Journal of Computational Mathematics
Vol.42, No.1, 2024, 248–270.
http://www.global-sci.org/jcm
doi:10.4208/jcm.2206-m2021-0354

# STRONG CONVERGENCE OF JUMP-ADAPTED IMPLICIT MILSTEIN METHOD FOR A CLASS OF NONLINEAR JUMP-DIFFUSION PROBLEMS*

Xu Yang
*School of Mathematics, China University of Mining and Technology,*
*Xuzhou 221116, China*
*Email: xuyang96@cumt.edu.cn*
Weidong Zhao[1)]
*School of Mathematics, Shandong University, Jinan 250100, China*
*Email: wdzhao@sdu.edu.cn*

**Abstract**

In this paper, we study the strong convergence of a jump-adapted implicit Milstein method for a class of jump-diffusion stochastic differential equations with non-globally Lipschitz drift coefficients. Compared with the regular methods, the jump-adapted methods can significantly reduce the complexity of higher order methods, which makes them easily implementable for scenario simulation. However, due to the fact that jump-adapted time discretization is path dependent and the stepsize is not uniform, this makes the numerical analysis of jump-adapted methods much more involved, especially in the non-globally Lipschitz setting. We provide a rigorous strong convergence analysis of the considered jump-adapted implicit Milstein method by developing some novel analysis techniques and optimal rate with order one is also successfully recovered. Numerical experiments are carried out to verify the theoretical findings.

*Mathematics subject classification:* 60H35, 65C30, 65L20.
*Key words:* Jump-diffusion, Jump-adapted implicit Milstein method, Poisson jumps, Strong convergence rate, Non-Lipschitz coefficients.

## 1. Introduction

In this paper, we consider numerical solution of jump-diffusion Itô stochastic differential equations of the form

$$\begin{cases} \mathrm{d}X_t = f(X_{t-})\mathrm{d}t + g(X_{t-})\mathrm{d}W_t + h(X_{t-})\mathrm{d}N_t,\, t \in \big(0,T\big], \\ X_0 = x_0, \end{cases} \tag{1.1}$$

where $T > 0$ is a fixed constant, $X_{t-}$ denotes $\lim_{s\uparrow t} X_s$, $f\colon \mathbb{R}^m \to \mathbb{R}^m$ is the drift coefficient, $g\colon \mathbb{R}^m \to \mathbb{R}^{m\times d}$ is the diffusion coefficient which is frequently written as $g = (g_{i,j})_{m\times d} = (g_1, g_2, \cdots, g_d)$ for $g_{i,j}\colon \mathbb{R}^m \to \mathbb{R}$ and $g_j\colon \mathbb{R}^m \to \mathbb{R}^m$, $j \in \{1,\dots,d\}$, $h\colon \mathbb{R}^m \to \mathbb{R}^m$ is the jump coefficient, with $m, d \in \mathbb{N}^+$. Here $W_t$ is a $d$-dimensional Wiener process and $N_t$ is a scalar Poisson process with intensity $\lambda > 0$, both defined on a complete probability space $(\Omega,\mathcal{F},\mathbb{P})$, with a normal filtration $\mathbb{F}\colon = \{\mathcal{F}_t\}_{t\in[0,T]}$, and they are independent with each other. Specific conditions on the coefficients $f, g, h$ and the initial value $x_0$ will be given in next section.

* Received December 15, 2021 / Revised version received January 12, 2022 / Accepted June 14, 2022 / Published online February 22, 2023 /
1) Corresponding author

Jump-diffusion stochastic differential equations (JSDEs) have found many applications in biology, physics, chemistry, engineering, finance, and insurance. Especially in the area of finance and insurance, jump-diffusion models are often used to describe the dynamic evolution of various state variables, such as stock indices, asset prices, interest rates, credit ratings, exchange rates or commodity prices. In these models, the jump driving part is often used to characterize and capture the event-driven uncertainties and the unexpected abrupt changes, such as corporate defaults, operational failures or insured accidents. For more details on the application of jump-diffusion stochastic differential equations in the financial field, we refer the reader to the books [9, 32].

Since most of JSDEs cannot be solved explicitly, it is necessary and important to develop discrete time approximations to study the behavior of jump-diffusion models and solve practical application problems. In view of the above reason, the study of numerical approximation of JSDEs has been attracting lots of attention. Generally, the discrete time approximations of JSDEs are divided into regular and jump-adapted schemes. Regular schemes employ time discretizations that do not include the jump times of the Poisson process, while jump-adapted schemes are based on jump-adapted time discretizations which include these jump times.

As far as the regular schemes are concerned, up to now, a lot of research results have been achieved, see [1, 3–5, 11, 14, 17–19, 32–34, 37] and the references therein. However, in the vast majority of the above mentioned research articles, a global Lipschitz assumption on the coefficients of JSDEs is often used. Note that the assumption of global Lipschitz continuity is very restrictive and many practical problems fail to satisfy such condition. In fact, the coefficients of numerous JSDEs used to describe financial models are either super-linear growth or sub-linear growth, which are obviously not globally Lipschitz continuous. Therefore, the theoretical analysis basis of many existing numerical methods has changed. This makes the scope of application of those numerical methods whose theoretical framework are based on global Lipschitz condition, greatly reduced, and those numerical methods even no longer be applicable. Therefore, in recent years, many scholars have begun to turn to dealing with numerical approximation of JSDEs under some weaker conditions, such as local Lipschitz condition, one-sided Lipschitz condition, and great progress has been made, see [6–8, 10, 12, 13, 15, 20–23, 25, 27, 35, 36, 38, 39, 41].

To avoid the generation of multiple stochastic integrals with respect to the Poisson process, jump-adapted approximations were first introduced in [31]. As the jump-adapted time discretizations include all jump times generated by the Poisson process, the form of the resulting schemes is much simpler than that of the regular schemes, significantly reducing the complexity of higher order schemes. Hence the jump-adapted time discretization makes these corresponding schemes easily implementable for scenario simulation. Up to now, various jump-adapted numerical methods have been formulated and analyzed in [2, 24, 26, 28, 29, 32] and references therein. Nevertheless, like the regular case, most of convergence results for jump-adapted approximations are always analyzed in the globally Lipschitz setting, and there exist only a very limited number of works devoted to the numerical study of jump-adapted schemes under the non-global Lipschitz condition. A transformed jump-adapted backward Euler method for a class of jump-diffusion financial models whose coefficients do not satisfy the global Lipschitz condition was studied in [40].

The main objective of this paper is to study the strong convergence of a jump-adapted implicit Milstein method for a class of general jump-diffusion models, for which we require that the drift coefficient is one-sided Lipschitz continuous, and the diffusion coefficient and the jump coefficient are globally Lipschitz continuous (see Assumption 2.1 in the next section). Due to the presence of non-globally Lipschitz drift coefficients, the numerical analysis developed in the existing works does not work in the present framework. In the error analysis, we do not measure the approximation error $E_k := Y_{t_k} - X_{t_k}$ directly. Based on certain error terms only getting involved with the exact solution processes, we first establish an upper mean-square error bounds for the discrepancy between the intermediate solutions, namely $E_{k-} := Y_{t_{k-}} - X_{t_{k-}}$. Then by taking properties of the jump-adapted grids into account and using some techniques from stochastic analysis, we finally prove the strong convergence for the considered jump-adapted implicit Milstein method and successfully recover the expected optimal mean-square convergence rate of order one. To our best knowledge, this is the first paper to consider the jump-adapted implicit Milstein method in the non-globally Lipschitz drift coefficients setting.

This paper is organized as follows. In Section 2, some notation and preliminaries are collected. The well-posedness of the considered problem (1.1) and properties of the exact solution are discussed. In Section 3 the jump-adapted implicit Milstein method is introduced and its strong convergence rate is analyzed. Numerical experiments confirming the previous findings are presented in Section 4. Finally, we give some brief remarks in Section 5.

## 2. Preliminaries

We first introduce some notation used throughout the paper. Letting $|\cdot|$ denotes both the Euclidean vector norm and the Hilbert-Schmidt matrix norm. Denote by $\langle \cdot, \cdot \rangle$ the inner product in $\mathbb{R}^m$. Let $x \vee y := \max\{x, y\}$ for any $x, y \in \mathbb{R}$. Let $\mathcal{F}_t^W := \sigma(W_s, 0 \leq s \leq t)$ denote the natural filtration generated by the Wiener process $W_t$ and $\mathcal{F}_t^N := \sigma(N_s, 0 \leq s \leq t)$ represent the natural filtration generated by the Poisson process $N_t$. Define $\mathcal{F}_t := \sigma\big(\mathcal{F}_s^W \bigcup \mathcal{F}_s^N, 0 \leq s \leq t\big)$, augmented by all $\mathbb{P}$-null sets of $\mathcal{F}$. Let $\mathbb{F}: = \{\mathcal{F}_t\}_{t\in[0,T]}$. Define the conditional expectations $\mathbb{E}^N\big[X\big] := \mathbb{E}\big[X|\mathcal{F}_T^N\big]$ and $\mathbb{E}^N\big[X|\mathcal{F}_t^W\big] := \mathbb{E}\big[X|\sigma\big(\mathcal{F}_t^W \cup \mathcal{F}_T^N\big)\big]$. From now on, we will work on the filtered probability space $\big(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}\big)$.

Next, we give some assumptions to guarantee the well-posedness of the considered problem (1.1).

**Assumption 2.1.** *The coefficients $f, g$, and $h$ are continuously differentiable. Moreover, the drift coefficient $f$ satisfies the one-sided Lipschitz condition, i.e., there exists a positive constant $K_1$ such that*

$$\langle x - y, f(x) - f(y)\rangle \leq K_1|x - y|^2, \quad \forall x, y \in \mathbb{R}^m, \tag{2.1}$$

*and the diffusion coefficient $g$ and jump coefficient $h$ satisfy the global Lipschitz conditions*

$$|g(x) - g(y)|^2 \leq K_2|x - y|^2, \quad \forall\ x, y \in \mathbb{R}^m, \tag{2.2}$$

$$|h(x) - h(y)|^2 \leq K_3|x - y|^2, \quad \forall\ x, y \in \mathbb{R}^m, \tag{2.3}$$

*where $K_2$ and $K_3$ are positive constants.*

**Remark 2.1.** By (2.1)-(2.3), we can easily get

$$\langle x, f(x)\rangle \vee |g(x)|^2 \vee |h(x)|^2 \leq L(1 + |x|^2), \quad \forall x \in \mathbb{R}^m, \tag{2.4}$$

where

$$L = \max\left\{\left(K_1 + \frac{1}{2}\right), \frac{1}{2}|f(0)|^2, 2K_2, 2|g(0)|^2, 2K_3, 2|h(0)|^2\right\}.$$

**Assumption 2.2.** *The initial value $X_0 = x_0$ satisfies that for any $p > 2$, there exists a positive constant $K_p$ such that*

$$\mathbb{E}\big[|x_0|^p\big] \le K_p. \tag{2.5}$$

**Lemma 2.1 ([13]).** *Under Assumptions 2.1-2.2, the problem (1.1) admits a unique solution $\{X_t, t \in [0,T]\}$ satisfying that for each $p > 2$, there is a constant $K = K(p,T)$ such that*

$$\mathbb{E}\Big[\sup_{t\in[0,T]} |X_t|^p\Big] \le K\Big(1 + \mathbb{E}\big[|x_0|^p\big]\Big). \tag{2.6}$$

**Corollary 2.1.** *By Lemma 2.1, we have, for each $p > 2$, there is a constant $\widetilde{K} = \tilde{K}(p,T)$ such that*

$$\sup_{t\in[0,T]} \mathbb{E}\big[|X_t|^p\big] \le \widetilde{K}\Big(1 + \mathbb{E}\big[|x_0|^p\big]\Big). \tag{2.7}$$

In the following, based on Assumption 2.1, we further impose some extra conditions on the drift and diffusion coefficients, which will be used in our convergence analysis.

**Assumption 2.3.** *Suppose there is a constant $K_4$ and a positive constant $\kappa \ge 1$ such that,*

$$\left|\left(\frac{\partial f}{\partial x}(x) - \frac{\partial f}{\partial x}(y)\right) z\right| \le K_4\big(1 + |x| + |y|\big)^{\kappa-1}|x-y||z|, \quad \forall x,y,z \in \mathbb{R}^m. \tag{2.8}$$

**Remark 2.2.** Assumption 2.3 implies that there is a constant $K_4 > 0$ such that

$$\left|\frac{\partial f}{\partial x}(x) z\right| \le K_4\big(1 + |x|\big)^{\kappa}|z|, \quad \forall x,z \in \mathbb{R}^m. \tag{2.9}$$

As a consequence, one also gets that there exists a constant $K_5 > 0$ such that

$$|f(x) - f(y)| \le K_5\big(1 + |x| + |y|\big)^{\kappa}|x-y|, \quad \forall x,y \in \mathbb{R}^m, \tag{2.10}$$

which implies

$$|f(x)| \le K_6\big(1 + |x|^{1+\kappa}\big), \quad \forall x \in \mathbb{R}^m, \tag{2.11}$$

where $K_6$ is a positive constant.

**Assumption 2.4.** *Suppose the diffusion coefficient $g$ further fulfils that there exist constants $K_7$ and $K_8$ such that,*

$$\sum_{j_1,j_2=1}^{d} |\mathcal{L}^{j_1} g_{j_2}(x) - \mathcal{L}^{j_1} g_{j_2}(y)|^2 \le K_7|x-y|^2, \quad \forall x,y \in \mathbb{R}^m, \tag{2.12}$$

$$\left|\left(\frac{\partial g_j}{\partial x}(x) - \frac{\partial g_j}{\partial x}(y)\right) z\right|^2 \le K_8|x-y|^2|z|^2, \qquad \forall x,y,z \in \mathbb{R}^m, \quad j \in \{1,\dots,d\}, \tag{2.13}$$

*where*

$$\mathcal{L}^{j_1} := \sum_{k=1}^{m} g_{k,j_1}\frac{\partial}{\partial x^k}, \quad j_1 \in \{1,\dots,d\}.$$

**Remark 2.3.** It follows from (2.13) that

$$\left|\frac{\partial g_j}{\partial x}(x) z\right|^2 \le K_9(1 + |x|^2)|z|^2, \quad \forall x,z \in \mathbb{R}^m, \quad j \in \{1,\dots,d\}, \tag{2.14}$$

where $K_9$ is a positive constant.

# 3. The Jump-Adapted Method and Its Strong Convergence Analysis

## 3.1. The jump-adapted method

Firstly, we construct a jump-adapted time partition

$$\mathcal{T} = \{0 = t_0 < t_1 < \cdots < t_{n_T} = T\},$$

produced by a superposition of the jump times $\{\tau_1, \tau_2, \ldots, \tau_{N_T}\}$ to a deterministic equidistant grid with time stepsize $\Delta t = \frac{T}{M}$. Here $n_T = \max\{n \in \{0, 1, \ldots\} \colon t_n \leq T\} < \infty$. In this way the jump-adapted time discretization including jump times is path-dependent and the maximum stepsize of the jump-adapted partition is $\Delta t$. On the grid $\mathcal{T}$, $\{X_t\}_{t\in[0,T]}$ can be expressed as

$$\begin{cases} X_{t_{k+1-}} = X_{t_k} + \displaystyle\int_{t_k}^{t_{k+1}} f(X_t)\mathrm{d}t + \int_{t_k}^{t_{k+1}} g(X_t)\mathrm{d}W_t, \\ X_{t_{k+1}} = X_{t_{k+1-}} + h(X_{t_{k+1-}})\Delta N_k \end{cases} \tag{3.1}$$

for $k = 0, 1, \ldots, n_T - 1$. Accordingly we propose the jump-adapted implicit Milstein method (JAIMM) for (1.1), defined through $Y_0 = X_0$ and for $k = 0, 1, \ldots, n_T - 1$, as

$$\begin{cases} Y_{t_{k+1-}} = Y_{t_k} + f(Y_{t_{k+1-}})\Delta t_k + g(Y_{t_k})\Delta W_k + \displaystyle\sum_{j_1,j_2=1}^{d} \mathcal{L}^{j_1} g_{j_2}(Y_{t_k}) I_{j_1,j_2}^{t_k,t_{k+1}}, \\ Y_{t_{k+1}} = Y_{t_{k+1-}} + h(Y_{t_{k+1-}})\Delta N_k, \end{cases} \tag{3.2}$$

where $\Delta t_k := t_{k+1} - t_k$, $\Delta W_k := W_{t_{k+1}} - W_{t_k}$, $\Delta N_k := N_{t_{k+1}} - N_{t_k} \in \{0, 1\}$, and

$$\mathcal{L}^{j_1} := \sum_{k=1}^{m} g_{k,j_1} \frac{\partial}{\partial x^k}, \quad I_{j_1,j_2}^{t_k,t_{k+1}} := \int_{t_k}^{t_{k+1}} \int_{t_k}^{s_2} \mathrm{d}W_{s_1}^{j_1} \mathrm{d}W_{s_2}^{j_2}, \quad j_1, j_2 \in \{1, \ldots, d\}. \tag{3.3}$$

A further closer look at (3.2) suggests

$$\begin{cases} Y_{t_{k+1}} = Y_{t_{k+1-}} + h(Y_{t_{k+1-}}), & \text{if } t_{k+1} \text{ is a jump time}, \\ Y_{t_{k+1}} = Y_{t_{k+1-}}, & \text{otherwise}. \end{cases} \tag{3.4}$$

## 3.2. Strong convergence analysis

Throughout the following analysis, by $C$ we denote a generic positive constant that may change between occurrences but is independent of $\Delta t$. Note that the method (3.2) is an implicit method and the question of existence and uniqueness arises. The following lemma gives a positive answer.

**Lemma 3.1.** *Under Assumption* 2.1, *if* $K_1\Delta t \leq c_0 < 1$ *with* $K_1$ *from* (2.1) *and* $c_0 \in (0, 1)$, *then the method* (3.2) *is well-defined.*

*Proof.* The proof is similar to that of Lemma 3.4 in [16] and Lemma 2.3 in [30], so we omit it here. $\square$

Note that within the interval $[t_k, t_{k+1})$, the problem (1.1) evolves as an SDE without jump. Then we have the following result.

**Lemma 3.2.** *For fixed $k$, let $s,t \in [t_k, t_{k+1})$. Then for any $q \geq 2$, it holds*

$$\mathbb{E}[|X_t - X_s|^q] \leq C(t-s)^{\frac{q}{2}},$$

*where $C$ is a positive constant.*

*Proof.* Suppose $t > s$. By virtue of Hölder's inequality, we deduce

$$\begin{aligned}
\mathbb{E}\big[|X_t - X_s|^q\big] &= \mathbb{E}\Big[\Big|\int_s^t f(X_r)\mathrm{d}r + \int_s^t g(X_r)\mathrm{d}W_r\Big|^q\Big] \\
&\leq 2^{q-1}\mathbb{E}\Big[\Big|\int_s^t f(X_r)\mathrm{d}r\Big|^q\Big] + 2^{q-1}\mathbb{E}\Big[\Big|\int_s^t g(X_r)\mathrm{d}W_r\Big|^q\Big] \\
&\leq C(t-s)^{q-1}\mathbb{E}\Big[\int_s^t \big|f(X_r)\big|^q \mathrm{d}r\Big] + C(t-s)^{\frac{q-2}{2}}\mathbb{E}\Big[\int_s^t \big|g(X_r)\big|^q \mathrm{d}r\Big] \\
&\leq C(t-s)^{q-1}\int_s^t \mathbb{E}\big[\big|f(X_r)\big|^q\big]\mathrm{d}r + C(t-s)^{\frac{q-2}{2}}\int_s^t \mathbb{E}\Big[\big|g(X_r)\big|^q\Big]\mathrm{d}r. \qquad (3.5)
\end{aligned}$$

Using (2.11) and (2.6), we get

$$\begin{aligned}
\mathbb{E}\big[\big|f(X_r)\big|^q\big] &= \mathbb{E}\Big[\big(\big|f(X_r)\big|^2\big)^{\frac{q}{2}}\Big] \leq C\mathbb{E}\Big[\big(1 + |X_r|^{2+2\kappa}\big)^{\frac{q}{2}}\Big] \\
&\leq C\mathbb{E}\big[1 + |X_r|^{(1+\kappa)q}\big] \leq C. \qquad (3.6)
\end{aligned}$$

Similarly, we can obtain

$$\mathbb{E}\big[\big|g(X_r)\big|^q\big] \leq C. \qquad (3.7)$$

Substituting (3.6)–(3.7) into (3.5) gives

$$\mathbb{E}[|X_t - X_s|^q] \leq C(t-s)^q + C(t-s)^{\frac{q}{2}} \leq C(t-s)^{\frac{q}{2}}, \qquad (3.8)$$

which completes the proof. □

In the same way as above, with the help of property of conditional mathematical expectation, we can also obtain the following result.

**Lemma 3.3.** *For $\forall k \in \{1,2,\ldots,n_T - 1\}$, let $s,t \in [t_k, t_{k+1})$. Then, for any $q \geq 2$, there exists a positive constant $C$ such that*

$$\mathbb{E}^N[|X_t - X_s|^q] \leq C\Phi_{N,q}(t-s)^{\frac{q}{2}},$$

*where*

$$\Phi_{N,q} := 1 + \mathbb{E}^N\Big[\sup_{0\leq t\leq T}|X_t|^{(1+\kappa)q}\Big] + \mathbb{E}^N\Big[\sup_{0\leq t\leq T}|X_t|^q\Big].$$

Based on (3.2), we get, for $k = 1,2,\ldots,n_T - 1$,

$$\begin{aligned}
Y_{t_{k+1-}} &= Y_{t_{k-}} + h(Y_{t_{k-}})\Delta N_{k-1} + f(Y_{t_{k+1-}})\Delta t_k \\
&\quad + g(Y_{t_k})\Delta W_k + \sum_{j_1,j_2=1}^{d}\mathcal{L}^{j_1}g_{j_2}(Y_{t_k})I_{j_1,j_2}^{t_k,t_{k+1}}. \qquad (3.9)
\end{aligned}$$

Likewise, (3.1) implies that, for $k = 1, 2, \ldots, n_T - 1$,

$$\begin{aligned}X_{t_{k+1-}} &= X_{t_{k-}} + h(X_{t_{k-}})\Delta N_{k-1} + \int_{t_k}^{t_{k+1}} f(X_t)\mathrm{d}t + \int_{t_k}^{t_{k+1}} g(X_t)\mathrm{d}W_t \\ &= X_{t_{k-}} + h(X_{t_{k-}})\Delta N_{k-1} + f(X_{t_{k+1-}})\Delta t_k + g(X_{t_k})\Delta W_k \\ &\quad + \sum_{j_1,j_2=1}^{d} \mathcal{L}^{j_1} g_{j_2}(X_{t_k}) I_{j_1,j_2}^{t_k,t_{k+1}} + R_{k+1},\end{aligned} \tag{3.10}$$

where the remainder term reads

$$\begin{aligned}R_{k+1} &= \int_{t_k}^{t_{k+1}} [f(X_t) - f(X_{t_{k+1-}})]\mathrm{d}t + \int_{t_k}^{t_{k+1}} [g(X_t) - g(X_{t_k})]\mathrm{d}W_t \\ &\quad - \sum_{j_1,j_2=1}^{d} \mathcal{L}^{j_1} g_{j_2}(X_{t_k}) I_{j_1,j_2}^{t_k,t_{k+1}}.\end{aligned} \tag{3.11}$$

Subtracting (3.10) from (3.9) leads to

$$\begin{aligned}Y_{t_{k+1-}} - X_{t_{k+1-}} &= Y_{t_{k-}} - X_{t_{k-}} + [h(Y_{t_{k-}}) - h(X_{t_{k-}})]\Delta N_{k-1} \\ &\quad + [f(Y_{t_{k+1-}}) - f(X_{t_{k+1-}})]\Delta t_k + [g(Y_{t_k}) - g(X_{t_k})]\Delta W_k \\ &\quad + \sum_{j_1,j_2=1}^{d} \big(\mathcal{L}^{j_1} g_{j_2}(Y_{t_k}) - \mathcal{L}^{j_1} g_{j_2}(X_{t_k})\big) I_{j_1,j_2}^{t_k,t_{k+1}} - R_{k+1}.\end{aligned} \tag{3.12}$$

Before proceeding, define

$$\begin{aligned}&E_k := Y_{t_k} - X_{t_k}, \quad E_{k-} := Y_{t_{k-}} - X_{t_{k-}}, \\ &\Delta f_k^{Y,X} := f(Y_{t_{k-}}) - f(X_{t_{k-}}), \\ &\Delta g_k^{Y,X} := g(Y_{t_k}) - g(X_{t_k}), \\ &\Delta(\mathcal{L}^{j_1} g_{j_2})_k^{Y,X} := \mathcal{L}^{j_1} g_{j_2}(Y_{t_k}) - \mathcal{L}^{j_1} g_{j_2}(X_{t_k}),\end{aligned}$$

and $\Delta h_k^{Y,X} := h(Y_{t_{k-}}) - h(X_{t_{k-}})$. Now we can rewrite (3.12) as follows:

$$\begin{aligned}E_{k+1-} - \Delta f_{k+1}^{Y,X}\Delta t_k &= E_{k-} + \Delta h_k^{Y,X}\Delta N_{k-1} + \Delta g_k^{Y,X}\Delta W_k \\ &\quad + \sum_{j_1,j_2=1}^{d} \Delta(\mathcal{L}^{j_1} g_{j_2})_k^{Y,X} I_{j_1,j_2}^{t_k,t_{k+1}} - R_{k+1},\end{aligned} \tag{3.13}$$

from which we arrive at

$$\begin{aligned}\big|E_{k+1-} - \Delta f_{k+1}^{Y,X}\Delta t_k\big|^2 &= |E_{k-}|^2 + \big|\Delta h_k^{Y,X}\Delta N_{k-1}\big|^2 + \big|\Delta g_k^{Y,X}\Delta W_k\big|^2 \\ &\quad + \left|\sum_{j_1,j_2=1}^{d} \Delta(\mathcal{L}^{j_1} g_{j_2})_k^{Y,X} I_{j_1,j_2}^{t_k,t_{k+1}}\right|^2 + |R_{k+1}|^2 \\ &\quad + 2\left\langle E_{k-}, \Delta h_k^{Y,X}\Delta N_{k-1}\right\rangle + 2\left\langle E_{k-}, \Delta g_k^{Y,X}\Delta W_k\right\rangle \\ &\quad + 2\left\langle E_{k-}, \sum_{j_1,j_2=1}^{d} \Delta(\mathcal{L}^{j_1} g_{j_2})_k^{Y,X} I_{j_1,j_2}^{t_k,t_{k+1}}\right\rangle\end{aligned}$$

$$
\begin{aligned}
&- 2\langle E_{k-}, R_{k+1}\rangle + 2\left\langle \Delta h_k^{Y,X}\Delta N_{k-1}, \Delta g_k^{Y,X}\Delta W_k\right\rangle \\
&+ 2\left\langle \Delta h_k^{Y,X}\Delta N_{k-1}, \sum_{j_1,j_2=1}^{d}\Delta(\mathcal{L}^{j_1}g_{j_2})_k^{Y,X} I_{j_1,j_2}^{t_k,t_{k+1}}\right\rangle \\
&- 2\left\langle \Delta h_k^{Y,X}\Delta N_{k-1}, R_{k+1}\right\rangle - 2\left\langle \Delta g_k^{Y,X}\Delta W_k, R_{k+1}\right\rangle \\
&+ 2\left\langle \Delta g_k^{Y,X}\Delta W_k, \sum_{j_1,j_2=1}^{d}\Delta(\mathcal{L}^{j_1}g_{j_2})_k^{Y,X} I_{j_1,j_2}^{t_k,t_{k+1}}\right\rangle \\
&- 2\left\langle \sum_{j_1,j_2=1}^{d}\Delta(\mathcal{L}^{j_1}g_{j_2})_k^{Y,X} I_{j_1,j_2}^{t_k,t_{k+1}}, R_{k+1}\right\rangle. \qquad (3.14)
\end{aligned}
$$

In order to estimate $\sup_{k=1,2,\dots,n_T}\mathbb{E}[|E_k|^2]$, we first establish some technical lemmas, which play the key roles in our error analysis.

**Lemma 3.4.** *Suppose that Assumptions* 2.1–2.2 *and Assumptions* 2.3–2.4 *hold. Let* $\Delta t \in \left(0, \frac{1}{4K_1}\right]$ *with* $K_1$ *defined in* (2.1)*. Then, for* $k \in \{0,1,2,\dots,n_T\}$*, it holds*

$$
\begin{aligned}
\mathbb{E}^N\left[|E_{k-}|^2\right] \le\ & C\exp(4N_TK_1\Delta t)\left(\prod_{j=1}^{n_T}\alpha_j\right)\sum_{i=0}^{n_T-1}\mathbb{E}^N\left[|R_{i+1}|^2\right] \\
&+ C\exp(4N_TK_1\Delta t)\left(\prod_{j=1}^{n_T}\alpha_j\right)\frac{1}{\Delta t}\sum_{i=1}^{n_T-1}\mathbb{E}^N\left[\left|\mathbb{E}^N\left[R_{i+1}|\mathcal{F}_{t_i}^W\right]\right|^2\right], \qquad (3.15)
\end{aligned}
$$

*where*

$$\alpha_j := 1 + C\Delta t + C\Delta N_{j-1} + C|\Delta N_{j-1}|^2 + C\Delta t|\Delta N_{j-1}|^2.$$

*Proof.* Taking the conditional expectations on both sides of (3.14) leads to

$$
\begin{aligned}
&\mathbb{E}^N\left[|E_{k+1-} - \Delta f_{k+1}^{Y,X}\Delta t_k|^2\right] \\
&= \mathbb{E}^N\left[|E_{k-}|^2\right] + \mathbb{E}^N\left[|\Delta h_k^{Y,X}\Delta N_{k-1}|^2\right] + \mathbb{E}^N\left[|\Delta g_k^{Y,X}\Delta W_k|^2\right] \\
&\quad + \mathbb{E}^N\left[\left|\sum_{j_1,j_2=1}^{d}\Delta(\mathcal{L}^{j_1}g_{j_2})_k^{Y,X} I_{j_1,j_2}^{t_k,t_{k+1}}\right|^2\right] \\
&\quad + \mathbb{E}^N\left[|R_{k+1}|^2\right] + 2\mathbb{E}^N\left[\langle E_{k-}, \Delta h_k^{Y,X}\Delta N_{k-1}\rangle\right] \\
&\quad + 2\mathbb{E}^N\left[\langle E_{k-}, \Delta g_k^{Y,X}\Delta W_k\rangle\right] - 2\mathbb{E}^N\left[\langle E_{k-}, R_{k+1}\rangle\right] \\
&\quad + 2\mathbb{E}^N\left[\left\langle E_{k-}, \sum_{j_1,j_2=1}^{d}\Delta(\mathcal{L}^{j_1}g_{j_2})_k^{Y,X} I_{j_1,j_2}^{t_k,t_{k+1}}\right\rangle\right] \\
&\quad + 2\mathbb{E}^N\left[\langle \Delta h_k^{Y,X}\Delta N_{k-1}, \Delta g_k^{Y,X}\Delta W_k\rangle\right] \\
&\quad + 2\mathbb{E}^N\left[\left\langle \Delta h_k^{Y,X}\Delta N_{k-1}, \sum_{j_1,j_2=1}^{d}\Delta(\mathcal{L}^{j_1}g_{j_2})_k^{Y,X} I_{j_1,j_2}^{t_k,t_{k+1}}\right\rangle\right] \\
&\quad + 2\mathbb{E}^N\left[\left\langle \Delta g_k^{Y,X}\Delta W_k, \sum_{j_1,j_2=1}^{d}\Delta(\mathcal{L}^{j_1}g_{j_2})_k^{Y,X} I_{j_1,j_2}^{t_k,t_{k+1}}\right\rangle\right]
\end{aligned}
$$

$$
\begin{aligned}
&-2\mathbb{E}^N\left[\left\langle \Delta h_k^{Y,X}\Delta N_{k-1}, R_{k+1}\right\rangle\right]-2\mathbb{E}^N\left[\left\langle \Delta g_k^{Y,X}\Delta W_k, R_{k+1}\right\rangle\right]\\
&-2\mathbb{E}^N\left[\left\langle \sum_{j_1,j_2=1}^{d}\Delta(\mathcal{L}^{j_1}g_{j_2})_k^{Y,X}I_{j_1,j_2}^{t_k,t_{k+1}}, R_{k+1}\right\rangle\right].
\end{aligned}
\tag{3.16}
$$

Using the independence of the Wiener process and the Poisson process, independent increment property of the Wiener process, and the properties of the conditional mathematical expectation, we have

$$
\begin{aligned}
\mathbb{E}^N\left[|\Delta g_k^{Y,X}\Delta W_k|^2\right] &= \mathbb{E}^N\left[\mathbb{E}^N\left[|\Delta g_k^{Y,X}\Delta W_k|^2|\mathcal{F}_{t_k}^W\right]\right]\\
&= \mathbb{E}^N\left[|\Delta g_k^{Y,X}|^2\mathbb{E}^N\left[|\Delta W_k|^2|\mathcal{F}_{t_k}^W\right]\right]\\
&= \mathbb{E}^N\left[|\Delta g_k^{Y,X}|^2\mathbb{E}\left[|\Delta W_k|^2\right]\right]\\
&= \mathbb{E}^N\left[|\Delta g_k^{Y,X}|^2\Delta t_k\right].
\end{aligned}
$$

In the same way, together with the following property of multiple stochastic integral with respect to the Wiener process:

$$
\mathbb{E}\left[I_{j_1,j_2}^{t_k,t_{k+1}}I_{j_3,j_4}^{t_k,t_{k+1}}\right]=\frac{1}{2}(\Delta t_k)^2\delta_{j_1,j_3}\delta_{j_2,j_4}\quad\text{for}\quad j_1,j_2,j_3,j_4\in\{1,\dots,d\},
$$

where $\delta_{i,j}$ is the Kronecker delta symbol, that is

$$
\delta_{i,j}=\begin{cases}1, & i=j,\\ 0, & i\neq j,\end{cases}
$$

we further arrive at

$$
\mathbb{E}^N\left[\left|\sum_{j_1,j_2=1}^{d}\Delta(\mathcal{L}^{j_1}g_{j_2})_k^{Y,X}I_{j_1,j_2}^{t_k,t_{k+1}}\right|^2\right]=\frac{1}{2}\sum_{j_1,j_2=1}^{d}\mathbb{E}^N\left[\left|\Delta(\mathcal{L}^{j_1}g_{j_2})_k^{Y,X}\right|^2(\Delta t_k)^2\right].
$$

Using again the properties of the conditional mathematical expectation and independent increment property of the Wiener process yields

$$
\begin{aligned}
\mathbb{E}^N\left[\left\langle E_{k-},\Delta g_k^{Y,X}\Delta W_k\right\rangle\right] &= \mathbb{E}^N\left[\mathbb{E}^N\left[\langle E_{k-},\Delta g_k^{Y,X}\Delta W_k\rangle|\mathcal{F}_{t_k}^W\right]\right]\\
&= \mathbb{E}^N\left[\langle E_{k-},\Delta g_k^{Y,X}\mathbb{E}^N\left[\Delta W_k|\mathcal{F}_{t_k}^W\right]\rangle\right]\\
&= \mathbb{E}^N\left[\langle E_{k-},\Delta g_k^{Y,X}\mathbb{E}[\Delta W_k]\rangle\right]=0.
\end{aligned}
$$

Similarly, we can obtain

$$
\begin{aligned}
&\mathbb{E}^N\left[\left\langle E_{k-},\sum_{j_1,j_2=1}^{d}\Delta(\mathcal{L}^{j_1}g_{j_2})_k^{Y,X}I_{j_1,j_2}^{t_k,t_{k+1}}\right\rangle\right]=0,\\
&\mathbb{E}^N\left[\left\langle \Delta h_k^{Y,X}\Delta N_{k-1},\Delta g_k^{Y,X}\Delta W_k\right\rangle\right]=0,\\
&\mathbb{E}^N\left[\left\langle \Delta h_k^{Y,X}\Delta N_{k-1},\sum_{j_1,j_2=1}^{d}\Delta(\mathcal{L}^{j_1}g_{j_2})_k^{Y,X}I_{j_1,j_2}^{t_k,t_{k+1}}\right\rangle\right]=0,\\
&\mathbb{E}^N\left[\left\langle \Delta g_k^{Y,X}\Delta W_k,\sum_{j_1,j_2=1}^{d}\Delta(\mathcal{L}^{j_1}g_{j_2})_k^{Y,X}I_{j_1,j_2}^{t_k,t_{k+1}}\right\rangle\right]=0.
\end{aligned}
$$

With these results in hand, we derive

$$\begin{aligned}
&\mathbb{E}^N\left[|E_{k+1-}-\Delta f_{k+1}^{Y,X}\Delta t_k|^2\right]\\
\le\,&\mathbb{E}^N\left[|E_{k-}|^2\right]+|\Delta N_{k-1}|^2\mathbb{E}^N\left[|\Delta h_k^{Y,X}|^2\right]+\Delta t\mathbb{E}^N\left[|\Delta g_k^{Y,X}|^2\right]\\
&+\frac{1}{2}(\Delta t)^2\sum_{j_1,j_2=1}^{d}\mathbb{E}^N\left[\left|\Delta(\mathcal{L}^{j_1}g_{j_2})_k^{Y,X}\right|^2\right]+\mathbb{E}^N\left[|R_{k+1}|^2\right]\\
&+2\Delta N_{k-1}\mathbb{E}^N\left[\left\langle E_{k-},\Delta h_k^{Y,X}\right\rangle\right]-2\mathbb{E}^N\left[\langle E_{k-},R_{k+1}\rangle\right]\\
&-2\Delta N_{k-1}\mathbb{E}^N\left[\left\langle \Delta h_k^{Y,X},R_{k+1}\right\rangle\right]-2\mathbb{E}^N\left[\left\langle\Delta g_k^{Y,X}\Delta W_k,R_{k+1}\right\rangle\right]\\
&-2\mathbb{E}^N\left[\left\langle\sum_{j_1,j_2=1}^{d}\Delta(\mathcal{L}^{j_1}g_{j_2})_k^{Y,X}I_{j_1,j_2}^{t_k,t_{k+1}},R_{k+1}\right\rangle\right].
\end{aligned}\tag{3.17}$$

Now we rewrite (3.17) as follows:

$$\begin{aligned}
&\mathbb{E}^N\left[|E_{k+1-}|^2\right]+\mathbb{E}^N\left[\left|\Delta f_{k+1}^{Y,X}\Delta t_k\right|^2\right]\\
\le\,&\mathbb{E}^N\left[|E_{k-}|^2\right]+|\Delta N_{k-1}|^2\mathbb{E}^N\left[\left|\Delta h_k^{Y,X}\right|^2\right]+\Delta t\mathbb{E}^N\left[\left|\Delta g_k^{Y,X}\right|^2\right]\\
&+\frac{1}{2}(\Delta t)^2\sum_{j_1,j_2=1}^{d}\mathbb{E}^N\left[|\Delta(\mathcal{L}^{j_1}g_{j_2})_k^{Y,X}|^2\right]+\mathbb{E}^N\left[|R_{k+1}|^2\right]\\
&+2\Delta N_{k-1}\mathbb{E}^N\left[\left\langle E_{k-},\Delta h_k^{Y,X}\right\rangle\right]-2\mathbb{E}^N\left[\langle E_{k-},R_{k+1}\rangle\right]\\
&-2\Delta N_{k-1}\mathbb{E}^N\left[\left\langle \Delta h_k^{Y,X},R_{k+1}\right\rangle\right]-2\mathbb{E}^N\left[\left\langle\Delta g_k^{Y,X}\Delta W_k,R_{k+1}\right\rangle\right]\\
&-2\mathbb{E}^N\left[\left\langle\sum_{j_1,j_2=1}^{d}\Delta(\mathcal{L}^{j_1}g_{j_2})_k^{Y,X}I_{j_1,j_2}^{t_k,t_{k+1}},R_{k+1}\right\rangle\right]+2\mathbb{E}^N\left[\left\langle E_{k+1-},\Delta f_{k+1}^{Y,X}\Delta t_k\right\rangle\right],
\end{aligned}\tag{3.18}$$

from which we directly obtain

$$\begin{aligned}
\mathbb{E}^N\left[|E_{k+1-}|^2\right]\le\,&\mathbb{E}^N\left[|E_{k-}|^2\right]+|\Delta N_{k-1}|^2\mathbb{E}^N\left[\left|\Delta h_k^{Y,X}\right|^2\right]+\Delta t\mathbb{E}^N\left[\left|\Delta g_k^{Y,X}\right|^2\right]\\
&+\frac{1}{2}(\Delta t)^2\sum_{j_1,j_2=1}^{d}\mathbb{E}^N\left[|\Delta(\mathcal{L}^{j_1}g_{j_2})_k^{Y,X}|^2\right]+\mathbb{E}^N\left[|R_{k+1}|^2\right]\\
&+2\Delta N_{k-1}\mathbb{E}^N\left[\left\langle E_{k-},\Delta h_k^{Y,X}\right\rangle\right]-2\mathbb{E}^N\left[\langle E_{k-},R_{k+1}\rangle\right]\\
&-2\Delta N_{k-1}\mathbb{E}^N\left[\left\langle \Delta h_k^{Y,X},R_{k+1}\right\rangle\right]-2\mathbb{E}^N\left[\left\langle\Delta g_k^{Y,X}\Delta W_k,R_{k+1}\right\rangle\right]\\
&-2\mathbb{E}^N\left[\left\langle\sum_{j_1,j_2=1}^{d}\Delta(\mathcal{L}^{j_1}g_{j_2})_k^{Y,X}I_{j_1,j_2}^{t_k,t_{k+1}},R_{k+1}\right\rangle\right]\\
&+2\mathbb{E}^N\left[\left\langle E_{k+1-},\Delta f_{k+1}^{Y,X}\Delta t_k\right\rangle\right].
\end{aligned}\tag{3.19}$$

Using the Cauchy-Schwarz inequality, the Young inequality, and using the Lipschitz condition on $h$ gives

$$\mathbb{E}^N\left[|E_{k+1-}|^2\right]\le\mathbb{E}^N\left[|E_{k-}|^2\right]+C\Delta N_{k-1}\mathbb{E}^N\left[|E_{k-}|^2\right]+C|\Delta N_{k-1}|^2\mathbb{E}^N\left[|E_{k-}|^2\right]$$

$$
\begin{aligned}
&+ C\Delta t\mathbb{E}^N\Big[\big|\Delta g_k^{Y,X}\big|^2\Big] + C(\Delta t)^2\sum_{j_1,j_2=1}^{d}\mathbb{E}^N\Big[\big|\Delta(\mathcal{L}^{j_1}g_{j_2})_k^{Y,X}\big|^2\Big]\\
&+ C\mathbb{E}^N\big[|R_{k+1}|^2\big] + C\Delta N_{k-1}\mathbb{E}^N\big[|R_{k+1}|^2\big] - 2\mathbb{E}^N\big[\langle E_{k-}, R_{k+1}\rangle\big]\\
&+ 2\mathbb{E}^N\Big[\big\langle E_{k+1-}, \Delta f_{k+1}^{Y,X}\Delta t_k\big\rangle\Big]. \qquad (3.20)
\end{aligned}
$$

In view of the Lipschitz condition on $g$ and $h$, we deduce

$$
\begin{aligned}
\big|\Delta g_k^{Y,X}\big|^2 &= \big|g(Y_{t_k}) - g(X_{t_k})\big|^2\\
&= \big|g\big(Y_{t_{k-}} + h(Y_{t_{k-}})\Delta N_{k-1}\big) - g\big(X_{t_{k-}} + h(X_{t_{k-}})\Delta N_{k-1}\big)\big|^2\\
&\le C\big|Y_{t_{k-}} + h(Y_{t_{k-}})\Delta N_{k-1} - X_{t_{k-}} - h(X_{t_{k-}})\Delta N_{k-1}\big|^2\\
&\le C\big|E_{k-}\big|^2 + C\big|\Delta h_k^{Y,X}\Delta N_{k-1}\big|^2\\
&\le C\big|E_{k-}\big|^2 + C\big|\Delta N_{k-1}\big|^2\big|E_{k-}\big|^2. \qquad (3.21)
\end{aligned}
$$

Similarly, by employing (2.12), we obtain

$$
\sum_{j_1,j_2=1}^{d}\big|\Delta(\mathcal{L}^{j_1}g_{j_2})_k^{Y,X}\big|^2 \le C\big|E_{k-}\big|^2 + C\big|\Delta N_{k-1}\big|^2\big|E_{k-}\big|^2. \qquad (3.22)
$$

Inserting (3.21) and (3.22) into (3.20), and using the condition (2.1) on $f$ give

$$
\begin{aligned}
\mathbb{E}^N\big[|E_{k+1-}|^2\big] \le\ &\mathbb{E}^N\big[|E_{k-}|^2\big] + C\Delta N_{k-1}\mathbb{E}^N\big[|E_{k-}|^2\big] + C|\Delta N_{k-1}|^2\mathbb{E}^N\big[|E_{k-}|^2\big]\\
&+ C\Delta t\mathbb{E}^N\big[|E_{k-}|^2\big] + C\Delta t|\Delta N_{k-1}|^2\mathbb{E}^N\big[|E_{k-}|^2\big]\\
&+ C\mathbb{E}^N\big[|R_{k+1}|^2\big] + C\Delta N_{k-1}\mathbb{E}^N\big[|R_{k+1}|^2\big] - 2\mathbb{E}^N\big[\langle E_{k-}, R_{k+1}\rangle\big]\\
&+ 2K_1\Delta t\mathbb{E}^N\big[|E_{k+1-}|^2\big], \qquad (3.23)
\end{aligned}
$$

which yields

$$
\begin{aligned}
&(1-2K_1\Delta t)\mathbb{E}^N\big[|E_{k+1-}|^2\big]\\
&\le \mathbb{E}^N\big[|E_{k-}|^2\big] + C\Delta N_{k-1}\mathbb{E}^N\big[|E_{k-}|^2\big] + C|\Delta N_{k-1}|^2\mathbb{E}^N\big[|E_{k-}|^2\big]\\
&\quad + C\Delta t\mathbb{E}^N\big[|E_{k-}|^2\big] + C\Delta t|\Delta N_{k-1}|^2\mathbb{E}^N\big[|E_{k-}|^2\big]\\
&\quad + C\mathbb{E}^N\big[|R_{k+1}|^2\big] + C\Delta N_{k-1}\mathbb{E}^N\big[|R_{k+1}|^2\big] - 2\mathbb{E}^N\big[\langle E_{k-}, R_{k+1}\rangle\big]. \qquad (3.24)
\end{aligned}
$$

With the aid of the properties of the conditional expectation and the Young inequality, we deduce

$$
\begin{aligned}
&(1-2K_1\Delta t)\mathbb{E}^N\big[|E_{k+1-}|^2\big]\\
&\le \mathbb{E}^N\big[|E_{k-}|^2\big] + C\Delta N_{k-1}\mathbb{E}^N\big[|E_{k-}|^2\big] + C|\Delta N_{k-1}|^2\mathbb{E}^N\big[|E_{k-}|^2\big]\\
&\quad + C\Delta t\mathbb{E}^N\big[|E_{k-}|^2\big] + C\Delta t|\Delta N_{k-1}|^2\mathbb{E}^N\big[|E_{k-}|^2\big]\\
&\quad + C\mathbb{E}^N\big[|R_{k+1}|^2\big] + C\Delta N_{k-1}\mathbb{E}^N\big[|R_{k+1}|^2\big] - 2\mathbb{E}^N\big[\big\langle E_{k-}, \mathbb{E}^N\big[R_{k+1}|\mathcal{F}_{t_k}^W\big]\big\rangle\big]\\
&\le \mathbb{E}^N\big[|E_{k-}|^2\big] + C\Delta N_{k-1}\mathbb{E}^N\big[|E_{k-}|^2\big] + C|\Delta N_{k-1}|^2\mathbb{E}^N\big[|E_{k-}|^2\big]\\
&\quad + C\Delta t\mathbb{E}^N\big[|E_{k-}|^2\big] + C\Delta t|\Delta N_{k-1}|^2\mathbb{E}^N\big[|E_{k-}|^2\big]\\
&\quad + C\mathbb{E}^N\big[|R_{k+1}|^2\big] + C\Delta N_{k-1}\mathbb{E}^N\big[|R_{k+1}|^2\big] + \Delta t\mathbb{E}^N\big[|E_{k-}|^2\big]\\
&\quad + \frac{1}{\Delta t}\mathbb{E}^N\Big[\big|\mathbb{E}^N\big[R_{k+1}|\mathcal{F}_{t_k}^W\big]\big|^2\Big]
\end{aligned}
$$

$$
\begin{aligned}
&\leq \big(1 + C\Delta t + C\Delta N_{k-1} + C|\Delta N_{k-1}|^2 + C\Delta t|\Delta N_{k-1}|^2\big)\mathbb{E}^N\left[|E_{k-}|^2\right] \\
&\quad + (C + C\Delta N_{k-1})\mathbb{E}^N\left[|R_{k+1}|^2\right] + \frac{1}{\Delta t}\mathbb{E}^N\left[\left|\mathbb{E}^N\left[R_{k+1}|\mathcal{F}^W_{t_k}\right]\right|^2\right]. \qquad (3.25)
\end{aligned}
$$

Define

$$
\begin{aligned}
&\alpha_k := 1 + C\Delta t + C\Delta N_{k-1} + C|\Delta N_{k-1}|^2 + C\Delta t|\Delta N_{k-1}|^2, \\
&\beta_k := (C + C\Delta N_{k-1}), \quad \gamma_\Delta := 1 - 2K_1\Delta t,
\end{aligned}
$$

we then have

$$
\begin{aligned}
\mathbb{E}^N\left[|E_{k+1-}|^2\right] &\leq \gamma_\Delta^{-1}\alpha_k\mathbb{E}^N\left[|E_{k-}|^2\right] + \gamma_\Delta^{-1}\beta_k\mathbb{E}^N\left[|R_{k+1}|^2\right] \\
&\quad + \frac{1}{\gamma_\Delta\Delta t}\mathbb{E}^N\left[\left|\mathbb{E}^N\left[R_{k+1}|\mathcal{F}^W_{t_k}\right]\right|^2\right]. \qquad (3.26)
\end{aligned}
$$

By the iteration of (3.26), one can deduce for all $k = 1, 2, \ldots, n_T - 1$ that

$$
\begin{aligned}
\mathbb{E}^N\left[|E_{k+1-}|^2\right] &\leq \left(\prod_{i=2}^{k+1}\gamma_\Delta^{-1}\right)\left(\prod_{j=1}^{k}\alpha_j\right)\mathbb{E}^N\left[|E_{1-}|^2\right] \\
&\quad + \sum_{i=1}^{k}\left(\prod_{l=i+1}^{k+1}\gamma_\Delta^{-1}\right)\left(\prod_{j=i+1}^{k}\alpha_j\right)\beta_i\mathbb{E}^N\left[|R_{i+1}|^2\right] \\
&\quad + \sum_{i=1}^{k}\left(\prod_{l=i+1}^{k+1}\gamma_\Delta^{-1}\right)\left(\prod_{j=i+1}^{k}\alpha_j\right)\frac{1}{\Delta t}\mathbb{E}^N\left[\left|\mathbb{E}^N\left[R_{i+1}|\mathcal{F}^W_{t_i}\right]\right|^2\right]. \qquad (3.27)
\end{aligned}
$$

From (3.13), we can derive

$$
E_{1-} = \Delta f_1^{Y,X}\Delta t_1 - R_1, \qquad (3.28)
$$

from which we have

$$
\begin{aligned}
|E_{1-}|^2 &= \left\langle \Delta f_1^{Y,X}\Delta t_1, E_{1-}\right\rangle - \langle R_1, E_{1-}\rangle \\
&\leq K_1\Delta t|E_{1-}|^2 + \frac{1}{2}|E_{1-}|^2 + \frac{1}{2}|R_1|^2. \qquad (3.29)
\end{aligned}
$$

Then it follows that

$$
|E_{1-}|^2 \leq \gamma_\Delta^{-1}|R_1|^2. \qquad (3.30)
$$

Substituting (3.30) into (3.27) gives

$$
\begin{aligned}
\mathbb{E}^N\left[|E_{k+1-}|^2\right] &\leq \left(\prod_{i=2}^{k+1}\gamma_\Delta^{-1}\right)\left(\prod_{j=1}^{k}\alpha_j\right)\gamma_\Delta^{-1}\mathbb{E}^N\left[|R_1|^2\right] \\
&\quad + \sum_{i=1}^{k}\left(\prod_{l=i+1}^{k+1}\gamma_\Delta^{-1}\right)\left(\prod_{j=i+1}^{k}\alpha_j\right)\beta_i\mathbb{E}^N\left[|R_{i+1}|^2\right] \\
&\quad + \sum_{i=1}^{k}\left(\prod_{l=i+1}^{k+1}\gamma_\Delta^{-1}\right)\left(\prod_{j=i+1}^{k}\alpha_j\right)\frac{1}{\Delta t}\mathbb{E}^N\left[\left|\mathbb{E}^N\left[R_{i+1}|\mathcal{F}^W_{t_i}\right]\right|^2\right] \\
&= \sum_{i=0}^{k}\left(\prod_{l=i+1}^{k+1}\gamma_\Delta^{-1}\right)\left(\prod_{j=i+1}^{k}\alpha_j\right)\beta_i\mathbb{E}^N\left[|R_{i+1}|^2\right] \\
&\quad + \sum_{i=1}^{k}\left(\prod_{l=i+1}^{k+1}\gamma_\Delta^{-1}\right)\left(\prod_{j=i+1}^{k}\alpha_j\right)\frac{1}{\Delta t}\mathbb{E}^N\left[\left|\mathbb{E}^N\left[R_{i+1}|\mathcal{F}^W_{t_i}\right]\right|^2\right] \qquad (3.31)
\end{aligned}
$$

with $\beta_0 := 1$. Note that $\beta_i \le C, i = 0, 1, \ldots, n_T$, then it follows for all $k = 1, \ldots, n_T$ that

$$\mathbb{E}^N\left[|E_{k-}|^2\right] \le C\left(\prod_{i=1}^{n_T}\gamma_\Delta^{-1}\right)\left(\prod_{j=1}^{n_T}\alpha_j\right)\sum_{i=0}^{n_T-1}\mathbb{E}^N\left[|R_{i+1}|^2\right] + \left(\prod_{i=1}^{n_T}\gamma_\Delta^{-1}\right)\left(\prod_{j=1}^{n_T}\alpha_j\right)\frac{1}{\Delta t}\sum_{i=1}^{n_T-1}\mathbb{E}^N\left[\left|\mathbb{E}^N\left[R_{i+1}|\mathcal{F}_{t_i}^W\right]\right|^2\right]. \tag{3.32}$$

Note that $2K_1\Delta t \in (0, 1/2]$, we further get

$$\begin{aligned}\prod_{i=1}^{n_T}\gamma_\Delta^{-1} &= (1-2K_1\Delta t)^{-n_T}\\ &\le \exp(4n_TK_1\Delta t) \le \exp(4K_1T)\exp(4N_TK_1\Delta t)\\ &\le C\exp(4N_TK_1\Delta t),\end{aligned} \tag{3.33}$$

where we have used the inequality $\frac{1}{1-x} \le \exp(2x)$ for $x \in [0, 1/2]$. Substituting this into (3.32) yields

$$\mathbb{E}^N\left[|E_{k-}|^2\right] \le C\exp(4N_TK_1\Delta t)\left(\prod_{j=1}^{n_T}\alpha_j\right)\sum_{i=0}^{n_T-1}\mathbb{E}^N\left[|R_{i+1}|^2\right] + C\exp(4N_TK_1\Delta t)\left(\prod_{j=1}^{n_T}\alpha_j\right)\frac{1}{\Delta t}\sum_{i=1}^{n_T-1}\mathbb{E}^N\left[\left|\mathbb{E}^N\left[R_{i+1}|\mathcal{F}_{t_i}^W\right]\right|^2\right]. \tag{3.34}$$

The proof ends. □

In the following, we are to establish the estimates of the remainder terms $\mathbb{E}^N[|R_{k+1}|^2]$ and $\mathbb{E}^N[|\mathbb{E}^N[R_{k+1}|\mathcal{F}_{t_k}^W]|^2]$.

**Lemma 3.5.** *Under the same assumptions of Lemma* 3.4*, we have*

$$\mathbb{E}^N\left[|R_{k+1}|^2\right] \le C\mathcal{X}_1(\Delta t)^3, \quad \mathbb{E}^N\left[\left|\mathbb{E}^N\left[R_{k+1}|\mathcal{F}_{t_k}^W\right]\right|^2\right] \le C\mathcal{X}_2(\Delta t)^4, \tag{3.35}$$

*where*

$$\begin{aligned}\mathcal{X}_1 &:= \mathbb{E}^N\Big[\sup_{0\le s\le T}|X_s|^{4+2\kappa}\Big] + \Big(1+\mathbb{E}^N\big[\sup_{0\le t\le T}|X_t|^4\big]\Big)^{\frac12}\Phi_{N,4}^{\frac12}\\ &\quad + \Big(1+\mathbb{E}^N\Big[\sup_{0\le t\le T}|X_t|^{4\kappa}\Big]\Big)^{\frac12}\Phi_{N,4}^{\frac12} + \Phi_{N,4} + \Phi_{N,2},\\ \mathcal{X}_2 &:= \mathbb{E}^N\Big[\sup_{0\le t\le T}|X_t|^{2\kappa}\Big] + \mathbb{E}^N\Big[\sup_{0\le t\le T}|X_t|^{2+4\kappa}\Big]\\ &\quad + \Big(1+\mathbb{E}^N\Big[\sup_{0\le t\le T}|X_t|^{4\kappa-4}\Big]\Big)^{\frac12}\Phi_{N,8}^{\frac12} + \Phi_{N,2}.\end{aligned}$$

*Proof.* Based on (3.11), we first have

$$\begin{aligned}\mathbb{E}^N\left[|R_{k+1}|^2\right] &\le 2\mathbb{E}^N\left[\left|\int_{t_k}^{t_{k+1}}[f(X_t)-f(X_{t_{k+1}-})]\mathrm{d}t\right|^2\right]\\ &\quad + 2\mathbb{E}^N\left[\left|\int_{t_k}^{t_{k+1}}[g(X_t)-g(X_{t_k})]\mathrm{d}W_t - \sum_{j_1,j_2=1}^{d}\mathcal{L}^{j_1}g_{j_2}(X_{t_k})I_{j_1,j_2}^{t_k,t_{k+1}}\right|^2\right]\\ &=: 2\mathbb{J}_1 + 2\mathbb{J}_2.\end{aligned} \tag{3.36}$$

Next, we deal with $\mathbb{J}_1$ and $\mathbb{J}_2$ separately. For $\mathbb{J}_1$, due to (2.10) we have

$$\begin{aligned}\mathbb{J}_1 &\le C\mathbb{E}^N\left[\Delta t_k\int_{t_k}^{t_{k+1}}\big|f(X_t)-f(X_{t_{k+1}-})\big|^2\mathrm{d}t\right]\\ &\le C\Delta t\mathbb{E}^N\left[\int_{t_k}^{t_{k+1}}\big|f(X_t)-f(X_{t_{k+1}-})\big|^2\mathrm{d}t\right]\\ &\le C\Delta t\mathbb{E}^N\left[\int_{t_k}^{t_{k+1}}\Big(1+\big|X_t\big|^{2\kappa}+\big|X_{t_{k+1}-}\big|^{2\kappa}\Big)\big|X_t-X_{t_{k+1}-}\big|^2\mathrm{d}t\right]\\ &\le C\Delta t\mathbb{E}^N\left[\sup_{0\le t\le T}(1+|X_t|^{2\kappa})\int_{t_k}^{t_{k+1}}\big|X_t-X_{t_{k+1}-}\big|^2\mathrm{d}t\right].\end{aligned}\tag{3.37}$$

Utilizing the Cauchy–Schwarz inequality and Lemma 3.3 yields

$$\begin{aligned}\mathbb{J}_1 &\le C\Delta t\left(1+\mathbb{E}^N\Big[\sup_{0\le t\le T}|X_t|^{4\kappa}\Big]\right)^{\frac12}\left(\mathbb{E}^N\left[\left(\int_{t_k}^{t_{k+1}}\big|X_t-X_{t_{k+1}-}\big|^2\mathrm{d}t\right)^2\right]\right)^{\frac12}\\ &\le C\Delta t\left(1+\mathbb{E}^N\Big[\sup_{0\le t\le T}|X_t|^{4\kappa}\Big]\right)^{\frac12}\left(\Delta t\mathbb{E}^N\left[\int_{t_k}^{t_{k+1}}\big|X_t-X_{t_{k+1}-}\big|^4\mathrm{d}t\right]\right)^{\frac12}\\ &\le C\left(1+\mathbb{E}^N\Big[\sup_{0\le t\le T}|X_t|^{4\kappa}\Big]\right)^{\frac12}\Phi_{N,4}^{\frac12}(\Delta t)^3.\end{aligned}\tag{3.38}$$

In order to estimate the term $\mathbb{J}_2$, we first observe that

$$\mathcal{L}^{j_1}g_{j_2}(x)=\sum_{k=1}^m g_{k.j_1}\frac{\partial g_{j_2}(x)}{\partial x^k}=\frac{\partial g_{j_2}}{\partial x}(x)g_{j_1}(x).$$

Therefore, it follows that

$$\begin{aligned}&\int_{t_k}^{t_{k+1}}[g(X_t)-g(X_{t_k})]\mathrm{d}W_t-\sum_{j_1,j_2=1}^d\mathcal{L}^{j_1}g_{j_2}(X_{t_k})I_{j_1,j_2}^{t_k,t_{k+1}}\\ =&\sum_{j_2=1}^d\int_{t_k}^{t_{k+1}}\left[g_{j_2}(X_t)-g_{j_2}(X_{t_k})-\sum_{j_1=1}^d\mathcal{L}^{j_1}g_{j_2}(X_{t_k})(W_t^{j_1}-W_{t_k}^{j_1})\right]\mathrm{d}W_t^{j_2}\\ =&\sum_{j_2=1}^d\int_{t_k}^{t_{k+1}}\left[g_{j_2}(X_t)-g_{j_2}(X_{t_k})-\sum_{j_1=1}^d\frac{\partial g_{j_2}}{\partial x}(X_{t_k})g_{j_1}(X_{t_k})(W_t^{j_1}-W_{t_k}^{j_1})\right]\mathrm{d}W_t^{j_2}.\end{aligned}\tag{3.39}$$

Let $t\in[t_k,t_{k+1})$, then we have

$$\begin{aligned}g_{j_2}(X_t)-g_{j_2}(X_{t_k})&=\frac{\partial g_{j_2}}{\partial x}(X_{t_k})(X_t-X_{t_k})+\widetilde{R}_{g_{j_2}}\\ &=\frac{\partial g_{j_2}}{\partial x}(X_{t_k})\left(\int_{t_k}^t f(X_s)\mathrm{d}s+\int_{t_k}^t g(X_s)\mathrm{d}W_s\right)+\widetilde{R}_{g_{j_2}},\end{aligned}\tag{3.40}$$

where

$$\widetilde{R}_{g_{j_2}}:=\int_0^1\left[\frac{\partial g_{j_2}}{\partial x}\big(X_{t_k}+s(X_t-X_{t_k})\big)-\frac{\partial g_{j_2}}{\partial x}(X_{t_k})\right](X_t-X_{t_k})\mathrm{d}s.\tag{3.41}$$

Inserting (3.40) into (3.39) immediately gives

$$\int_{t_k}^{t_{k+1}}[g(X_t)-g(X_{t_k})]\mathrm{d}W_t-\sum_{j_1,j_2=1}^d\mathcal{L}^{j_1}g_{j_2}(X_{t_k})I_{j_1,j_2}^{t_k,t_{k+1}}$$

$$= \sum_{j_2=1}^{d} \int_{t_k}^{t_{k+1}} \left[ \frac{\partial g_{j_2}}{\partial x}(X_{t_k}) \Big( \int_{t_k}^{t} f(X_s) \mathrm{d}s + \int_{t_k}^{t} [g(X_s) - g(X_{t_k})] \mathrm{d}W_s \Big) + \widetilde{R}_{g_{j_2}} \right] \mathrm{d}W_t^{j_2}. \tag{3.42}$$

Taking the Itô isometry into account, we arrive at

$$\begin{aligned}
\mathbb{J}_2 &= \sum_{j_2=1}^{d} \mathbb{E}^N \Big[ \int_{t_k}^{t_{k+1}} \Big| \frac{\partial g_{j_2}}{\partial x}(X_{t_k}) \Big( \int_{t_k}^{t} f(X_s) \mathrm{d}s + \int_{t_k}^{t} [g(X_s) - g(X_{t_k})] \mathrm{d}W_s \Big) + \widetilde{R}_{g_{j_2}} \Big|^2 \mathrm{d}t \Big] \\
&\le 3 \sum_{j_2=1}^{d} \mathbb{E}^N \Big[ \int_{t_k}^{t_{k+1}} \Big| \frac{\partial g_{j_2}}{\partial x}(X_{t_k}) \int_{t_k}^{t} f(X_s) \mathrm{d}s \Big|^2 \mathrm{d}t \Big] + 3 \sum_{j_2=1}^{d} \mathbb{E}^N \Big[ \int_{t_k}^{t_{k+1}} \big| \widetilde{R}_{g_{j_2}} \big|^2 \mathrm{d}t \Big] \\
&\quad + 3 \sum_{j_2=1}^{d} \mathbb{E}^N \Big[ \int_{t_k}^{t_{k+1}} \Big| \frac{\partial g_{j_2}}{\partial x}(X_{t_k}) \int_{t_k}^{t} [g(X_s) - g(X_{t_k})] \mathrm{d}W_s \Big|^2 \mathrm{d}t \Big] \\
&=: 3\mathbb{J}_{21} + 3\mathbb{J}_{22} + 3\mathbb{J}_{23}.
\end{aligned} \tag{3.43}$$

Next, we handle $\mathbb{J}_{21}$, $\mathbb{J}_{22}$, and $\mathbb{J}_{23}$ one by one. By means of the Cauchy–Schwarz inequality, (2.14), (2.11), and (2.7), we derive

$$\begin{aligned}
\mathbb{J}_{21} &= \sum_{j_2=1}^{d} \mathbb{E}^N \Big[ \int_{t_k}^{t_{k+1}} \Big| \int_{t_k}^{t} \frac{\partial g_{j_2}}{\partial x}(X_{t_k}) f(X_s) \mathrm{d}s \Big|^2 \mathrm{d}t \Big] \\
&\le \sum_{j_2=1}^{d} \mathbb{E}^N \Big[ \int_{t_k}^{t_{k+1}} (t - t_k) \int_{t_k}^{t} \Big| \frac{\partial g_{j_2}}{\partial x}(X_{t_k}) f(X_s) \Big|^2 \mathrm{d}s \mathrm{d}t \Big] \\
&\le C(\Delta t)^2 \sum_{j_2=1}^{d} \int_{t_k}^{t_{k+1}} \Big( \Phi_{N,2} + \sup_{0 \le s \le T} \mathbb{E}^N \big[ |X_s|^{4+2\kappa} \big] \Big) \mathrm{d}t \\
&\le C \bigg( \Phi_{N,2} + \mathbb{E}^N \Big[ \sup_{0 \le s \le T} |X_s|^{4+2\kappa} \Big] \bigg) (\Delta t)^3.
\end{aligned}$$

Using again the Cauchy–Schwarz inequality, (2.13) and the result of Lemma 3.3, we obtain

$$\begin{aligned}
\mathbb{J}_{22} &= \sum_{j_2=1}^{d} \mathbb{E}^N \Big[ \int_{t_k}^{t_{k+1}} \Big| \int_0^1 \Big[ \frac{\partial g_{j_2}}{\partial x}\big(X_{t_k} + s(X_t - X_{t_k})\big) - \frac{\partial g_{j_2}}{\partial x}(X_{t_k}) \Big] (X_t - X_{t_k}) \mathrm{d}s \Big|^2 \mathrm{d}t \Big] \\
&\le C \sum_{j_2=1}^{d} \mathbb{E}^N \Big[ \int_{t_k}^{t_{k+1}} \int_0^1 \Big| \frac{\partial g_{j_2}}{\partial x}\big(X_{t_k} + s(X_t - X_{t_k})\big) - \frac{\partial g_{j_2}}{\partial x}(X_{t_k}) \Big|^2 \Big| X_t - X_{t_k} \Big|^2 \mathrm{d}s \mathrm{d}t \Big] \\
&\le C \sum_{j_2=1}^{d} \int_{t_k}^{t_{k+1}} \mathbb{E}^N \big[ \big| X_t - X_{t_k} \big|^4 \big] \mathrm{d}t \le C \Phi_{N,4} (\Delta t)^3.
\end{aligned}$$

With the help of the Itô isometry, (2.14) and the Höder inequality, we have

$$\begin{aligned}
\mathbb{J}_{23} &= \sum_{j_2=1}^{d} \mathbb{E}^N \Big[ \int_{t_k}^{t_{k+1}} \int_{t_k}^{t} \Big| \frac{\partial g_{j_2}}{\partial x}(X_{t_k}) \Big|^2 \Big| g(X_s) - g(X_{t_k}) \Big|^2 \mathrm{d}s \mathrm{d}t \Big] \\
&\le C \sum_{j_2=1}^{d} \int_{t_k}^{t_{k+1}} \mathbb{E}^N \Big[ \int_{t_k}^{t} \big(1 + |X_{t_k}|^2\big) \big| X_s - X_{t_k} \big|^2 \mathrm{d}s \Big] \mathrm{d}t \\
&\le C \sum_{j_2=1}^{d} \int_{t_k}^{t_{k+1}} \mathbb{E}^N \Big[ \big(1 + \sup_{0 \le t \le T} |X_t|^2\big) \int_{t_k}^{t} \big| X_s - X_{t_k} \big|^2 \mathrm{d}s \Big] \mathrm{d}t
\end{aligned}$$

$$\leq C\Big(1+\mathbb{E}^N\Big[\sup_{0\leq t\leq T}|X_t|^4\Big]\Big)^{\frac{1}{2}}\sum_{j_2=1}^{d}\int_{t_k}^{t_{k+1}}\Big(\mathbb{E}^N\Big[\Big(\int_{t_k}^{t_{k+1}}\big|X_s-X_{t_k}\big|^2\mathrm{d}s\Big)^2\Big]\Big)^{\frac{1}{2}}\mathrm{d}t,$$

from which we obtain

$$\begin{aligned}\mathbb{J}_{23}&\leq C\Big(1+\mathbb{E}^N\big[\sup_{0\leq t\leq T}|X_t|^4\big]\Big)^{\frac{1}{2}}\sum_{j_2=1}^{d}\int_{t_k}^{t_{k+1}}\Big(\Delta t\mathbb{E}^N\Big[\int_{t_k}^{t_{k+1}}\big|X_s-X_{t_k}\big|^4\mathrm{d}s\Big]\Big)^{\frac{1}{2}}\mathrm{d}t\\&\leq C\Big(1+\mathbb{E}^N\Big[\sup_{0\leq t\leq T}|X_t|^4\Big]\Big)^{\frac{1}{2}}\Phi_{N,4}^{\frac{1}{2}}(\Delta t)^3.\end{aligned}$$

Using these estimates along with (3.43) results in

$$\mathbb{J}_2\leq C\Big(\mathbb{E}^N\Big[\sup_{0\leq s\leq T}|X_s|^{4+2\kappa}\Big]+\Big(1+\mathbb{E}^N\Big[\sup_{0\leq t\leq T}|X_t|^4\Big]\Big)^{\frac{1}{2}}\Phi_{N,4}^{\frac{1}{2}}+\Phi_{N,4}+\Phi_{N,2}\Big)(\Delta t)^3. \tag{3.44}$$

Combining (3.38) and (3.44), we end up with

$$\mathbb{E}^N\big[|R_{k+1}|^2\big]\leq C\mathcal{X}_1(\Delta t)^3, \tag{3.45}$$

where

$$\begin{aligned}\mathcal{X}_1:=\mathbb{E}^N\Big[\sup_{0\leq s\leq T}|X_s|^{4+2\kappa}\Big]+\Big(1+\mathbb{E}^N\big[\sup_{0\leq t\leq T}|X_t|^4\big]\Big)^{\frac{1}{2}}\Phi_{N,4}^{\frac{1}{2}}\\+\Big(1+\mathbb{E}^N\Big[\sup_{0\leq t\leq T}|X_t|^{4\kappa}\Big]\Big)^{\frac{1}{2}}\Phi_{N,4}^{\frac{1}{2}}+\Phi_{N,4}+\Phi_{N,2}.\end{aligned}$$

Now we turn to bound the term $\mathbb{E}^N[|\mathbb{E}^N[R_{k+1}|\mathcal{F}_{t_k}^W]|^2]$. Since the stochastic integral vanishes under conditional expectation, we observe that

$$\mathbb{E}^N\Big[\big|\mathbb{E}^N\big[R_{k+1}|\mathcal{F}_{t_k}^W\big]\big|^2\Big]\leq\mathbb{E}^N\left[\left|\mathbb{E}^N\left[\int_{t_k}^{t_{k+1}}[f(X_t)-f(X_{t_{k+1}-})]\mathrm{d}t\big|\mathcal{F}_{t_k}^W\right]\right|^2\right]. \tag{3.46}$$

Note that

$$\begin{aligned}&\mathbb{E}^N\left[\int_{t_k}^{t_{k+1}}\big[f(X_t)-f(X_{t_{k+1}-})\big]\mathrm{d}t\big|\mathcal{F}_{t_k}^W\right]\\&=-\mathbb{E}^N\left[\int_{t_k}^{t_{k+1}}\frac{\partial f}{\partial x}(X_t)\Big(\int_t^{t_{k+1}}f(X_s)\mathrm{d}s+\int_t^{t_{k+1}}g(X_s)\mathrm{d}W_s\Big)+\widetilde{R}_f\,\mathrm{d}t\big|\mathcal{F}_{t_k}^W\right]\\&=-\mathbb{E}^N\left[\int_{t_k}^{t_{k+1}}\int_t^{t_{k+1}}\frac{\partial f}{\partial x}(X_t)f(X_s)\mathrm{d}s\,\mathrm{d}t\big|\mathcal{F}_{t_k}^W\right]-\mathbb{E}^N\left[\int_{t_k}^{t_{k+1}}\widetilde{R}_f\,\mathrm{d}t\big|\mathcal{F}_{t_k}^W\right],\end{aligned} \tag{3.47}$$

where

$$\widetilde{R}_f:=-\int_0^1\left[\frac{\partial f}{\partial x}\big(X_t+s(X_{t_{k+1}-}-X_t)\big)-\frac{\partial f}{\partial x}(X_t)\right](X_{t_{k+1}-}-X_t)\mathrm{d}s. \tag{3.48}$$

Thus, it follows immediately that

$$\begin{aligned}\mathbb{E}^N\Big[\big|\mathbb{E}^N\big[R_{k+1}|\mathcal{F}_{t_k}^W\big]\big|^2\Big]\leq 2\mathbb{E}^N\left[\left|\mathbb{E}^N\left[\int_{t_k}^{t_{k+1}}\int_t^{t_{k+1}}\frac{\partial f}{\partial x}(X_t)f(X_s)\mathrm{d}s\,\mathrm{d}t\big|\mathcal{F}_{t_k}^W\right]\right|^2\right]\\+2\mathbb{E}^N\left[\left|\mathbb{E}^N\left[\int_{t_k}^{t_{k+1}}\widetilde{R}_f\,\mathrm{d}t\big|\mathcal{F}_{t_k}^W\right]\right|^2\right]=:\mathbb{J}_3+\mathbb{J}_4.\end{aligned} \tag{3.49}$$

Taking into account (2.9) and (2.11) yields

$$
\begin{aligned}
\mathbb{J}_3 &\leq C(\Delta t)^2 \mathbb{E}^N \left[ \int_{t_k}^{t_{k+1}} \int_t^{t_{k+1}} \left| \frac{\partial f}{\partial x}(X_t) f(X_s) \right|^2 \mathrm{d}s \, \mathrm{d}t \right] \\
&\leq C(\Delta t)^2 \int_{t_k}^{t_{k+1}} \int_t^{t_{k+1}} \left( \Phi_{N,2} + \mathbb{E}^N \Big[ \sup_{0 \leq t \leq T} |X_t|^{2\kappa} \Big] + \mathbb{E}^N \Big[ \sup_{0 \leq t \leq T} |X_t|^{2+4\kappa} \Big] \right) \mathrm{d}s \, \mathrm{d}t \\
&\leq C \left( \Phi_{N,2} + \mathbb{E}^N \Big[ \sup_{0 \leq t \leq T} |X_t|^{2\kappa} \Big] + \mathbb{E}^N \Big[ \sup_{0 \leq t \leq T} |X_t|^{2+4\kappa} \Big] \right) (\Delta t)^4. \qquad (3.50)
\end{aligned}
$$

For the term $\mathbb{J}_4$, with the aid of (2.8) and Lemma 3.3, we obtain

$$
\begin{aligned}
\mathbb{J}_4 &\leq C \Delta t \mathbb{E}^N \left[ \int_{t_k}^{t_{k+1}} \int_0^1 \left| \frac{\partial f}{\partial x}(X_t + s(X_{t_{k+1}-} - X_t)) - \frac{\partial f}{\partial x}(X_t) \right|^2 \left| X_{t_{k+1}-} - X_t \right|^2 \mathrm{d}s \, \mathrm{d}t \right] \\
&\leq C \Delta t \mathbb{E}^N \left[ \int_{t_k}^{t_{k+1}} \int_0^1 \left( 1 + |sX_{t_{k+1}-} + (1-s)X_t|^{2\kappa-2} + |X_t|^{2\kappa-2} \right) \left| X_{t_{k+1}-} - X_t \right|^4 \mathrm{d}s \, \mathrm{d}t \right] \\
&\leq C \Delta t \mathbb{E}^N \left[ \left( 1 + \sup_{0 \leq t \leq T} |X_t|^{2\kappa-2} \right) \int_{t_k}^{t_{k+1}} \left| X_{t_{k+1}-} - X_t \right|^4 \mathrm{d}t \right] \\
&\leq C \Delta t \left( 1 + \mathbb{E}^N \Big[ \sup_{0 \leq t \leq T} |X_t|^{4\kappa-4} \Big] \right)^{\frac{1}{2}} \left( \mathbb{E}^N \left[ \left( \int_{t_k}^{t_{k+1}} \left| X_{t_{k+1}-} - X_t \right|^4 \mathrm{d}t \right)^2 \right] \right)^{\frac{1}{2}} \\
&\leq C \left( 1 + \mathbb{E}^N \Big[ \sup_{0 \leq t \leq T} |X_t|^{4\kappa-4} \Big] \right)^{\frac{1}{2}} \Phi_{N,8}^{\frac{1}{2}} (\Delta t)^4. \qquad (3.51)
\end{aligned}
$$

Substituting (3.50) and (3.51) into (3.49) gives

$$
\mathbb{E}^N \Big[ \left| \mathbb{E}^N \left[ R_{k+1} | \mathcal{F}_{t_k}^W \right] \right|^2 \Big] \leq C \mathcal{X}_2 (\Delta t)^4, \qquad (3.52)
$$

where

$$
\mathcal{X}_2 := \mathbb{E}^N \Big[ \sup_{0 \leq t \leq T} |X_t|^{2\kappa} \Big] + \mathbb{E}^N \Big[ \sup_{0 \leq t \leq T} |X_t|^{2+4\kappa} \Big] + \left( 1 + \mathbb{E}^N \Big[ \sup_{0 \leq t \leq T} |X_t|^{4\kappa-4} \Big] \right)^{\frac{1}{2}} \Phi_{N,8}^{\frac{1}{2}} + \Phi_{N,2}.
$$

This completes the proof. □

Now, we are ready to prove the main strong convergence result.

**Theorem 3.1.** *Suppose that Assumptions* 2.1–2.2 *and Assumptions* 2.3–2.4 *are fulfilled and let* $\Delta t \in (0, \frac{1}{4K_1}]$ *with* $K_1$ *defined in* (2.1). *Then, there exists a positive constant* $C$ *(independent of* $\Delta t$*) such that*

$$
\sup_{k=1,\ldots,n_T} \mathbb{E}\big[ |Y_{t_k} - X_{t_k}|^2 \big] \leq C(\Delta t)^2. \qquad (3.53)
$$

*Proof.* Based on Lemmas 3.4 and 3.5, taking the conditional expectations on both sides of (3.15) and using the properties of conditional mathematical expectation and the Poisson process, we arrive at

$$
\mathbb{E}\big[ |E_{k-}|^2 \big] \leq C \mathbb{E} \left[ \exp(4 N_T K_1 \Delta t) \left( \prod_{j=1}^{n_T} \alpha_j \right) \sum_{i=0}^{n_T - 1} \mathcal{X}_1 (\Delta t)^3 \right.
$$

$$
\begin{aligned}
&+ C\mathbb{E}\left[\exp(4N_TK_1\Delta t)\left(\prod_{j=1}^{n_T}\alpha_j\right)\frac{1}{\Delta t}\sum_{i=1}^{n_T-1}\mathcal{X}_2(\Delta t)^4\right] \\
\le\, & C(\Delta t)^2\mathbb{E}\left[\exp(4N_TK_1\Delta t)(1+2C+2C\Delta t)^{N_T}(M+N_T)\Delta t\mathcal{X}_1\right] \\
&+ C(\Delta t)^2\mathbb{E}\left[\exp(4N_TK_1\Delta t)(1+2C+2C\Delta t)^{N_T}(M+N_T)\Delta t\mathcal{X}_2\right] \\
\le\, & C(\Delta t)^2\left(\mathbb{E}\left[\exp(4N_TK_1\Delta t)(1+2C+2C\Delta t)^{2N_T}(M+N_T)^2\Delta t^2\right]\right)^{\frac{1}{2}}\left(\mathbb{E}\left[\mathcal{X}_1^2\right]\right)^{\frac{1}{2}} \\
&+ C(\Delta t)^2\left(\mathbb{E}\left[\exp(8N_TK_1\Delta t)(1+2C+2C\Delta t)^{2N_T}(M+N_T)^2\Delta t^2\right]\right)^{\frac{1}{2}}\left(\mathbb{E}\left[\mathcal{X}_2^2\right]\right)^{\frac{1}{2}}.
\end{aligned}
$$

On the other hand, we first have

$$
\begin{aligned}
&\mathbb{E}\left[\exp(8N_TK_1\Delta t)(1+2C+2C\Delta t)^{2N_T}(M+N_T)^2\Delta t^2\right] \\
\le & \sum_{j=0}^{\infty}\exp(8jK_1T)(1+2C+2CT)^{2j}(1+j)^2T^2\frac{(\lambda T)^j}{j!}\exp(-\lambda T) < \infty.
\end{aligned} \tag{3.54}
$$

Using the properties of conditional mathematical expectation and Lemma 2.1 gives

$$
\begin{aligned}
\mathbb{E}\left[\mathcal{X}_1^2\right] \le\, & \mathbb{E}\Bigg[\Bigg(\mathbb{E}^N\Big[\sup_{0\le s\le T}|X_s|^{4+2\kappa}\Big] + \Big(1+\mathbb{E}^N\big[\sup_{0\le t\le T}|X_t|^4\big]\Big)^{\frac{1}{2}}\Phi_{N,4}^{\frac{1}{2}} \\
&+ \Big(1+\mathbb{E}^N\Big[\sup_{0\le t\le T}|X_t|^{4\kappa}\Big]\Big)^{\frac{1}{2}}\Phi_{N,4}^{\frac{1}{2}} + \Phi_{N,4} + \Phi_{N,2}\Bigg)^2\Bigg] \\
\le\, & C\Bigg(1+\mathbb{E}\Big[\sup_{0\le t\le T}|X_t|^{8+4\kappa}\Big] + \mathbb{E}\Big[\sup_{0\le t\le T}|X_t|^{8+8\kappa}\Big] + \mathbb{E}\Big[\sup_{0\le t\le T}|X_t|^{4+4\kappa}\Big] \\
&+ \mathbb{E}\Big[\sup_{0\le t\le T}|X_t|^{4}\Big] + \mathbb{E}\Big[\sup_{0\le t\le T}|X_t|^{8}\Big] + \mathbb{E}\Big[\sup_{0\le t\le T}|X_t|^{8\kappa}\Big]\Bigg) < \infty.
\end{aligned} \tag{3.55}
$$

Similarly, we can get $\mathbb{E}[\mathcal{X}_2^2] < \infty$. We thus have the estimate

$$
\mathbb{E}\left[|E_{k-}|^2\right] \le C(\Delta t)^2. \tag{3.56}
$$

Note that

$$
\begin{aligned}
|E_k| &= \left|Y_{t_{k-}} + h(Y_{t_{k-}})\Delta N_{k-1} - X_{t_{k-}} - h(X_{t_{k-}})\Delta N_{k-1}\right| \\
&\le \left|E_{k-}\right| + \left|\Delta h_{k-}^{Y,X}\right| \le C\left|E_{k-}\right|.
\end{aligned} \tag{3.57}
$$

This together with (3.56) finally completes the proof. □

## 4. Numerical Results

We present here a numerical study to illustrate strong convergence rates of the JAIMM. As test problems we first consider

$$
\mathrm{d}X_t = f\big(X_{t-}\big)\mathrm{d}t + g\big(X_{t-}\big)\mathrm{d}W_t + h\big(X_{t-}\big)\mathrm{d}N_t, \quad t\in(0,T], \quad X_0 = x_0, \tag{4.1}
$$

where $f(x) := x - x^{\kappa}$ with the integer $\kappa \ge 3$ being odd, $g(x) = \sigma x$, $h(x) = \mu x$, and $x_0$ is deterministic for simplicity. It is easy to check that conditions in Assumptions 2.1-2.2 are all fulfilled. In Fig. 4.1, we plot two one-path simulations of (4.1) for two sets of parameters.

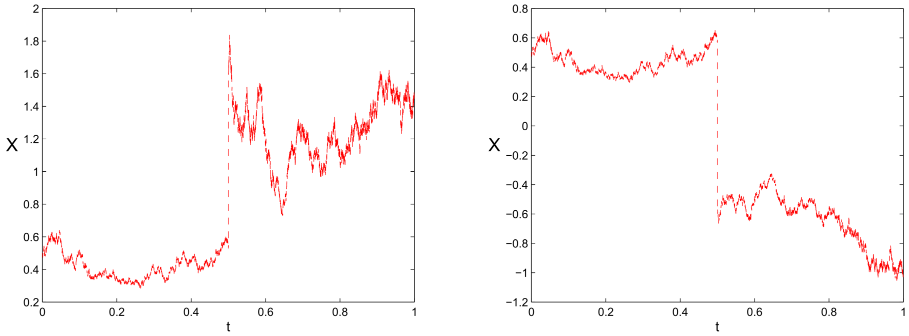


Fig. 4.1. One path simulation of exact solution to (4.1). Left: $\kappa = 3, \sigma = 1, \mu = 2$, and $T = 1, x_0 = 0.5, \lambda = 1$; Right: $\kappa = 5, \sigma = 1, \mu = -2$ and $T = 1, x_0 = 0.5, \lambda = 1$.

In the sequent experiments with (4.1) admitting the following parameters:

- $\kappa = 3, \quad \sigma = 2, \quad \mu = 1,$
- $\kappa = 5, \quad \sigma = 2, \quad \mu = 1,$
- $\kappa = 7, \quad \sigma = 2, \quad \mu = 1,$
- $\kappa = 9, \quad \sigma = 2, \quad \mu = 1,$

we are to test approximation errors in terms of means of absolute errors $\varepsilon = \mathbb{E}[|X_T - Y_T|]$. As usual, the expectation is approximated by the Monte-Carlo approximation

$$e := \hat{\mathbb{E}}[|X_T - Y_T|] := \frac{1}{N_{mc}} \sum_{i=1}^{N_{mc}} \left| X_T^{(i)} - Y_T^{(i)} \right|,$$

where the positive integer $N_{mc}$ is the sample times in numerical tests, $Y_T^{(i)}$ is the numerical approximation solution at the time $t_{n_T} = T$ by JAIMM at the $i$-th sampling. Note that the $\hat{\mathbb{E}}[|X_T - Y_T|]$ is the Monte-Carlo approximation of the mathematical expectation $\mathbb{E}[|X_T - Y_T|]$. We plot the achieved accuracy versus stepsizes in logarithmic scale. $N_{mc} = 5000$ Brownian and Poisson paths have been simulated with initial value $x_0 = 0.5$, $T = 1$, and $\lambda = 1$ as the intensity of Poisson process.

We list computational errors $e$ for various choices of parameters in Table 4.1, from which one can detect that the approximation errors decrease at a slope of order one as the stepsize $\Delta t$ decreases. To clearly display the convergence rates, we plot in Figs. 4.2-4.3 the achieved errors

Table 4.1: Numerical results for (4.1) with $\sigma = 2$, $\mu = 1$.

| $\Delta t$ | $\kappa = 3$ | $\kappa = 5$ | $\kappa = 7$ | $\kappa = 9$ |
|---|---|---|---|---|
| $2^{-7}$ | 0.0194 | 0.0292 | 0.0390 | 0.0486 |
| $2^{-8}$ | 0.0096 | 0.0142 | 0.0193 | 0.0243 |
| $2^{-9}$ | 0.0046 | 0.0069 | 0.0093 | 0.0119 |
| $2^{-10}$ | 0.0021 | 0.0032 | 0.0044 | 0.0056 |
| $2^{-11}$ | 0.0010 | 0.0015 | 0.0020 | 0.0026 |

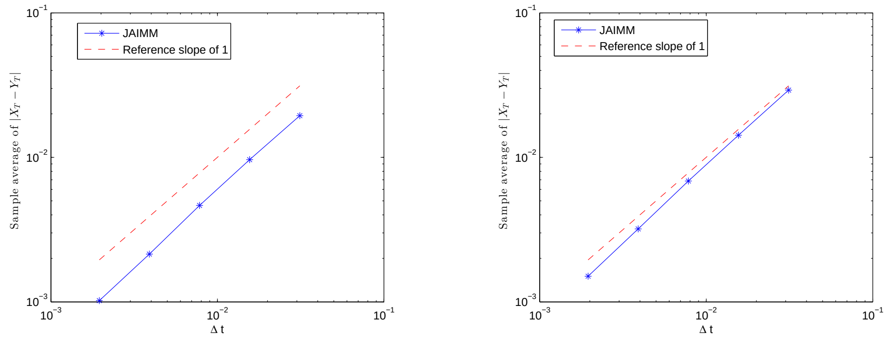


Fig. 4.2. Numerical results for (4.1) with $T = 1$. Computational errors versus stepsize $\Delta t$ on a log-log scale. Left: $\kappa = 3, \sigma = 2, \mu = 1$, and $\lambda = 1$; Right: $\kappa = 5, \sigma = 2, \mu = 1$, and $\lambda = 1$.

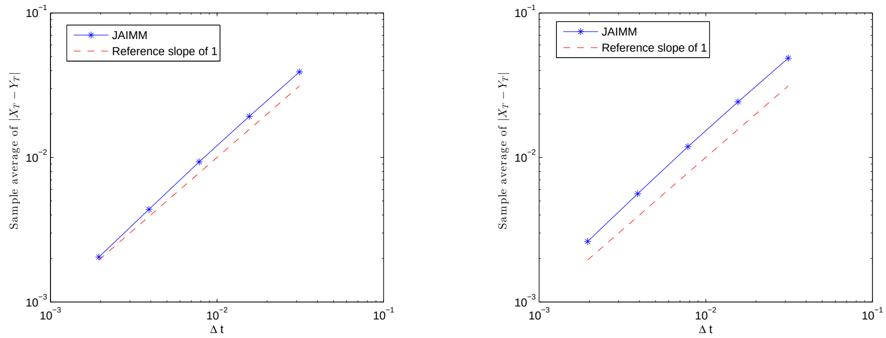


Fig. 4.3. Numerical results for (4.1) with $T = 1$. Computational errors versus stepsize $\Delta t$ on a log-log scale. Left: $\kappa = 7, \sigma = 2, \mu = 1$, and $\lambda = 1$; Right: $\kappa = 9, \sigma = 2, \mu = -1$, and $\lambda = 1$.

versus stepsizes in logarithmic scale. As predicted, the slopes of the errors (solid lines) and the reference dashed line match well, which indicates that the proposed scheme shows a strong convergence rate of order one.

Next, we consider a 2-dimensional nonlinear jump-diffusion SDEs as follows:

$$\mathrm{d}X_t = f\big(X_{t-}\big)\mathrm{d}t + g\big(X_{t-}\big)\mathrm{d}W_t + h\big(X_{t-}\big)\mathrm{d}N_t, \quad t \in (0, T], \quad X_0 = \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0.1 \end{bmatrix}, \tag{4.2}$$

where $f\colon \mathbb{R}^2 \to \mathbb{R}^2$, $g\colon \mathbb{R}^2 \to \mathbb{R}^{2\times 2}$, $h\colon \mathbb{R}^2 \to \mathbb{R}^2$ are given by

$$f(\mathbf{x}) = \begin{bmatrix} x^1 - (x^1)^5 \\ \frac{1}{2}\sin(x^1) + x^2 \end{bmatrix}, \quad g(\mathbf{x}) = \begin{bmatrix} \sigma_1 x^1 & 0 \\ 0 & \sigma_2 x^2 \end{bmatrix}, \quad h(\mathbf{x}) = \begin{bmatrix} \mu_1 x^1 \\ \mu_2 x^2 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x^1 \\ x^2 \end{bmatrix} \in \mathbb{R}^2.$$

Here $\sigma_1, \sigma_2, \mu_1, \mu_2$ are parameters that will be given in the following numerical tests. In Fig. 4.4, we plot two one-path simulations of (4.2) for two sets of parameters. Like the 1-dimensional cases, we plot in Figs. 4.5-4.6 the achieved errors versus stepsizes ($\Delta t = 2^{-i}, i = 7, 8, 9, 10, 11$) in logarithmic scale with four sets of parameters. As is shown, the slopes of the errors (solid lines) and the reference dashed line match well, which indicates that the proposed scheme admits a strong convergence rate of order one.

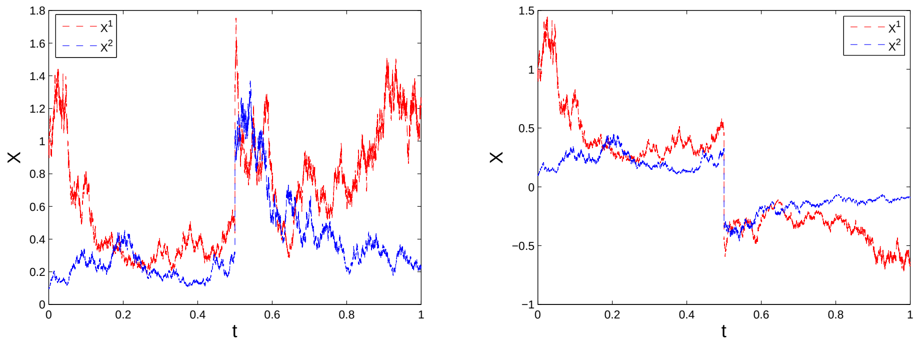


Fig. 4.4. One path simulation of exact solution to (4.2). Left: $\sigma_1 = 2, \sigma_2 = 4, \mu_1 = \mu_2 = 2$, and $T = 1, \lambda = 1$; Right: $\sigma_1 = 2, \sigma_2 = 4, \mu_1 = \mu_2 = -2$ and $T = 1, \lambda = 1$.

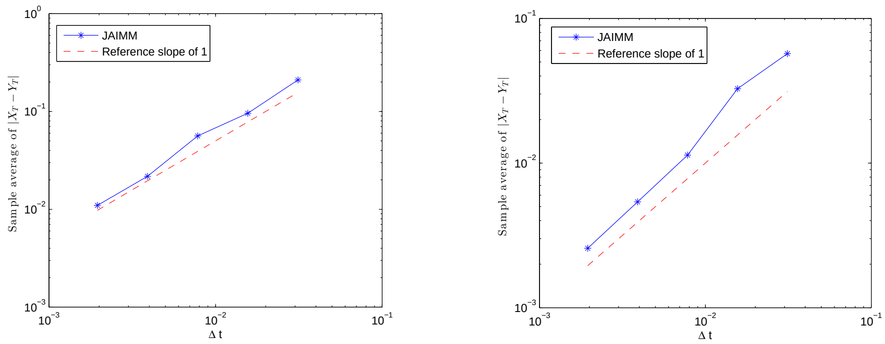


Fig. 4.5. Numerical results for (4.2) with $T = 1$. Computational errors versus stepsize $\Delta t$ on a log-log scale. Left: $\sigma_1 = 2, \sigma_2 = 4, \mu_1 = \mu_2 = 2$, and $\lambda = 1$; Right: $\sigma_1 = 2, \sigma_2 = 4, \mu_1 = \mu_2 = -2$, and $\lambda = 1$.

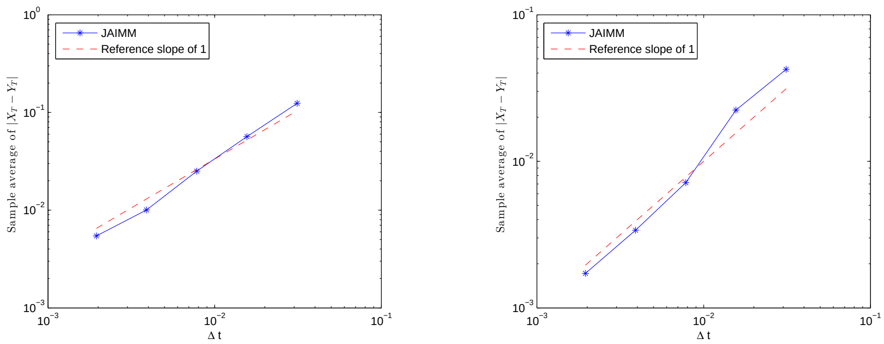


Fig. 4.6. Numerical results for (4.2) with $T = 1$. Computational errors versus stepsize $\Delta t$ on a log-log scale. Left: $\sigma_1 = 3, \sigma_2 = 5, \mu_1 = \mu_2 = 2$, and $\lambda = 1$; Right: $\sigma_1 = 3, \sigma_2 = 5, \mu_1 = \mu_2 = -2$, and $\lambda = 1$.

## 5. Conclusion Remarks

In this paper, strong convergence of a jump-adapted implicit Milstein method for a class of nonlinear jump-diffusion problems, of which the drift coefficients are one-sided Lipschitz continuous and the diffusion and jump coefficients are globally Lipschitz continuous, has been rigorously analysed. The optimal strong convergence rate of order one also has been recovered.

In the future, we are to carry out further numerical analysis of jump-adapted methods for more general models.

**Acknowledgments.** We would like to thank the referees for the detailed and helpful comments, which improve the paper a lot.

This research is partially supported by the National Natural Science Foundation of China (Grant Nos. 11901565, 12071261, 11831010, 11871068), by the Science Challenge Project (No. TZ2018001), and by National Key R&D Plan of China (Grant No. 2018YFA0703900).

## References

[1] M. Baran, Approximation for solutions of Lévy-type stochastic differential equations, *Stoch. Anal. Appl.*, **27** (2009), 924–961.

[2] N. Bruti-Liberati and E. Platen, Strong approximations of stochastic differential equations with jumps, *J. Comput. Appl. Math.*, **205** (2007), 982–1001.

[3] E. Buckwar and M.G. Riedler, Runge-Kutta methods for jump-diffusion differential equations, *J. Comput. Appl. Math.*, **236** (2011), 1155–1182.

[4] G.D. Chalmers and D.J. Higham, Asymptotic stability of a jump-diffusion equation and its numerical approximation, *SIAM J. Sci. Comput.*, **31** (2008/09), 1141–1155.

[5] F. Carbonell and J.C. Jimenez, Weak local linear discretizations for stochastic differential equations with jumps, *J. Appl. Probab.*, **45** (2008), 201–210.

[6] Y. Chen, A. Xiao, and W. Wang, Numerical solutions of SDEs with Markovian switching and jumps under non-Lipschitz conditions, *J. Comput. Appl. Math.*, **360** (2019), 41–54.

[7] Z. Chen and S. Gan, Convergence and stability of the backward Euler method for jump-diffusion SDEs with super-linearly growing diffusion and jump coefficients, *J. Comput. Appl. Math.*, **363** (2020), 350–369.

[8] Z. Chen, S. Gan, and X. Wang, Mean-square approximations of Lévy noise driven SDEs with super-linearly growing diffusion and jump coefficients, *Discrete Contin. Dyn. Syst. Ser. B*, **24** (2019), 4513–4545.

[9] R. Cont and P. Tankov, Financial modelling with jump processes, *Chapman & Hall/CRC Financial Mathematics Series*, Chapman & Hall/CRC, Boca Raton, FL, 2004.

[10] K. Dareiotis, C. Kumar, and S. Sabanis, On tamed Euler approximations of SDEs driven by Lévy noise with applications to delay equations, *SIAM J. Numer. Anal.*, **54** (2016), 1840–1872.

[11] A. Gardoń, The order of approximations for solutions of Itô-type stochastic differential equations with jumps, *Stochastic Anal. Appl.*, **22** (2004), 679–699.

[12] A. Ghasemifard and M. Tahmasebi, Multilevel path simulation to jump-diffusion process with superlinear drift, *Appl. Numer. Math.*, **144** (2019), 176–189.

[13] D.J. Higham and P.E. Kloeden, Numerical methods for nonlinear stochastic differential equations with jumps, *Numer. Math.*, **101** (2005), 101–119.

[14] D.J. Higham and P.E. Kloeden, Convergence and stability of implicit methods for jump-diffusion systems, *Int. J. Numer. Anal. Model.*, **3** (2006), 125–140.

[15] D.J. Higham and P.E. Kloeden, Strong convergence rates for backward Euler on a class of nonlinear jump-diffusion problems, *J. Comput. Appl. Math.*, **205** (2007), 949–956.

[16] D.J. Higham, X. Mao, and A. M. Stuart, Strong convergence of Euler-type methods for nonlinear stochastic differential equations, *SIAM J. Numer. Anal.*, **40** (2002), 1041–1063.

[17] L. Hu, S. Gan, and X. Wang, Asymptotic stability of balanced methods for stochastic jump-diffusion differential equations, *J. Comput. Appl. Math.*, **238** (2013), 126–143.

[18] A. Kał uża and P. Przybył owicz, Optimal global approximation of jump-diffusion SDEs via path-independent step-size control, *Appl. Numer. Math.*, **128** (2018), 24–42.

[19] K. Kashima and R. Kawai, An optimization approach to weak approximation of stochastic differential equations with jumps, *Appl. Numer. Math.*, **61** (2011), 641–650.

[20] C. Kumar, On Milstein-type scheme for SDE driven by Lévy noise with super-linear coefficients, *Discrete Contin. Dyn. Syst. Ser. B*, **26** (2021), 1405–1446.

[21] C. Kumar and S. Sabanis, On explicit approximations for Lévy driven SDEs with super-linear diffusion coefficients, *Electron. J. Probab.*, 22 , pp. Paper No. **73** (2017), 19.

[22] C. Kumar and S. Sabanis, On tamed Milstein schemes of SDEs driven by Lévy noise, *Discrete Contin. Dyn. Syst. Ser. B*, **22** (2017), 421–463.

[23] F. Kühn and R. L. Schilling, Strong convergence of the Euler-Maruyama approximation for a class of Lévy-driven SDEs, *Stochastic Process. Appl.*, **129** (2019), 2654–2680.

[24] A. Kohatsu-Higa and P. Tankov, Jump-adapted discretization schemes for Lévy-driven SDEs, *Stochastic Process. Appl.*, **120** (2010), 2258–2285.

[25] Q. Ma, D. Ding, and X. Ding, Mean-square dissipativity of several numerical methods for stochastic differential equations with jumps, *Appl. Numer. Math.*, **82** (2014), 44–50.

[26] Y. Maghsoodi, Mean square efficient numerical solution of jump-diffusion stochastic differential equations, *Sankhyā Ser. A*, **58** (1996), 25–47.

[27] W. Mao, S. You, and X. Mao, On the asymptotic stability and numerical analysis of solutions to nonlinear stochastic differential equations with jumps, *J. Comput. Appl. Math.*, **301** (2016), 1–15.

[28] R. Mikulevicius, On the rate of convergence of simple and jump-adapted weak Euler schemes for Lévy driven SDEs, *Stochastic Process. Appl.*, **122** (2012), 2730–2757.

[29] E. Mordecki, A. Szepessy, R. Tempone, and G. E. Zouraris, Adaptive weak approximation of diffusions with jumps, *SIAM J. Numer. Anal.*, **46** (2008), 1732–1768.

[30] A. Neuenkirch and L. Szpruch, First order strong approximations of scalar SDEs defined in a domain, *Numer. Math.*, **128** (2014), 103–136.

[31] E. Platen, An approximation method for a class of Itô processes with jump component, *Litovsk. Mat. Sb.*, **22** (1982), 124–136.

[32] E. Platen and N. Bruti-Liberati, Numerical solution of stochastic differential equations with jumps in finance, vol. 64 of Stochastic Modelling and Applied Probability, Springer-Verlag, Berlin, 2010.

[33] P. Przybył owicz, Optimal sampling design for global approximation of jump diffusion stochastic differential equations, *Stochastics*, **91** (2019), 235–264.

[34] Q. Ren and H. Tian, Compensated $\theta$-Milstein methods for stochastic differential equations with Poisson jumps, *Appl. Numer. Math.*, **150** (2020), 27–37.

[35] A. Tambue and J. D. Mukam, Strong convergence and stability of the semi-tamed and tamed Euler schemes for stochastic differential equations with jumps under non-global Lipschitz condition, *Int. J. Numer. Anal. Model.*, **16** (2019), 847–872.

[36] X. Wang, Mean-square convergence rates of implicit Milstein type methods for SDEs with non-Lipschitz coefficients, arXiv:2007.15733, *Adv. Comput. Math.*, (2023).

[37] X. Wang and S. Gan, Compensated stochastic theta methods for stochastic differential equations with jumps, *Appl. Numer. Math.*, **60** (2010), 877–887.

[38] X. Wang, J. Wu, and B. Dong, Mean-square convergence rates of stochastic theta methods for SDEs under a coupled monotonicity condition, *BIT*, **60** (2020), 759–790.

[39] F. Wu, X. Mao, and K. Chen, Strong convergence of Monte Carlo simulations of the mean-reverting square root process with jump, *Appl. Math. Comput.*, **206** (2008), 494–505.

[40] X. Yang and X. Wang, A transformed jump-adapted backward Euler method for jump-extended CIR and CEV models, *Numer. Algorithms*, **74** (2017), 39–57.

[41] X. Yang and W. Zhao, Strong convergence analysis of split-step $\theta$-scheme for nonlinear stochastic differential equations with jumps, *Adv. Appl. Math. Mech.*, **8** (2016), 1004–1022.