|\Delta L(\tau_i)|^2}\right) - Y_{\gamma^0(\tau_i-)} - f(Y_{\gamma^0(\tau_i-)})\Delta L(\tau_i)\right) \\
&\times I_{[\gamma^0(\tau_i-),+\infty)}(t) + \int_0^t b(Y(s))d\varsigma^0(s) + \int_0^t f(Y(s))dL_{\varsigma^0(s)} + \frac{1}{2}\int_0^t ff'(Y(s))d[L]^c_{\varsigma^0(s)}.
\end{aligned} \tag{3.6}
$$

Since Y^ϵ, Y are continuous, we infer that Y^ϵ converges in distribution to Y under the compact uniform topology. Finally, using a similar argument in [12, Theorem 6.5], we conclude that Y^ϵ converges in probability to Y under the compact uniform topology, thereby completing the proof. □

Remark 3.1. *With the added assumption that b' is bounded and Lipschitz, the proof of Theorem 6.5 in [12] can be slightly modified to prove that for almost all t, $X^\epsilon(t)$ converges in probability to $X(t)$. By comparison, Lemma 3.1 reaches the same conclusion without the added assumption as above.*

The following lemma characterizes the convergence behavior of γ^ϵ.

Lemma 3.2. *As ϵ tends to 0, γ^ϵ converges almost surely to γ^0 under the Skorokhod M_1-topology.*

Proof. We first prove that γ^ϵ converges in probability to γ^0 under the Skorokhod M_1-topology. It suffices to verify that V^ϵ converges in probability to $[L]^d$ under the Skorokhod M_1-topology, where

$$
V^\epsilon(t) := \frac{1}{\epsilon}\int_{t-\epsilon}^t [L]^d(s)ds.
$$

To this end, by [19, Theorem 22, Page 66], the quadratic variation process $[L]^d$ of the semimartingale L^d is a càdlàg, increasing and adapted process, which implies that the mapping $t \mapsto \frac{1}{\epsilon}\int_{t-\epsilon}^t [L]^d(s)ds$ is monotone. So, by the definition of ρ (see (A.1)), we have $\rho(V^\epsilon, \delta) = 0$, which implies that for any fixed $\varepsilon > 0$, $\lim_{\delta\to 0+}\limsup_\epsilon \mathbb{P}(\rho(V^\epsilon, \delta) > \varepsilon) = 0$. Moreover, one verifies that for any t, $V^\epsilon(t)$ converges in probability to $[L]^d(t)$. With the preparations as above, we invoke Proposition A.1 to conclude that V^ϵ converges in probability $[L]^d$ under the Skorokhod M_1-topology.

Now we turn to prove that γ^ϵ converges almost surely to γ^0 under the Skorokhod M_1-topology. Using the fact that $[L]^d(t)$ is monotone in t and the definition of V^ϵ, we have that for any $0 = \epsilon_\infty < \epsilon_2 < \epsilon_1$,

$$[L]^d(t) = V^0(t) = V^{\epsilon_\infty}(t) \geq V^{\epsilon_2}(t) \geq V^{\epsilon_1}(t).$$

It then follows from the definition of $d_{M_1,T}$ that

$$d_{M_1,T}(V^{\epsilon_2}, [L]^d) \leq d_{M_1,T}(V^{\epsilon_1}, [L]^d),$$

that is to say, for any fixed time T, almost all ω in Ω, $d_{M_1,T}(V^\epsilon, [L]^d)$ is monotonically increasing in ϵ. Now, applying the proven fact V^ϵ converges in probability to the semimartingale $[L]^d$ under the Skorokhod M_1-topology and [2, Lemma 2.5.4], we conclude that V^ϵ converges almost surely to the semimartingale $[L]^d$ under the Skorokhod M_1-topology, which implies that γ^ϵ converges almost surely to γ^0 under the Skorokhod M_1-topology, as desired. □

Henceforth, letting $\hat{X}(t) = Y_{\gamma^0(t)}$, we prove the following two lemmas.

Lemma 3.3. *As ϵ tends to 0, X^ϵ converges in probability to $\hat{X}$ under the Skorokhod M_1-topology.*

Proof. The lemma immediately follows from Lemma 3.1, Lemma 3.2 and [23, Theorem 13.2.3]. □

Lemma 3.4. *$\hat{X}$ is the unique solution to the equation (1.3), and therefore $\hat{X} \equiv X$.*

Proof. Since $\hat{X}(t) = Y_{\gamma^0(t)}$, by the equation (3.6), we have

$$\begin{aligned}
\hat{X}(t) = X_0 + \sum_i \Bigg(\varphi\left(f\Delta L(\tau_i), Y_{\gamma^0(\tau_i-)}, \frac{\gamma^0(t) \wedge \gamma^0(\tau_i) - \gamma^0(\tau_i-)}{|\Delta L(\tau_i)|^2} \right) - Y_{\gamma^0(\tau_i-)} - f(Y_{\gamma^0(\tau_i-)})\Delta L(\tau_i) \Bigg) \\
\times I_{[\gamma^0(\tau_i-),+\infty)}(\gamma^0(t)) + \int_0^{\gamma^0(t)} b(Y(s))d\varsigma^0(s) + \int_0^{\gamma^0(t)} f(Y(s))dL_{\varsigma^0(s)} + \frac{1}{2}\int_0^{\gamma^0(t)} ff'(Y(s))d[L]^c_{\varsigma^0(s)} \\
= X_0 + \sum_i \left(\varphi(f\Delta L(\tau_i), Y_{\gamma^0(\tau_i-)}, 1) - Y_{\gamma^0(\tau_i-)} - f(Y_{\gamma^0(\tau_i-)})\Delta L(\tau_i) \right) I_{[\gamma^0(\tau_i-),+\infty)}(\gamma^0(t)) \\
+ \int_0^t b(\hat{X}(s))ds + \int_0^t f(\hat{X}(s-))dL(s) + \frac{1}{2}\int_0^t ff'(\hat{X}(s))d[L]^c(s) \\
= X_0 + \sum_{0<s\leq t} \left(\varphi(f\Delta L(s), \hat{X}(s-), 1) - \hat{X}(s-) - f(\hat{X}(s-))\Delta L(s) \right) + \int_0^t b(\hat{X}(s))ds \\
+ \int_0^t f(\hat{X}(s-))dL(s) + \frac{1}{2}\int_0^t ff'(\hat{X}(s))d[L]^c(s) \\
= X_0 + \int_0^t b(\hat{X}(s))ds + \int_0^t f(\hat{X}(s-)) \diamond dL(s),
\end{aligned}$$

where in the last equality, we have used the alternative definition of Marcus canonical equation in (1.4). So, $\hat{X}$ is the solution to the equation (1.3), which, together with the uniqueness of the solution to the equation (1.3), implies that $\hat{X} \equiv X$.

□

With all the lemmas as above, we are finally ready to prove Theorem 1.1.

Proof of Theorem 1.1. It follows from Lemma 3.1 that Y^ϵ converges in probability to Y under the compact uniform topology. Moreover, it follows from Lemma 3.3 that X^ϵ converges in probability to $\hat{X}$ under the Skorokhod M_1-topology. The theorem then follows from Lemma 3.4, which asserts $\hat{X} \equiv X$. □

Appendix

A Skorokhod topologies

Throughout this section, we fix $T > 0$.

The following J_1-metric has been defined by Skorokhod [21]:

$$d_{J_1,T}(x,y) = \inf_{\lambda \in \Lambda} \left\{ \sup_{0 \le t \le T} |x(t) - y((\lambda(t))| + \sup_{s,t \in [0,T], s \neq t} \left| \log \frac{\lambda(s) - \lambda(t)}{s - t} \right| \right\}, \quad x, y \in D([0,T], \mathbb{R}),$$

where Λ is the set of all the strictly increasing continuous functions mapping $[0,T]$ onto itself. The topology on $D([0,T], \mathbb{R})$ induced by the J_1-metric is called the Skorokhod J_1-topology.

Skorokhod [21] also defined the M_1-metric using the notion of completed graph of a function. More precisely, for any $x \in D([0,T], \mathbb{R})$, the *completed graph* of x, denoted by Γ_x, is defined as

$$\Gamma_x := \{(t, z) \in [0,T] \times \mathbb{R} : z \in [[x(t-), x(t)]]\},$$

where $x(0-)$ is interpreted as $x(0)$, $[[z_1, z_2]]$ is the line segment connecting z_1 and z_2, i.e.,

$$[[z_1, z_2]] = \{z \in \mathbb{R} : z = az_1 + (1-a)z_2 \text{ for some } a \in [0,1]\}.$$

Note that Γ_x can be parametrically represented by the following continuous function

$$(r, u) : [0,1] \to \Gamma_z, \quad (r,u)(0) = (0, z(0)), (r,u)(1) = (T, z(T)),$$

which is nondecreasing with respect to the following order on Γ_x:

$$(t_1, z_1) \le (t_2, z_2) \Leftrightarrow t_1 < t_2 \quad \text{or} \quad (t_1 = t_2 \text{ and } |x(t_1-) - z_1| \le |x(t_2-) - z_2|).$$

Skorokhod [21] defined the M_1-metric as follows:

$$d_{M_1,T}(x,y) = \inf_{(r_1,u_1) \in \Pi(x), (r_2,u_2) \in \prod(y)} \{ \sup_{0 \le t \le 1} |r_1(t) - r_2(t)| + \sup_{0 \le t \le 1} |u_1(t) - u_2(t)| \}, \quad x, y \in D([0,T], \mathbb{R}),$$

where $\Pi(\cdot)$ denotes the set of all parametric representations of an element in $D([0,T], \mathbb{R})$. The topology on $D([0,T], \mathbb{R})$ induced by the M_1-metric is called the Skorokhod M_1-topology.

Noting that the limit of a sequence of continuous functions under either the uniform or the Skorokhod J_1-topology is continuous, we remark that, when approximating a càdlàg function using continuous functions, the Skorokhod M_1-topology can be particularly useful. For example, for any $n \ge 1$, let

$$x(t) = I_{[1/2,1]}(t), \quad x^n(t) = n(t - 1/2 + 1/n) I_{[1/2 - 1/n, 1/2]}(t) + I_{[1/2,1]}(t), \quad 0 \le t \le 1.$$

One can verify that, as n tends to infinity, $x^n(t) \to x(t)$ in $D([0,1], \mathbb{R})$ under Skorokhod M_1-topology but not under the Skorokhod J_1-topology.

The following theorem is well known; see, e.g., [17, Theorem 3.2].

Theorem A.1. *Let $W = \{W(t); 0 \leq t \leq T\}$ and $W^n = \{W^n(t); 0 \leq t \leq T\}, n = 1, 2, \dots,$ be stochastically continuous càdlàg stochastic processes. Then, as n tends to infinity, W^n converges in probability to W under the Skorokhod M_1-topology if and only if*

1) *for any $t \in [0, T]$, $W^n(t)$ converges in probability to $W(t)$;*

2) *and for any fixed $\varepsilon > 0$,*

$$\lim_{\delta \to 0+} \limsup_{n} \mathbb{P}(\rho(W^n, \delta) > \varepsilon) = 0,$$

where $\rho : D([0, T], \mathbb{R}) \times \mathbb{R} \to \mathbb{R}$ is defined as

$$\rho(x, \delta) := \sup_{0 \leq t_1 < t < t_2 \leq 1; t_2 - t_1 < \delta} \inf_{a \in [0,1]} |x(t) - (ax(t_1) + (1-a)x(t_2))|. \tag{A.1}$$

References

[1] D. Applebaum, *Lévy Processes and Stochastic Calculus,* Cambridge University Press, second edition, Cambridge, UK, 2009.

[2] R. B. Ash, C. A. Doléans-Dade, *Probability and measure theory,* Academic Press, second edition, San Diego, 2000.

[3] P. Billingsley, *Convergence of probability measures,* John Wiley Sons, 2013.

[4] T. Fujiwara and H. Kunita, Canonical SDE's based on semimartingales with spatial parameters, Part 1 Stochastic flows of diffeomorphisms, *Kyushu J. Math.,* vol. 53, pp. 265-300, 1999.

[5] M. Hairer, E. Pardoux, A Wong-Zakai theorem for stochastic PDEs, *J. Math. Soc. Japan,* vol. 67, no. 4, pp. 1551-1604, 2015.

[6] R. Hintze and I. Pavlyukevich, Small noise asymptotics and first passage times of integrated Ornstein-Uhlenbeck processes driven by α-stable Lévy processes, *Bernoulli,* vol. 20, no. 1, pp. 265-281, 2014.

[7] N. Ikeda and S. Watanabe, *Stochastic differential equations and diffusion processes*, second edition, North-Holland, 1989.

[8] J. Jacod, A. N. Shiryaev, *Limit theorems for stochastic processes,* Springer, Berlin, Heidelberg, 2003.

[9] H. Kunita, Stochastic differenntial equations with jumps and stochastic flows of diffeomorphisms, In *Itô's stochastic calculus and probability theory* (Eds. N. Ikeda and K. Itô), Springer, Tokyo, pp. 197-211, 1996.

[10] H. Kunita, Stochastic differential equations based on Lévy processes and stochastic flows of diffeomorphisms, *Real and stochastic analysis* (Eds. M. M. Rao), Birkhüser, Boston, MA, pp. 305-373, 2004.

[11] T. G. Kurtz, Random time changes and convergence in distribution under the Meyer-Zheng conditions. *The Annals of Probability,* vol. 19, pp. 1010-1034, 1991.

[12] T. G. Kurtz, E. Pardoux and P. Protter, Stratonovich stochastic differential equations driven by general semimartingales, *Ann. Inst. Henri Poincaré Probab. Stat.,* vol. 23, pp. 351-377, 1995.

[13] T. G. Kurtz, P. Protter, Weak limit theorems for stochastic integrals and stochastic differential equations, *The Annals of Probability,* vol. 19, pp. 1035-1070, 1991.

[14] T. G. Kurtz, P. Protter, Weak convergence of stochastic integrals and differential equations, *Probabilistic models for nonlinear partial differential equations,* Springer, Berlin, Heidelberg, pp. 1-41, 1996.

[15] S. I. Marcus, Modeling and analysis of stochastic differential equations driven by point processes, *IEEE Transactions on Information Theory,* vol. 24, no.2, pp. 164-172, 1978.

[16] S. I. Marcus, Modelling and approximation of Stochastic differenntial equations driven by semimaringales, *Stochastics*, vol. 4, pp. 223-245, 1981.

[17] I. Pavlyukevich, M. Riedle, Non-standard Skorokhod convergence of Lévy-driven convolution integrals in Hilbert spaces, *Stochastic Analysis and Applications,* vol. 33, no. 2, pp. 271-305, 2015.

[18] D. Pollard, *Convergence of stochastic processes,* Springer Science Business Media, 2012.

[19] P. Protter, *Stochastic integration and differential equations,* Springer, Berlin, Heidelberg, 2005.

[20] A. A. Puhalskii, W.Whitt, Functional large deviation principles for first-passage-time processes. *The Annals of Applied Probability,* vol. 7, no. 2, pp. 362-381, 1997.

[21] A. V. Skorokhod, Limit theorems for stochastic processes, *Theory of Probability and Its Applications,* vol. 1, no. 3, pp. 261-290, 1956.

[22] G. Tessitore and J. Zabczyk, Wong-Zakai approximation of stochastic evolution equations, *J. Evol. Equ.,* vol. 6, no. 4, pp. 621-655, 2006.

[23] W. Whitt, *Stochastic-process limits: an introduction to stochastic-process limits and their application to queues,* Springer Science Business Media, 2002.

[24] E. Wong, M. Zakai, On the relation between ordinary and stochastic differential equations, *Internat. J. Engrg. Sci.,* 1965, 3: 213-229.

[25] E. Wong, M. Zakai, On the convergence of ordinary integrals to stochastic integrals, *Ann. Math. Statist.,* vol. 36, pp. 1560-1564, 1965.

[26] X. Zhang, Derivative formulas and gradient estimates for SDEs driven by α-stable processes, *Stochastic Processes and their Applications,* vol. 123, no. 4, pp. 1213-1228, 2013.