INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 23, Number 4, Pages 551–575

doi: 10.4208/ijnam2026-1024

A FIRST ORDER POSITIVITY-PRESERVING SCHEME FOR A CLASS OF NONLINEAR JUMP-DIFFUSION PROBLEMS

QING CHEN, XU YANG*, AND WEIDONG ZHAO

Abstract. In this study, we propose and analyze a first-order positivity-preserving numerical scheme, referred to as the jump-adapted drift-diffusion double implicit Milstein method, for solving a class of nonlinear jump-diffusion problems. The proposed method offers a more simple structure than the classic Milstein method for jump-diffusion problems, making it easier to implement on computers. By addressing the key challenges in strong convergence analysis stemming from non-Lipschitz continuity, weaker temporal regularity, and stochastic time partitioning, we rigorously prove that the proposed method achieves strong convergence with an optimal mean-square rate of order one. Moreover, under certain reasonable assumptions, we further demonstrate the positivity preservation of the proposed numerical method. Finally, numerical experiments are carried out to substantiate the validity of our theoretical results.

Key words. Non-globally Lipschitz condition, jump-adapted method, Poisson jumps, mean-square convergence, positivity-preserving.

1. Introduction

We consider numerical approximation of the following Itô type jump-diffusion stochastic differential equations (JSDEs):

$$\begin{cases} \mathrm{d}X_t = f(X_{t-})\mathrm{d}t + g(X_{t-})\mathrm{d}W_t + h(X_{t-})\mathrm{d}N_t, \quad t \in (0,T], \\ \\ X_0 = x_0, \end{cases} \tag{1}$$

where $T > 0$ is a fixed constant, $X_{t-} := \lim_{s\uparrow t} X_s$, $f\colon \mathbb{R}^m \to \mathbb{R}^m$ is the drift coefficient function, $g\colon \mathbb{R}^m \to \mathbb{R}^{m\times d}$ is the diffusion coefficient function, 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 $g_j\colon \mathbb{R}^m \to \mathbb{R}^m$, $j \in \{1, 2, \cdots, d\}$, and $h\colon \mathbb{R}^m \to \mathbb{R}^m$ is the jump coefficient function, 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 filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ with normal filtration $\mathbb{F} := \big\{\mathcal{F}_t : t \in [0,T]\big\}$, which 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.

Stochastic differential equations with jumps are widely used to model the interplay between continuous evolution and abrupt changes in intricate phenomena across various fields, such as finance, insurance, physics, biology, and engineering. We refer the reader to [2, 8, 10, 28] and the references therein for further detailed discussion about properties of JSDEs and their applications. Given the inherent challenges in deriving closed-form analytical solutions for JSDEs, it is of paramount importance and practical necessity to explore and develop numerical methods capable of addressing these equations. Typically, the discrete-time approximations for JSDEs are categorized into two types: regular methods and jump-adapted methods. Regular methods utilize time discretizations that exclude the jump times inherent

Received by the editors on June 11, 2025 and accepted on August 31, 2025.
2000 *Mathematics Subject Classification.* 60H35, 65C30, 60H10.
*Corresponding author.

in the Poisson process, whereas jump-adapted methods rely on jump-adapted time discretizations that are specifically tailored to incorporate all jump times. For the study of regular methods, a large amount of numerical methods have been developed and analyzed, see [1, 3, 4, 5, 6, 7, 9, 11, 12, 13, 15, 17, 18, 20, 21, 24, 23, 28, 29, 30, 31, 33, 34, 36] and the references therein.

As far as the jump-adapted methods are concerned, since jump-adapted time discretizations incorporate all jump times produced by the Poisson process, they effectively prevent the emergence of multiple stochastic integrals concerning the Poisson process. Consequently, the resultant schemes exhibit a considerably simpler form compared to regular methods, leading to a substantial reduction in the complexity of higher-order schemes. As a result, the jump-adapted time discretization renders these corresponding methods highly implementable for scenario simulation. Nevertheless, due to the simulation of the jump impact at the correct jump times in the jump-adapted scheme, jump-adapted schemes are more accurate than their regular counterparts [28]. Following the initial introduction of jump-adapted approximations in [27], numerous jump-adapted numerical methods have been developed and studied, as discussed in [19, 25, 26, 28] and the references therein. However, in the strong convergence analysis of the majority of the aforementioned jump-adapted numerical methods, a global Lipschitz assumption is frequently imposed on the coefficients of JSDEs. This assumption significantly narrows the applicability of these numerical methods. In fact, many real-world JSDEs have coefficients that are super-linear or sub-linear in growth, which clearly do not satisfy the global Lipschitz condition. This makes the existing error analysis and convergence results that based on global Lipschitz condition for numerical methods unreliable or inapplicable. In view of the advantages of jump-adapted numerical methods and the current state of research, in recent years, many researchers have shifted their attention to constructing and analyzing jump-adapted numerical methods in the non-global Lipschitz setting [22, 35, 37, 38]. A transformed jump-adapted backward Euler method for jump-extended CIR and CEV models is proposed in [35]. A transformed jump-adapted backward Euler method for the Ait-Sahalia model is proposed in [22]. In [37], under the non-global Lipschitz condition, the strong convergence of a jump-adapted backward Euler method for a family of jump-diffusion stochastic differential equations is established. In [38], the strong convergence of a jump-adapted drift-implicit Milstein method for a class of nonlinear jump-diffusion problems with non-global Lipschitz coefficients is investigated.

In this paper, we focus on developing high-order jump-adapted numerical methods for jump-diffusion stochastic differential equations with non-globally Lipschitz coefficients. Specifically, we propose a jump-adapted drift-diffusion double implicit Milstein method for JSDEs (1), assuming that the drift coefficient is one-sided Lipschitz continuous, and the diffusion and jump coefficients are globally Lipschitz continuous (see Assumption 2.1 in the next section). Compared with the classical Milstein method for JSDEs, our proposed jump-adapted drift-diffusion double implicit Milstein method avoids computing multiple stochastic integrals related to the Poisson process, resulting in a simpler formulation and easier implementation. However, due to the path-dependent nature of jump-adapted time discretization and non-uniform step sizes, the numerical analysis becomes more intricate, especially under non-globally Lipschitz conditions. Drawing inspiration from the work in [35], our strong error analysis adopts an indirect approach by refraining from direct evaluation of the approximation error $E_k := Y_{t_k} - X_{t_k}$. Instead, we first derive rigorous upper mean-square error bounds for the pre-jump discrepancy between intermediate solutions, denoted as $E_{k-} := Y_{t_{k-}} - X_{t_{k-}}$. Leveraging the inherent structure of jump-adapted temporal grids and advanced stochastic analytical techniques, we rigorously prove that the proposed jump-adapted drift-diffusion double implicit Milstein method converges strongly and successfully attain the anticipated optimal mean-square convergence rate of order one. Moreover, under some reasonable assumptions, we further demonstrate that the proposed numerical method possesses the advantage of positivity preservation, which is crucial in many practical applications.

The remainder of this article is organized as follows. In the next section, we present the necessary preliminary knowledge, including basic notations, fundamental assumptions, and the existence and uniqueness of solutions to the problems under consideration, as well as the proposed numerical method. In Section 3, Strong convergence and positivity preservation results of the proposed numerical method are rigorously established. In Section 4, we present numerical experiments to support our theoretical results. Finally, concluding remarks are provided in Section 5.

2. Preliminaries and the proposed numerical method

2.1. Preliminaries. We first introduce some notations and conventions. Let $T \in (0, \infty)$, $d, m \in \mathbb{N}^+$, where $\mathbb{N}^+$ represent the set of all positive integers. For any $a, b, c \in \mathbb{R}$, $a \vee b \vee c$ denotes the maximum of a, b and c. Let $|\cdot|$ and $\langle \cdot, \cdot \rangle$ denote the Euclidean norm and the inner product of vectors in $\mathbb{R}^m$, respectively. Accordingly, we write $|A| := \sqrt{\text{trace}(A^T A)}$ as the trace norm of a matrix $A \in \mathbb{R}^{m \times d}$. And $x \in \mathbb{R}^m$ is positive if and only if each component $x_i, i = 1, 2, \cdots, m$ is positive.

Let $\widetilde{\mathcal{F}}_t^W := \sigma(W_s, s \leq t)$ and $\widetilde{\mathcal{F}}_t^N := \sigma(N_s, s \leq t)$ represent the natural filtration generated by the Wiener process W_t and the Poisson process N_t, respectively. Denote by $\mathcal{N}$ the collection of all $\mathbb{P}$-null sets of $\mathcal{F}$. Define $\mathcal{F}_t^W := \sigma\big(\widetilde{\mathcal{F}}_t^W \bigcup \mathcal{N}\big)$, $\mathcal{F}_t^N := \sigma\big(\widetilde{\mathcal{F}}_t^N \bigcup \mathcal{N}\big)$, and $\mathcal{F}_t := \sigma\big(\mathcal{F}_t^W \bigcup \mathcal{F}_t^N \bigcup \mathcal{N}\big)$. Given the probability space $\big(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}\big)$ with $\mathbb{F} := \big\{\mathcal{F}_t : t \in [0, T]\big\}$, we use $\mathbb{E}^N[X] := \mathbb{E}[X|\mathcal{F}_T^N]$ to denote the conditional expectation and $\mathbb{E}^N\big[X|\mathcal{F}_t^W\big] := \mathbb{E}\Big[X|\sigma\big(\mathcal{F}_t^W \cup \mathcal{F}_T^N\big)\Big]$. And define $\|X\|_{L_N^r} := (\mathbb{E}^N[|X|^r])^{\frac{1}{r}}$ and $\|X\|_{L^r} := (\mathbb{E}[|X|^r])^{\frac{1}{r}}$.

To guarantee the well-posedness of the considered problem (1), we make some assumptions as follows.

Assumption 2.1. *The coefficient functions f, g and h from* (1) *are continuously differentiable, and there exist positive constants L_1, L_2, L_3 such that*

$$\langle x - y, f(x) - f(y) \rangle \leq L_1 |x - y|^2, \quad \forall x, y \in \mathbb{R}^m, \tag{2}$$

$$|g(x) - g(y)|^2 \leq L_2 |x - y|^2, \quad \forall x, y \in \mathbb{R}^m, \tag{3}$$

$$|h(x) - h(y)|^2 \leq L_3 |x - y|^2, \quad \forall x, y \in \mathbb{R}^m. \tag{4}$$

Remark 2.1. *From* (2)-(4)*, it is easy to get*

$$\langle x, f(x) \rangle \vee |g(x)|^2 \vee |h(x)|^2 \leq L(1 + |x|^2), \quad \forall x \in \mathbb{R}^m, \tag{5}$$

where $L = \max\big\{(L_1 + \frac{1}{2}), \frac{1}{2}|f(0)|^2, 2L_2, 2|g(0)|^2, 2L_3, 2|h(0)|^2\big\}$.

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*

$$\mathbb{E}\big[|x_0|^p\big] \leq K_p. \tag{6}$$

Lemma 2.1 ([11]). *Under Assumptions 2.1–2.2, the problem* (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*

$$\mathbb{E}\Big[\sup_{t\in[0,T]} |X_t|^p\Big] \leq K\Big(1 + \mathbb{E}\big[|x_0|^p\big]\Big). \tag{7}$$

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

$$\sup_{t\in[0,T]} \mathbb{E}\big[|X_t|^p\big] \leq \widetilde{K}\Big(1 + \mathbb{E}\big[|x_0|^p\big]\Big). \tag{8}$$

Following [14, 32], in the sequel, we impose some additional conditions on the drift and diffusion coefficients to facilitate our strong convergence analysis.

Assumption 2.3. *Suppose there is a constant* $L_4 > 0$ *and a positive constant* $\gamma \geq 1$ *such that*

$$\Big|\big(\tfrac{\partial f}{\partial x}(x) - \tfrac{\partial f}{\partial x}(y)\big)z\Big| \leq L_4\big(1 + |x| + |y|\big)^{\gamma-1}|x-y||z|, \quad \forall x, y, z \in \mathbb{R}^m. \tag{9}$$

Remark 2.2. *Assumption 2.3 implies that there is a constant* $L_5 > 0$ *such that*

$$\Big|\tfrac{\partial f}{\partial x}(x)z\Big| \leq L_5\big(1 + |x|\big)^{\gamma}|z|, \quad \forall x, z \in \mathbb{R}^m. \tag{10}$$

As a consequence, one also gets that there exists a constant $L_6 > 0$ *such that*

$$|f(x) - f(y)| \leq L_6\big(1 + |x| + |y|\big)^{\gamma}|x-y|, \quad \forall x, y \in \mathbb{R}^m, \tag{11}$$

which infers

$$|f(x)| \leq L_7\big(1 + |x|^{1+\gamma}\big), \quad \forall x \in \mathbb{R}^m, \tag{12}$$

where L_7 *is a positive constant.*

Assumption 2.4. *Suppose the diffusion coefficient* g *further fulfils that there exist constants* $L_8 > 0$, $L_9 > 0$, *and* $L_{10} > 0$ *such that*

$$\Big\langle x - y, \sum_{j=1}^{d} \mathcal{L}^j g_j(x) - \mathcal{L}^j g_j(y)\Big\rangle \geq 0, \quad \forall x, y \in \mathbb{R}^m, \tag{13}$$

$$\sum_{j_1,j_2=1}^{d} |\mathcal{L}^{j_1} g_{j_2}(x) - \mathcal{L}^{j_1} g_{j_2}(y)|^2 \leq L_8|x-y|^2, \quad \forall x, y \in \mathbb{R}^m, \tag{14}$$

$$\Big|\big(\tfrac{\partial g_j}{\partial x}(x) - \tfrac{\partial g_j}{\partial x}(y)\big)z\Big|^2 \leq L_9|x-y|^2|z|^2, \quad \forall x, y, z \in \mathbb{R}^m,\ j \in \{1, 2, \cdots, d\}, \tag{15}$$

$$\Big|\Big(\frac{\partial \mathcal{L}^j g_j}{\partial x}(x) - \frac{\partial \mathcal{L}^j g_j}{\partial x}(y)\Big)z\Big| \leq L_{10}|x-y||z|, \quad \forall x, y, z \in \mathbb{R}^m, \tag{16}$$

where

$$\mathcal{L}^{j_1} := \sum_{k=1}^{m} g_{k,j_1}\frac{\partial}{\partial x^k}, \quad j_1 \in \{1, 2, \cdots, d\}.$$

Remark 2.3. *It follows from* (14) *and* (15) *that*

$$\sum_{j=1}^{d} |\mathcal{L}^j g_j(x) - \mathcal{L}^j g_j(y)|^2 \leq L_{11}|x-y|^2, \quad \forall x, y \in \mathbb{R}^m, \tag{17}$$

$$\left|\frac{\partial g_j}{\partial x}(x)z\right|^2 \leq L_{12}(1+|x|^2)|z|^2, \quad \forall x, z \in \mathbb{R}^m,\ j \in \{1, 2, \cdots, d\}, \tag{18}$$

where L_{11} *and* L_{12} *are positive constants.*

2.2. Jump-adapted drift-diffusion double implicit Milstein method. In order to give the jump-adapted numerical method for (1), we first construct the jump-adapted time partition

$$\mathcal{T} = \big\{0 = t_0 \leq t_1 \leq ... \leq t_{n_T} = T\big\}, n_T = \max\Big\{n \in \big\{0, 1, ...\big\} : t_{n_T} \leq T\Big\} < \infty,$$

produced by a superposition of the jump times $\{\tau_1, \tau_2, ..., \tau_l\}$ ($l \in \mathbb{N}, l < \infty$) to a deterministic equidistant grid with time step-size $\Delta t = \frac{T}{M}$ with $M \in \mathbb{N}$, which means that the time mesh $\mathcal{T}$ is path-dependent and the maximum step-size of it equals Δt.

Note that on the jump-adapted mesh $\mathcal{T}$, $\big\{X_t\big\}_{t \in [0,T]}$ in (1) evolves in every interval $[t_k, t_{k+1}]$ for $\forall k = 0, 1, ..., n_T - 1$ as

$$X_{t_{k+1}-} = X_{t_k} + \int_{t_k}^{t_{k+1}} f(X_t)\mathrm{d}t + \int_{t_k}^{t_{k+1}} g(X_t)\mathrm{d}W_t, \tag{19}$$

$$X_{t_{k+1}} = X_{t_{k+1}-} + h(X_{t_{k+1}-})\Delta N_k, \tag{20}$$

where $\Delta N_k := N_{t_{k+1}} - N_{t_k}$.

Now, based on (19)–(20), the jump-adapted drift-diffusion double implicit Milstein method for (1) is proposed as follows:

Scheme 2.1. *Let* $Y_{t_0} = X_0 = x_0$ *and for* $k = 0, 1, 2, ..., n_T - 1$,

$$\begin{aligned} Y_{t_{k+1}-} =& Y_{t_k} + f\left(Y_{t_{k+1}-}\right)\Delta t_k + g\left(Y_k\right)\Delta W_k + \sum_{j_1,j_2=1}^{d} \mathcal{L}^{j_1} g_{j_2}(Y_{t_k}) I_{j_1,j_2}^{t_k,t_{k+1}} \\ &+ \frac{1}{2}\sum_{j=1}^{d} \mathcal{L}^j g_j(Y_{t_k})\Delta t_k - \frac{1}{2}\sum_{j=1}^{d} \mathcal{L}^j g_j(Y_{t_{k+1}-})\Delta t_k, \end{aligned} \tag{21}$$

$$Y_{t_{k+1}} = Y_{t_{k+1}-} + h(Y_{t_{k+1}-})\Delta N_k, \tag{22}$$

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

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

Remark 2.4. *Note that when the jump coefficient* $h \equiv 0$*, the jump-adapted drift-diffusion double implicit Milstein method reduced to the drift-diffusion double implicit Milstein method for SDEs without jumps, which has been well studied in* [14, 32].

Observe that the equation (21) in Scheme 2.1 is implicit, which poses the question of existence and uniqueness. The following lemma gives a positive answer.

Lemma 2.2. *Let Assumption 2.1 and 2.4 be fulfilled and let* $\Delta t \in \big(0, \frac{1}{2(1+L_1)}\big)$ *with* L_1 *from* (2)*, then* (21) *is uniquely solvable, which implies that Scheme 2.1 is well-defined.*

The proof is similar to that of Lemma 2.3 in [14], and we omit it here.

Remark 2.5. *Due to the jump-adapted structure, which eliminates the need for multiple stochastic integrals with respect to the Poisson process, Scheme 2.1 adopts a simpler form compared to the regular Milstein scheme, thereby facilitating its straightforward implementation in scenario simulation.*

Before closing this subsection, we make the convention that throughout the following analysis, by C we denote a generic positive constant that may change between occurrences but is independent of Δt.

3. Strong convergence and positivity preservation analysis

3.1. Strong convergence analysis. In this subsection, we are devoted to investigating the strong convergence of the proposed jump-adapted drift-diffusion double implicit Milstein method. For the convenience of discussion below, we define $E_k := X_{t_k} - Y_{t_k}$, $E_{k-} := X_{t_k-} - Y_{t_k-}$, $\Delta f_{k-}^{X,Y} := f(X_{t_{k-}}) - f(Y_{t_{k-}})$, $\Delta g_k^{X,Y} := g(X_{t_k}) - g(Y_{t_k})$, $\Delta h_{k-}^{X,Y} := h(X_{t_k-}) - h(Y_{t_k-})$, $\Delta(\mathcal{L}^{j_1} g_{j_2})_k^{X,Y} := \mathcal{L}^{j_1} g_{j_2}\big(X_{t_k})\big) - \mathcal{L}^{j_1} g_{j_2}\big(Y_{t_k}\big)$, $\Delta(\mathcal{L}^{j} g_{j})_k^{X,Y} := \mathcal{L}^{j} g_{j}\big(X_{t_k})\big) - \mathcal{L}^{j} g_{j}\big(Y_{t_k}\big)$ and $\Delta(\mathcal{L}^{j} g_{j})_{k-}^{X,Y} := \mathcal{L}^{j} g_{j}\big(X_{t_k-})\big) - \mathcal{L}^{j} g_{j}\big(Y_{t_k-}\big)$.

We begin by introducing a lemma that shows the local temporal regularity of the solution process X_t to the problem (1), which is crucial for our strong convergence analysis.

Lemma 3.1 ([38])**.** *Suppose that all assumptions used in Lemma 2.1 hold and X_t is the analytic solution of (1). Then for any fixed $k \in \{0, 1, 2, \ldots, n_T - 1\}$ and for $\forall s, t \in [t_k, t_{k+1})$, there exists a positive constant C such that for any $q \geq 2$*

$$\|X_t - X_s\|_{L_N^q}^q \leq C\Phi_{N,\gamma,q}|t-s|^{\frac{q}{2}},$$

where $\Phi_{N,\gamma,q} := 1 + \mathbb{E}^N\left[\sup_{0\leq t\leq T} |X_t|^{(1+\gamma)q}\right]$ *and γ comes from (9).*

Next, we are going to establish the error equation. To this end, by using (19)–(20), we first derive

$$\begin{aligned} X_{t_{k+1-}} &= X_{t_{k-}} + h(X_{t_{k-}})\Delta N_{k-1} + f(X_{t_{k+1-}})\Delta t_k + g(X_{t_k})\Delta W_k \\ &\quad + \sum_{j_1,j_2=1}^{d} \mathcal{L}^{j_1} g_{j_2}(X_{t_k}) I_{j_1,j_2}^{t_k,t_{k+1}} + \frac{1}{2}\sum_{j=1}^{d} \mathcal{L}^{j} g_{j}(X_{t_k})\Delta t_k \\ &\quad - \frac{1}{2}\sum_{j=1}^{d} \mathcal{L}^{j} g_{j}(X_{t_{k+1-}})\Delta t_k + M_{k+1}, \quad k = 0, 1, 2, ..., n_T - 1, \end{aligned} \tag{23}$$

where the remainder term M_{k+1} is defined as follows:

$$\begin{aligned} M_{k+1} := &\int_{t_k}^{t_{k+1}} [f(X_t) - f(X_{t_{k+1-}})]\mathrm{d}t + \int_{t_k}^{t_{k+1}} [g(X_t) - g(Y_{t_k})]\mathrm{d}W_t \\ &- \sum_{j_1,j_2=1}^{d} \mathcal{L}^{j_1} g_{j_2}(X_{t_k}) I_{j_1,j_2}^{t_k,t_{k+1}} - \frac{1}{2}\sum_{j=1}^{d} \mathcal{L}^{j} g_{j}(X_{t_k})\Delta t_k \\ &+ \frac{1}{2}\sum_{j=1}^{d} \mathcal{L}^{j} g_{j}(X_{t_{k+1-}})\Delta t_k. \end{aligned}$$

Furthermore, by substituting (22) into (21), we obtain the following recursive equation for $k = 0, 1, 2, ..., n_T - 1$,

$$\begin{aligned} Y_{t_{k+1-}} &= Y_{t_{k-}} + h(Y_{t_{k-}})\Delta N_{k-1} + f(Y_{t_{k+1-}})\Delta t_k + g(Y_{t_k})\Delta W_k \\ &\quad + \sum_{j_1,j_2=1}^{d} \mathcal{L}^{j_1} g_{j_2}(Y_{t_k}) I_{j_1,j_2}^{t_k,t_{k+1}} + \frac{1}{2}\sum_{j=1}^{d} \mathcal{L}^j g_j(Y_{t_k})\Delta t_k \\ &\quad - \frac{1}{2}\sum_{j=1}^{d} \mathcal{L}^j g_j(Y_{t_{k+1-}})\Delta t_k. \end{aligned} \tag{24}$$

Subtracting (24) from (23) and applying the previous defined notations, we obtain the error equation:

$$\begin{aligned} E_{k+1-} &= E_{k-} + \Delta h_{k-}^{X,Y} \Delta N_{k-1} + \Delta f_{k+1-}^{X,Y} \Delta t_k + \Delta g_k^{X,Y} \Delta W_k \\ &\quad + \sum_{j_1,j_2=1}^{d} \Delta(\mathcal{L}^{j_1} g_{j_2})_k^{X,Y} I_{j_1,j_2}^{t_k,t_{k+1}} + \frac{1}{2}\sum_{j=1}^{d} \Delta(\mathcal{L}^j g_j)_k^{X,Y}\Delta t_k \\ &\quad - \frac{1}{2}\sum_{j=1}^{d} \Delta(\mathcal{L}^j g_j)_{k+1-}^{X,Y}\Delta t_k + M_{k+1}. \end{aligned} \tag{25}$$

Now, we present a relationship between approximation error E_k and the left-hand limit error E_{k-}.

Lemma 3.2. *Suppose that conditions in Assumption 2.1 hold, then we have the following estimate*

$$|E_k|^2 \le \big(2 + 2L_3|\Delta N_{k-1}|^2\big)|E_{k-}|^2. \tag{26}$$

Proof. Subtract (22) from (20), square both sides of the resulting equation and use the globally Lipschitz condition (4) on h to get

$$\begin{aligned} |E_k|^2 &= |X_{t_{k-}} + h(X_{t_{k-}})\Delta N_{k-1} - Y_{t_{k-}} - h(Y_{t_{k-}})\Delta N_{k-1}|^2 \\ &\le 2|E_{k-}|^2 + 2|\Delta N_{k-1}|^2|\Delta h_{k-}^{Y,X}|^2 \\ &\le \big(2 + 2L_3|\Delta N_{k-1}|^2\big)|E_{k-}|^2. \end{aligned} \tag{27}$$

This completes the proof. ☐

Remark 3.1. *In the following, we do not directly measure the approximation error E_k. Instead, we first estimate E_{k-} and then derive the desired estimate for E_k by applying the aforementioned lemma.*

Lemma 3.3. *Let Assumptions 2.1–2.4 be fulfilled and $\Delta t \in (0, \frac{1}{1+L_8+4L_1}]$, then there exists a constant $C > 0$, independent of Δt, such that*

$$\begin{aligned} \|E_{k-}\|_{L_N^2}^2 &\le C\exp(2N_T\rho\Delta t)\Xi_T \sum_{i=0}^{n_T-1} \|M_{i+1}\|_{L_N^2}^2 \\ &\quad + C\exp(2N_T\rho\Delta t)\Xi_T \frac{1}{\Delta t}\sum_{i=0}^{n_T-1} \left\|E^N\left[M_{i+1}|\mathcal{F}_{t_i}^W\right]\right\|_{L_N^2}^2, \end{aligned} \tag{28}$$

where $\rho := 2L_1 + \frac{1}{2} + \frac{1}{2}L_8$ and $\Xi_T = \prod_{j=1}^{n_T}\eta_j$ with $\eta_i := 1 + C\Delta N_{i-1} + C\Delta t$.

Proof. It follows from (25)that

$$
\begin{aligned}
&E_{k+1-} - \Delta f^{X,Y}_{k+1-}\Delta t_k + \frac{1}{2}\sum_{j=1}^{d}\Delta(\mathcal{L}^j g_j)^{X,Y}_{k+1-}\Delta t_k \\
&= E_{k-} + \Delta h^{X,Y}_{k-}\Delta N_{k-1} + \Delta g^{X,Y}_{k}\Delta W_k \\
&\quad + \sum_{j_1,j_2=1}^{d}\Delta(\mathcal{L}^{j_1} g_{j_2})^{X,Y}_{k} I^{t_k,t_{k+1}}_{j_1,j_2} + \frac{1}{2}\sum_{j=1}^{d}\Delta(\mathcal{L}^j g_j)^{X,Y}_{k}\Delta t_k \\
&\quad + M_{k+1}.
\end{aligned} \tag{29}
$$

Squaring both sides of the above equation yields

$$
\begin{aligned}
&\Big|E_{k+1-} - \Delta f^{X,Y}_{k+1-}\Delta t_k + \frac{1}{2}\sum_{j=1}^{d}\Delta(\mathcal{L}^j g_j)^{X,Y}_{k+1-}\Delta t_k\Big|^2 \\
&= |E_{k-}|^2 + |\Delta h^{Y,X}_{k-}\Delta N_{k-1}|^2 + |\Delta g^{Y,X}_{k}\Delta W_k|^2 \\
&\quad + \Big|\sum_{j_1,j_2=1}^{d}\Delta(\mathcal{L}^{j_1} g_{j_2})^{Y,X}_{k} I^{t_k,t_{k+1}}_{j_1,j_2}\Big|^2 \\
&\quad + \Big|\frac{1}{2}\sum_{j=1}^{d}\Delta(\mathcal{L}^j g_j)^{Y,X}_{k}\Delta t_k\Big|^2 + |M_{k+1}|^2 \\
&\quad + \sum_{i=1}^{4}\mathcal{H}_i,
\end{aligned} \tag{30}
$$

where

$$
\begin{aligned}
\mathcal{H}_1 := {}& 2\big\langle E_{k-}, \Delta h^{Y,X}_{k-}\Delta N_{k-1}\big\rangle + 2\big\langle E_{k-}, \Delta g^{Y,X}_{k}\Delta W_k\big\rangle \\
&+ 2\Big\langle E_{k-}, \sum_{j_1,j_2=1}^{d}\Delta(\mathcal{L}^{j_1} g_{j_2})^{Y,X}_{k} I^{t_k,t_{k+1}}_{j_1,j_2}\Big\rangle \\
&+ \Big\langle E_{k-}, \sum_{j=1}^{d}\Delta(\mathcal{L}^j g_j)^{Y,X}_{k}\Delta t_k\Big\rangle \\
&+ 2\big\langle E_{k-}, M_{k+1}\big\rangle, \\
\mathcal{H}_2 := {}& 2\big\langle h^{Y,X}_{k-}\Delta N_{k-1}, \Delta g^{Y,X}_{k}\Delta W_k\big\rangle + 2\big\langle \Delta h^{Y,X}_{k-}\Delta N_{k-1}, M_{k+1}\big\rangle \\
&+ 2\Big\langle h^{Y,X}_{k-}\Delta N_{k-1}, \sum_{j_1,j_2=1}^{d}\Delta(\mathcal{L}^{j_1} g_{j_2})^{Y,X}_{k} I^{t_k,t_{k+1}}_{j_1,j_2}\Big\rangle \\
&+ \Big\langle h^{Y,X}_{k-}\Delta N_{k-1}, \sum_{j=1}^{d}\Delta(\mathcal{L}^j g_j)^{Y,X}_{k}\Delta t_k\Big\rangle, \\
\mathcal{H}_3 := {}& 2\Big\langle \Delta g^{Y,X}_{k}\Delta W_k, \sum_{j_1,j_2=1}^{d}\Delta(\mathcal{L}^{j_1} g_{j_2})^{Y,X}_{k} I^{t_k,t_{k+1}}_{j_1,j_2}\Big\rangle \\
&+ \Big\langle \Delta g^{Y,X}_{k}\Delta W_k, \sum_{j=1}^{d}\Delta(\mathcal{L}^j g_j)^{Y,X}_{k}\Delta t_k\Big\rangle + 2\big\langle \Delta g^{Y,X}_{k}\Delta W_k, M_{k+1}\big\rangle,
\end{aligned}
$$

and

$$
\begin{aligned}
\mathcal{H}_4 :=& \Big\langle \sum_{j_1,j_2=1}^{d} \Delta(\mathcal{L}^{j_1} g_{j_2})_k^{Y,X} I_{j_1,j_2}^{t_k,t_{k+1}}, \sum_{j=1}^{d} \Delta(\mathcal{L}^{j} g_{j})_k^{Y,X} \Delta t_k \Big\rangle \\
&+ 2\Big\langle \sum_{j_1,j_2=1}^{d} \Delta(\mathcal{L}^{j_1} g_{j_2})_k^{Y,X} I_{j_1,j_2}^{t_k,t_{k+1}}, M_{k+1} \Big\rangle \\
&+ \Big\langle \sum_{j=1}^{d} \Delta(\mathcal{L}^{j} g_{j})_k^{Y,X} \Delta t_k, M_{k+1} \Big\rangle.
\end{aligned}
$$

By taking the conditional expectation on both sides of (30), we arrive at

$$
\begin{aligned}
&\Big\| E_{k+1-} - \Delta f_{k+1-}^{X,Y} \Delta t_k + \frac{1}{2} \sum_{j=1}^{d} \Delta(\mathcal{L}^{j} g_{j})_{k+1-}^{X,Y} \Delta t_k \Big\|_{L_N^2}^2 \\
&= \|E_{k-}\|_{L_N^2}^2 + \|\Delta h_{k-}^{Y,X} \Delta N_{k-1}\|_{L_N^2}^2 + \|\Delta g_k^{Y,X} \Delta W_k\|_{L_N^2}^2 \\
&\quad + \Big\| \sum_{j_1,j_2=1}^{d} \Delta(\mathcal{L}^{j_1} g_{j_2})_k^{Y,X} I_{j_1,j_2}^{t_k,t_{k+1}} \Big\|_{L_N^2}^2 + \Big\| \frac{1}{2} \sum_{j=1}^{d} \Delta(\mathcal{L}^{j} g_{j})_k^{Y,X} \Delta t_k \Big\|_{L_N^2}^2 \\
&\quad + \|M_{k+1}\|_{L_N^2}^2 + \sum_{i=1}^{4} \mathbb{E}^N\big[\mathcal{H}_i\big],
\end{aligned}
$$

from which we deduce

$$
\begin{aligned}
\|E_{k+1-}\|_{L_N^2}^2 \le& \|E_{k-}\|_{L_N^2}^2 + \|\Delta h_{k-}^{Y,X} \Delta N_{k-1}\|_{L_N^2}^2 \\
&+ \|\Delta g_k^{Y,X} \Delta W_k\|_{L_N^2}^2 + \Big\| \sum_{j_1,j_2=1}^{d} \Delta(\mathcal{L}^{j_1} g_{j_2})_k^{Y,X} I_{j_1,j_2}^{t_k,t_{k+1}} \Big\|_{L_N^2}^2 \\
&+ \Big\| \frac{1}{2} \sum_{j=1}^{d} \Delta(\mathcal{L}^{j} g_{j})_k^{Y,X} \Delta t_k \Big\|_{L_N^2}^2 + \|M_{k+1}\|_{L_N^2}^2 \\
&+ 2\mathbb{E}^N\Big[\big\langle E_{k+1-}, \Delta f_{k+1-}^{X,Y} \Delta t_k \big\rangle\Big] \\
&- \mathbb{E}^N\Big[\Big\langle E_{k+1-}, \sum_{j=1}^{d} \Delta(\mathcal{L}^{j_1} g_{j_2})_{k+1-}^{X,Y} \Delta t_k \Big\rangle\Big] \\
&+ \sum_{i=1}^{4} \mathbb{E}^N\big[\mathcal{H}_i\big].
\end{aligned}
\tag{31}
$$

By adopting a similar methodology presented in [32, 38] and using the properties of conditional expectation, Wiener process, Poisson process, and stochastic integral with respect to Wiener process, along with the independence of Wiener process and

Poisson process, it then follows from (31) that

$$
\begin{aligned}
\|E_{k+1-}\|^2_{L^2_N} \le{}& \|E_{k-}\|^2_{L^2_N} + |\Delta N_{k-1}|^2 \|\Delta h^{Y,X}_{k-}\|^2_{L^2_N} \\
&+ \Delta t \|\Delta g^{Y,X}_k\|^2_{L^2_N} + \frac{1}{2}(\Delta t)^2 \mathbb{E}^N\Bigg[\sum_{j_1,j_2=1}^d \big|\Delta(\mathcal{L}^{j_1} g_{j_2})^{Y,X}_k\big|^2\Bigg] \\
&+ (\Delta t)^2 \mathbb{E}^N\Bigg[\sum_{j=1}^d \big|\Delta(\mathcal{L}^{j} g_{j})^{Y,X}_k\big|^2\Bigg] + \|M_{k+1}\|^2_{L^2_N} \\
&+ 2\Delta t_k \mathbb{E}^N\Big[\big\langle E_{k+1-}, \Delta f^{X,Y}_{k+1-}\big\rangle\Big] \\
&- \Delta t_k \mathbb{E}^N\Bigg[\Bigg\langle E_{k+1-}, \sum_{j=1}^d \Delta(\mathcal{L}^{j_1} g_{j_2})^{X,Y}_{k+1-}\Bigg\rangle\Bigg] \\
&+ 2\Delta N_{k-1} \mathbb{E}^N\big[\big\langle E_{k-}, \Delta h^{Y,X}_{k-}\big\rangle\big] + \Delta t_k \mathbb{E}^N\Bigg[\Bigg\langle E_{k-}, \sum_{j=1}^d \Delta(\mathcal{L}^{j} g_{j})^{Y,X}_k\Bigg\rangle\Bigg] \\
&+ 2\mathbb{E}^N\big[\langle E_{k-}, \mathbb{E}^N\big[M_{k+1}|\mathcal{F}^W_{t_k}\big]\rangle\big] \\
&+ \Delta t_k \Delta N_{k-1} \mathbb{E}^N\Bigg[\Bigg\langle \Delta h^{Y,X}_{k-}, \sum_{j=1}^d \Delta(\mathcal{L}^{j} g_{j})^{Y,X}_k\Bigg\rangle\Bigg] \\
&+ 2\Delta N_{k-1} \mathbb{E}^N\big[\big\langle \Delta h^{Y,X}_{k-}, M_{k+1}\big\rangle\big] + 2\mathbb{E}^N\big[\big\langle \Delta g^{Y,X}_k \Delta W_k, M_{k+1}\big\rangle\big] \\
&+ 2\mathbb{E}^N\Bigg[\Bigg\langle \sum_{j_1,j_2=1}^d \Delta(\mathcal{L}^{j_1} g_{j_2})^{Y,X}_k I^{t_k,t_{k+1}}_{j_1,j_2}, M_{k+1}\Bigg\rangle\Bigg] \\
&+ \mathbb{E}^N\Bigg[\Bigg\langle \sum_{j=1}^d \Delta(\mathcal{L}^{j} g_{j})^{Y,X}_k \Delta t_k, M_{k+1}\Bigg\rangle\Bigg].
\end{aligned}
$$

Utilizing the Cauchy-Schwarz inequality and the Young inequality, we further derive

$$
\begin{aligned}
\|E_{k+1-}\|^2_{L^2_N} \le{}& (1 + \Delta N_{k-1} + 2\Delta t)\|E_{k-}\|^2_{L^2_N} \\
&+ (2\Delta N_{k-1} + \Delta t \Delta N_{k-1} + |\Delta N_{k-1}|^2)\|\Delta h^{Y,X}_{k-}\|^2_{L^2_N} \\
&+ 2\Delta t \|\Delta g^{Y,X}_k\|^2_{L^2_N} + (\Delta t)^2 \mathbb{E}^N\Bigg[\sum_{j_1,j_2=1}^d \big|\Delta(\mathcal{L}^{j_1} g_{j_2})^{Y,X}_k\big|^2\Bigg] \\
&+ (\frac{1}{2}\Delta t + \frac{1}{2}\Delta t \Delta N_{k-1} + \frac{3}{2}(\Delta t)^2)\mathbb{E}^N\Bigg[\sum_{j=1}^d \big|\Delta(\mathcal{L}^{j} g_{j})^{Y,X}_k\big|^2\Bigg] \\
&+ 2\mathbb{E}^N\Big[\big\langle E_{k+1-}, \Delta f^{X,Y}_{k+1-}\Delta t_k\big\rangle\Big] + \frac{1}{2}\Delta t \|E_{k+1-}\|^2_{L^2_N} \\
&+ \frac{1}{2}\Delta t \mathbb{E}^N\Bigg[\sum_{j=1}^d \big|\Delta(\mathcal{L}^{j} g_{j})^{X,Y}_{k+1-}\big|^2\Bigg] + (C + C\Delta N_{k-1})\|M_{k+1}\|^2_{L^2_N} \\
&+ \frac{1}{\Delta t}\left\|E^N\big[M_{k+1}|\mathcal{F}^W_{t_k}\big]\right\|^2_{L^2_N}.
\end{aligned}
$$

Now, with the help of Assumptions 2.1 and 2.4, we get

$$\begin{aligned}
\|E_{k+1-}\|^2_{L^2_N} \leq{}& (1+\Delta N_{k-1}+2\Delta t)\|E_{k-}\|^2_{L^2_N}\\
&+ (2\Delta N_{k-1}+\Delta t\Delta N_{k-1}+|\Delta N_{k-1}|^2)L_3\|E_{k-}\|^2_{L^2_N}+2L_2\Delta t\|E_k|^2_{L^2_N}\\
&+ L_8(\Delta t)^2\|E_k|^2_{L^2_N}+L_8(\Delta t+\frac{1}{2}\Delta t\Delta N_{k-1}+(\Delta t)^2)\|E_k\|^2_{L^2_N}\\
&+ 2L_1\Delta t\|E_{k+1-}\|^2_{L^2_N}+\frac{1}{2}\Delta t\|E_{k+1-}\|^2_{L^2_N}\\
&+ \frac{1}{2}L_8\Delta t\|E_{k+1-}\|^2_{L^2_N}+(C+C\Delta N_{k-1})\|M_{k+1}\|^2_{L^2_N}\\
&+ \frac{1}{\Delta t}\left\|E^N\left[M_{k+1}|\mathcal{F}^W_{t_k}\right]\right\|^2_{L^2_N}.
\end{aligned}$$

This together with Lemma 3.2 further ensures that

$$\begin{aligned}
&\Big(1-\frac{4L_1+L_8+1}{2}\Delta t\Big)\|E_{k+1-}\|^2_{L^2_N}\\
&\leq (1+C\Delta N_{k-1}+C\Delta t+C|\Delta N_{k-1}|^2+C\Delta t\Delta N_{k-1})\|E_{k-}\|^2_{L^2_N}\\
&\quad+(C+C\Delta N_{k-1})\|M_{k+1}\|^2_{L^2_N}+\frac{1}{\Delta t}\mathbb{E}^N\left[|\mathbb{E}^N\left[M_{k+1}|\mathcal{F}^W_{t_k}\right]|^2\right]\\
&\leq (1+C\Delta N_{k-1}+C\Delta t)\|E_{k-}\|^2_{L^2_N}+(C+C\Delta N_{k-1})\|M_{k+1}\|^2_{L^2_N}\\
&\quad+\frac{1}{\Delta t}\left\|E^N\left[M_{k+1}|\mathcal{F}^W_{t_k}\right]\right\|^2_{L^2_N}.
\end{aligned}$$

Define $\eta_k := 1+C\Delta N_{k-1}+C\Delta t$ and $\nu_k := C+C\Delta N_{k-1}$, then we have

$$\begin{aligned}
&\Big(1-\frac{4L_1+L_8+1}{2}\Delta t\Big)\|E_{k+1-}\|^2_{L^2_N}\\
&\leq \eta_k\|E_{k-}\|^2_{L^2_N}+\nu_k\|M_{k+1}\|^2_{L^2_N}\\
&\quad+\frac{1}{\Delta t}\left\|E^N\left[M_{k+1}|\mathcal{F}^W_{t_k}\right]\right\|^2_{L^2_N}.
\end{aligned}$$

Noting that $\Gamma_\Delta := 1-\frac{4L_1+L_8+1}{2}\Delta t>0$ ensures

$$\begin{aligned}
\|E_{k+1-}\|^2_{L^2_N} \leq{}& \Gamma_\Delta^{-1}\eta_k\|E_{k-}\|^2_{L^2_N}+\Gamma_\Delta^{-1}\nu_k\|M_{k+1}\|^2_{L^2_N}\\
&+\frac{\Gamma_\Delta^{-1}}{\Delta t}\left\|E^N\left[M_{k+1}|\mathcal{F}^W_{t_k}\right]\right\|^2_{L^2_N}.
\end{aligned}\tag{32}$$

Based on (32) and by continuous iteration, we deduce for all $k=1,2,...,n_T-1$

$$\begin{aligned}
\|E_{k+1-}\|^2_{L^2_N} \leq{}& \Big(\prod_{i=2}^{k+1}\Gamma_\Delta^{-1}\Big)\Big(\prod_{j=1}^{k}\eta_j\Big)\|E_{1-}\|^2_{L^2_N}\\
&+\sum_{i=1}^{k}\Big(\prod_{l=i+1}^{k+1}\Gamma_\Delta^{-1}\Big)\Big(\prod_{j=i+1}^{k}\eta_j\Big)\nu_i\mathbb{E}^N[|M_{i+1}|^2]\\
&+\sum_{i=1}^{k}\Big(\prod_{l=i+1}^{k+1}\Gamma_\Delta^{-1}\Big)\Big(\prod_{j=i+1}^{k}\eta_j\Big)\frac{1}{\Delta t}\left\|E^N\left[M_{i+1}|\mathcal{F}^W_{t_i}\right]\right\|^2_{L^2_N}.
\end{aligned}\tag{33}$$

By (25), we derive

$$E_{1-}=\Delta f_1^{Y,X}\Delta t_0-\frac{1}{2}\sum_{i=1}^{d}\Delta(\mathcal{L}^jg_j)_{1-}^{Y,X}\Delta t_0-M_1,$$

which leads to

$$\begin{aligned}|E_{1-}|^2 &= \langle \Delta f_1^{Y,X}\Delta t_0, E_{1-}\rangle - \langle \frac{1}{2}\sum_{i=1}^{d}\Delta(\mathcal{L}^j g_j)_{1-}^{Y,X}\Delta t_0, E_{1-}\rangle - \langle M_1, E_{1-}\rangle \\ &\leq L_1\Delta t|E_{1-}|^2 + \frac{1}{4}\Delta t|E_{1-}|^2 + \frac{1}{4}L_8\Delta t|E_{1-}|^2 + \frac{1}{2}|E_{1-}|^2 + \frac{1}{2}|M_1|^2.\end{aligned}$$

Then it follows that

$$|E_{1-}|^2 \leq \Gamma_\Delta^{-1}|M_1|^2. \tag{34}$$

Noting that $\nu_i \leq 2C, i = 0, 1, 2, ..., n_T$, we plug (34) into (33) to get

$$\begin{aligned}\|E_{k+1-}\|_{L_N^2}^2 \leq& \Big(\prod_{i=2}^{k+1}\Gamma_\Delta^{-1}\Big)\Big(\prod_{j=1}^{k}\eta_j\Big)\Gamma_\Delta^{-1}\|M_1\|_{L_N^2}^2 \\ &+\sum_{i=1}^{k}\Big(\prod_{l=i+1}^{k+1}\Gamma_\Delta^{-1}\Big)\Big(\prod_{j=i+1}^{k}\eta_j\Big)\nu_i\|M_{i+1}\|_{L_N^2}^2 \\ &+\sum_{i=1}^{k}\Big(\prod_{l=i+1}^{k+1}\Gamma_\Delta^{-1}\Big)\Big(\prod_{j=i+1}^{k}\eta_j\Big)\frac{1}{\Delta t}\mathbb{E}^N[|\mathbb{E}^N[M_{i+1}|\mathcal{F}_{t_i}^W]|^2] \\ \leq& C\sum_{i=0}^{k}\Big(\prod_{l=i+1}^{k+1}\Gamma_\Delta^{-1}\Big)\Big(\prod_{j=i+1}^{k}\eta_j\Big)\|M_{i+1}\|_{L_N^2}^2 \\ &+\sum_{i=1}^{k}\Big(\prod_{l=i+1}^{k+1}\Gamma_\Delta^{-1}\Big)\Big(\prod_{j=i+1}^{k}\eta_j\Big)\frac{1}{\Delta t}\mathbb{E}^N[|\mathbb{E}^N[M_{i+1}|\mathcal{F}_{t_i}^W]|^2].\end{aligned}$$

Then, by defining $\Psi_T := \prod_{i=1}^{n_T}\Gamma_\Delta^{-1}$ and $\Xi_T := \prod_{j=1}^{n_T}\eta_j$, we obtain for $k = 0, 1, 2, ..., n_T$

$$\begin{aligned}\|E_{k-}\|_{L_N^2}^2 \leq& C\sum_{i=0}^{n_T-1}\Psi_T\Xi_T\|M_{i+1}\|_{L_N^2}^2 \\ &+\sum_{i=1}^{n_T-1}\Psi_T\Xi_T\frac{1}{\Delta t}\left\|E^N\left[M_{i+1}|\mathcal{F}_{t_i}^W\right]\right\|_{L_N^2}^2.\end{aligned}\tag{35}$$

Define $\rho := 2L_1 + \frac{1}{2} + \frac{1}{2}L_8$. Note that $\Delta t \in (0, \frac{1}{1+L_8+4L_1}]$, then $0 < \rho\Delta t \leq \frac{1}{2}$. We now employ the inequality $\frac{1}{1-x} \leq \exp(2x)$ for $x \in (0, \frac{1}{2}]$ to infer that

$$\begin{aligned}\Psi_T = (1-\rho\Delta t)^{-n_T} &\leq \exp(2n_T\rho\Delta t) \leq \exp[(M+N_T)(2\rho\Delta t)] \\ &\leq \exp(2\rho T)\exp(2N_T\rho\Delta t) \leq C\exp(2N_T\rho\Delta t).\end{aligned}\tag{36}$$

Substituting (36) into (35) concludes the proof. □

Remark 3.2. *Drawing from the above lemma, to establish a clear and explicit convergence rate, it is sufficient to obtain estimates for the remainder terms. Therefore, moving forward, our focus will be on estimating* $\|M_{k+1}\|_{L_N^2}^2$ *and* $\left\|E^N\left[M_{k+1}|\mathcal{F}_{t_k}^W\right]\right\|_{L_N^2}^2$ *separately.*

Lemma 3.4. *Let Assumptions 2.1–2.4 hold. Then it holds for* $k \in \{0, 1, 2, ..., n_T\}$ *that*

$$\|M_{k+1}\|_{L_N^2}^2 \leq C\Phi_{N,\gamma,4}(\Delta t)^3,$$

where the definition of $\Phi_{N,\gamma,4}$ *comes from Lemma 3.1.*

Proof. Recalling the definition of M_{k+1} in (23) and applying the elementary inequality $(a+b+c)^3 \leq 3a^2+3b^2+3c^2$, we obtain

(37)

$$\begin{aligned}\|M_{k+1}\|^2_{L^2_N} &\leq 3\left\|\int_{t_k}^{t_{k+1}} \left[f(X_t) - f(X_{t_{k+1}-})\right] \mathrm{d}t\right\|^2_{L^2_N} \\ &\quad + 3\left\|\int_{t_k}^{t_{k+1}} [g(X(t)) - g(X_{t_k})]\,\mathrm{d}W_t - \sum_{j_1,j_2=1}^{d} \mathcal{L}^{j_1} g_{j_2}(X_{t_k}) I^{t_k,t_{k+1}}_{j_1,j_2}\right\|^2_{L^2_N} \\ &\quad + 3\left\|-\frac{1}{2}\sum_{j=1}^{d} \mathcal{L}^j g_j(X_{t_k})\Delta t_k + \frac{1}{2}\sum_{j=1}^{d} \mathcal{L}^j g_j(X_{t_{k+1}-})\Delta t_k\right\|^2_{L^2_N} \\ &:= \mathbb{I}_1 + \mathbb{I}_2 + \mathbb{I}_3.\end{aligned}$$

Subsequently, we deal with the above three items one by one. For the term $\mathbb{I}_1$, by applying the Cauchy–Schwarz inequality, the polynomial growth condition (11) on f, and Lemma 3.1, we have

$$\begin{aligned}\mathbb{I}_1 &\leq C\Delta t \mathbb{E}^N\left[\int_{t_k}^{t_{k+1}} (1 + |X_t|^{2\gamma} + |X_{t_{k+1}-}|^{2\gamma})|X_t - X_{t_{k+1}-}|^2 \mathrm{d}t\right] \\ &\leq C\Delta t \mathbb{E}^N\left[\sup_{0\leq t\leq T}(1+|X_t|^{2\gamma})\int_{t_k}^{t_{k+1}} |X_t - X_{t_{k+1}-}|^2 \mathrm{d}t\right] \\ &\leq C\Delta t\left(1 + \mathbb{E}^N\left[\sup_{0\leq t\leq T}|X_t|^{4\gamma}\right]\right)^{\frac{1}{2}}\left\|\int_{t_k}^{t_{k+1}} |X_t - X_{t_{k+1}-}|^2 \mathrm{d}t\right\|_{L^2_N} \\ &\leq C(\Delta t)^{\frac{3}{2}}\left(1 + \mathbb{E}^N\left[\sup_{0\leq t\leq T}|X_t|^{4\gamma}\right]\right)^{\frac{1}{2}}\left(\int_{t_k}^{t_{k+1}} \left\|X_t - X_{t_{k+1}-}\right\|^4_{L^4_N}\,\mathrm{d}t\right)^{\frac{1}{2}} \\ &\leq C\Phi_{N,\gamma,4}(\Delta t)^3.\end{aligned} \tag{38}$$

Now, we turn to bound $\mathbb{I}_2$. First, by virtue of Taylor's formula, we deduce that

$$\begin{aligned}&\int_{t_k}^{t_{k+1}} [g(X_t) - g(X_{t_k})]\,\mathrm{d}W_t - \sum_{j_1,j_2=1}^{d} \mathcal{L}^{j_1} g_{j_2}(X_{t_k}) I^{t_k,t_{k+1}}_{j_1,j_2} \\ &= \sum_{j_2=1}^{d}\int_{t_k}^{t_{k+1}} \Big[g_{j_2}(X_t) - g_{j_2}(X_{t_k}) - \sum_{j_1=1}^{d} \mathcal{L}^{j_1} g_{j_2}(X_{t_k})(W^{j_1}_t - W^{j_1}_{t_k})\Big]\mathrm{d}W^{j_2}_t \\ &= \sum_{j_2=1}^{d}\int_{t_k}^{t_{k+1}} \Big[g_{j_2}(X_t) - g_{j_2}(X_{t_k}) - \sum_{j_1=1}^{d} \frac{\partial g_{j_2}}{\partial x}(X_{t_k}) g_{j_1}(X_{t_k})(W^{j_1}_t - W^{j_1}_{t_k})\Big]\mathrm{d}W^{j_2}_t \\ &= \sum_{j_2=1}^{d}\int_{t_k}^{t_{k+1}} \left[\frac{\partial g_{j_2}}{\partial x}(X_{t_k})\left(\int_{t_k}^{t} f(X_s)\mathrm{d}s + \int_{t_k}^{t} [g(X_s) - g(X_{t_k})]\,dW_s\right) + R_{g_{j_2}}\right]\mathrm{d}W^{j_2}_t,\end{aligned}$$

where

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

Taking conditional expectation and applying the Itô isometry in conjunction with the elementary inequality $(a+b+c)^3 \leq 3a^2+3b^2+3c^2$ yields

(39)

$$\begin{aligned}\mathbb{I}_2 &= \sum_{j_2=1}^{d}\int_{t_k}^{t_{k+1}}\left\|\frac{\partial g_{j_2}}{\partial x}(X_{t_k})\left(\int_{t_k}^{t} f(X_s)\mathrm{d}s+\int_{t_k}^{t}[g(X_s)-g(X_{t_k})]\,\mathrm{d}W_s\right)+R_{g_{j_2}}\right\|_{L_N^2}^2\mathrm{d}t\\
&\leq 3\sum_{j_2=1}^{d}\int_{t_k}^{t_{k+1}}\left\|\frac{\partial g_{j_2}}{\partial x}(X_{t_k})\int_{t_k}^{t} f(X_s)\mathrm{d}s\right\|_{L_N^2}^2\mathrm{d}t+3\sum_{j_2=1}^{d}\int_{t_k}^{t_{k+1}}\|R_{g_{j_2}}\|_{L_N^2}^2\,\mathrm{d}t\\
&\quad+3\sum_{j_2=1}^{d}\int_{t_k}^{t_{k+1}}\left\|\frac{\partial g_{j_2}}{\partial x}(X_{t_k})\int_{t_k}^{t}[g(X_s)-g(X_{t_k})]\,\mathrm{d}W_s\right\|_{L_N^2}^2\mathrm{d}t\\
&:= \mathbb{J}_1+\mathbb{J}_2+\mathbb{J}_3.\end{aligned}$$

Using the Cauchy–Schwarz inequality, (18) and (12) gives

$$\begin{aligned}\mathbb{J}_1 &\leq C\sum_{j_2=1}^{d}\int_{t_k}^{t_{k+1}}(t-t_k)\int_{t_k}^{t}\left\|\frac{\partial g_{j_2}}{\partial x}(X_{t_k})f(X_s)\right\|_{L_N^2}^2\mathrm{d}s\mathrm{d}t\\
&\leq C(\Delta t)^2\sum_{j_2=1}^{d}\int_{t_k}^{t_{k+1}}\left(1+\left\|\sup_{0\leq t\leq T}|X_t|\right\|_{L_N^{2\gamma+4}}^{2\gamma+4}\right)\mathrm{d}t\\
&\leq C\Phi_{N,\gamma,4}(\Delta t)^3.\end{aligned}$$

For $\mathbb{J}_2$, we employ the Cauchy–Schwarz inequality, (15), and Lemma 3.1 to get

$$\begin{aligned}\mathbb{J}_2 &\leq C\sum_{j_2=1}^{d}\int_{t_k}^{t_{k+1}}\int_0^1\left\|\frac{\partial g_{j_2}}{\partial x}(X_{t_k}+s(X_t-X_{t_k}))-\frac{\partial g_{j_2}}{\partial x}(X_{t_k})(X_t-X_{t_k})\right\|_{L_N^2}^2\mathrm{d}s\mathrm{d}t\\
&\leq C\sum_{j_2=1}^{d}\int_{t_k}^{t_{k+1}}\|X_t-X_{t_k}\|_{L_N^4}^4\,\mathrm{d}t \leq C\Phi_{N,\gamma,4}(\Delta t)^3.\end{aligned}$$

With regard to $\mathbb{J}_3$, in light of the Itô isometry, (18), (3), the Cauchy–Schwarz inequality, and Lemma 3.1, we arrive at

$$\begin{aligned}\mathbb{J}_3 &\leq C\sum_{j_2=1}^{d}\int_{t_k}^{t_{k+1}}\mathbb{E}^N\left[\int_{t_k}^{t}(1+|X_{t_k}|^2)|X_s-X_{t_k}|^2\mathrm{d}s\right]\mathrm{d}t\\
&\leq C\sum_{j_2=1}^{d}\int_{t_k}^{t_{k+1}}\mathbb{E}^N\left[\left(1+\sup_{0\leq t\leq T}|X_t|^2\right)\int_{t_k}^{t_{k+1}}|X_s-X_{t_k}|^2\mathrm{d}s\right]\\
&\leq C\left(1+\mathbb{E}^N\Big[\sup_{0\leq t\leq T}|X_t|^4\Big]\right)^{\frac{1}{2}}\sum_{j_2=1}^{d}\int_{t_k}^{t_{k+1}}\left(\Delta t\int_{t_k}^{t_{k+1}}\|X_s-X_{t_k}\|_{L_N^4}^4\,\mathrm{d}s\right)^{\frac{1}{2}}\mathrm{d}t\\
&\leq C\left(1+\mathbb{E}^N\Big[\sup_{0\leq t\leq T}|X_t|^4\Big]\right)^{\frac{1}{2}}\Phi_{N,\gamma,4}^{1/2}(\Delta t)^3\leq C\Phi_{N,\gamma,4}(\Delta t)^3.\end{aligned}$$

Compiling the aforementioned estimates pertaining to $\mathbb{J}_1$, $\mathbb{J}_2$, and $\mathbb{J}_3$ results in

$$\mathbb{I}_2 \leq C\Phi_{N,\gamma,4}(\Delta t)^3. \tag{40}$$

For $\mathbb{I}_3$, by applying (17) and Lemma 3.1, we obtain

$$\begin{aligned}\mathbb{I}_3 &\le C(\Delta t)^2\mathbb{E}^N\left[\sum_{j=1}^{d}\left|\mathcal{L}^j g_j(X_{t_{k+1-}}) - \mathcal{L}^j g_j(X_{t_k})\right|^2\right]\\ &\le C(\Delta t)^2\|X_{t_{k+1-}} - X_{t_k})\|^2_{L^2_N} \le C\Phi_{N,\gamma,2}(\Delta t)^3.\end{aligned} \tag{41}$$

Plugging (38), (40) and (41) into (37) completes the proof. □

Lemma 3.5. *Under the same conditions used in Lemma 3.4, it holds for all* $k \in \{0, 1, 2, ..., n_T\}$ *that*

$$\left\|E^N\left[M_{k+1}|\mathcal{F}^W_{t_k}\right]\right\|^2_{L^2_N} \le C\Phi_{N,\gamma,8}(\Delta t)^4,$$

where the difinaiton of $\Phi_{N,\gamma,8}$ *comes from Lemma 3.1.*

Proof. As the stochastic integral vanishes under the conditional expectation, it then follows by using the elementary inequality $(a + b)^2 \le 2a^2 + 2b^2$ that

$$\begin{aligned}&\left\|\mathbb{E}^N\left[M_{k+1}|\mathcal{F}^W_{t_k}\right]\right\|^2_{L^2_N}\\ &\le 2\left\|\mathbb{E}^N\left[\int_{t_k}^{t_{k+1}}\left[f(X_t) - f(X_{t_{k+1-}})\right]\mathrm{d}t\middle|\mathcal{F}^W_{t_k}\right]\right\|^2_{L^2_N}\\ &\quad + 2\left\|\mathbb{E}^N\left[\frac{1}{2}\sum_{j=1}^{d}\mathcal{L}^j g_j(X_{t_{k+1-}})\Delta t_k - \frac{1}{2}\sum_{j=1}^{d}\mathcal{L}^j g_j(X_{t_k})\Delta t_k\middle|\mathcal{F}^W_{t_k}\right]\right\|^2_{L^2_N}\\ &:= \mathbb{I}_4 + \mathbb{I}_5.\end{aligned} \tag{42}$$

In the following, we cope with the above two items separately. Use Taylor's formula and the property of stochastic integral to infer that

$$\begin{aligned}&\mathbb{E}^N\left[\int_{t_k}^{t_{k+1}}\left[f(X_t) - f(X_{t_{k+1-}})\right]\mathrm{d}t\middle|\mathcal{F}^W_{t_k}\right]\\ &= -\mathbb{E}^N\left[\int_{t_k}^{t_{k+1}}\left[\frac{\partial f}{\partial x}(X_t)\left(\int_t^{t_{k+1}} f(X_s)\mathrm{d}s + \int_t^{t_{k+1}} g(X_s)dW_s\right) + R_f\right]\mathrm{d}t\middle|\mathcal{F}^W_{t_k}\right]\\ &= -\mathbb{E}^N\left[\int_{t_k}^{t_{k+1}}\int_t^{t_{k+1}}\frac{\partial f}{\partial x}(X_t)f(X_s)\mathrm{d}s\mathrm{d}t\middle|\mathcal{F}^W_{t_k}\right] - \mathbb{E}^N\left[\int_{t_k}^{t_{k+1}} R_f\mathrm{d}t\middle|\mathcal{F}^W_{t_k}\right],\end{aligned} \tag{43}$$

where

$$R_f := \int_0^1\left[\frac{\partial f}{\partial x}(X_t + s(X_{t_{k+1-}} - X_t)) - \frac{\partial f}{\partial x}(X_t)\right](X_{t_{k+1-}} - X_t)\ \mathrm{d}s. \tag{44}$$

By using again the elementary inequality $(a + b)^2 \le 2a^2 + 2b^2$, we derive

$$\begin{aligned}\mathbb{I}_4 &\le 4\left\|\mathbb{E}^N\left[\int_{t_k}^{t_{k+1}}\int_t^{t_{k+1}}\left[\frac{\partial f}{\partial x}(X_t)f(X_s)\mathrm{d}s\mathrm{d}t\right]\middle|\mathcal{F}^W_{t_k}\right]\right\|^2_{L^2_N}\\ &\quad + 4\left\|\mathbb{E}^N\left[\int_{t_k}^{t_{k+1}} R_f\mathrm{d}t\middle|\mathcal{F}^W_{t_k}\right]\right\|^2_{L^2_N}.\end{aligned} \tag{45}$$

Then, we use the property of conditional expectation, the Cauchy–Schwarz inequality, (10) and (12) to obtain

$$
\begin{aligned}
&\left\|\mathbb{E}^N\left[\int_{t_k}^{t_{k+1}}\int_t^{t_{k+1}}\left[\frac{\partial f}{\partial x}(X_t)f(X_s)\mathrm{d}s\mathrm{d}t\right]\middle|\mathcal{F}_{t_k}^W\right]\right\|_{L_N^2}^2\\
&\le C(\Delta t)^2\int_{t_k}^{t_{k+1}}\int_t^{t_{k+1}}\left\|\frac{\partial f}{\partial x}(X_t)f(X_s)\right\|_{L_N^2}^2\mathrm{d}s\mathrm{d}t\\
&\le C(\Delta t)^2\int_{t_k}^{t_{k+1}}\int_t^{t_{k+1}}\left\|(1+|X_t|)^\gamma(1+|X_s|)^{\gamma+1}\right\|_{L_N^2}^2\mathrm{d}s\mathrm{d}t\\
&\le C(\Delta t)^2\int_{t_k}^{t_{k+1}}\int_t^{t_{k+1}}\left(1+\mathbb{E}^N\left[\sup_{0\le t\le T}|X_t|^{4\gamma+2}\right]\right)\mathrm{d}s\mathrm{d}t\\
&\le C\Phi_{N,\gamma,4}(\Delta t)^4.
\end{aligned}\tag{46}
$$

With the help of the property of conditional expectation, (9), the Cauchy–Schwarz inequality together with Lemma 3.1, we obtain

$$
\begin{aligned}
&\left\|\mathbb{E}^N\left[\int_{t_k}^{t_{k+1}}R_f\mathrm{d}t\middle|\mathcal{F}_{t_k}^W\right]\right\|_{L_N^2}^2\\
&\le C\Delta t\mathbb{E}^N\left[\int_{t_k}^{t_{k+1}}\int_0^1\left|\left[\frac{\partial f}{\partial x}(X_t+s(X_{t_{k+1-}}-X_t))\right.\right.\right.\\
&\quad\left.\left.\left.-\frac{\partial f}{\partial x}(X_t)\right](X_{t_{k+1-}}-X_t)\right|^2\mathrm{d}s\mathrm{d}t\right]\\
&\le C\Delta t\mathbb{E}^N\left[\int_{t_k}^{t_{k+1}}\int_0^1 s(1+|sX_{t_{k+1-}}+(1-s)X_t|^{2\gamma-2}+|X_t|^{2\gamma-2})\right.\\
&\quad\left.\times|X_{t_{k+1-}}-X_t|^4\mathrm{d}s\mathrm{d}t\right]\\
&\le C\Delta t\left(1+\mathbb{E}^N\left[\sup_{0\le t\le T}|X_t|^{4\gamma-4}\right]\right)^{\frac12}\left\|\int_{t_k}^{t_{k+1}}|X_{t_{k+1-}}-X_t|^4\mathrm{d}t\right\|_{L_N^2}\\
&\le C\left(1+\mathbb{E}^N\left[\sup_{0\le t\le T}|X_t|^{4\gamma-4}\right]\right)^{\frac12}\Phi_{N,\gamma,8}^{1/2}(\Delta t)^4\le C\Phi_{N,\gamma,8}(\Delta t)^4.
\end{aligned}\tag{47}
$$

Plugging (46) and (47) into (45) gives

$$
\mathbb{I}_4\le C\Phi_{N,\gamma,8}(\Delta t)^4.\tag{48}
$$

Following the same estimation approach as above, we first derive

$$
\begin{aligned}
&\mathbb{E}^N\left[\mathcal{L}^jg_j(X_{t_{k+1-}})\Delta t_k-\mathcal{L}^jg_j(X_{t_k})\Delta t_k|\mathcal{F}_{t_k}^W\right]\\
&=\mathbb{E}^N\left[\int_{t_k}^{t_{k+1}}\mathcal{L}^jg_j(X_{t_{k+1-}})-\mathcal{L}^jg_j(X_{t_k})\,\mathrm{d}t|\mathcal{F}_{t_k}^W\right]\\
&=-\mathbb{E}^N\left[\int_{t_k}^{t_{k+1}}\left[\frac{\partial\mathcal{L}^jg_j}{\partial x}(X_{t_k})\left(\int_{t_k}^{t_{k+1}}f(X_s)\,\mathrm{d}s+\int_{t_k}^{t_{k+1}}g(X_s)\,\mathrm{d}W_s\right)\right.\right.\\
&\quad\left.\left.+R_{\mathcal{L}^jg_j}\right]\,\mathrm{d}t|\mathcal{F}_{t_k}^W\right]\\
&=-\mathbb{E}^N\left[\int_{t_k}^{t_{k+1}}\int_{t_k}^{t_{k+1}}\frac{\partial\mathcal{L}^jg_j}{\partial x}(X_{t_k})f(X_s)\,\mathrm{d}s\,\mathrm{d}t|\mathcal{F}_{t_k}^W\right]-\mathbb{E}^N\left[\int_{t_k}^{t_{k+1}}R_{\mathcal{L}^jg_j}\,\mathrm{d}t|\mathcal{F}_{t_k}^W\right],
\end{aligned}
$$

where

$$R_{\mathcal{L}^j g_j} := \int_0^1 \left[\frac{\partial \mathcal{L}^j g_j}{\partial x}(X_{t_k} + s(X_{t_{k+1-}} - X_{t_k})) - \frac{\partial \mathcal{L}^j g_j}{\partial x}(X_{t_k})\right](X_{t_{k+1-}} - X_{t_k})\, \mathrm{d}s.$$

Then, by utilizing the property of conditional expectation, the Cauchy–Schwarz inequality, (18), (12), (15) and Lemma 3.1, we can derive

$$\mathbb{I}_5 \leq C\Phi_{N,\gamma,4}(\Delta t)^4. \tag{49}$$

Inserting (48) and (49) into (42) yields desired result and the proof is completed. □

Lemma 3.6. *Let Assumptions 2.1–2.4 hold and* $\Delta t \in (0, \frac{1}{1+L_8+4L_1}]$, *then there exists a constant* $C > 0$, *independent of* Δt, *such that*

$$\|E_{k-}\|_{L^2}^2 \leq C(\Delta t)^2. \tag{50}$$

Proof. By Lemma 3.3, 3.4, and 3.5, we have

$$\begin{aligned}
\|E_{k-}\|_{L_N^2}^2 &\leq C\exp(2N_T\rho\Delta t)\Xi_T \sum_{i=0}^{n_T-1} \Phi_{N,\gamma,4}(\Delta t)^3 \\
&\quad + C\exp(2N_T\rho\Delta t)\Xi_T \sum_{i=0}^{n_T-1} \Phi_{N,\gamma,8}(\Delta t)^3 \\
&\leq C(\Delta t)^2\varphi(N_T)\left(\prod_{j=1}^{M}\eta_j\right)\Phi_{N,\gamma,8},
\end{aligned} \tag{51}$$

where $\varphi(N_T) := \exp(2N_T\rho T)(1+C+CT)^{N_T}(1+N_T)$. Taking expectation and using the property of conditional expectation, together with the Hölder inequality gives

$$\|E_{k-}\|_{L^2}^2 \leq C(\Delta t)^2 \big\|\varphi(N_T)\big\|_{L_2} \left\|\prod_{j=1}^{M}\eta_j\right\|_{L_4} \big\|\Phi_{N,\gamma,8}\big\|_{L_4}. \tag{52}$$

Employing the properties of Poisson process yields

$$\big\|\varphi(N_T)\big\|_{L_2}^2 \leq C\sum_{j=0}^{\infty}\exp(4j\rho T)(1+C+CT)^{2j}(1+j)^2\frac{(\lambda T)^j}{j!}\exp(-\lambda T) < \infty.$$

Apply the property of independent increment of Poisson process to derive

$$\left\|\prod_{j=1}^{M}\eta_j\right\|_{L_4}^4 \leq (1+C\Delta t)^{4M} < \infty.$$

By the definition of $\Phi_{N,\gamma,q}$, we have

$$\big\|\Phi_{N,\gamma,8}\big\|_{L_2}^2 \leq C\left(1+\mathbb{E}\left[\sup_{0\leq t\leq T}|X_t|^{(1+\gamma)16}\right]\right) < \infty.$$

Therefore, it immediately follows that

$$\|E_{k-}\|_{L_2}^2 \leq C(\Delta t)^2.$$

The proof is now complete. □

Equipped with the preceding lemmas, we now present the main strong convergence result, which establishes the optimal mean-square convergence rate of the proposed scheme.

Theorem 3.1. *Assume all conditions of Lemma 3.6 are satisfied. Let Y_{t_k} and X_{t_k} denote the solutions to the numerical scheme defined by* (21)-(22) *and the exact solution of* (1) *at time $t = t_k, k = 1, 2, ..., n_T$, respectively. Then, there exists a constant $C > 0$, independent of Δt, such that*

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

Proof. By applying the relationship established in Lemma 3.2 and the estimate provided by Lemma 3.6, we obtain the desired result. This completes the proof. □

3.2. Positivity preservation analysis. In this subsection, we focus on the positivity preservation of the proposed jump-adapted drift-diffusion double implicit Milstein method. First, we impose the following assumption on the solution to the corresponding SDEs without jumps:

$$\mathrm{d}X_t = f(X_t)\mathrm{d}t + g(X_t)\mathrm{d}W_t, \quad t \in (0, T], \quad X_0 = x_0 > 0. \tag{53}$$

Assumption 3.1. *The solution of* (53) *is positive with probability* 1.

To ensure that the considered problem (1) admits a positive solution, we further impose the following assumptions on the jump coefficient h.

Assumption 3.2. *Suppose that the coefficient function h in* (1) *satisfies*

$$x + h(x) > 0, \quad \forall x > 0. \tag{54}$$

Now, we are in the position to give the following results for the underlying problem with jumps.

Lemma 3.7. *Under Assumptions 2.1–2.2 and Assumptions 3.1–3.2, and for $x_0 > 0$, the problem* (1) *admits a unique solution, which remains positive with probability* 1.

Proof. Let $\tau_n, n = 0, 1, \cdots$ be the n-th jump arrival times and $\tau_0 = 0$. Within the interval $[\tau_n, \tau_{n+1})$ between jumps, the problem (1) evolves as an SDE without jump:

$$X_t = X_{\tau_n} + \int_{\tau_n}^{t} f(X_s)\mathrm{d}s + \int_{\tau_n}^{t} g(X_s)\mathrm{d}W_s, \quad t \in [\tau_n, \tau_{n+1}). \tag{55}$$

Owing to the jump at τ_{n+1}, we have

$$\Delta X_{\tau_{n+1}} := X_{\tau_{n+1}} - X_{\tau_{n+1}-} = h\big(X_{\tau_{n+1}-}\big). \tag{56}$$

Thus,

$$X_{\tau_{n+1}} = X_{\tau_{n+1}-} + h(X_{\tau_{n+1}-}), \quad n = 0, 1, \cdots. \tag{57}$$

In particular, for $t \in [0, \tau_1)$, (55) becomes

$$X_t = x_0 + \int_0^t f(X_s)\mathrm{d}s + \int_0^t g(X_s)\mathrm{d}W_s, \quad x_0 > 0. \tag{58}$$

Thanks to Assumption 3.1, with probability one $X_t > 0, t \in [0, \tau_1)$. At the instance $t = \tau_1$, $X_{\tau_1} = X_{\tau_1-} + h(X_{\tau_1-}) > 0$ under the condition that $x + h(x) > 0$ for $x > 0$. Repeating this procedure in the subsequent intervals $[\tau_n, \tau_{n+1})$ for $n = 1, 2, \cdots$ gives the required assertion. □

Before presenting the positivity preservation result for the proposed jump-adapted drift-diffusion double implicit Milstein method, we introduce an additional assumption [14] concerning the diffusion coefficient g, as detailed below.

Assumption 3.3. *The coefficient g satisfies the following conditions:*

$$\sum_{j=1}^{d} \mathcal{L}^j g_j(x) > 0, \quad \text{for} \quad \forall x > 0 \tag{59}$$

and

$$\left\langle x, \sum_{j=1}^{d} \mathcal{L}^j g_j(x) \right\rangle > \frac{1}{2}|g(x)|^2 > 0, \quad \text{for} \quad \forall x > 0. \tag{60}$$

Theorem 3.2. *Suppose that all conditions in Assumptions 2.1–2.4 and Assumptions 3.1-3.3 hold, if the initial value $x_0 > 0$, then the proposed the jump-adapted drift-diffusion double implicit Milstein method* (21)-(22) *is positivity preserving, i.e., with probability one, $Y_{t_k} > 0$ implies $Y_{t_{k+1}} > 0$.*

Proof. When the jump coefficient h is identically zero ($h \equiv 0$), the proposed jump-adapted double implicit Milstein method degenerates into the drift-diffusion double implicit Milstein method for stochastic differential equations (53). This specialized method has been rigorously analyzed in prior works [14, 32] and is proven to preserve positivity under Assumptions 2.1–2.4, 3.1, and 3.3, given a positive initial condition.

Bear this in mind, it then follows that given $Y_{t_k} > 0$, one can infer that $Y_{t_{k+1-}} > 0$ since $Y_{t_{k+1-}}$ is produced through the drift-diffusion double implicit Milstein method with the initial value $Y_{t_0} = x_0 > 0$. A more careful examination of (22) indicates that

$$\begin{cases} Y_{t_{k+1}} = Y_{t_{k+1-}} + h(Y_{t_{k+1-}}), & \text{if } t_{k+1} \text{ is a jump time;} \\ Y_{t_{k+1}} = Y_{t_{k+1-}}, & \text{otherwise.} \end{cases} \tag{61}$$

Then, Assumption 3.2 suffices to guarantee $Y_{t_{k+1}} > 0$ by taking (61) into account. Consequently, we can conclude that the scheme (21)-(22) is able to reproduces the positivity of (1). □

Remark 3.3. *Building on previous discussions, positivity-preserving numerical methods for SDEs without jumps show great potential to evolve into jump-adapted ones for tackling jumpdiffusion problems.*

4. Numerical results

In this section we will present some numerical experiments to verify the previous theoretical findings. We consider the following nonlinear jump-diffusion problems:

$$\mathrm{d}X_t = \left(X_{t-} - X_{t-}^{\beta}\right)\mathrm{d}t + \mu X_{t-}\mathrm{d}W_t + \nu X_{t-}\mathrm{d}N_t, \; t \in (0, T], \; X_0 = x_0, \tag{62}$$

where the constant $\beta \geq 3$ is odd, $\mu = 2$, and $\nu = 1$. It is easy to check that conditions in Assumptions 2.1 – 2.2 are all fulfilled. In the sequent experiments with (62), we are to test approximation errors in terms of means square errors $\varepsilon = \sqrt{\mathbb{E}[|X_T - Y_T|^2]}$. As usual, the expectation is approximated by the Monte-Carlo approximation:

$$e := \sqrt{\hat{\mathbb{E}}\left[|X_T - Y_T|^2\right]} := \sqrt{\frac{1}{N_{mc}} \sum_{i=1}^{N_{mc}} |X_T^{(i)} - Y_T^{(i)}|^2},$$

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 the jump-adapted drift-diffusion double implicit Milstein method at the ith sampling. We plot the achieved accuracy versus stepsizes in logarithmic scale. $N_{mc} = 1000$ Wiener process and

Poisson process 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 as the stepsize Δt decreases, which implies that the proposed jump-adapted drift-diffusion double implicit Milstein method is convergent.

TABLE 4.1. Numerical results for (62) with $\mu = 2$, $\nu = 1$.

Δt	$\beta = 3$	$\beta = 5$	$\beta = 7$	$\beta = 9$	$\beta = 11$	$\beta = 13$
2^{-7}	0.0121	0.0145	0.0203	0.0289	0.0370	0.0437
2^{-8}	0.0058	0.0070	0.0105	0.0150	0.0197	0.0244
2^{-9}	0.0029	0.0036	0.0057	0.0079	0.0106	0.0137
2^{-10}	0.0013	0.0016	0.0026	0.0037	0.0047	0.0059
2^{-11}	0.0007	0.0008	0.0013	0.0017	0.0020	0.0023

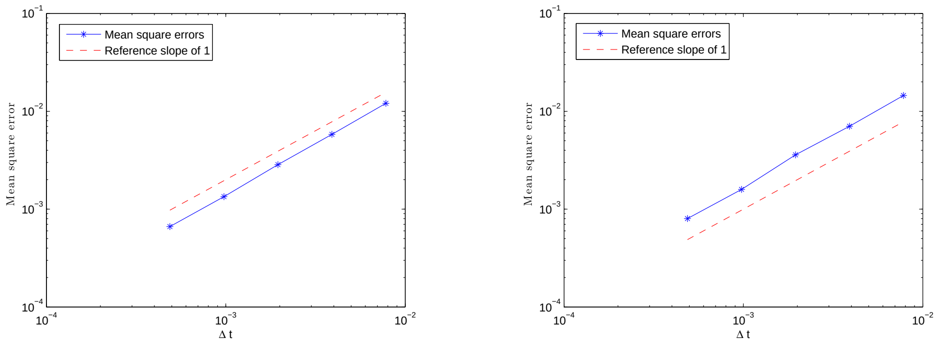

FIGURE 4.1. Numerical results for (62): Computational errors versus stepsize Δt on a log-log scale. Left: $\beta = 3$; right: $\beta = 5$.

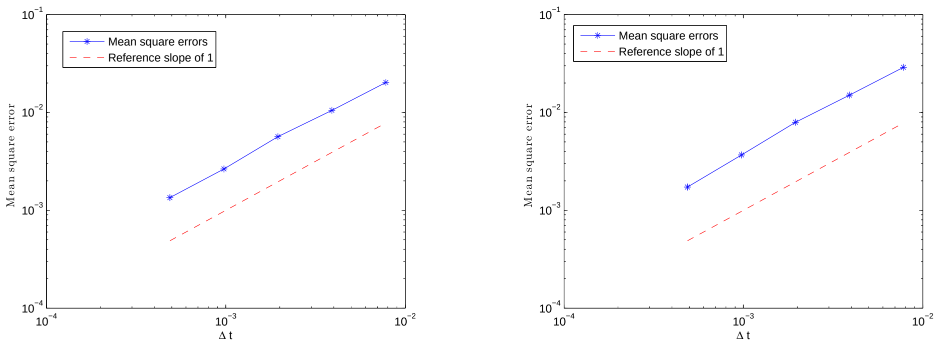

FIGURE 4.2. Numerical results for (62): Computational errors versus stepsize Δt on a log-log scale. Left: $\beta = 7$; right: $\beta = 9$.

To clearly display the convergence rates, we plot in Figures 4.3 – 4.5 the achieved errors versus stepsizes in logarithmic scale. As predicted, the slopes of the errors

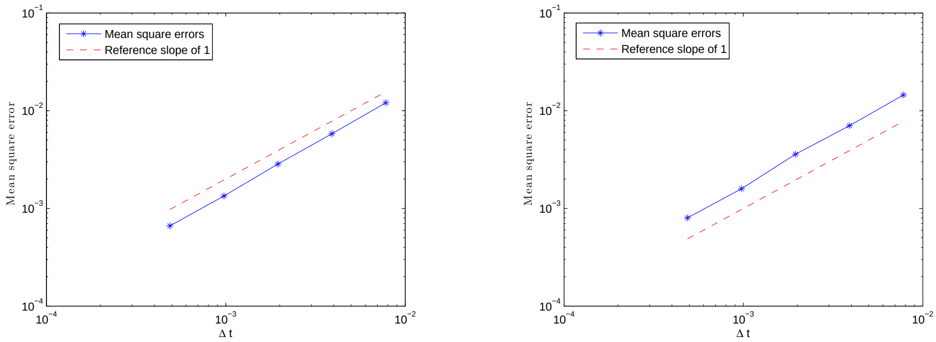

FIGURE 4.3. Numerical results for (62): Computational errors versus stepsize Δt on a log-log scale. Left: $\beta = 3$; right: $\beta = 5$.

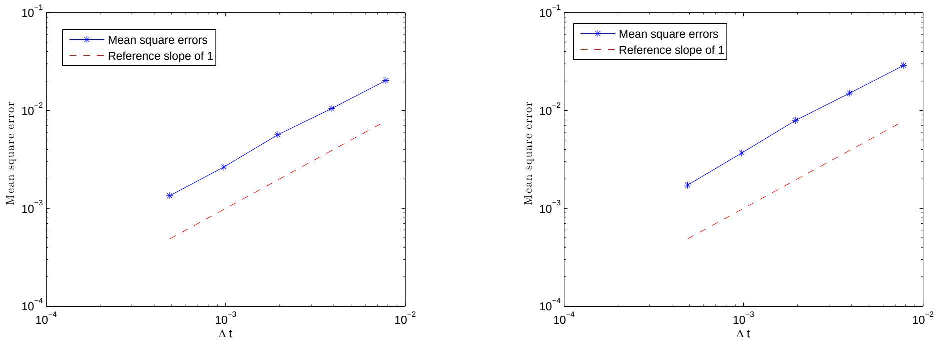

FIGURE 4.4. Numerical results for (62): Computational errors versus stepsize Δt on a log-log scale. Left: $\beta = 7$; right: $\beta = 9$.

(solid lines) and the reference dashed line match well, which indicates that the proposed scheme shows a strong convergence rate of order one. These numerical results are consistent with our strong convergence result in the preceding section.

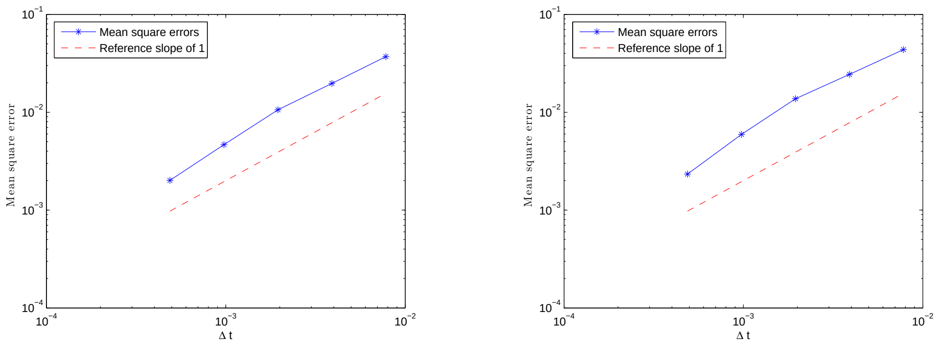

FIGURE 4.5. Numerical results for (62): Computational errors versus stepsize Δt on a log-log scale. Left: $\beta = 11$; right: $\beta = 13$.

As the second part of numerical experiments, in what follows, we are to test the positivity preserving property of the proposed scheme. More precisely, we present in Table 4.2 the percentage of negative paths when simulating the following nonlinear problems:

$$\mathrm{d}X_t = \Big(\big(\alpha+\frac{1}{2}\sigma^2\big)X_{t-}-\vartheta X_{t-}^3\Big)\mathrm{d}t+\sigma X_{t-}\mathrm{d}W_t+\kappa X_{t-}\mathrm{d}N_t,\ t\in(0,T],\ X_0=x_0>0, \tag{63}$$

where $\alpha \geq 0, \vartheta, \sigma > 0$ and $\kappa > -1$. Note that when $\kappa = 0$, equation (63) reduces to the stochastic Ginzburg-Landau equation considered in [16], whose solution is explicitly known as

$$X_t = \frac{x_0 \exp(\alpha t + \sigma W_t)}{\sqrt{1 + 2x_0^2\vartheta \int_0^t \exp(2\alpha s + 2\sigma W_s)ds}} > 0. \tag{64}$$

Since $\kappa > -1$ and $x + \kappa x > 0$ for $x > 0$, equation (63) admits a unique positive solution almost surely.

In Figure 4.6, we use the proposed scheme with a small step-size $\Delta t_{exact} = 2^{-15}$ to plot two one-path simulations of (62) for two sets of parameters.

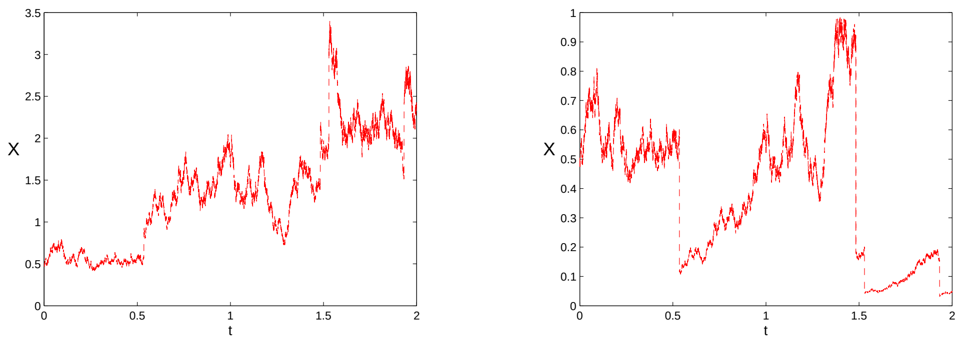

FIGURE 4.6. One path simulation of (63). Left: $\alpha = 2, \sigma = 1, \vartheta = 1, \kappa = 0.5, \lambda = 3, x_0 = 0.5, T = 2$; right: $\alpha = 2, \sigma = 1, \vartheta = 1, \kappa = -0.8, \lambda = 3, x_0 = 0.5, T = 2$.

TABLE 4.2. Numerical results for (63) with $\alpha = 4, \sigma = 1, \vartheta = 1, \kappa = -0.8, \lambda = 100, x_0 = 0.5$ and with $N_{mc} = 1000$ simulated paths.

Time T	Stepsize Δt	JAMM	JAIMM	JADIMM
$T = 4$	$\Delta t = \frac{1}{4}$	92.6%	18.6%	0%
	$\Delta t = \frac{1}{8}$	66.6%	12.9%	0%
	$\Delta t = \frac{1}{16}$	13.8%	7.4%	0%
$T = 8$	$\Delta t = \frac{1}{4}$	100%	38.1%	0%
	$\Delta t = \frac{1}{8}$	93.4%	26.6%	0%
	$\Delta t = \frac{1}{16}$	27.6%	14.9%	0%
$T = 16$	$\Delta t = \frac{1}{4}$	100%	75.5%	0%
	$\Delta t = \frac{1}{8}$	100%	52.9%	0%
	$\Delta t = \frac{1}{16}$	51.4%	29.9%	0%

Here, for comparison, we employ the jump-adapted Milstein method (JAMM) from [28], the jump-adapted drift implicit Milstein method (JAIMM) from [38], and the proposed jump-adapted drift-diffusion double implicit Milstein method (JADIMM) to solve (63).

As shown in Table 4.2, JAMM and JAIMM fail to preserve positive solutions. An increase in the integration interval leads to a higher percentage of negative paths for these methods. However, the proposed JADIMM consistently maintains positivity across all step sizes and time intervals. This is in line with the previous theoretical findings.

5. Conclusion remarks

In this work, a jump-adapted drift-diffusion double implicit Milstein method has been proposed and analyzed for solving a class of jump-diffusion stochastic differential equations. The proposed scheme features a simplified algorithmic structure while maintaining high accuracy, representing a significant improvement over the classical Milstein method for jump-diffusion systems. By addressing the challenges posed by non-Lipschitz continuity, weaker temporal regularity, and stochastic time partitioning, a rigorous strong convergence result with optimal mean-square convergence rate of order one has been established. Furthermore, under some mild assumptions, the proposed scheme has been proved to possess the ability to preserve positivity, which is crucial for practical applications. Numerical experiments have been conducted to validate the theoretical findings, demonstrating the effectiveness and efficiency of the method. In the future, we plan to further develop and analyze efficient jump-adapted numerical methods to more general models driven by broader classes of jump processes.

Acknowledgments

We would like to thank the reviewers for valuable suggestions and comments, which greatly improved the paper. This research is partially supported by the National Natural Science Foundation of China (under Grant No. 12371398, 12071261).

References

[1] M. Baran, Approximation for solutions of Lévy-type stochastic differential equations, Stoch. Anal. Appl., 27 (2009), 924–961.
[2] T. Björk, Point processes and jump diffusions—an introduction with finance applications, Cambridge University Press, Cambridge, 2021.
[3] N. Bruti-Liberati and E. Platen, Strong approximations of stochastic differential equations with jumps, J. Comput. Appl. Math., 205 (2007), 982–1001.
[4] E. Buckwar and M. G. Riedler, Runge-Kutta methods for jump-diffusion differential equations, J. Comput. Appl. Math., 236 (2011), 1155–1182.
[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] 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.
[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] R. Cont and P. Tankov, Financial modelling with jump processes, Chapman & Hall/CRC Financial Mathematics Series, Chapman & Hall/CRC, Boca Raton, FL, 2004.
[9] A. Gardoń, The order of approximations for solutions of Itô-type stochastic differential equations with jumps, Stochastic Anal. Appl., 22 (2004), 679–699.
[10] I. I. Gikhman and A. V. Skorokhod, The theory of stochastic processes. III, Classics in Mathematics, Springer, Berlin, 2007.
[11] D. J. Higham and P. E. Kloeden, Numerical methods for nonlinear stochastic differential equations with jumps, Numer. Math., 101 (2005), 101–119.

[12] 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.

[13] 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.

[14] D. J. Higham, X. Mao, and L. Szpruch, Convergence, non-negativity and stability of a new Milstein scheme with applications to finance, Discrete Contin. Dyn. Syst. Ser. B, 18 (2013), 2083–2100.

[15] 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.

[16] M. Hutzenthaler, A. Jentzen, and P. E. Kloeden, Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 467 (2011), 1563–1576.

[17] A. Kałuża and P. Przybyłowicz, Optimal global approximation of jump-diffusion SDEs via path-independent step-size control, Appl. Numer. Math., 128 (2018), 24–42.

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

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

[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 tamed Milstein schemes of SDEs driven by Lévy noise, Discrete Contin. Dyn. Syst. Ser. B, 22 (2017), 421–463.

[22] Z. Lei, S. Gan, and J. Liu, First order strong approximation of Ait–Sahalia-type interest rate model with Poisson jumps, Numer. Algorithms, 94 (2023), 93–130.

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

[24] 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.

[25] 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.

[26] 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.

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

[28] 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.

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

[30] Q. Ren and H. Tian, Mean-square convergence and stability of two-step Milstein methods for stochastic differential equations with Poisson jumps, Comput. Appl. Math., 41 (2022), Paper No. 125, 18.

[31] 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.

[32] X. Wang, Mean-square convergence rates of implicit Milstein type methods for SDEs with non-Lipschitz coefficients, Adv. Comput. Math., 49 (2023), Paper No. 37.

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

[34] Y. Wang, Z. Chen, M. Niu, and Y. Niu, Mean-square convergence and stability of compensated stochastic theta methods for jump-diffusion SDEs with super-linearly growing coefficients, Calcolo, 60 (2023), Paper No. 30, 32.

[35] 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.

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

[37] X. Yang and W. Zhao, Strong convergence analysis of jump-adapted backward Euler method under non-globally Lipschitz condition, Math. Numer. Sin., 44 (2022), 163–177.

[38] X. Yang and W. Zhao, Strong convergence of jump-adapted implicit Milstein method for a class of nonlinear jump-diffusion problems, J. Comput. Math., 42 (2024), 248–270.

School of Mathematics, China University of Mining and Technology, Xuzhou, Jiangsu 221116, China

E-mail: `cq202122@163.com and xuyang96@cumt.edu.cn`

School of Mathematics, Shandong University, Jinan, Shandong 250100, China

E-mail: `wdzhao@sdu.edu.cn`