

# NUMERICAL METHODS FOR MEAN-FIELD STOCHASTIC DIFFERENTIAL EQUATIONS WITH JUMPS \*

YABING SUN\* AND WEIDONG ZHAO†

**Abstract.** In this paper, we are devoted to the numerical methods for mean-field stochastic differential equations with jumps (MSDEJs). First by using the mean-field Itô formula [Sun, Yang and Zhao, Numer. Math. Theor. Meth. Appl., 10 (2017), pp. 798–828], we develop the Itô formula and construct the Itô-Taylor expansion for MSDEJs. Then based on the Itô-Taylor expansion, we propose the strong order  $\gamma$  and the weak order  $\eta$  Itô-Taylor schemes for MSDEJs. The strong and weak convergence rates  $\gamma$  and  $\eta$  of the strong and weak Itô-Taylor schemes are theoretically proved, respectively. Finally some numerical tests are also presented to verify our theoretical conclusions.

**Key words.** Mean-field stochastic differential equations with jumps, Itô formula, Itô-Taylor expansion, Itô-Taylor schemes, error estimates.

**AMS subject classifications.** 60H35, 65C20, 60H10

**1. Introduction.** Let  $(\Omega, \mathcal{F}, \mathbb{F}, P)$  be a complete filtered probability space with  $\mathbb{F} = \{\mathcal{F}_t\}_{0 \leq t \leq T}$  being the filtration of the following two mutually independent stochastic processes:

- the  $m$ -dimensional Brownian motion:  $W = (W_t)_{0 \leq t \leq T}$ ;
- the Poisson random measure on  $\mathbb{E} \times [0, T]$ :  $\{\mu(A \times [0, t]), A \in \mathcal{E}, 0 \leq t \leq T\}$ , where  $\mathbb{E} = \mathbb{R}^q \setminus \{0\}$  and  $\mathcal{E}$  is its Borel field.

Suppose that  $\mu$  has the intensity measure  $\nu(de, dt) = \lambda(de)dt$ , where  $\lambda$  is a  $\sigma$ -finite measure on  $(\mathbb{E}, \mathcal{E})$  satisfying  $\int_{\mathbb{E}} (1 \wedge |e|^2) \lambda(de) < +\infty$ . Then we have the compensated Poisson random measure

$$\tilde{\mu}(de, dt) = \mu(de, dt) - \lambda(de)dt,$$

such that  $\{\tilde{\mu}(A \times [0, t]) = (\mu - \nu)(A \times [0, t])\}_{0 \leq t \leq T}$  is a martingale for any  $A \in \mathcal{E}$  with  $\lambda(A) < \infty$ . Moreover, let  $F$  be the distribution of the jump size, then it holds that

$$\lambda(de) = \dot{\lambda}F(de),$$

where  $\dot{\lambda} = \lambda(\mathbb{E}) < \infty$  is the intensity of the Poisson process  $N_t = \mu(\mathbb{E} \times [0, t])$ , which counts the number of jumps of  $\mu$  occurring in  $[0, t]$ . Then, the Poisson measure  $\mu$  generates a sequence of pairs  $\{(\tau_i, Y_i), i = 1, 2, \dots, N_T\}$  with  $\{\tau_i \in [0, T], i = 1, 2, \dots, N_T\}$  representing the jump times of the Poisson process  $N_t$  and  $\{Y_i \in \mathbb{E}, i = 1, 2, \dots, N_T\}$  the corresponding jump sizes satisfying  $Y_i \stackrel{iid}{\sim} F$ . For more details of the Poisson random measure or Lévy measure, the readers are referred to [10, 33].

We consider the following mean-field stochastic differential equation with jumps

---

\*This work is partially supported by the science challenge Project (No. TZ2018001), the NSF of China (under Grant Nos. 11571351, 11571206, 11831010, 11871068) and the China Postdoctoral Science Foundation (No. 2019TQ0073).

†Corresponding author

\*College of Science, National University of Defense Technology, Changsha, Hunan 410073, China. Email: sunybly@163.com

†School of Mathematics & Institute of Finance, Shandong University, Jinan, Shandong 250100, China. Email: wdzhao@sdu.edu.cn

(MSDEJs) on  $(\Omega, \mathcal{F}, \mathbb{F}, P)$

$$(1.1) \quad \begin{aligned} X_t^{t_0, \xi} &= \xi + \int_{t_0}^t \mathbb{E}[b(s, X_s^{t_0, \xi'}, x)]|_{x=X_s^{t_0, \xi}} ds + \int_{t_0}^t \mathbb{E}[\sigma(s, X_s^{t_0, \xi'}, x)]|_{x=X_s^{t_0, \xi}} dW_s \\ &+ \int_{t_0}^t \int_{\mathbb{E}} \mathbb{E}[c(s, X_{s-}^{t_0, \xi'}, x, e)]|_{x=X_{s-}^{t_0, \xi}} \mu(de, ds), \quad 0 \leq t_0 \leq t \leq T, \end{aligned}$$

where  $t_0$  and  $T$  are, respectively, the deterministic initial and terminal time; the initial condition  $\xi$  is  $\mathcal{F}_{t_0}$  measurable;  $b : [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ ,  $\sigma : [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times m}$ , and  $c : [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{E} \rightarrow \mathbb{R}^d$  are the so called drift, diffusion and jump coefficients, respectively. Here the superscript  $t_0, \xi$  indicates that the MSDEJ (1.1) starts from the time-space point  $(t_0, \xi)$ , and  $X_t^{t_0, \xi'}$  is the solution of the MSDEJ (1.1) with  $\xi = \xi'$ . In general,  $\xi$  and  $\xi'$  are different.

Mean-field stochastic differential equations (MSDEs), also called McKean-Vlasov SDEs, was first studied by Kac [30, 31] in the 1950s. Since then, MSDEs have been encountered and intensively investigated in many areas such as kinetic gas theory [2, 26, 38], quantum mechanics [27], quantum chemistry [34], McKean-Vlasov type partial differential equations (PDEs) [4, 5, 19, 25], mean-field games [8, 9, 12, 20] and mean-field backward stochastic differential equations (MBSDEs) [3, 4, 6, 36, 37]. In the last decade, MSDEJs have also received much attention because of its wild applications in the research on nonlocal PDEs [16], MBSDEs with jumps [23, 24], economics and finance [14], and mean-field control and mean-field games with jumps [13, 29, 39]. Therefore, it is important and necessary to study the numerical solutions of MSDEJs.

Compared with the well developed theory of numerical methods for stochastic differential equations with jumps (SDEJs) (see [15, 21, 33] and references therein), little attention has been paid to the numerical methods for MSDEJs. In this work, we aim to propose the general Itô-Taylor schemes for solving MSDEJs. The authors studied the mean-field Itô formula and proposed the general Itô-Taylor schemes for MSDEs in [35]. By using the mean-field Itô formula, we first develop the Itô formula for MSDEJs, then based on which, we construct the Itô-Taylor expansion for MSDEJs and further propose the Itô-Taylor schemes of strong order  $\gamma$  and weak order  $\eta$  for solving MSDEJs. Taking  $\gamma = 0.5, 1.0$  and  $\eta = 2.0$ , we obtain the Euler scheme, the strong order 1.0 Taylor scheme, and the weak order 2.0 Taylor scheme, respectively. Moreover, the rigorous error estimates indicate that the order of strong convergence of the strong order  $\gamma$  Taylor scheme is  $\gamma$  and the order of weak convergence of the weak order  $\eta$  Taylor scheme is  $\eta$ . Some numerical tests are carried out to show the efficiency and the accuracy of the proposed schemes for solving MSDEJs and to verify our theoretical conclusions. The numerical results are consistent with our theoretical ones and show that the efficiency of the proposed schemes depends on the level of the intensity of the Poisson random measure.

## 2. Preliminaries.

**2.1. Existence and uniqueness of solution of MSDEJs.** In this subsection, we state a standard result on the existence and uniqueness of the strong solutions of the MSDEJ (1.1). For this end, we set the following assumptions on  $b$ ,  $\sigma$  and  $c$ .

- (A1)  $b(\cdot, x', x)$ ,  $\sigma(\cdot, x', x)$  and  $c(\cdot, x', x, e)$  are deterministic continuous processes, for any fixed  $(x', x, e) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{E}$ .

(A2) There exists a positive constant  $L$  such that

$$\begin{aligned} & |b(t, x', x) - b(t, y', y)| + |\sigma(t, x', x) - \sigma(t, y', y)| \\ & \leq L(|x' - y'| + |x - y|), \quad \text{for all } t \in [0, T] \text{ and } x, x', y, y' \in \mathbb{R}^d. \end{aligned}$$

(A3) There exists a function  $\rho : \mathbf{E} \rightarrow \mathbb{R}^+$  satisfying  $\int_{\mathbf{E}} \rho^2(e) \lambda(de) < +\infty$ , such that

$$\begin{aligned} & |c(t, x', x, e) - c(t, y', y, e)| \leq \rho(e)(|x' - y'| + |x - y|), \\ & |c(t, 0, 0, e)| \leq \rho(e), \quad \text{for all } t \in [0, T], \quad x, x', y, y' \in \mathbb{R}^d \text{ and } e \in \mathbf{E}. \end{aligned}$$

(A4) There exists a constant  $K > 0$  such that

$$\begin{aligned} & |b(t, x', x)| + |\sigma(t, x', x)| \leq K(1 + |x| + |x'|), \\ & |c(t, x', x, e)| \leq \rho(e)(1 + |x'| + |x|), \end{aligned}$$

for all  $t \in [0, T]$ ,  $x, x' \in \mathbb{R}^d$ , and  $e \in \mathbf{E}$ .

Now we state the existence and uniqueness of the solution of the MSDEJ (1.1) and some useful properties in the following theorem [16, 23].

**THEOREM 2.1.** *Suppose that the coefficients  $b$ ,  $\sigma$  and  $c$  of the MSDEJ (1.1) satisfy the assumptions (A1) – (A4), and the initial data  $\xi$  and  $\xi'$  satisfy*

$$\mathbb{E}[|\xi|^2 + |\xi'|^2] < +\infty,$$

then the MSDEJ (1.1) admits a unique strong solution  $X_t^{t_0, \xi}$  on  $[t_0, T]$  with

$$(2.1) \quad \sup_{t_0 \leq t \leq T} \mathbb{E} \left[ |X_t^{t_0, \xi}|^2 \right] < +\infty.$$

In addition, for any  $p \geq 2$ , there exists a  $C_p \in \mathbb{R}^+$  such that for any initial time  $t_0 \in [0, T]$  and  $\mathcal{F}_{t_0}$  measurable  $\xi_1, \xi_2 \in L^p$ ,

$$(2.2a) \quad \mathbb{E} \left[ \sup_{t_0 \leq s \leq T} |X_s^{t_0, \xi_1}|^p \middle| \mathcal{F}_{t_0} \right] \leq C_p (1 + |\xi_1|^p),$$

$$(2.2b) \quad \mathbb{E} \left[ \sup_{t_0 \leq s \leq T} |X_s^{t_0, \xi_1} - X_s^{t_0, \xi_2}|^p \middle| \mathcal{F}_{t_0} \right] \leq C_p (|\xi_1 - \xi_2|^p),$$

$$(2.2c) \quad \mathbb{E} \left[ \sup_{t_0 \leq s \leq t_0 + \delta} |X_s^{t_0, \xi_1} - \xi_1|^p \middle| \mathcal{F}_{t_0} \right] \leq C_p (1 + |\xi_1|^p) \delta,$$

a.s. for all  $\delta > 0$  with  $t_0 + \delta \leq T$ , where  $X_s^{t_0, \xi_1}$  and  $X_s^{t_0, \xi_2}$  are the solutions of (1.1) with initial conditions  $\xi_1$  and  $\xi_2$ , respectively. Here the constant  $C_p$  in (2.2) only depends on  $L$ ,  $K$  and  $\rho(e)$ .

**2.2. The Markov property.** In this subsection, we present the Markov property of the solutions of MSDEJs, which will play a key role in our error estimates. For simplicity, we let  $t_0 = 0$  and denote by  $X_0 = \xi$  and  $X_t = X_t^{0, \xi}$ , then the MSDEJ (1.1) becomes

$$(2.3) \quad \begin{aligned} X_t = X_0 & + \int_0^t \mathbb{E}[b(s, X_s^{0, \xi'}, x)]_{x=X_s} ds + \int_0^t \mathbb{E}[\sigma(s, X_s^{0, \xi'}, x)]_{x=X_s} dW_s \\ & + \int_0^t \int_{\mathbf{E}} \mathbb{E}[c(s, X_{s-}^{0, \xi'}, x, e)]_{x=X_{s-}} \mu(de, ds). \end{aligned}$$

Let  $X_t^{s,x}$  be the solution of the MSDEJ (2.3) starting from the point  $(s, x)$ , i.e.,

$$\begin{aligned} X_t^{s,x} &= x + \int_s^t \mathbb{E}[b(r, X_r^{0,\xi'}, x)]|_{x=X_r^{s,x}} dr + \int_s^t \mathbb{E}[\sigma(r, X_r^{0,\xi'}, x)]|_{x=X_r^{s,x}} dW_r \\ &\quad + \int_s^t \int_{\mathbb{E}} \mathbb{E}[c(r, X_{r-}^{0,\xi'}, x, e)]|_{x=X_{r-}^{s,x}} \mu(de, dr) \end{aligned}$$

for  $0 \leq s \leq t \leq T$ . Then by Theorem 2.1, we have

$$(2.4) \quad X_t^{s, X_s} = X_t.$$

Now we state the Markov property of the solutions of MSDEJs as below.

**THEOREM 2.2** (The Markov property). *Let  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  be a bounded Borel measurable function. Then for the solution  $X_t$  of (2.3), it holds that*

$$(2.5) \quad \mathbb{E}[f(X_t)|\mathcal{F}_s] = \mathbb{E}[f(X_t^{s, X_s})] = \mathbb{E}[f(X_t^{s, y})]|_{y=X_s},$$

where  $0 \leq s \leq t \leq T$ .

*Proof.* By using the relationship (2.4), the proof of Theorem 2.2 is similar to that of Theorem 7.1.2 in [32]. So we omit it here.  $\square$

**2.3. The equivalent form of MSDEJs.** By using the relationship  $\tilde{\mu}(de, dt) = \mu(de, dt) - \lambda(de)dt$ , the MSDEJ (1.1) can be written as

$$(2.6) \quad \begin{aligned} X_t^{t_0, \xi} &= \xi + \int_{t_0}^t \mathbb{E}[\tilde{b}(s, X_s^{t_0, \xi'}, x)]|_{x=X_s^{t_0, \xi}} ds + \int_{t_0}^t \mathbb{E}[\sigma(s, X_s^{t_0, \xi'}, x)]|_{x=X_s^{t_0, \xi}} dW_s \\ &\quad + \int_{t_0}^t \int_{\mathbb{E}} \mathbb{E}[c(s, X_{s-}^{t_0, \xi'}, x, e)]|_{x=X_{s-}^{t_0, \xi}} \tilde{\mu}(de, ds), \end{aligned}$$

where the compensated drift coefficient  $\tilde{b}$  is defined by

$$(2.7) \quad \tilde{b}(t, x', x) = b(t, x', x) + \int_{\mathbb{E}} c(t, x', x, e) \lambda(de).$$

Note that by (2.7) and the assumptions (A1) – (A4), we can conclude that  $\tilde{b}$  satisfies the Lipschitz condition

$$|\tilde{b}(t, x', x) - \tilde{b}(t, y', y)| \leq C(|x' - y'| + |x - y|),$$

as well as the linear growth condition

$$|\tilde{b}(t, x', x)| \leq C(1 + |x'| + |x|),$$

for  $t \in [0, T]$  and  $x, y, x', y' \in \mathbb{R}^d$ . Here  $C$  is a constant depending on  $L, K$  and  $\rho(e)$ .

Based on the two equivalent forms of the MSDEJs (1.1) and (2.6), we will derive two different types of Itô-Taylor schemes for solving MSDEJs.

**3. The Itô formula and Itô-Taylor expansion.** In this section, we develop the Itô formula and Itô-Taylor expansion for MSDEJs, which are the foundation for proposing the Itô-Taylor schemes for MSDEJs.

**3.1. Itô's formula for MSDEJs.** In this subsection, based on the mean-field Itô formula [35], we rigorously prove the Itô's formula for MSDEJs.

Let  $X_t$  be a  $d$ -dimensional Itô process satisfying the MSDE

$$(3.1) \quad dX_t = b^\beta(t, X_t)dt + \sigma^\beta(t, X_t)dW_t, \quad 0 \leq t \leq T$$

with  $\beta_t$  a  $d$ -dimensional Itô process defined as

$$(3.2) \quad d\beta_t = \psi_t dt + \varphi_t dW_t,$$

where  $\psi_t$  and  $\varphi_t$  are two progressively measurable processes such that  $\int_0^T |\psi_t| dt < +\infty$  and  $\int_0^T \text{Tr}[\varphi_s \varphi_s^\top] dt < +\infty$ . Here  $\text{Tr}[A]$  denotes the trace of a matrix  $A$ .

Moreover, we define  $f^\beta(t, x)$  and  $g^\beta(t, x, e)$  by

$$f^\beta(t, x) = \mathbb{E}[f(t, \beta_t, x)], \quad g^\beta(t, x, e) = \mathbb{E}[g(t, \beta_t, x, e)],$$

for functions  $f(t, x', x) : \mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$  and  $g(t, x, x', e) : \mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{E} \rightarrow \mathbb{R}$ . Then we have the following Itô's formula for the MSDE (3.1).

**THEOREM 3.1** (Mean-field Itô formula [35]). *Let  $X_t$  and  $\beta_t$  be  $d$ -dimensional Itô processes satisfying (3.1) and (3.2), respectively, and function  $f = f(t, x', x) \in C^{1,2,2}$ . Then  $f^\beta(t, X_t)$  is an Itô process and satisfies*

$$(3.3) \quad f^\beta(t, X_t) = f^\beta(0, X_0) + \int_0^t L^0 f^\beta(s, X_s) ds + \int_0^t \vec{L}^1 f^\beta(s, X_s) dW_s,$$

where  $L^0$  and  $\vec{L}^1$  are defined by

$$(3.4) \quad \begin{aligned} L^0 f^\beta(s, x) &= \frac{\partial f^\beta}{\partial s}(s, x) + \nabla_x f^\beta(s, x) b^\beta(s, x) \\ &\quad + \frac{1}{2} \text{Tr}[\nabla_{xx} f^\beta(s, x) (\sigma^\beta(s, x)) (\sigma^\beta(s, x))^\top], \\ \vec{L}^1 f^\beta(s, x) &= \nabla_x f^\beta(s, x) \sigma^\beta(s, x) = (L^1 f^\beta(s, x), \dots, L^m f^\beta(s, x)) \end{aligned}$$

with

$$\begin{aligned} L^j f^\beta(t, x) &= \sum_{k=1}^d \frac{\partial f^\beta}{\partial x^k}(t, x) \sigma_{kj}^\beta(t, x), \quad j = 1, \dots, m, \\ \frac{\partial f^\beta}{\partial s}(s, x) &= \mathbb{E} \left[ \frac{\partial f}{\partial s}(s, \beta_s, x) + \nabla_{x'} f(s, \beta_s, x) \psi_s + \frac{1}{2} \text{Tr} [f_{x'x'}(s, \beta_s, x) \varphi_s \varphi_s^\top] \right], \end{aligned}$$

and

$$\nabla_x f^\beta(s, x) = \mathbb{E}[\nabla_x f(s, \beta_s, x)], \quad f_{xx}^\beta(s, x) = \mathbb{E}[f_{xx}(s, \beta_s, x)],$$

where  $\nabla_x f = \left( \frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_d} \right)$  is a  $d$ -dimensional row vector,  $f_{xx} = \left( \frac{\partial^2 f}{\partial x_i \partial x_j} \right)_{d \times d}$  a  $d \times d$  matrix.

Now let  $X_t$  be a  $d$ -dimensional Itô process with jumps satisfying the MSDEJ

$$(3.5) \quad dX_t = b^\beta(t, X_t)dt + \sigma^\beta(t, X_t)dW_t + \int_{\mathbb{E}} c^\beta(t-, X_{t-}, e) \mu(de, dt), \quad 0 \leq t \leq T$$

with  $\beta_t$  a  $d$ -dimensional Itô process with jumps defined by

$$(3.6) \quad d\beta_t = \psi_t dt + \varphi_t dW_t + \int_{\mathbf{E}} h_t \mu(de, dt),$$

where  $h_t$  is a progressively measurable process such that  $\int_{\mathbf{E}} |h_t| \lambda(de) < +\infty$ . Here by the definition of  $g^\beta(t, x, e)$ , we have

$$(3.7) \quad \begin{cases} c^\beta(t, x, e) = \mathbb{E}[c(t, \beta_t, x, e)], \\ c^\beta(t-, x, e) = \mathbb{E}[c(t, \beta_{t-}, x, e)]. \end{cases}$$

**REMARK 3.1.** *By the property of the Poisson process  $N_t = \mu(\mathbf{E} \times [0, t])$ , for any fixed  $t \in [0, T]$ , with probability 1,  $t$  is not a jump time [10]. Then by (3.6), it holds that  $\beta_t = \beta_{t-}$  a.s., which leads to*

$$\mathbb{E}[c(t, \beta_t, x, e)] = \mathbb{E}[c(t, \beta_{t-}, x, e)],$$

that is,

$$(3.8) \quad c^\beta(t, x, e) = c^\beta(t-, x, e).$$

Now we state the Itô's formula for MSDEJs in the following theorem.

**THEOREM 3.2** (Mean-field Itô formula with jumps). *Let  $X_t$  and  $\beta_t$  be two  $d$ -dimensional Itô processes with jumps defined by (3.5) and (3.6), respectively, and function  $f = f(t, x', x) \in C^{1,2,2}$ . Then  $f^\beta(t, X_t)$  is an Itô process with jumps and satisfies*

$$(3.9) \quad \begin{aligned} f^\beta(t, X_t) = & f^\beta(0, X_0) + \int_0^t L^0 f^\beta(s, X_s) ds + \int_0^t \overrightarrow{L}^1 f^\beta(s, X_s) dW_s \\ & + \int_0^t \int_{\mathbf{E}} L_e^{-1} f^\beta(s, X_{s-}) \mu(de, ds), \end{aligned}$$

where  $L^0$  and  $\overrightarrow{L}^1$  are defined by (3.4), and

$$(3.10) \quad L_e^{-1} f^\beta(s, x) = f^\beta(s, x + c^\beta(s, x, e)) - f^\beta(s, x).$$

*Proof.* For simplicity, we consider the case  $d = m = q = 1$ . The general case can be obtained similarly.

Assume that the Poisson random measure  $\mu$  generates a sequence of pairs  $\{(\tau_i, Y_i), i = 1, 2, \dots, N_t\}$ , where  $N_t = \mu(\mathbf{E} \times [0, t])$  represents the total number of jumps of  $\mu$  up to time  $t$ , and  $(\tau_i, Y_i)$  are the  $i$ th jump time and jump size, respectively. Then we can write the MSDEJ (3.5) as

$$(3.11) \quad X_t = X_0 + \int_0^t b^\beta(s, X_s) ds + \int_0^t \sigma^\beta(s, X_s) dW_s + \sum_{i=1}^{N_t} c^\beta(\tau_i, X_{\tau_i-}, Y_i),$$

where  $c^\beta(\tau_i, X_{\tau_i-}, Y_i) = \mathbb{E}[c(\tau_i, \beta_{\tau_i}, x, e)]|_{(x,e)=(X_{\tau_i-}, Y_i)}$ .

Let  $\tau_0 = 0$  and  $\tau_{N_t+1} = t$ , and we have

$$\begin{aligned}
 f^\beta(t, X_t) - f^\beta(0, X_0) &= \sum_{i=0}^{N_t} (f^\beta(\tau_{i+1}, X_{\tau_{i+1}}) - f^\beta(\tau_i, X_{\tau_i})) \\
 (3.12) \qquad &= \sum_{i=0}^{N_t} (f^\beta(\tau_{i+1}, X_{\tau_{i+1}}) - f^\beta(\tau_{i+1}^-, X_{\tau_{i+1}^-})) \\
 &\quad + \sum_{i=0}^{N_t} (f^\beta(\tau_{i+1}^-, X_{\tau_{i+1}^-}) - f^\beta(\tau_i, X_{\tau_i})).
 \end{aligned}$$

Note that,  $X_t$  is a MSDE on each time interval  $[\tau_i, \tau_{i+1})$  for  $i = 0, \dots, N_t$ , then by the mean-field Itô formula (3.3), we obtain

$$(3.13) \qquad f^\beta(\tau_{i+1}^-, X_{\tau_{i+1}^-}) - f^\beta(\tau_i, X_{\tau_i}) = \int_{\tau_i}^{\tau_{i+1}^-} L^0 f^\beta(s, X_s) ds + \int_{\tau_i}^{\tau_{i+1}^-} L^1 f^\beta(s, X_s) dW_s.$$

According to Remark 3.1, at each jump time  $\tau_i$ ,  $i = 1, \dots, N_t$ ,  $f^\beta(t, X_t)$  has a jump

$$(3.14) \qquad f^\beta(\tau_i, X_{\tau_i}) - f^\beta(\tau_i^-, X_{\tau_i^-}) = f^\beta(\tau_i, X_{\tau_i^-} + c^\beta(\tau_i, X_{\tau_i^-}, Y_i)) - f^\beta(\tau_i, X_{\tau_i^-}).$$

Then by (3.12) - (3.14) we get

$$\begin{aligned}
 f^\beta(t, X_t) - f^\beta(0, X_0) &= \sum_{i=0}^{N_t} (f^\beta(\tau_{i+1}^-, X_{\tau_{i+1}^-}) - f^\beta(\tau_i, X_{\tau_i})) \\
 &\quad + \sum_{i=1}^{N_t} (f^\beta(\tau_i, X_{\tau_i}) - f^\beta(\tau_i^-, X_{\tau_i^-})) \\
 &= \sum_{i=0}^{N_t} \left( \int_{\tau_i}^{\tau_{i+1}^-} L^0 f^\beta(s, X_s) ds + \int_{\tau_i}^{\tau_{i+1}^-} L^1 f^\beta(s, X_s) dW_s \right) \\
 &\quad + \sum_{i=1}^{N_t} (f^\beta(\tau_i, X_{\tau_i^-} + c^\beta(\tau_i, X_{\tau_i^-}, Y_i)) - f^\beta(\tau_i, X_{\tau_i^-})) \\
 &= \int_0^t L^0 f^\beta(s, X_s) ds + \int_0^t \vec{L}^1 f^\beta(s, X_s) dW_s \\
 &\quad + \int_0^t \int_{\mathbb{E}} L_e^{-1} f^\beta(s, X_{s-}) \mu(de, ds),
 \end{aligned}$$

where  $L_e^{-1}$  is defined by (3.10). We complete the proof.  $\square$

By using the relationship  $\tilde{\mu}(de, dt) = \mu(de, dt) - \lambda(de)dt$ , we write (3.5) as

$$(3.15) \qquad dX_t = \tilde{b}^\beta(t, X_t)dt + \sigma^\beta(t, X_t)dW_t + \int_{\mathbb{E}} c^\beta(t, X_{t-}, e)\tilde{\mu}(de, dt), \quad 0 \leq t \leq T,$$

where  $\tilde{b}^\beta(t, x) = b^\beta(t, x) + \int_{\mathbb{E}} c^\beta(t, x, e)\lambda(de)$ . Based on Theorem 3.2, we have the Itô's formula for the equivalent MSDEJ (3.15) as below.

**PROPOSITION 3.1.** *Let  $X_t$  and  $\beta_t$  be two  $d$ -dimensional Itô processes with jumps defined by (3.15) and (3.6), respectively, and function  $f = f(t, x', x) \in C^{1,2,2}$ . Then*

$f^\beta(t, X_t)$  is an Itô process with jumps and satisfies

$$(3.16) \quad \begin{aligned} f^\beta(t, X_t) = & f^\beta(0, X_0) + \int_0^t \tilde{L}^0 f^\beta(s, X_s) ds + \int_0^t \overrightarrow{L}^1 f^\beta(s, X_s) dW_s \\ & + \int_0^t \int_{\mathbb{E}} L_e^{-1} f^\beta(s, X_{s-}) \tilde{\mu}(de, ds), \end{aligned}$$

where

$$(3.17) \quad \tilde{L}^0 f^\beta(s, x) = L^0 f^\beta(s, x) + \int_{\mathbb{E}} L_e^{-1} f^\beta(s, x) \lambda(de).$$

Note that (3.16) is an equivalent form of the Itô's formula (3.9) for MSDEJs. Moreover, when  $f$  is independent of  $x'$ , the Itô's formulas (3.9) and (3.16) for MSDEJs reduce to the ones for standard SDEJs [10, 33]. Hence the Itô's formulas for MSDEJs can be seen as a generalization of the ones for SDEJs.

**3.2. Itô-Taylor expansion for MSDEJs.** In this subsection, by utilizing Itô's formula, we construct the Itô-Taylor expansion for MSDEJs. To proceed, we introduce multiple Itô integrals and coefficient functions as below.

**3.2.1. Multiple Itô integrals.** In this subsection, we introduce two types of multiple stochastic integrals.

(A) Multi-indices

Let  $\alpha = (j_1, \dots, j_l)$  be a multi-index with  $j_i \in \{-1, 0, 1, \dots, m\}$ ,  $i = 1, \dots, l$ . Set  $l(\alpha) = l$  to be the length of  $\alpha$ , and let  $\mathcal{M}$  be the set of all multi-indices, i.e.,

$$\mathcal{M} = \left\{ (j_1, j_2, \dots, j_l) : j_i \in \{-1, 0, 1, \dots, m\}, i \in \{1, 2, \dots, l\}, l \in \mathcal{N}^+ \right\} \cup \{v\},$$

where  $v$  is the multi-index of length zero, i.e.,  $l(v) = 0$ . For a given  $\alpha \in \mathcal{M}$  with  $l(\alpha) \geq 1$ ,  $-\alpha$  and  $\alpha-$  are two multi-indices obtained by deleting the first and the last component of  $\alpha$ , respectively. We also denote by

$$\begin{aligned} n(\alpha) &: \text{the number of the components of } \alpha \text{ equal to } 0, \\ s(\alpha) &: \text{the number of the components of } \alpha \text{ equal to } -1. \end{aligned}$$

Moreover, for a given  $\alpha \in \mathcal{M}$ , let  $e = (e_1, \dots, e_{s(\alpha)})$  denote a vector  $e \in \mathbb{E}^{s(\alpha)}$ .

(B) Multiple integrals

For a given  $\alpha \in \mathcal{M}$ , we define the multiple Itô integral operator  $I_\alpha$  on the adapted right continuous processes  $\{f = f(t, e_1, \dots, e_{s(\alpha)}), t \geq 0\}$  with left limits by

$$(3.18) \quad I_\alpha[f(\cdot)]_{\rho, \tau} := \begin{cases} f(\tau), & l = 0, \\ \int_\rho^\tau I_{\alpha-}[f(\cdot)]_{\rho, s} ds, & l \geq 1 \text{ and } j_l = 0, \\ \int_\rho^\tau I_{\alpha-}[f(\cdot)]_{\rho, s} dW_s^{j_l}, & l \geq 1 \text{ and } j_l \geq 1, \\ \int_\rho^\tau \int_{\mathbb{E}} I_{\alpha-}[f(\cdot)]_{\rho, s-} \mu(de_{s(\alpha)}, ds), & l \geq 1 \text{ and } j_l = -1, \end{cases}$$

where  $\rho$  and  $\tau$  are two stopping times satisfying  $0 \leq \rho \leq \tau \leq T$ , a.s. and all the

integrals exist. For instance,

$$\begin{aligned} I_v[f(\cdot)]_{0,t} &= f(t), \quad I_{(-1)}[f(\cdot)]_{0,t} = \int_0^t \int_{\mathbf{E}} f(s-, e) \mu(de, ds), \\ I_{(0)}[f(\cdot)]_{0,t} &= \int_0^t f(s) ds, \quad I_{(1)}[f(\cdot)]_{0,t} = \int_0^t f(s) dW_s^1, \\ I_{(-1,-1)}[f(\cdot)]_{0,t} &= \int_0^t \int_{\mathbf{E}} \int_0^{s_2^-} \int_{\mathbf{E}} f(s_1-, e_1, e_2) \mu(de_1, ds_1) \mu(de_2, ds_2). \end{aligned}$$

(C) Compensated multiple integrals

Replace  $\mu$  with  $\tilde{\mu}$  in (3.18), and we get the compensated multiple Itô integral

$$\tilde{I}_\alpha[f(\cdot)]_{\rho,\tau} := \begin{cases} f(\tau), & l = 0, \\ \int_\rho^\tau \tilde{I}_{\alpha-}[f(\cdot)]_{\rho,s} ds, & l \geq 1 \text{ and } j_l = 0, \\ \int_\rho^\tau \tilde{I}_{\alpha-}[f(\cdot)]_{\rho,s} dW_s^{j_l}, & l \geq 1 \text{ and } j_l \geq 1, \\ \int_\rho^\tau \int_{\mathbf{E}} \tilde{I}_{\alpha-}[f(\cdot)]_{\rho,s-} \tilde{\mu}(de_{s(\alpha)}, ds), & l \geq 1 \text{ and } j_l = -1, \end{cases}$$

where  $f(\cdot) = f(\cdot, e_1, \dots, e_{s(\alpha)})$ . For instance,

$$\begin{aligned} \tilde{I}_\alpha[f(\cdot)]_{\rho,\tau} &= I_\alpha[f(\cdot)]_{\rho,\tau}, \quad \text{when } s(\alpha) = 0, \\ \tilde{I}_{(1,-1)}[f(\cdot)]_{0,t} &= \int_0^t \int_{\mathbf{E}} \int_0^{s_2^-} f(s_1, e) dW_{s_1}^1 \tilde{\mu}(de, ds_2). \end{aligned}$$

**3.2.2. Coefficient functions.** For a given function

$$f(t, x', x, e) : \mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbf{E} \rightarrow \mathbb{R},$$

by (3.4) and (3.10), we have

$$\begin{aligned} (3.19) \quad L^0 f^\beta(s, x, e) &= \frac{\partial f^\beta}{\partial s}(s, x, e) + \sum_{k=1}^d b_k^\beta(s, x) \frac{\partial f^\beta}{\partial x^k}(t, x, e) \\ &+ \frac{1}{2} \sum_{i,k=1}^d \sum_{j=1}^d \sigma_{ij}^\beta(t, x) \sigma_{kj}^\beta(s, x) \frac{\partial^2 f^\beta}{\partial x^i \partial x^k}(t, x, e), \end{aligned}$$

with

$$\frac{\partial f^\beta}{\partial s}(s, x, e) = \mathbb{E} \left[ \frac{\partial f}{\partial s}(s, \beta_s, x, e) + \nabla_{x'} f(s, \beta_s, x, e) \psi_s + \frac{1}{2} \text{Tr} [f_{x'x'}(s, \beta_s, x, e) \varphi_s \varphi_s^\top] \right],$$

and

$$\begin{aligned} (3.20) \quad L^j f^\beta(t, x, e) &= \sum_{k=1}^d \frac{\partial f^\beta}{\partial x^k}(t, x, e) \sigma_{kj}^\beta(t, x), \quad j = 1, \dots, m, \\ L_{e_2}^{-1} f^\beta(s, x, e_1) &= f^\beta(s, x + c^\beta(s, x, e_2), e_1) - f^\beta(s, x, e_1). \end{aligned}$$

Then by (3.17), we obtain

$$(3.21) \quad \tilde{L}^0 f^\beta(s, x, e) = L^0 f^\beta(s, x, e) + \int_{\mathbf{E}} L_{e_1}^{-1} f^\beta(t, x, e) \lambda(de_1).$$

Based on (3.19) - (3.21), we present the following two types of coefficient functions and hierarchical and remainder sets.

## (C) Itô coefficient functions

For a given  $\alpha = (j_1, \dots, j_l) \in \mathcal{M}$  and a smooth function  $f(t, x', x)$ , we define the coefficient function  $f_\alpha^\beta$  by

$$(3.22) \quad f_\alpha^\beta(t, x, e) := \begin{cases} f^\beta(t, x), & l = 0, \\ L^{j_1} f_{-\alpha}^\beta(t, x, e_1, \dots, e_{s(-\alpha)}), & l \geq 1 \text{ and } j_1 \geq 0, \\ L_{e_{s(\alpha)}}^{-1} f_{-\alpha}^\beta(t, x, e_1, \dots, e_{s(-\alpha)}), & l \geq 1 \text{ and } j_1 = -1, \end{cases}$$

where  $e = (e_1, \dots, e_{s(\alpha)}) \in \mathbf{E}^{s(\alpha)}$ . The dependence on  $e$  in (3.22) is introduced by the repeated application of the operator  $L_e^{-1}$  in (3.20). Take  $m = d = q = 1$  and let  $f(t, x', x) = x$ , then we can deduce the following examples

$$\begin{aligned} f_{(0)}^\beta(t, x) &= b^\beta(t, x), & f_{(1)}^\beta(t, x) &= \sigma^\beta(t, x), & f_{(-1)}^\beta(t, x, e) &= c^\beta(t, x, e), \\ f_{(-1, -1)}^\beta(t, x, e_1, e_2) &= L_{e_2}^{-1} c^\beta(t, x, e_1) = c^\beta(t, x + c^\beta(t, x, e_2), e_1) - c^\beta(t, x, e_1). \end{aligned}$$

## (D) Compensated Itô coefficient functions

By replacing  $L^0$  with  $\tilde{L}^0$  in (3.22), we get the compensated Itô coefficient functions

$$(3.23) \quad \tilde{f}_\alpha^\beta(t, x, e) := \begin{cases} f^\beta(t, x), & l = 0, \\ \tilde{L}^0 \tilde{f}_{-\alpha}^\beta(t, x, e_1, \dots, e_{s(-\alpha)}), & l \geq 1 \text{ and } j_l = 0, \\ L^{j_l} \tilde{f}_{-\alpha}^\beta(t, x, e_1, \dots, e_{s(-\alpha)}), & l \geq 1 \text{ and } j_l \geq 1, \\ L_{e_{s(\alpha)}}^{-1} \tilde{f}_{-\alpha}^\beta(t, x, e_1, \dots, e_{s(-\alpha)}), & l \geq 1 \text{ and } j_l = -1, \end{cases}$$

For instance, let  $f(t, x', x) = x$  and we have

$$\begin{aligned} \tilde{f}_\alpha^\beta(t, x, e) &= f_\alpha^\beta(t, x, e), & \text{when } n(\alpha) &= 0, \\ \tilde{f}_{(-1, 0)}^\beta(t, x, e) &= L_e^{-1} \tilde{b}^\beta(t, x) = b^\beta(t, x + c^\beta(t, x, e)) - b^\beta(t, x) \\ &\quad + \int_{\mathbf{E}} (c^\beta(t, x + c^\beta(t, x, e), e_1) - c^\beta(t, x, e_1)) \lambda(de_1). \end{aligned}$$

Here we have assumed that the functions  $b, \sigma, c$  and  $f$  satisfy all the smoothness and integrability conditions needed in the definitions of (3.22) and (3.23).

## (E) Hierarchical and remainder sets

We call a subset  $\mathcal{A} \subset \mathcal{M}$  a hierarchical set if it satisfies

$$\mathcal{A} \neq \emptyset, \quad \sup_{\alpha \in \mathcal{A}} l(\alpha) < \infty, \quad \text{and } -\alpha \in \mathcal{A} \text{ for each } \alpha \in \mathcal{A} \setminus \{v\};$$

and its remainder set  $\mathcal{B}(\mathcal{A})$  is defined by

$$\mathcal{B}(\mathcal{A}) = \{\alpha \in \mathcal{M} \setminus \mathcal{A} : -\alpha \in \mathcal{A}\}.$$

Take  $m = 1$  for instance and we give two hierarchical sets

$$\begin{aligned} \mathcal{A}_0 &= \{v, (-1), (0), (1)\}, \\ \mathcal{A}_1 &= \{v, (-1), (0), (1), (1, 1), (1, -1), (-1, 1), (-1, -1)\}, \end{aligned}$$

and their remainder sets are

$$\begin{aligned} \mathcal{B}(\mathcal{A}_0) &= \{(-1, -1), (0, -1), (1, -1), (-1, 0), (0, 0), (1, 0), (-1, 1), (0, 1), (1, 1)\}, \\ \mathcal{B}(\mathcal{A}_1) &= \{(0, -1), (-1, 0), (0, 0), (1, 0), (0, 1), (-1, 1, 1), (0, 1, 1), \\ &\quad (-1, 1, -1), (0, 1, -1), (1, 1, -1), (-1, -1, 1), (0, -1, 1), \\ &\quad (-1, 1, 1), (1, -1, 1), (-1, -1, -1), (0, -1, -1), (1, -1, -1)\}. \end{aligned}$$

**3.2.3. The Itô-Taylor expansion.** In this subsection, by using the Itô's formula (3.9) and (3.16) for MSDEJs, we present the Itô-Taylor expansions of

$$f^\beta(t, X_t) = \mathbb{E}[f(t, \beta_t, x)]|_{x=X_t}$$

for the solution  $X_t$  of the MSDEJ (3.5) with  $\beta_t$  defined by (3.6) satisfying

$$\mathbb{E} \left[ \int_0^T \int_{\mathbb{E}} |\beta_t|^2 \lambda(de) dt \right] < \infty.$$

Now we state the Itô-Taylor expansions for MSDEJs in the following theorem.

**THEOREM 3.3.** *Let  $\rho$  and  $\tau$  be two stopping times with  $0 \leq \rho \leq \tau \leq T$ , a.s.. Then for a given hierarchical set  $\mathcal{A} \subset \mathcal{M}$  and a function  $f(t, x', x) : \mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ , we have the Itô-Taylor expansion*

$$(3.24) \quad f^\beta(\tau, X_\tau) = \sum_{\alpha \in \mathcal{A}} I_\alpha [f_\alpha^\beta(\rho, X_\rho)]_{\rho, \tau} + \sum_{\alpha \in \mathcal{B}(\mathcal{A})} I_\alpha [f_\alpha^\beta(\cdot, X_\cdot)]_{\rho, \tau},$$

and the compensated Itô-Taylor expansion

$$(3.25) \quad f^\beta(\tau, X_\tau) = \sum_{\alpha \in \mathcal{A}} \tilde{I}_\alpha [\tilde{f}_\alpha^\beta(\rho, X_\rho)]_{\rho, \tau} + \sum_{\alpha \in \mathcal{B}(\mathcal{A})} \tilde{I}_\alpha [\tilde{f}_\alpha^\beta(\cdot, X_\cdot)]_{\rho, \tau},$$

provided that all of the coefficient functions  $f_\alpha^\beta$  and  $\tilde{f}_\alpha^\beta$  are well defined and all of the multiple Itô integrals exist.

*Proof.* By an iterated application of the Itô's formulas (3.9) and (3.16), the proof of the above theorem is analogous to the ones of the Itô-Taylor expansions for standard SDEs [17] and SDEJs [33]. So we omit it here.  $\square$

We list some remarks for the Itô-Taylor expansions for MSDEJs as below.

- For notational simplicity, we have suppressed the dependence on  $e \in \mathbf{E}^{s(\alpha)}$  in the coefficients  $f_\alpha$  and  $\tilde{f}_\alpha$  in (3.24) and (3.25).
- When  $f$  is independent of  $x'$ , the Itô-Taylor expansions (3.24) and (3.25) for MSDEJs reduce to the ones for standard SDEJs [33]. Hence, the Itô-Taylor expansions for MSDEJs can be seen as a generalization of the ones for SDEJs.

**4. Itô-Taylor schemes for MSDEJs.** In this section, based on the Itô-Taylor expansions (3.24) and (3.25), we propose the general Itô-Taylor schemes for solving the MSDEJ (1.1).

Without loss of generality, we let  $t_0 = 0$  and  $\xi = \xi'$  in (1.1). Then omit the superscript  $t_0, \xi$ , and we get

$$(4.1) \quad \begin{aligned} X_t = X_0 &+ \int_0^t \mathbb{E}[b(s, X_s, x)]|_{x=X_s} ds + \int_0^t \mathbb{E}[\sigma(s, X_s, x)]|_{x=X_s} dW_s \\ &+ \int_0^t \int_{\mathbb{E}} \mathbb{E}[c(s, X_s, x, e)]|_{x=X_{s-}} \mu(de, ds). \end{aligned}$$

Note that the MSDEJ (4.1) has an equivalent form

$$\begin{aligned} X_t = X_0 &+ \int_0^t \mathbb{E}[\tilde{b}(s, X_s, x)]|_{x=X_s} ds + \int_0^t \mathbb{E}[\sigma(s, X_s, x)]|_{x=X_s} dW_s \\ &+ \int_0^t \int_{\mathbb{E}} \mathbb{E}[c(s, X_s, x, e)]|_{x=X_{s-}} \tilde{\mu}(de, ds), \end{aligned}$$

where  $\tilde{b}$  is defined by (2.7).

By choosing different hierarchical sets  $\mathcal{A}$  in the Itô-Taylor expansion (3.24), we shall derive the two types of strong order  $\gamma$  and weak order  $\eta$  Itô-Taylor schemes for solving the MSDEJ (4.1). To this end, we take a uniform time partition on  $[0, T]$ :

$$0 = t_0 < t_1 < \cdots < t_N = T,$$

where  $t_{k+1}$  is  $\mathcal{F}_{t_k}$ -measurable for  $k = 0, 1, \dots, N-1$ .

Let  $X_k$  be the approximation of the solution  $X_t$  of (4.1) at time  $t = t_k$ , and denote by

$$\begin{aligned} f^{X_k}(t_k, X_k) &= \mathbb{E}[f(t_k, X_k, x)]|_{x=X_k}, \\ g^{X_k}(t_k, X_k, e) &= \mathbb{E}[g(t_k, X_k, x, e)]|_{x=X_k}, \end{aligned}$$

for  $f(t, x', x) : \mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$  and  $g(t, x', x, e) : \mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbf{E} \rightarrow \mathbb{R}$ .

**4.1. Strong Itô-Taylor schemes.** To construct the strong Itô-Taylor schemes for the MSDEJ (4.1), for  $\gamma = 0.5, 1.0, 1.5, \dots$ , we define the hierarchical set  $\mathcal{A}_\gamma$  by

$$\mathcal{A}_\gamma = \left\{ \alpha \in \mathcal{M} : l(\alpha) + n(\alpha) \leq 2\gamma \text{ or } l(\alpha) = n(\alpha) = \gamma + \frac{1}{2} \right\}$$

and denote its remainder set by  $\mathcal{B}(\mathcal{A}_\gamma)$ . Take  $f(t, x', x) = x$  and let  $\beta_t = X_t$ , then by Theorem 3.3, for  $k = 0, 1, \dots, N-1$ , we have the Itô-Taylor expansion

$$(4.2) \quad X_{t_{k+1}} = \sum_{\alpha \in \mathcal{A}_\gamma} I_\alpha [f_\alpha^X(t_k, X_{t_k})]_{t_k, t_{k+1}} + \sum_{\alpha \in \mathcal{B}(\mathcal{A}_\gamma)} I_\alpha [f_\alpha^X(\cdot, X)]_{t_k, t_{k+1}},$$

and the compensated Itô-Taylor expansion

$$(4.3) \quad X_{t_{k+1}} = \sum_{\alpha \in \mathcal{A}_\gamma} \tilde{I}_\alpha [\tilde{f}_\alpha^X(t_k, X_{t_k})]_{t_k, t_{k+1}} + \sum_{\alpha \in \mathcal{B}(\mathcal{A}_\gamma)} \tilde{I}_\alpha [\tilde{f}_\alpha^X(\cdot, X)]_{t_k, t_{k+1}}.$$

By removing the remainder term in (4.2), we propose the following general strong order  $\gamma$  Itô-Taylor scheme for solving the MSDEJ (4.1).

**Scheme 4.1** (Strong order  $\gamma$  Itô-Taylor scheme).

$$(4.4) \quad X_{k+1} = \sum_{\alpha \in \mathcal{A}_\gamma} I_\alpha [f_\alpha^{X_k}(t_k, X_k)]_{t_k, t_{k+1}}.$$

Similarly, we propose the compensated strong order  $\gamma$  Itô-Taylor scheme.

**Scheme 4.2** (Compensated strong order  $\gamma$  Itô-Taylor scheme).

$$(4.5) \quad X_{k+1} = \sum_{\alpha \in \mathcal{A}_\gamma} \tilde{I}_\alpha [\tilde{f}_\alpha^{X_k}(t_k, X_k)]_{t_k, t_{k+1}}.$$

Based on Schemes 4.1 and 4.2, by taking  $\gamma = 0.5$  and  $1.0$ , we will give some specific strong Taylor schemes for MSDEJs in the following subsections.

**4.1.1. The Euler scheme.** Take  $\gamma = 0.5$  in Scheme 4.1, and we have

$$\mathcal{A}_{0.5} = \{v, (-1), (0), (1)\}$$

and

$$\mathcal{B}(\mathcal{A}_{0.5}) = \{(-1, -1), (0, -1), (1, -1), (-1, 0), (0, 0), (1, 0), (-1, 1), (0, 1), (1, 1)\}.$$

Then by the Itô-Taylor expansion (4.2), for  $k = 0, \dots, N - 1$ , we obtain

$$(4.6) \quad \begin{aligned} X_{t_{k+1}} &= X_{t_k} + b^X(t_k, X_{t_k}) \int_{t_k}^{t_{k+1}} ds + \sigma^X(t_k, X_{t_k}) \int_{t_k}^{t_{k+1}} dW_s \\ &\quad + \int_{t_k}^{t_{k+1}} \int_{\mathbb{E}} c^X(t_k, X_{t_k}, e) \mu(de, ds) + R_1, \end{aligned}$$

where

$$\begin{aligned} R_1 &= \int_{t_k}^{t_{k+1}} \int_{t_k}^s L^0 b^X(z, X_z) dz ds + \int_{t_k}^{t_{k+1}} \int_{t_k}^s L^1 b^X(z, X_z) dW_z ds \\ &\quad + \int_{t_k}^{t_{k+1}} \int_{t_k}^s L^0 \sigma^X(z, X_z) dz dW_s + \int_{t_k}^{t_{k+1}} \int_{t_k}^s L^1 \sigma^X(z, X_z) dW_z dW_s \\ &\quad + \int_{t_k}^{t_{k+1}} \int_{\mathbb{E}} \int_{t_k}^{s-} L^0 c^X(z, X_z, e) dz \mu(de, ds) \\ &\quad + \int_{t_k}^{t_{k+1}} \int_{t_k}^s \int_{\mathbb{E}} L_e^{-1} b^X(z, X_{z-}) \mu(de, dz) ds \\ &\quad + \int_{t_k}^{t_{k+1}} \int_{\mathbb{E}} \int_{t_k}^{s-} L^1 c^X(z, X_z, e) dW_z \mu(de, ds) \\ &\quad + \int_{t_k}^{t_{k+1}} \int_{t_k}^s \int_{\mathbb{E}} L_e^{-1} \sigma^X(z, X_{z-}) \mu(de, dz) dW_s \\ &\quad + \int_{t_k}^{t_{k+1}} \int_{\mathbb{E}} \int_{t_k}^{s-} \int_{\mathbb{E}} L_{e_2}^{-1} c^X(z, X_{z-}, e_1) \mu(de_1, dz) \mu(de_2, ds). \end{aligned}$$

Remove the remainder term  $R_1$  in (4.6), and we get the strong order 0.5 Itô-Taylor scheme for solving the MSDEJ (4.1)

$$(4.7) \quad \begin{aligned} X_{k+1} &= X_k + b^{X_k}(t_k, X_k) \Delta t_k + \sigma^{X_k}(t_k, X_k) \Delta W_k + \int_{t_k}^{t_{k+1}} \int_{\mathbb{E}} c^{X_k}(t_k, X_k, e) \mu(de, dt) \\ &= X_k + b^{X_k}(t_k, X_k) \Delta t_k + \sigma^{X_k}(t_k, X_k) \Delta W_k + \sum_{i=N_{t_k}+1}^{N_{t_{k+1}}} c^{X_k}(t_k, X_k, Y_i), \end{aligned}$$

which is the so-called Euler scheme. Here

$$\Delta t_k = t_{k+1} - t_k \quad \text{and} \quad \Delta W_k = W_{t_{k+1}} - W_{t_k},$$

and  $N_t = \mu(\mathbb{E} \times [0, t])$  is a Poisson process counting the number of jumps of  $\mu$  up to time  $t$  and  $(\tau_i, Y_i)$  are the  $i$ th jump time and jump size.

**4.1.2. The strong order 1.0 Itô-Taylor scheme.** Take  $\gamma = 1$  in Scheme 4.1, and we have

$$\mathcal{A}_1 = \{v, (-1), (0), (1), (1, 1), (1, -1), (-1, 1), (-1, -1)\}$$

and

$$\begin{aligned} \mathcal{B}(\mathcal{A}_1) = \{ & (0, -1), (-1, 0), (0, 0), (1, 0), (0, 1), (-1, 1, 1), (0, 1, 1), \\ & (-1, 1, 1), (-1, 1, -1), (0, 1, -1), (1, 1, -1), (-1, -1, 1), \\ & (0, -1, 1), (1, -1, 1), (-1, -1, -1), (0, -1, -1), (1, -1, -1)\}. \end{aligned}$$

Then by (4.4), we get the strong order 1.0 Itô-Taylor scheme (4.8)

$$\begin{aligned} X_{k+1} = & X_k + b^{X_k}(t_k, X_k)\Delta t_k + \sigma^{X_k}(t_k, X_k)\Delta W_k + \int_{t_k}^{t_{k+1}} \int_{\mathbb{E}} c^{X_k}(t_k, X_k, e)\mu(de, dt) \\ & + \int_{t_k}^{t_{k+1}} \int_{t_k}^s L^1 \sigma^{X_k}(t_k, X_k) dW_z dW_s \\ & + \int_{t_k}^{t_{k+1}} \int_{\mathbb{E}} \int_{t_k}^{s^-} L^1 c^{X_k}(t_k, X_k, e) dW_z \mu(de, ds) \\ & + \int_{t_k}^{t_{k+1}} \int_{t_k}^s \int_{\mathbb{E}} L_e^{-1} \sigma^{X_k}(t_k, X_k) \mu(de, dz) dW_s \\ & + \int_{t_k}^{t_{k+1}} \int_{\mathbb{E}} \int_{t_k}^{s^-} \int_{\mathbb{E}} L_{e_1}^{-1} c^{X_k}(t_k, X_k, e_2) \mu(de_1, dz) \mu(de_2, ds). \end{aligned}$$

Combining with the Itô's formula (3.9) for MSDEJs and the properties of jump times, the scheme (4.8) can be written as

$$\begin{aligned} X_{k+1} = & X_k + b^{X_k}(t_k, X_k)\Delta t_k + \sigma^{X_k}(t_k, X_k)\Delta W_k + \sum_{i=N_{t_k}+1}^{N_{t_{k+1}}} c^{X_k}(t_k, X_k, Y_i) \\ & + \frac{1}{2} L^1 \sigma^{X_k}(t_k, X_k) ((\Delta W_k)^2 - \Delta t_k) + \sum_{i=N_{t_k}+1}^{N_{t_{k+1}}} L^1 c^{X_k}(t_k, X_k, Y_i) (W_{\tau_i} - W_{t_k}) \\ & + \sum_{i=N_{t_k}+1}^{N_{t_{k+1}}} \left( \sigma^{X_k}(t_k, X_k + c^{X_k}(t_k, X_k, Y_i)) - \sigma^{X_k}(t_k, X_k) \right) (W_{t_{k+1}} - W_{\tau_i}) \\ & + \sum_{i=N_{t_k}+1}^{N_{t_{k+1}}} \sum_{j=N_{t_k}+1}^{N_{\tau_i}-} \left( c^{X_k}(t_k, X_k + c^{X_k}(t_k, X_k, Y_j), Y_i) - c^{X_k}(t_k, X_k, Y_i) \right), \end{aligned}$$

which is readily applicable for scenario simulation.

Based on Scheme 4.2, we can get the compensated Euler scheme and the compensated strong order 1.0 Itô-Taylor scheme in the same way. It is also worth noting that the compensated Euler scheme and the Euler scheme are the same.

**4.2. Weak Itô-Taylor schemes.** To construct the weak Itô-Taylor schemes for MSDEJs, for  $\eta = 1.0, 2.0, \dots$ , we define the hierarchical set  $\Gamma_\eta$  by

$$(4.9) \quad \Gamma_\eta = \{\alpha \in \mathcal{M} : l(\alpha) \leq \eta\}$$

and denote its remainder set by  $\mathcal{B}(\Gamma_\eta)$ . Take  $f(t, x', x) = x$  and let  $\beta_t = X_t$ , then by Theorem 3.3, for  $k = 0, 1, \dots, N-1$ , we have

$$(4.10) \quad X_{t_{k+1}} = \sum_{\alpha \in \Gamma_\eta} I_\alpha [f_\alpha^X(t_k, X_{t_k})]_{t_k, t_{k+1}} + \sum_{\alpha \in \mathcal{B}(\Gamma_\eta)} I_\alpha [f_\alpha^X(\cdot, X_\cdot)]_{t_k, t_{k+1}},$$

and

$$(4.11) \quad X_{t_{k+1}} = \sum_{\alpha \in \Gamma_\eta} \tilde{I}_\alpha [\tilde{f}_\alpha^X(t_k, X_{t_k})]_{t_k, t_{k+1}} + \sum_{\alpha \in \mathcal{B}(\Gamma_\eta)} \tilde{I}_\alpha [\tilde{f}_\alpha^X(\cdot, X_\cdot)]_{t_k, t_{k+1}}.$$

Remove the remainder term in (4.10), and we propose the following general weak order  $\eta$  Itô-Taylor scheme for solving the MSDEJ (4.1).

**Scheme 4.3** (Weak order  $\eta$  Itô-Taylor scheme).

$$(4.12) \quad X_{k+1} = \sum_{\alpha \in \Gamma_\eta} I_\alpha [f_\alpha^{X_k}(t_k, X_k)]_{t_k, t_{k+1}}.$$

Similarly, we propose the compensated weak order  $\eta$  Itô-Taylor scheme.

**Scheme 4.4** (Compensated weak order  $\eta$  Itô-Taylor scheme).

$$(4.13) \quad X_{k+1} = \sum_{\alpha \in \Gamma_\eta} \tilde{I}_\alpha [\tilde{f}_\alpha^{X_k}(t_k, X_k)]_{t_k, t_{k+1}}.$$

Based on Schemes 4.3 and 4.4, by taking  $\eta = 1.0$  and  $2.0$ , we will present some specific weak Taylor schemes for MSDEJs.

**4.2.1. The Euler scheme.** Taking  $\eta = 1.0$  in (4.9) leads to

$$\Gamma_{1.0} = \{v, (-1), (0), (1)\}$$

and

$$\mathcal{B}(\Gamma_{1.0}) = \{(-1, -1), (0, -1), (1, -1), (-1, 0), (0, 0), (1, 0), (-1, 1), (0, 1), (1, 1)\}.$$

Then by Scheme 4.3, we get the Euler scheme (4.7)

$$X_{k+1} = X_k + b^{X_k}(t_k, X_k)\Delta t_k + \sigma^{X_k}(t_k, X_k)\Delta W_k + \sum_{i=N_{t_k}+1}^{N_{t_{k+1}}} c^{X_k}(t_k, X_k, Y_i),$$

which is also the weak order 1.0 Itô-Taylor scheme.

**4.2.2. The weak order 2.0 Itô-Taylor scheme.** Taking  $\eta = 2.0$  in (4.9) gives

$$\Gamma_{2.0} = \{v, (-1), (0), (1), (1, 1), (1, -1), (-1, 1), (-1, -1), (0, 0), (1, 0), (0, 1), (-1, 0), (0, -1)\},$$

and

$$\begin{aligned} \mathcal{B}(\Gamma_{2.0}) = \{ & (-1, 1, 1), (0, 1, 1), (-1, 1, 1), (-1, 1, -1), (0, 1, -1), (1, 1, -1), (-1, -1, 1), \\ & (0, -1, 1), (1, -1, 1), (-1, -1, -1), (0, -1, -1), (1, -1, -1), (-1, 0, 0), \\ & (0, 0, 0), (1, 0, 0), (-1, 1, 0), (0, 1, 0), (1, 1, 0), (-1, 0, 1), (0, 0, 1), (1, 0, 1), \\ & (-1, -1, 0), (0, -1, 0), (1, -1, 0), (-1, 0, -1), (0, 0, -1), (1, 0, -1)\}. \end{aligned}$$

Then by Scheme 4.3, we get the weak order 2.0 Itô-Taylor scheme

$$\begin{aligned}
X_{k+1} &= X_k + b^{X_k}(t_k, X_k)\Delta t_k + \sigma^{X_k}(t_k, X_k)\Delta W_k + \int_{t_k}^{t_{k+1}} \int_{\mathbb{E}} c^{X_k}(t_k, X_k, e)\mu(de, dt) \\
&\quad + \int_{t_k}^{t_{k+1}} \int_{t_k}^s L^1 \sigma^{X_k}(t_k, X_k) dW_z dW_s \\
&\quad + \int_{t_k}^{t_{k+1}} \int_{\mathbb{E}} \int_{t_k}^{s-} L^1 c^{X_k}(t_k, X_k, e) dW_z \mu(de, ds) \\
&\quad + \int_{t_k}^{t_{k+1}} \int_{t_k}^s \int_{\mathbb{E}} L_e^{-1} \sigma^{X_k}(t_k, X_k) \mu(de, dz) dW_s \\
&\quad + \int_{t_k}^{t_{k+1}} \int_{\mathbb{E}} \int_{t_k}^{s-} \int_{\mathbb{E}} L_{e_1}^{-1} c^{X_k}(t_k, X_k, e_2) \mu(de_1, dz) \mu(de_2, ds) \\
&\quad + \int_{t_k}^{t_{k+1}} \int_{t_k}^s L^0 b^{X_k}(t_k, X_k) dz ds + \int_{t_k}^{t_{k+1}} \int_{t_k}^s L^1 b^{X_k}(t_k, X_k) dW_z ds \\
&\quad + \int_{t_k}^{t_{k+1}} \int_{t_k}^s L^0 \sigma^{X_k}(t_k, X_k) dz dW_s + \int_{t_k}^{t_{k+1}} \int_{\mathbb{E}} \int_{t_k}^{s-} L^0 c^{X_k}(t_k, X_k, e) dz \mu(de, ds) \\
(4.14) \quad &\quad + \int_{t_k}^{t_{k+1}} \int_{t_k}^s \int_{\mathbb{E}} L_e^{-1} b^{X_k}(t_k, X_k) \mu(de, dz) ds.
\end{aligned}$$

Combining with the Itô's formula (3.9) for MSDEJs and the properties of jump times, we can rewrite (4.14) as

$$\begin{aligned}
X_{k+1} &= X_k + b^{X_k}(t_k, X_k)\Delta t_k + \sigma^{X_k}(t_k, X_k)\Delta W_k + \sum_{i=N_{t_k}+1}^{N_{t_{k+1}}} c^{X_k}(t_k, X_k, Y_i) \\
&\quad + \frac{1}{2} L^1 \sigma^{X_k}(t_k, X_k) ((\Delta W_k)^2 - \Delta t_k) + \sum_{i=N_{t_k}+1}^{N_{t_{k+1}}} L^1 c^{X_k}(t_k, X_k, Y_i) (W_{\tau_i} - W_{t_k}) \\
&\quad + \sum_{i=N_{t_k}+1}^{N_{t_{k+1}}} \left( \sigma^{X_k}(t_k, X_k + c^{X_k}(t_k, X_k, Y_i)) - \sigma^{X_k}(t_k, X_k) \right) (W_{t_{k+1}} - W_{\tau_i}) \\
&\quad + \sum_{i=N_{t_k}+1}^{N_{t_{k+1}}} \sum_{j=N_{t_k}+1}^{N_{\tau_i}-} \left( c^{X_k}(t_k, X_k + c^{X_k}(t_k, X_k, Y_j), Y_i) - c^{X_k}(t_k, X_k, Y_i) \right) \\
&\quad + \frac{1}{2} L^0 b^{X_k}(t_k, X_k) (\Delta t_k)^2 + L^1 b^{X_k}(t_k, X_k) \Delta Z_k \\
&\quad + L^0 \sigma^{X_k}(t_k, X_k) (\Delta W_k \Delta t_k - \Delta Z_k) + \sum_{i=N_{t_k}+1}^{N_{t_{k+1}}} L^0 c^{X_k}(t_k, X_k, Y_i) (\tau_i - t_k) \\
&\quad + \sum_{i=N_{t_k}+1}^{N_{t_{k+1}}} \left( b^{X_k}(t_k, X_k + c^{X_k}(t_k, X_k, Y_i)) - b^{X_k}(t_k, X_k) \right) (t_{k+1} - \tau_i),
\end{aligned}$$

where  $\Delta Z_k$  is a random variable defined by

$$\Delta Z_k = \int_{t_k}^{t_{k+1}} \int_{t_k}^s dW_z ds = \Delta W_k \Delta t_k - \int_{t_k}^{t_{k+1}} \int_{t_k}^s dz dW_s.$$

Similarly, based on Scheme 4.4, we can obtain the compensated weak order 2.0 Itô-Taylor scheme for solving the MSDEJ (4.1).

Now, we illustrate how to use the proposed Itô-Taylor schemes to solve the MSDEJ (1.1) with different initial values  $\xi$  and  $\xi'$ . Without loss of generality, we set  $(t_0, \xi') = (0, x_0)$  and  $(t_0, \xi) = (0, X_0)$ . Let  $X_n^{x_0}$  and  $X_n^{X_0}$  denote the numerical solutions of the Itô-Taylor schemes with initial values of  $x_0$  and  $X_0$ , respectively. Then take the Euler scheme (4.7) for instance, and we can rewrite it as

$$(4.15) \quad \begin{aligned} X_{k+1}^{X_0} &= X_k^{X_0} + b^{X_k^{x_0}}(t_k, X_k^{X_0})\Delta t_k + \sigma^{X_k^{x_0}}(t_k, X_k^{X_0})\Delta W_k \\ &+ \sum_{i=N_{t_k}+1}^{N_{t_{k+1}}} c^{X_k^{x_0}}(t_k, X_k^{X_0}, Y_i), \end{aligned}$$

where  $f^{X_k^{x_0}}(t_k, x, e)$  denote  $\mathbb{E}[f(t_k, X_k^{x_0}, x, e)]$  for  $f = b, \sigma$  and  $c$ . Now we can apply the scheme (4.15) to solve the MSDEJ (1.1) by the following two procedures

- i) Take  $X_0 = x_0$  and we solve (1.1) by the scheme (4.15) to obtain  $\{X_n^{x_0}\}_{n=0}^N$ ;
- ii) Based on  $\{X_n^{x_0}\}_{n=0}^N$ , we get  $\{X_n^{X_0}\}_{n=0}^N$  by using the scheme (4.15) again.

REMARK 4.1. *Note that when using the strong order 1.0 Taylor scheme (4.8) and the weak order 2.0 Taylor scheme (4.14) to solve MSDEJs, we need the knowledge of the exact locations of jump times on the time interval  $[0, T]$ . Hence, the efficiency of the schemes depends on the intensity of the Poisson measure  $\mu$ . And the readers are referred to [10, 33] for details of sampling the jump times of the Poisson measure  $\mu$ .*

**5. Error estimates for strong Taylor schemes.** In this section, based on the relationship between the local and global convergence rates, we shall prove the error estimates of the strong order  $\gamma$  Itô-Taylor Scheme 4.1 and the compensated strong order  $\gamma$  Itô-Taylor Scheme 4.2.

**5.1. The general error estimate theorem.** Let  $\{X_{t,X}(s)\}_{t \leq s \leq T}$  be the solution of the MSDEJ (4.1) starting from the point  $(t, X)$ , that is

$$(5.1) \quad \begin{aligned} X_{t,X}(t+h) &= X + \int_t^{t+h} b^{X_{t,X}}(s, X_{t,X}(s))ds + \int_t^{t+h} \sigma^{X_{t,X}}(s, X_{t,X}(s))dW_s \\ &+ \int_t^{t+h} \int_{\mathbb{E}} c^{X_{t,X}}(s, X_{t,X}(s-), e)\mu(de, ds), \end{aligned}$$

where  $h \in [0, T-t]$ . Let  $\bar{X}_{t,X}(t+h)$  be the one-step approximation of  $X_{t,X}(t+h)$ , and  $\bar{X}_{0,X_0}(t_k)$  is the corresponding solution of the one-step scheme

$$(5.2) \quad \bar{X}_{0,X_0}(t_k) = \bar{X}_{t_{k-1}, \bar{X}_{0,X_0}(t_{k-1})}(t_k),$$

with  $\bar{X}_{0,X_0}(0) = X_0$ . For simplicity, we denote  $X_{0,X_0}(t_k)$  by  $X(t_k)$  and  $\bar{X}_{0,X_0}(t_k)$  by  $\bar{X}_k$ . Then the one-step scheme (5.2) becomes

$$(5.3) \quad \bar{X}_k = \bar{X}_{t_{k-1}, \bar{X}_{k-1}}(t_k).$$

To present the general error estimate theorem for the one-step scheme (5.2), we first give the following two lemmas.

LEMMA 5.1. *Let  $X_{t,X}(s)$  and  $X_{t,Y}(s)$  be the solutions of the MSDEJ (5.1) with initial conditions  $X_{t,X}(t) = X$  and  $X_{t,Y}(t) = Y$ , respectively. Let  $Z = X_{t,X}(t+h) - X_{t,Y}(t+h) - (X - Y)$ , then under the assumptions (A1) – (A4), we have*

$$(5.4a) \quad \mathbb{E}[|Z|^2] \leq C\mathbb{E}[|X - Y|^2]h,$$

$$(5.4b) \quad \mathbb{E}[|X_{t,X}(t+h) - X_{t,Y}(t+h)|^2] \leq (1 + Ch)\mathbb{E}[|X - Y|^2],$$

where  $C$  is a positive constant depending on  $\lambda(\mathbf{E})$ , the Lipschitz constant  $L$  and the function  $\rho(e)$  in assumption (A3).

*Proof.* For any  $0 \leq \theta \leq h$ , based on (5.1), by the Itô's formula (3.9) for MSDEJs and the Itô's isometry formula, we get

$$\begin{aligned} & \mathbb{E}[|X_{t,X}(t+\theta) - X_{t,Y}(t+\theta)|^2] \\ = & \mathbb{E}[|X - Y|^2] + \mathbb{E}\left[\int_t^{t+\theta} |\sigma^{X_{t,X}}(s, X_{t,X}(s)) - \sigma^{X_{t,Y}}(s, X_{t,Y}(s))|^2 ds\right] \\ & + 2\mathbb{E}\left[\int_t^{t+\theta} (X_{t,X}(s) - X_{t,Y}(s)) \left(b^{X_{t,X}}(s, X_{t,X}(s)) - b^{X_{t,Y}}(s, X_{t,Y}(s))\right) ds\right] \\ & + \mathbb{E}\left[\int_t^{t+\theta} \int_{\mathbf{E}} \left(|X_{t,X}(s-) - X_{t,Y}(s-) + c^{X_{t,X}}(s, X_{t,X}(s-), e) \right. \right. \\ & \quad \left. \left. - c^{X_{t,Y}}(s, X_{t,Y}(s-), e)\right|^2 - |X_{t,X}(s-) - X_{t,Y}(s-)|^2\right) \mu(de, ds)\right] \\ = & \mathbb{E}[|X - Y|^2] + \mathbb{E}\left[\int_t^{t+\theta} |\sigma^{X_{t,X}}(s, X_{t,X}(s)) - \sigma^{X_{t,Y}}(s, X_{t,Y}(s))|^2 ds\right] \\ & + 2\mathbb{E}\left[\int_t^{t+\theta} (X_{t,X}(s) - X_{t,Y}(s)) \left(b^{X_{t,X}}(s, X_{t,X}(s)) - b^{X_{t,Y}}(s, X_{t,Y}(s))\right) ds\right] \\ & + \mathbb{E}\left[\int_t^{t+\theta} \int_{\mathbf{E}} \left(|c^{X_{t,X}}(s, X_{t,X}(s), e) - c^{X_{t,Y}}(s, X_{t,Y}(s), e)\right|^2 \right. \\ & \quad \left. + 2|X_{t,X}(s) - X_{t,Y}(s)| |c^{X_{t,X}}(s, X_{t,X}(s), e) - c^{X_{t,Y}}(s, X_{t,Y}(s), e)|\right) \lambda(de) ds\right]. \end{aligned}$$

Then under the assumption (A2) and (A3), we deduce

$$\begin{aligned} & \mathbb{E}[|X_{t,X}(t+\theta) - X_{t,Y}(t+\theta)|^2] \\ \leq & \mathbb{E}[|X - Y|^2] + 4L^2 \int_t^{t+\theta} \mathbb{E}[|X_{t,X}(s) - X_{t,Y}(s)|^2] ds \\ & + 4L \int_t^{t+\theta} \mathbb{E}[|X_{t,X}(s) - X_{t,Y}(s)|^2] ds \\ & + 8K_1 \int_t^{t+\theta} \mathbb{E}[|X_{t,X}(s) - X_{t,Y}(s)|^2] ds \\ \leq & \mathbb{E}[|X - Y|^2] + (4L^2 + 4L + 8K_1) \int_t^{t+\theta} \mathbb{E}[|X_{t,X}(s) - X_{t,Y}(s)|^2] ds, \end{aligned}$$

where  $K_1 = \int_{\mathbf{E}} \rho^2(s) \lambda(de) \vee \lambda(\mathbf{E})$ . Then by the Gronwall lemma [17], we obtain

$$(5.5) \quad \mathbb{E}[|X_{t,X}(t+\theta) - X_{t,Y}(t+\theta)|^2] \leq e^{4(L^2+L+2K_1)h} \mathbb{E}[|X - Y|^2],$$

which leads to the inequality (5.4b).

By the definition of  $Z$ , we have

$$(5.6) \quad \begin{aligned} Z &= \int_t^{t+h} \left( b^{X_{t,X}}(s, X_{t,X}(s)) - b^{X_{t,Y}}(s, X_{t,Y}(s)) \right) ds \\ &\quad + \int_t^{t+h} \left( \sigma^{X_{t,X}}(s, X_{t,X}(s)) - \sigma^{X_{t,Y}}(s, X_{t,Y}(s)) \right) dW_s \\ &\quad + \int_t^{t+h} \int_{\mathbf{E}} \left( c^{X_{t,X}}(s, X_{t,X}(s-), e) - c^{X_{t,Y}}(s, X_{t,Y}(s-), e) \right) \mu(de, ds) \\ &= \int_t^{t+h} \left( \tilde{b}^{X_{t,X}}(s, X_{t,X}(s)) - \tilde{b}^{X_{t,Y}}(s, X_{t,Y}(s)) \right) ds \\ &\quad + \int_t^{t+h} \left( \sigma^{X_{t,X}}(s, X_{t,X}(s)) - \sigma^{X_{t,Y}}(s, X_{t,Y}(s)) \right) dW_s \\ &\quad + \int_t^{t+h} \int_{\mathbf{E}} \left( c^{X_{t,X}}(s, X_{t,X}(s-), e) - c^{X_{t,Y}}(s, X_{t,Y}(s-), e) \right) \tilde{\mu}(de, ds). \end{aligned}$$

Taking square on both sides of (5.6) and taking  $\mathbb{E}[\cdot]$  on the derived equation, we get

$$(5.7) \quad \begin{aligned} \mathbb{E}[|Z|^2] &\leq 3\mathbb{E}\left[ \left| \int_t^{t+h} \left( \tilde{b}^{X_{t,X}}(s, X_{t,X}(s)) - \tilde{b}^{X_{t,Y}}(s, X_{t,Y}(s)) \right) ds \right|^2 \right] \\ &\quad + 3\mathbb{E}\left[ \int_t^{t+h} \left| \sigma^{X_{t,X}}(s, X_{t,X}(s)) - \sigma^{X_{t,Y}}(s, X_{t,Y}(s)) \right|^2 ds \right] \\ &\quad + 3\mathbb{E}\left[ \int_t^{t+h} \int_{\mathbf{E}} \left| c^{X_{t,X}}(s, X_{t,X}(s), e) - c^{X_{t,Y}}(s, X_{t,Y}(s), e) \right|^2 \lambda(de) ds \right]. \end{aligned}$$

Then by using (5.5) and (5.7), we deduce

$$\begin{aligned} \mathbb{E}[|Z|^2] &\leq 12(L^2 + K_1^2)h\mathbb{E}\left[ \int_t^{t+h} \left( \mathbb{E}[|X_{t,X}(s) - X_{t,Y}(s)|^2] + |X_{t,X}(s) - X_{t,Y}(s)|^2 \right) ds \right] \\ &\quad + 6L^2\mathbb{E}\left[ \int_t^{t+h} \left( \mathbb{E}[|X_{t,X}(s) - X_{t,Y}(s)|^2] + |X_{t,X}(s) - X_{t,Y}(s)|^2 \right) ds \right] \\ &\quad + 6K_1\mathbb{E}\left[ \int_t^{t+h} \left( \mathbb{E}[|X_{t,X}(s) - X_{t,Y}(s)|^2] + |X_{t,X}(s) - X_{t,Y}(s)|^2 \right) ds \right] \\ &\leq 12(2L^2 + 2K_1^2 + K_1)(1+h)\left[ \int_t^{t+h} \mathbb{E}[|X_{t,X}(s) - X_{t,Y}(s)|^2] ds \right] \\ &\leq 12(2L^2 + 2K_1^2 + K_1)(1+h)\left[ \int_t^{t+h} \mathbb{E}[|X - Y|^2] \times e^{4(L^2+L+2K_1)h} ds \right] \\ &\leq 12(2L^2 + 2K_1^2 + K_1)(1+h)e^{4(L^2+L+2K_1)h} \mathbb{E}[|X - Y|^2]h, \end{aligned}$$

which proves (5.4a). The proof ends.  $\square$

LEMMA 5.2. *Under the assumptions (A1) – (A4), for  $k = 0, \dots, N - 1$ , we have*

$$(5.8) \quad \mathbb{E} \left[ \left| \mathbb{E} [X_{t_k, \bar{X}_k}(t_{k+1}) - \bar{X}_k | \mathcal{F}_{t_k}] \right|^2 \right] \leq C(1 + \mathbb{E} [|\bar{X}_k|^2])h^2,$$

where  $C$  is a positive constant depending on  $\lambda(\mathbf{E})$ , the function  $\rho(e)$  in (A3) and the linear growth constant  $K$  in (A4).

*Proof.* By (5.1), we get

$$\begin{aligned} \mathbb{E} [X_{t_k, \bar{X}_k}(t_{k+1}) - \bar{X}_k | \mathcal{F}_{t_k}] &= \mathbb{E} \left[ \int_{t_k}^{t_{k+1}} b^{X_{t_k, \bar{X}_k}}(s, X_{t_k, \bar{X}_k}(s)) ds \middle| \mathcal{F}_{t_k} \right] \\ &\quad + \mathbb{E} \left[ \int_{t_k}^{t_{k+1}} \int_{\mathbf{E}} c^{X_{t_k, \bar{X}_k}}(s, X_{t_k, \bar{X}_k}(s), e) \lambda(de) ds \middle| \mathcal{F}_{t_k} \right] \\ &= \mathbb{E} \left[ \int_{t_k}^{t_{k+1}} \tilde{b}^{X_{t_k, \bar{X}_k}}(s, X_{t_k, \bar{X}_k}(s)) ds \middle| \mathcal{F}_{t_k} \right]. \end{aligned}$$

Then

$$(5.9) \quad \begin{aligned} &\mathbb{E} \left[ \left| \mathbb{E} [X_{t_k, \bar{X}_k}(t_{k+1}) - \bar{X}_k | \mathcal{F}_{t_k}] \right|^2 \right] \\ &\leq \mathbb{E} \left[ \mathbb{E} \left[ \left| \int_{t_k}^{t_{k+1}} \tilde{b}^{X_{t_k, \bar{X}_k}}(s, X_{t_k, \bar{X}_k}(s)) ds \right|^2 \middle| \mathcal{F}_{t_k} \right] \right] \\ &\leq h \int_{t_k}^{t_{k+1}} \mathbb{E} \left[ \mathbb{E} \left[ \left| \tilde{b}(s, X_{t_k, \bar{X}_k}(s), x) \right|^2 \middle| x = X_{t_k, \bar{X}_k}(s) \right] \right] ds. \end{aligned}$$

Using the assumption (A4), it is easy to have

$$(5.10) \quad \left| \tilde{b}(s, X_{t_k, \bar{X}_k}(s), x) \right|^2 \leq C(1 + |X_{t_k, \bar{X}_k}(s)|^2 + |x|^2),$$

where  $C$  depends on  $\lambda(\mathbf{E})$ ,  $\rho(e)$  and  $K$ . Then by Theorem 2.1 and (5.10), we deduce

$$(5.11) \quad \begin{aligned} &\mathbb{E} \left[ \mathbb{E} \left[ \left| \tilde{b}(s, X_{t_k, \bar{X}_k}(s), x) \right|^2 \middle| x = X_{t_k, \bar{X}_k}(s) \right] \right] \\ &\leq C \left( 1 + 2\mathbb{E} [ |X_{t_k, \bar{X}_k}(s)|^2 ] \right) \leq C \left( 1 + \mathbb{E} [ |\bar{X}_k|^2 ] \right). \end{aligned}$$

Combining with (5.9) and (5.11), we obtain

$$\begin{aligned} &\mathbb{E} \left[ \left| \mathbb{E} [X_{t_k, \bar{X}_k}(t_{k+1}) - \bar{X}_k | \mathcal{F}_{t_k}] \right|^2 \right] \\ &\leq h \int_{t_k}^{t_{k+1}} C(1 + \mathbb{E} [ |\bar{X}_k|^2 ]) ds \leq C(1 + \mathbb{E} [ |\bar{X}_k|^2 ])h^2, \end{aligned}$$

which completes the proof.  $\square$

Now, based on the Lemmas 5.1 and 5.2, we give the general error estimate theorem for the one-step scheme (5.2).

THEOREM 5.1. *Let  $X_{t, X}(t+h)$  be defined as (5.1). If  $\bar{X}_{t, X}(t+h)$  satisfies*

$$(5.12a) \quad \left| \mathbb{E} [X_{t, X}(t+h) - \bar{X}_{t, X}(t+h) | \mathcal{F}_t] \right| \leq C^*(1 + \mathbb{E} [ |X|^2 ] + |X|^2)^{\frac{1}{2}} h^{p_1},$$

$$(5.12b) \quad \left( \mathbb{E} [ |X_{t, X}(t+h) - \bar{X}_{t, X}(t+h)|^2 | \mathcal{F}_t ] \right)^{\frac{1}{2}} \leq C^*(1 + \mathbb{E} [ |X|^2 ] + |X|^2)^{\frac{1}{2}} h^{p_2},$$

where  $t \in [0, T - h]$ ,  $p_1$  and  $p_2$  are parameters satisfying  $p_2 \geq \frac{1}{2}$  and  $p_1 \geq p_2 + \frac{1}{2}$ , and  $C^* > 0$  is a constant independent of  $h$ ,  $X_{t,X}(t+h)$  and  $\bar{X}_{t,X}(t+h)$ . Then for  $k = 1, \dots, N$ , it holds that

$$(5.13) \quad (\mathbb{E}[|X(t_k) - \bar{X}_k|^2])^{\frac{1}{2}} \leq C(1 + \mathbb{E}[|X_0|^2])^{\frac{1}{2}} h^{p_2 - \frac{1}{2}},$$

where  $C$  is a constant independent of  $h$ ,  $X_{t,X}(t+h)$  and  $\bar{X}_{t,X}(t+h)$ .

*Proof.* By Theorem 2.1, Lemma 5.2 and the discrete Gronwall lemma [28], it is easy to prove that for all  $k = 0, \dots, N$ ,

$$(5.14) \quad \mathbb{E}[|\bar{X}_k|^2] \leq C(1 + \mathbb{E}[|X_0|^2]).$$

Then based on Lemmas 5.1 and 5.2 and the inequality (5.14), the proof of Theorem 5.1 is similar to that of Theorem 4.1 in [35]. So we omit it here.  $\square$

REMARK 5.1. Theorem 5.1 implies that when the weak local error estimate of the one-step scheme (5.2) is of order  $p_1$  and its strong local error estimate is of order  $p_2$ , then the global strong order of the scheme (5.2) is  $p_2 - \frac{1}{2}$ .

**5.2. The error estimates for strong Taylor schemes.** In this subsection, utilizing Theorem 5.1, we prove the error estimates of Schemes 4.1 and 4.2 to reveal the orders of strong convergence of strong Taylor schemes.

Let  $W_t^0 = t$  and  $\alpha = (i_1, i_2, \dots, i_k) \in \mathcal{M}$  be a given multi-index. Then we have the following two lemmas.

LEMMA 5.3. Let  $f^\beta$  and  $\beta$  be defined by (3.1) and (3.6), respectively. Assume that  $\tilde{f}_\alpha^\beta$  and  $\tilde{I}_\alpha[\tilde{f}_\alpha^\beta(\cdot)]_{t,t+h}$  exist and  $|\tilde{f}_\alpha^\beta(t, x)| \leq C(1 + \mathbb{E}[|\beta_t|^2] + |x|^2)^{1/2}$ . Then

$$(5.15) \quad \mathbb{E}\left[\left(\tilde{I}_\alpha[\tilde{f}_\alpha^{X_{t,x}}(\cdot, X_{t,X}(\cdot))]_{t,t+h}\right)^2 \middle| \mathcal{F}_t\right] \leq CMh^{l(\alpha) + n(\alpha)},$$

and

$$(5.16) \quad \left|\mathbb{E}\left[\tilde{I}_\alpha[\tilde{f}_\alpha^{X_{t,x}}(\cdot, X_{t,X}(\cdot))]_{t,t+h} \middle| \mathbb{F}_t\right]\right| \begin{cases} = 0, & \text{if } l(\alpha) \neq n(\alpha), \\ \leq CMh^{l(\alpha)}, & \text{if } l(\alpha) = n(\alpha), \end{cases}$$

where  $M = (1 + \mathbb{E}[|X|^2] + |X|^2)^{1/2}$ .

*Proof.* If  $\alpha = v$ , i.e.,  $l(\alpha) + n(\alpha) = 0$ , we get

$$\begin{aligned} & \mathbb{E}\left[\left(\tilde{I}_\alpha[\tilde{f}_\alpha^{X_{t,x}}(\cdot, X_{t,X}(\cdot))]_{t,t+h}\right)^2 \middle| \mathcal{F}_t\right] \\ &= \mathbb{E}\left[|\tilde{f}_\alpha^{X_{t,x}}(t+h, X_{t,X}(t+h))|^2 \middle| \mathcal{F}_t\right] \\ &\leq C\mathbb{E}\left[1 + \mathbb{E}[|X_{t,X}(t+h)|^2] + |X_{t,X}(t+h)|^2 \middle| \mathcal{F}_t\right] \\ &\leq C(1 + \mathbb{E}[|X|^2] + |X|^2), \end{aligned}$$

which leads to (5.15) with  $p(\alpha) = l(\alpha) + n(\alpha) = 0$ .

Now we consider  $l(\alpha) > 0$ . If  $i_k \neq 0$ , by the Itô's isometry formula, we have

$$(5.17) \quad \begin{aligned} & \mathbb{E}\left[\left(\tilde{I}_\alpha[\tilde{f}_\alpha^{X_{t,x}}(\cdot, X_{t,X}(\cdot))]_{t,t+h}\right)^2 \middle| \mathcal{F}_t\right] \\ &= \begin{cases} \int_t^{t+h} \mathbb{E}\left[\left(\tilde{I}_{\alpha-}[\tilde{f}_\alpha^{X_{t,x}}(\cdot, X_{t,X}(\cdot))]_{t,s}\right)^2 \middle| \mathcal{F}_t\right] ds, & \text{if } i_k = 1, \\ \int_t^{t+h} \int_{\mathbb{E}} \mathbb{E}\left[\left(\tilde{I}_{\alpha-}[\tilde{f}_\alpha^{X_{t,x}}(\cdot, X_{t,X}(\cdot))]_{t,s}\right)^2 \middle| \mathcal{F}_t\right] \lambda(de) ds, & \text{if } i_k = -1. \end{cases} \end{aligned}$$

If  $i_k = 0$ , by the Holder's inequality, we obtain

$$\begin{aligned}
(5.18) \quad & \mathbb{E} \left[ \left( \tilde{I}_\alpha [ \tilde{f}_\alpha^{X_t, X} (\cdot, X_{t, X}(\cdot)) ]_{t, t+h} \right)^2 \middle| \mathcal{F}_t \right] \\
&= \mathbb{E} \left[ \left| \int_t^{t+h} \tilde{I}_{\alpha-} [ \tilde{f}_\alpha^{X_t, X} (\cdot, X_{t, X}(\cdot)) ]_{t, s} ds \right|^2 \middle| \mathcal{F}_t \right] \\
&\leq h \int_t^{t+h} \mathbb{E} \left[ \left( \tilde{I}_{\alpha-} [ \tilde{f}_\alpha^{X_t, X} (\cdot, X_{t, X}(\cdot)) ]_{t, s} \right)^2 \middle| \mathcal{F}_t \right] ds.
\end{aligned}$$

Then combining with (5.17) and (5.18), we deduce the recurrence relation

$$p(\alpha) = p(\alpha-) + (1 + \mathbb{I}_{\{i_k=0\}}) = \sum_{j=1}^k (1 + \mathbb{I}_{\{i_j=0\}}) = l(\alpha) + n(\alpha),$$

which implies that (5.15) holds true.

Similarly, we can prove (5.16). The proof ends.  $\square$

LEMMA 5.4. *Let  $f^\beta$  and  $\beta$  be defined by (3.1) and (3.6), respectively. Assume that  $f_\alpha^\beta$  and  $I_\alpha [f^\beta(\cdot)]_{t, t+h}$  exist and  $|f_\alpha^\beta(t, x)| \leq C(1 + \mathbb{E}[|\beta_t|^2] + |x|^2)^{1/2}$ . Then*

$$(5.19) \quad \mathbb{E} \left[ \left( I_\alpha [f_\alpha^{X_t, X} (\cdot, X_{t, X}(\cdot))]_{t, t+h} \right)^2 \middle| \mathcal{F}_t \right] \leq CMh^{l(\alpha)+n(\alpha)},$$

and

$$(5.20) \quad \left| \mathbb{E} \left[ I_\alpha [f_\alpha^{X_t, X} (\cdot, X_{t, X}(\cdot))]_{t, t+h} \middle| \mathbb{F}_t \right] \right| \begin{cases} = 0, & \text{if } l(\alpha) \neq n(\alpha) + s(\alpha), \\ \leq CMh^{l(\alpha)}, & \text{if } l(\alpha) = n(\alpha) + s(\alpha), \end{cases}$$

where  $M = (1 + \mathbb{E}[|X|^2] + |X|^2)^{1/2}$ .

*Proof.* By Lemma 5.3 and the relationship  $\tilde{\mu}(de, dt) = \mu(de, dt) - \lambda(de)dt$ , it is easy to prove (5.19) and (5.20). The proof ends.  $\square$

Based on Theorem 5.1 and Lemma 5.4, we prove the error estimate of the strong order  $\gamma$  Taylor scheme in the following theorem.

THEOREM 5.2. *Let  $X(t)$  and  $\bar{X}_k$  be the solutions of the MSDEJ (4.1) and the strong order  $\gamma$  Taylor scheme 4.1, respectively. Let  $f(t, x', x) = x$  and assume that  $f^{X_t, X}(s, X_{t, X}(s))$  has the Itô-Taylor expansion (3.24) with  $\mathcal{A} = \mathcal{A}_\gamma$  and*

$$|f_\alpha^\beta(t, x)| \leq C(1 + \mathbb{E}[|\beta_t|^2] + |x|^2)^{1/2}, \quad (t, x) \in [0, T] \times \mathbb{R}^d$$

for all  $\alpha \in \mathcal{A}_\gamma \cup \mathcal{B}(\mathcal{A}_\gamma)$  with  $\beta$  defined by (3.6). Then it holds that

$$\max_{k \in \{1, 2, \dots, N\}} \mathbb{E}[|X_{t_k} - \bar{X}_k|^2] \leq C(1 + \mathbb{E}[|X_0|^2]) (\Delta t)^{2\gamma},$$

where  $\Delta t = \Delta t_k = T/N$  for  $k = 0, 1, \dots, N-1$ .

*Proof.* By the Itô-Taylor expansion (3.24), we have

$$(5.21) \quad X_{t, X}(t+h) = X + \sum_{\alpha \in \mathcal{A}_\gamma \setminus v} I_\alpha [f_\alpha^{X_t, X}(t, X)]_{t, t+h} + R^\gamma,$$

where

$$R^\gamma = \sum_{\alpha \in \mathcal{B}(\mathcal{A}_\gamma)} I_\alpha [f_\alpha^{X_{t,X}}(\cdot, X_{t,X}(\cdot))]_{t,t+h}.$$

Moreover, the strong order  $\gamma$  Taylor scheme 4.1 can be written as

$$(5.22) \quad \bar{X}_{t,X}(t+h) = X + \sum_{\alpha \in \mathcal{A}_\gamma \setminus v} I_\alpha [f_\alpha^{X_{t,X}}(t, X)]_{t,t+h}.$$

Then we subtract (5.22) from (5.21) and obtain

$$(5.23) \quad R^\gamma = X_{t,X}(t+h) - \bar{X}_{t,X}(t+h).$$

According to Lemma 5.4, we deduce

$$\begin{aligned} \mathbb{E}[|R^\gamma|^2 | \mathcal{F}_t] &= \mathbb{E}\left[\left|\sum_{\alpha \in \mathcal{B}(\mathcal{A}_\gamma)} I_\alpha [f_\alpha^{X_{t,X}}(\cdot, X_{t,X}(\cdot))]_{t,t+h}\right|^2 \middle| \mathbb{F}_t\right] \\ &\leq C \sum_{\alpha \in \mathcal{B}(\mathcal{A}_\gamma)} \mathbb{E}\left[|I_\alpha [f_\alpha^{X_{t,X}}(\cdot, X_{t,X}(\cdot))]_{t,t+h}|^2 \middle| \mathcal{F}_t\right] \\ &\leq C \sum_{\alpha \in \mathcal{B}(\mathcal{A}_\gamma)} (1 + \mathbb{E}[|X|^2] + |X|^2) h^{l(\alpha)+n(\alpha)} \\ &\leq C(1 + \mathbb{E}[|X|^2] + |X|^2) h^{2p_2}, \end{aligned}$$

where  $2p_2 = \min_{\alpha \in \mathcal{B}(\mathcal{A}_\gamma)} \{l(\alpha) + n(\alpha)\}$ . Since  $\mathcal{A}_\gamma = \{\alpha \in \mathcal{M} : l(\alpha) + n(\alpha) \leq 2\gamma \text{ or } l(\alpha) = n(\alpha) = \gamma + \frac{1}{2}\}$ , then we get

$$p_2 = \frac{1}{2} \min_{\alpha \in \mathcal{B}(\mathcal{A}_\gamma)} \{l(\alpha) + n(\alpha)\} = \gamma + \frac{1}{2}.$$

Now we prove  $p_1 \geq p_2 + \frac{1}{2}$ . By Lemma 5.4, we can deduce

$$\begin{aligned} \left|\mathbb{E}[R^\gamma | \mathcal{F}_t]\right| &= \left|\mathbb{E}\left[\sum_{\alpha \in \mathcal{B}(\mathcal{A}_\gamma)} I_\alpha [f_\alpha^{X_{t,X}}(\cdot, X_{t,X}(\cdot))]_{t,t+h} \middle| \mathcal{F}_t\right]\right| \\ &\leq \sum_{\alpha \in \mathcal{B}(\mathcal{A}_\gamma)} \left|\mathbb{E}\left[I_\alpha [f_\alpha^{X_{t,X}}(\cdot, X_{t,X}(\cdot))]_{t,t+h} \middle| \mathcal{F}_t\right]\right| \\ &= \sum_{\substack{\alpha \in \mathcal{B}(\mathcal{A}_\gamma) \\ l(\alpha) = n(\alpha) + s(\alpha)}} \left|\mathbb{E}\left[I_\alpha [f_\alpha^{X_{t,X}}(\cdot, X_{t,X}(\cdot))]_{t,t+h} \middle| \mathcal{F}_t\right]\right| \\ &\leq C \sum_{\substack{\alpha \in \mathcal{B}(\mathcal{A}_\gamma) \\ l(\alpha) = n(\alpha) + s(\alpha)}} (1 + \mathbb{E}[|X|^2] + |X|^2)^{1/2} h^{l(\alpha)} \\ &\leq C(1 + \mathbb{E}[|X|^2] + |X|^2)^{1/2} h^{p_1}, \end{aligned}$$

where  $p_1 = \min_{\alpha \in \mathcal{B}(\mathcal{A}_\gamma)} \{l(\alpha) : l(\alpha) = n(\alpha) + s(\alpha)\}$ . Simple calculation yields

$$p_1 = \min_{\alpha \in \mathcal{B}(\mathcal{A}_\gamma)} \{l(\alpha) : l(\alpha) = n(\alpha) + s(\alpha)\} = \begin{cases} \gamma + 1, & \gamma = 1, 2, \dots \\ \gamma + \frac{3}{2}, & \gamma = 0.5, 1.5, \dots, \end{cases}$$

which implies that  $p_1 \geq p_2 + \frac{1}{2}$ . Then by Theorem 5.1, we complete the proof.  $\square$

From Theorem 5.1, we come to the conclusion that the order of strong convergence of the strong order  $\gamma$  Taylor scheme 4.1 is  $\gamma$ . Moreover, by using Theorem 5.1 and Lemma 5.3, we can prove that the order of strong convergence of the compensated strong order  $\gamma$  Taylor scheme 4.2 is also  $\gamma$ .

COROLLARY 5.1. *Let  $X(t)$  and  $\bar{X}_k$  be the solutions of the MSDEJ (4.1) and the compensated strong order  $\gamma$  Taylor scheme 4.2, respectively. Let  $f(t, x', x) = x$  and assume that  $f^{X_{t,x}}(s, X_{t,X}(s))$  has the Itô-Taylor expansion (3.25) with  $\mathcal{A} = \mathcal{A}_\gamma$  and*

$$|\tilde{f}_\alpha^\beta(t, x)| \leq C(1 + \mathbb{E}[|\beta_t|^2] + |x|^2)^{1/2}, \quad (t, x) \in [0, T] \times \mathbb{R}^d$$

for all  $\alpha \in \mathcal{A}_\gamma \cup \mathcal{B}(\mathcal{A}_\gamma)$  with  $\beta$  defined by (3.6). Then it holds that

$$\max_{k \in \{1, 2, \dots, N\}} \mathbb{E}[|X_{t_k} - \bar{X}_k|^2] \leq C(1 + \mathbb{E}[|X_0|^2]) (\Delta t)^{2\gamma}.$$

**6. Error estimates for weak Taylor schemes.** In this section, we focus on the error estimates of the weak order  $\eta$  Itô-Taylor Scheme 4.3 and the compensated weak order  $\eta$  Itô-Taylor Scheme 4.4. For this purpose, we first present some useful lemmas as below.

Let  $C_P^{k, 2k, 2k}([0, T] \times \mathbb{R}^d \times \mathbb{R}^d; \mathbb{R})$  be the set of functions  $\varphi(t, x', x) : [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$  such that all their derivatives with respect to  $t$ ,  $x'$  and  $x$  up to  $k$ ,  $2k$  and  $2k$ , respectively, are continuous and of polynomial growth.

For a given  $\eta \in \{1, 2, \dots\}$  and function  $g \in C_P^{2\eta}(\mathbb{R}^d; \mathbb{R})$ , we define

$$(6.1) \quad u(s, y) = \mathbb{E}[g(X_T^{s,y})],$$

where  $(s, y) \in [0, T] \times \mathbb{R}^d$  and  $X_t^{s,y}$ ,  $s \leq t \leq T$ , is the solution of the MSDEJ (4.1) starting from  $(s, y)$ . Then we get

$$(6.2) \quad u(0, X_0) = \mathbb{E}[g(X_T^{0, X_0})] = \mathbb{E}[g(X_T)].$$

Then we have the following Kolmogorov backward equation for MSDEJs.

LEMMA 6.1 (Kolmogorov backward equation). *Assume that the coefficients of the MSDEJ (4.1) have the components  $b^k, \sigma^{k,j}, c^k \in C_P^{\eta, 2\eta, 2\eta}([0, T] \times \mathbb{R}^d \times \mathbb{R}^d; \mathbb{R})$  for  $1 \leq k \leq d$  and  $1 \leq j \leq m$  with uniformly bounded derivatives. Then the functional  $u$  define in (6.1) is the unique solution of the nonlocal Kolmogorov backward partial integral differential equations*

$$(6.3) \quad \begin{cases} \tilde{L}^0 u(s, y) = 0, & (s, y) \in [0, T] \times \mathbb{R}^d, \\ u(T, y) = g(y), & y \in \mathbb{R}^d, \end{cases}$$

where  $\tilde{L}^0$  is defined by (3.17). Moreover, we have

$$(6.4) \quad u(s, \cdot) \in C_P^{2\eta}(\mathbb{R}^d; \mathbb{R}), \quad 0 \leq s \leq T.$$

*Proof.* The proof of Lemma 6.1 is similar to that of Lemma 12.3.1 in [33]. So we omit it here. The readers are referred to [33] for more details.  $\square$

Based on the Kolmogorov backward equation, we shall prove the error estimates of Schemes 4.3 and 4.4 to reveal the orders of weak convergence of weak Taylor schemes. To proceed, we introduce the following two lemmas.

For a given  $x \in \mathbb{R}$  and  $p \in \mathbb{N}$ , we denote by  $[x]$  the integer part of  $x$  and  $\mathcal{A}_p$  the set of multi-indices  $\alpha = (j_1, \dots, j_l)$  of length  $l \leq p$  with components  $j_i \in \{-1, 0\}$ , for  $i \in \{1, \dots, l\}$ .

LEMMA 6.2. Let  $X_t$  be the solution of the MSDEJ (4.1), and  $\rho$  and  $\tau$  be two stopping times with  $\tau$  being  $\mathcal{F}_\rho$ -measurable and  $0 \leq \rho \leq \tau \leq T$  a.s.. Given  $\alpha \in \mathcal{M}$ , let  $p = l(\alpha) - \lfloor \frac{l(\alpha) + n(\alpha)}{2} \rfloor$  and  $f(t, x) \in C^{p, 2p}([\rho, \tau] \times \mathbb{R}^d; \mathbb{R})$  be a  $\mathcal{F}_t$ -adapted process such that for any  $\alpha \in \mathcal{A}_p$

$$\mathbb{E}[(f_\alpha(t, X_t))^2 | \mathcal{F}_\rho] \leq K, \quad \text{a.s., } t \in [\rho, \tau],$$

for some constant  $K$ . Moreover, for an adapted process  $g(\cdot) = g(\cdot, e)$  with  $e \in \mathbf{E}^{s(\alpha)}$ , if  $\mathbb{E}[g(t, e)^2 | \mathcal{F}_\rho] < +\infty$  a.s. for  $t \in [\rho, \tau]$ , then

$$(6.5) \quad \left| \mathbb{E} \left[ f(\tau, X_\tau) I_\alpha[g(\cdot)]_{\rho, \tau} | \mathcal{F}_\rho \right] \right| \leq C_1 (\tau - \rho)^{l(\alpha)},$$

and

$$(6.6) \quad \left| \mathbb{E} \left[ f(\tau, X_\tau) \tilde{I}_\alpha[g(\cdot)]_{\rho, \tau} | \mathcal{F}_\rho \right] \right| \leq C_2 (\tau - \rho)^{l(\alpha)},$$

where the positive constants  $C_1$  and  $C_2$  do not depend on  $(\tau - \rho)$ .

*Proof.* It is obvious that (6.5) holds for  $|\alpha| = 0$ . Suppose that (6.5) holds for all  $|\alpha| \leq l$ . Now we consider  $\alpha = (j_1, \dots, j_{l+1})$  with  $j_{l+1} = -1$  and obtain

$$(6.7) \quad I_\alpha[g(\cdot)]_{\rho, \tau} = \int_\rho^\tau \int_{\mathbf{E}} I_{\alpha-}[g(\cdot)]_{\rho, s-} \mu(de, ds).$$

Then by the Itô's formula (3.9), we get

$$(6.8) \quad \begin{aligned} f(\tau, X_\tau) I_\alpha[g(\cdot)]_{\rho, \tau} &= \int_\rho^\tau L^0 f(s, X_s) I_\alpha[g(\cdot)]_{\rho, s} ds + \int_\rho^\tau \overrightarrow{L}^1 f(s, X_s) I_\alpha[g(\cdot)]_{\rho, s} dW_s \\ &+ \int_\rho^\tau \int_{\mathbf{E}} \left( f(s, X_s) I_\alpha[g(\cdot)]_{\rho, s} - f(s, X_{s-}) I_\alpha[g(\cdot)]_{\rho, s-} \right) \mu(de, ds). \end{aligned}$$

When  $s$  is a jump time, we have

$$\begin{aligned} f(s, X_s) &= L_e^{-1} f(s, X_{s-}) + f(s, X_{s-}), \\ I_\alpha[g(\cdot)]_{\rho, s} &= I_\alpha[g(\cdot)]_{\rho, s-} + I_{\alpha-}[g(\cdot)]_{\rho, s-}. \end{aligned}$$

By inserting the above equations into (6.8), we deduce

$$(6.9) \quad \begin{aligned} f(\tau, X_\tau) I_\alpha[g(\cdot)]_{\rho, \tau} &= \int_\rho^\tau L^0 f(s, X_s) I_\alpha[g(\cdot)]_{\rho, s} ds + \int_\rho^\tau \overrightarrow{L}^1 f(s, X_s) I_\alpha[g(\cdot)]_{\rho, s} dW_s \\ &+ \int_\rho^\tau \int_{\mathbf{E}} \left( f(s, X_s) I_{\alpha-}[g(\cdot)]_{\rho, s-} + L_e^{-1} f(s, X_{s-}) I_\alpha[g(\cdot)]_{\rho, s-} \right) \mu(de, ds), \end{aligned}$$

which leads to

$$(6.10) \quad \begin{aligned} \left| \mathbb{E} \left[ f(\tau, X_\tau) I_\alpha[g(\cdot)]_{\rho, \tau} | \mathcal{F}_\rho \right] \right| &= \left| \int_\rho^\tau \mathbb{E} \left[ L^0 f(s, X_s) I_\alpha[g(\cdot)]_{\rho, s} | \mathcal{F}_\rho \right] ds \right. \\ &+ \int_\rho^\tau \int_{\mathbf{E}} \mathbb{E} \left[ f(s, X_s) I_{\alpha-}[g(\cdot)]_{\rho, s} | \mathcal{F}_\rho \right] \lambda(de) ds \\ &+ \left. \int_\rho^\tau \int_{\mathbf{E}} \mathbb{E} \left[ L_e^{-1} f(s, X_{s-}) I_\alpha[g(\cdot)]_{\rho, s} | \mathcal{F}_\rho \right] \lambda(de) ds \right| \\ &\leq C (\tau - \rho)^{l+1} + \int_\rho^\tau \left| \mathbb{E} \left[ L^0 f(s, X_s) I_\alpha[g(\cdot)]_{\rho, s} | \mathcal{F}_\rho \right] \right| ds \\ &+ \int_\rho^\tau \int_{\mathbf{E}} \left| \mathbb{E} \left[ L_e^{-1} f(s, X_{s-}) I_\alpha[g(\cdot)]_{\rho, s} | \mathcal{F}_\rho \right] \right| \lambda(de) ds. \end{aligned}$$

Moreover, by the conditions of Lemma 6.2, we have

$$\begin{aligned} L^0 f(t, x) &\in C^{p-1, 2(p-1)}, & L_e^{-1} f(t, x) &\in C^{p, 2p}, \\ \mathbb{E}[(L^0 f_\alpha(t, X_t))^2 | \mathcal{F}_\rho] &\leq K, & \mathbb{E}[(L_e^{-1} f_\alpha(t, X_t))^2 | \mathcal{F}_\rho] &\leq K, \end{aligned}$$

and

$$\mathbb{E}[(I_{(\alpha(1))}[g(\cdot)]_{\rho, t})^2 | \mathcal{F}_\rho] < +\infty, \quad \text{for } t \in [\rho, \tau].$$

Then we can repeatedly apply (6.10)  $p$  times to get

$$\begin{aligned} (6.11) \quad & \left| \mathbb{E} \left[ f(\tau, X_\tau) I_\alpha[g(\cdot)]_{\rho, \tau} | \mathcal{F}_\rho \right] \right| \\ & \leq \sum_{\beta \in \mathcal{A}_p, l(\beta)=p} \int_\rho^\tau \cdots \int_\rho^{s_2} \int_{\mathbb{E}} \cdots \int_{\mathbb{E}} \left| \mathbb{E} \left[ f_\beta(s_1, X_{s_1}) I_\alpha[g(\cdot)]_{\rho, s_1} | \mathcal{F}_\rho \right] \right| \\ & \quad \cdot \lambda(de_1) \cdots \lambda(de_{s(\beta)}) ds_1 \cdots ds_p + C(\tau - \rho)^{l+1}. \end{aligned}$$

Using the conditions of Lemma 6.2, for  $s \in [\rho, \tau]$ , we deduce

$$\mathbb{E} \left[ |I_\alpha[g(\cdot)]_{\rho, s}|^2 | \mathcal{F}_\rho \right] \leq C(s - \rho)^{l(\alpha) + n(\alpha)},$$

which implies that

$$\begin{aligned} (6.12) \quad & \left| \mathbb{E} \left[ f_\beta(s, X_s) I_\alpha[g(\cdot)]_{\rho, s} | \mathcal{F}_\rho \right] \right|^2 \leq \mathbb{E} \left[ |f_\beta(s, X_s)|^2 | \mathcal{F}_\rho \right] \mathbb{E} \left[ |I_\alpha[g(\cdot)]_{\rho, s}|^2 | \mathcal{F}_\rho \right] \\ & \leq C(s - \rho)^{l(\alpha) + n(\alpha)}. \end{aligned}$$

Then by (6.11) and (6.12), we obtain

$$\begin{aligned} (6.13) \quad & \left| \mathbb{E} \left[ f(\tau, X_\tau) I_\alpha[g(\cdot)]_{\rho, \tau} | \mathcal{F}_\rho \right] \right| \\ & \leq C \int_\rho^\tau \cdots \int_\rho^{s_2} (s_1 - \rho)^{\frac{l(\alpha) + n(\alpha)}{2}} ds_1 \cdots ds_p + C(\tau - \rho)^{l+1} \\ & \leq C(\tau - \rho)^{p + \frac{l(\alpha) + n(\alpha)}{2}} + C(\tau - \rho)^{l+1}. \end{aligned}$$

Since  $p = l(\alpha) - \lfloor \frac{l(\alpha) + n(\alpha)}{2} \rfloor$ , then by (6.13), we have

$$\left| \mathbb{E} \left[ f(\tau, X_\tau) I_\alpha[g(\cdot)]_{\rho, \tau} | \mathcal{F}_\rho \right] \right| \leq C(\tau - \rho)^{l+1}.$$

Similarly, we can prove (6.5) for  $\alpha = (j_1, \dots, j_{l+1})$  with  $j_{l+1} = 0$  or  $1$ . Then by using (6.5) and the relationship  $\tilde{\mu}(de, ds) = \mu(de, ds) - \lambda(de)ds$ , we can get (6.6). The proof ends.  $\square$

**LEMMA 6.3.** *Let  $\rho$  and  $\tau$  be two stopping times with  $\tau$  being  $\mathcal{F}_\rho$ -measurable and  $0 \leq \rho \leq \tau \leq T$ , a.s.. Given  $\alpha \in \mathcal{M}$  and  $\{g(t, e), t \in [\rho, \tau]\}$  with  $e \in \mathbb{E}^{s(\alpha)}$  is an adapted process. If  $g(t, e)$  is  $2^{s(\alpha)+3}q$  integrable for a given  $q \in \mathbb{N}^+$ , then for any square integrable adapted process  $\{h(t), t \in [\rho, \tau]\}$ , it holds that*

$$(6.14) \quad \left| \mathbb{E} \left[ h(\tau) |I_\alpha[g(\cdot)]_{\rho, \tau}|^{2q} | \mathcal{F}_\rho \right] \right| \leq C_1(\tau - \rho)^{q(l(\alpha) + n(\alpha) - s(\alpha)) + s(\alpha)},$$

and

$$(6.15) \quad \left| \mathbb{E} \left[ h(\tau) |\tilde{I}_\alpha[g(\cdot)]_{\rho, \tau}|^{2q} | \mathcal{F}_\rho \right] \right| \leq C_2(\tau - \rho)^{q(l(\alpha) + n(\alpha) - s(\alpha)) + s(\alpha)},$$

where the positive constants  $C_1$  and  $C_2$  do not depend on  $(\tau - \rho)$ .

*Proof.* The proof of Lemma 6.3 can be found in Lemma 3.2 in [21] and Lemma 4.5.5 in [33]. We omit it here.  $\square$

Based on Lemmas 6.1 - 6.3, we now prove the error estimates of the weak order  $\eta$  Itô-Taylor Scheme 4.3 in the following theorem.

**THEOREM 6.1.** *Let  $X_t$  and  $X_k$  be the solutions of the MSDEJ (4.1) and the weak order  $\eta$  Itô-Taylor scheme 4.3, respectively. Assume that  $\mathbb{E}[|X_0|^q] < \infty$  for  $q \geq 1$  and  $b^k, \sigma^{k,j}, c^k \in C_P^{\eta+1, 2(\eta+1), 2(\eta+1)}([0, T] \times \mathbb{R}^d \times \mathbb{R}^d; \mathbb{R})$  are Lipschitz continuous for  $1 \leq k \leq d$  and  $1 \leq j \leq m$ . Let the coefficients  $f_\alpha$  with  $f(t, x', x) = x$  satisfy*

$$(6.16) \quad |f_\alpha^\beta(t, x)| \leq K(1 + \mathbb{E}[|\beta_t|] + |x|), \quad (t, x) \in [0, T] \times \mathbb{R}^d$$

for all  $\alpha \in \Gamma_\eta \cup \mathcal{B}(\Gamma_\eta)$  with  $K > 0$  being a constant and  $\beta_t$  defined by (3.6). Then for any function  $g \in C_P^{2(\eta+1)}(\mathbb{R}^d; \mathbb{R})$ , it holds that

$$(6.17) \quad |\mathbb{E}[g(X_T) - g(X_N)]| \leq C(\Delta t)^\eta,$$

where  $C$  is a positive constant independent of  $\Delta t$ .

*Proof.* For simplicity, we consider  $d = m = 1$ . The proof of the general case is similar. According to (6.2) and (6.3), it holds that

$$(6.18) \quad \begin{aligned} H &= |\mathbb{E}[g(X_N)] - \mathbb{E}[g(X_T)]| \\ &= |\mathbb{E}[u(T, X_N)] - u(0, X_0)|. \end{aligned}$$

By the Itô-Taylor expansion (3.24) and (6.3), we deduce

$$(6.19) \quad \mathbb{E}[u(t, X_t^{s,y}) - u(s, y) | \mathcal{F}_s] = 0$$

for any  $0 \leq s \leq t \leq T$  and  $y \in \mathbb{R}^d$ . Then by (6.4), (6.18) and (6.19), we get

$$(6.20) \quad \begin{aligned} H &= \left| \mathbb{E} \left[ \sum_{k=1}^N (u(t_k, X_k) - u(t_{k-1}, X_{k-1})) \right] \right| \\ &= \left| \mathbb{E} \left[ \sum_{k=1}^N (u(t_k, X_k) - u(t_k, X_{t_k}^{t_{k-1}, X_{k-1}})) \right] \right| \\ &\leq H_1 + H_2, \end{aligned}$$

where

$$(6.21) \quad H_1 = \left| \mathbb{E} \left[ \sum_{k=1}^N \frac{\partial u}{\partial y}(t_k, X_{t_k}^{t_{k-1}, X_{k-1}})(X_k - X_{t_k}^{t_{k-1}, X_{k-1}}) \right] \right|,$$

$$(6.22) \quad \begin{aligned} H_2 &= \left| \mathbb{E} \left[ \sum_{k=1}^N \frac{1}{2} \frac{\partial^2 u}{\partial y^2}(t_k, X_{t_k}^{t_{k-1}, X_{k-1}} + \theta_k (X_k - X_{t_k}^{t_{k-1}, X_{k-1}})) \right. \right. \\ &\quad \left. \left. \times (X_k - X_{t_k}^{t_{k-1}, X_{k-1}})^2 \right] \right| \end{aligned}$$

with the parameter  $\theta_k \in (0, 1)$ .

By Theorem 3.3 and Scheme 4.3, we have

$$(6.23) \quad X_{t_k}^{t_{k-1}, X_{k-1}} - X_k = \sum_{\alpha \in \mathcal{B}(\Gamma_\eta)} I_\alpha \left[ f_\alpha^X(\cdot, X_{t_{k-1}, t_k}^{t_{k-1}, X_{k-1}}) \right]_{t_{k-1}, t_k}.$$

Utilizing the estimate (5.14) and the condition (6.16), one can prove that for every  $p \in \{1, 2, \dots\}$ , there exist constants  $C$  and  $r$  such that for every  $q \in \{1, \dots, p\}$

$$(6.24) \quad \mathbb{E} \left[ \max_{0 \leq n \leq N} |X_n|^{2q} \right] \leq C(1 + |X_0|^{2r}).$$

Since  $l(\alpha) = \eta + 1$  for  $\alpha \in \mathcal{B}(\eta)$ , then we obtain  $2\eta + 1 \geq 2p$  for  $p = l(\alpha) - \lfloor \frac{l(\alpha) + n(\alpha)}{2} \rfloor$ . Then by (6.4) and (6.24), for  $\alpha \in \mathcal{A}_p$  and  $k = 1, \dots, N$ , we deduce

$$(6.25) \quad \mathbb{E} \left[ (U_\alpha(t_k, X_{t_k}))^2 | \mathcal{F}_\rho \right] < +\infty, \quad a.s.,$$

$$(6.26) \quad \mathbb{E} \left[ (V(t_k, X_{t_k}))^2 | \mathcal{F}_\rho \right] < +\infty, \quad a.s.,$$

where

$$U(t_k, X_{t_k}) = \frac{\partial u}{\partial y}(t_k, X_{t_k}^{t_{k-1}, X_{k-1}}),$$

$$V(t_k, X_{t_k}) = \frac{\partial^2 u}{\partial y^2}(t_k, X_{t_k}^{t_{k-1}, X_{k-1}} + \theta_k(X_k - X_{t_k}^{t_{k-1}, X_{k-1}})).$$

Moreover, by Theorem 2.1 and the condition (6.16), we get

$$(6.27) \quad \mathbb{E} \left[ |f_\alpha^X(z, X_z^{t_{k-1}, X_{k-1}})|^2 | \mathcal{F}_{t_{k-1}} \right] \leq C(1 + \mathbb{E}[|X_0|^2] + |X_0|^2).$$

Then based on (6.23), (6.25) and (6.27), by Lemma 6.2, we have

$$(6.28) \quad H_1 \leq \mathbb{E} \left[ \sum_{k=1}^N \sum_{\{\alpha: l(\alpha) = \eta + 1\}} \left| \mathbb{E} \left[ \frac{\partial u}{\partial y}(t_k, X_{t_k}^{t_{k-1}, X_{k-1}}) I_\alpha \left[ f_\alpha^X(\cdot, X_{t_{k-1}, t_k}^{t_{k-1}, X_{k-1}}) \right]_{t_{k-1}, t_k} \middle| \mathcal{F}_{t_{k-1}} \right] \right|^2 \right]$$

$$\leq C \mathbb{E} \left[ \sum_{k=1}^N \sum_{\{\alpha: l(\alpha) = \eta + 1\}} (t_k - t_{k-1})^{\eta+1} \right] \leq C(\Delta t_k)^\eta.$$

Similarly, based on (6.23), (6.26) and (6.27), by Lemma 6.3 with  $q = 1$ , we deduce

$$(6.29) \quad H_2 \leq C \mathbb{E} \left[ \sum_{k=1}^N \sum_{\{\alpha: l(\alpha) = \eta + 1\}} \mathbb{E} \left[ \left| \frac{\partial^2 u}{\partial y^2}(t_k, X_{t_k}^{t_{k-1}, X_{k-1}} + \theta_k(X_k - X_{t_k}^{t_{k-1}, X_{k-1}})) \right|^2 \right] \right]$$

$$\times \left| I_\alpha \left[ f_\alpha^X(\cdot, X_{t_{k-1}, t_k}^{t_{k-1}, X_{k-1}}) \right]_{t_{k-1}, t_k} \right|^2 \Big| \mathcal{F}_{t_{k-1}} \Big]$$

$$\leq C(\Delta t_k)^\eta.$$

Then by using (6.18), (6.28) and (6.29), we get (6.17). The proof ends.  $\square$

From Theorem 6.1, we can conclude that the order of weak convergence of the weak order  $\eta$  Taylor scheme 4.3 is  $\eta$ . Moreover, based on Lemmas 6.1 - 6.3, we can prove that the order of weak convergence of the compensated weak order  $\eta$  Taylor scheme 4.4 is also  $\eta$ .

COROLLARY 6.1. *Let  $X_t$  and  $X_k$  be the solutions of the MSDEJ (4.1) and the compensated weak order  $\eta$  Taylor scheme 4.4, respectively. Assume that  $\mathbb{E}[|X_0|^q] < \infty$  for  $q \geq 1$  and  $b^k, \sigma^{k,j}, c^k \in C_P^{\eta+1, 2(\eta+1), 2(\eta+1)}([0, T] \times \mathbb{R}^d \times \mathbb{R}^d; \mathbb{R})$  are Lipschitz continuous for  $1 \leq k \leq d$  and  $1 \leq j \leq m$ . Let  $\tilde{f}_\alpha$  with  $f(t, x', x) = x$  satisfy*

$$|\tilde{f}_\alpha^\beta(t, x)| \leq K(1 + \mathbb{E}[|\beta_t|] + |x|), \quad (t, x) \in [0, T] \times \mathbb{R}^d$$

for all  $\alpha \in \Gamma_\eta \cup \mathcal{B}(\Gamma_\eta)$  with  $K > 0$  being a constant and  $\beta_t$  defined by (3.6). Then for any function  $g \in C_P^{2(\eta+1)}(\mathbb{R}^d; \mathbb{R})$ , it holds that

$$|\mathbb{E}[g(X_T) - g(X_N)]| \leq C(\Delta t)^\eta,$$

where  $C$  is a positive constant independent of  $\Delta t$ .

**7. Numerical examples.** In this section, we carry out some numerical tests to verify our theoretical conclusions and to show the efficiency and the accuracy of the proposed schemes for solving MSDEJs. For each example, we shall test the Euler scheme (4.7), the strong order 1.0 Taylor scheme (4.8), and the weak order 2.0 Taylor scheme (4.14), respectively.

For simplicity, we adopt the uniform time partition, and the time partition number  $N$  is given by  $N = \frac{T}{\Delta t}$ . We denote by  $\mathbb{E}[|X_T - X_N|]$  and  $|\mathbb{E}[X_T - X_N]|$  the strong errors and the weak errors between the exact solution  $X_t$  of the MSDEJ (4.1) at time  $t = T$  and the numerical solution  $X_n$  of the proposed schemes at  $n = N$ . The Monte Carlo method is used to approximate the expectation  $\mathbb{E}[\cdot]$  appeared in coefficients and errors with sample times  $M$ . The ‘‘exact’’ solution of the MSDEJs is identified with the numerical one using a small step-size  $\Delta t_{\text{exact}} = 2^{-12}$ . Moreover, we will test the efficiency of our schemes with respect to the level of the intensity  $\lambda$  of the Poisson measure  $\mu$  by the magnitudes of the sample times  $M$  and the running time (RT) for different values of  $\lambda$ .

In what follows, we denote by Euler, S-1.0 and W-2.0 the Euler scheme, the strong order 1.0 Taylor scheme, and the weak order 2.0 Taylor scheme, respectively. The convergence rate (CR) with respect to  $\Delta t$  is obtained by using linear least square fitting to the numerical errors. In all the tests, we set  $T = 1.0$ . The unit of RT is the second.

EXAMPLE 7.1. *Consider the following MSDEJ with  $X_0 = x_0$ :*

$$(7.1) \quad dX_s = a(\mathbb{E}[X_s] + X_s)ds + bX_s dW_s + \int_{\mathbb{E}} ce(\mathbb{E}[X_s] + X_{s-})\mu(de, ds),$$

where  $a, b$  and  $c$  are constants.

We set  $a = 1.25$ ,  $b = 0.75$ ,  $c = 0.25$ , and  $X_0 = 0.1$ . Assume that the jump sizes  $\{Y_i, i = 1, \dots, N_T\}$  satisfy  $Y_i \stackrel{iid}{\sim} U(-\frac{1}{2}, \frac{1}{2})$ , which is the uniform distribution on  $[-\frac{1}{2}, \frac{1}{2}]$ . And we use the Euler scheme (4.7), the strong order 1.0 Taylor scheme (4.8) and the weak order 2.0 Taylor scheme (4.14) to solve (7.1), respectively. We have listed the errors and convergence rates of the schemes (4.7), (4.8) and (4.14) for different intensity  $\lambda$  in Tables 1 - 3, respectively.

Table 1: Errors and convergence rates of the Euler scheme

Euler								
$N$	16	32	64	128	256	CR	M	RT
$\lambda$	$\mathbb{E}[ X(T) - X_N ]$							
0.1	2.739E-01	1.405E-01	6.630E-02	4.517E-02	3.377E-02	0.768	45	2.94
0.5	1.839E-01	9.217E-02	4.957E-02	3.212E-02	2.063E-02	0.783	65	4.18
1.0	2.174E-01	1.236E-01	7.031E-02	4.236E-02	2.479E-02	0.781	75	4.87
2.0	2.106E-01	1.173E-01	5.759E-02	3.800E-02	2.610E-02	0.765	85	5.46
3.0	1.479E-01	8.452E-02	5.922E-02	3.605E-02	2.614E-02	0.623	100	6.14
	$\mathbb{E}[X(T) - X_N]$							
0.1	2.192E-01	1.222E-01	6.122E-02	2.696E-02	1.034E-02	1.099	100	6.07
0.5	2.987E-01	1.510E-01	7.575E-02	4.380E-02	1.907E-02	0.972	200	12.94
1.0	3.229E-01	1.699E-01	8.444E-02	4.397E-02	1.934E-02	1.007	300	24.51
2.0	3.401E-01	1.880E-01	9.122E-02	4.380E-02	2.007E-02	1.027	500	59.56
3.0	2.996E-01	1.640E-01	8.564E-02	4.028E-02	1.946E-02	0.992	900	107.35

Table 2: Errors and convergence rates of the strong order 1.0 Taylor scheme

S-1.0								
$N$	16	32	64	128	256	CR	M	RT
$\lambda$	$\mathbb{E}[ X(T) - X_N ]$							
0.1	1.679E-01	8.939E-02	4.560E-02	2.292E-02	1.126E-02	0.976	100	6.76
0.5	2.027E-01	1.107E-01	5.738E-02	2.859E-02	1.408E-02	0.965	200	13.63
1.0	2.084E-01	1.138E-01	5.930E-02	2.962E-02	1.449E-02	0.963	400	41.21
2.0	2.138E-01	1.164E-01	6.046E-02	3.053E-02	1.477E-02	0.964	800	99.20
3.0	2.006E-01	1.092E-01	5.665E-02	2.877E-02	1.404E-02	0.960	1000	130.32
	$\mathbb{E}[X(T) - X_N]$							
0.1	1.679E-01	8.939E-02	4.560E-02	2.292E-02	1.126E-02	0.976	100	6.76
0.5	2.027E-01	1.107E-01	5.738E-02	2.859E-02	1.408E-02	0.965	200	13.63
1.0	2.084E-01	1.138E-01	5.927E-02	2.959E-02	1.447E-02	0.964	400	41.21
2.0	2.138E-01	1.164E-01	6.045E-02	3.049E-02	1.473E-02	0.965	800	99.20
3.0	2.006E-01	1.092E-01	5.657E-02	2.867E-02	1.394E-02	0.962	1000	130.32

Table 3: Errors and convergence rates of the weak order 2.0 Taylor scheme

W-2.0								
$N$	8	16	32	64	128	CR	M	RT
$\lambda$	$\mathbb{E}[ X(T) - X_N ]$							
0.1	3.686E-02	1.329E-02	6.709E-03	2.672E-03	1.271E-03	1.203	100	7.28
0.5	5.150E-02	1.722E-02	6.706E-03	3.670E-03	1.849E-03	1.183	200	16.15
1.0	4.686E-02	1.681E-02	6.918E-03	3.749E-03	2.165E-03	1.104	500	78.36
2.0	5.022E-02	1.974E-02	8.703E-03	4.663E-03	2.825E-03	1.039	800	114.51
3.0	4.808E-02	1.841E-02	9.511E-03	5.843E-03	3.511E-03	0.921	1500	228.86
	$\mathbb{E}[X(T) - X_N]$							
0.1	3.438E-02	1.013E-02	2.618E-03	5.500E-04	8.881E-05	2.140	2800	450.22
0.5	4.033E-02	1.157E-02	2.537E-03	7.168E-04	1.165E-04	2.089	3000	481.91
1.0	3.529E-02	9.424E-03	2.526E-03	5.719E-04	1.246E-04	2.033	3500	567.14
2.0	3.945E-02	1.101E-02	2.587E-03	3.550E-04	1.099E-04	2.193	4500	826.78
3.0	3.836E-02	1.057E-02	2.542E-03	6.279E-04	4.323E-05	2.366	5500	1201.71

The numerical results listed in Tables 1 - 3 show that the Euler scheme (4.7), the strong order 1.0 Taylor scheme (4.8) and the weak order 2.0 Taylor scheme (4.14) are stable and accurate for solving the linear MSDEJ (7.1). Moreover, we can draw the following conclusions.

1. The orders of strong convergence of the Euler scheme, the strong order 1.0 Taylor scheme and the weak order 2.0 Taylor scheme are 0.5, 1.0 and 1.0, respectively;
2. The orders of weak convergence of the Euler scheme, the strong order 1.0 Taylor scheme and the weak order 2.0 Taylor scheme are 1.0, 1.0 and 2.0, respectively;
3. The efficiency of the schemes depends on the level of the intensity  $\dot{\lambda}$  of the Poisson measure  $\mu$ . As the intensity  $\dot{\lambda}$  increases, the sample times  $M$  and the running time RT increase.

All of the conclusions above are consistent with our theoretical results.

EXAMPLE 7.2. Consider the following general nonlinear MSDEJ

$$(7.2) \quad \begin{aligned} dX_t^{0, X_0} = & \left( (X_t^{0, X_0})^{5/3} + 2\dot{\lambda}^2 \mathbb{E}[X_t^{0, x_0}] \right) dt + \frac{1}{2} \mathbb{E}[X_t^{0, x_0}] dW_t \\ & + \int_{\mathbb{E}} \frac{e}{2(1 + \dot{\lambda}^2)} \left( X_t^{0, X_0} + \mathbb{E}[(X_t^{0, x_0})^2] \right) \mu(de, dt), \end{aligned}$$

where  $x_0$  and  $X_0$  are initial values and  $\dot{\lambda}$  is the intensity of the Poisson measure  $\mu$ .

Let the jump sizes satisfy  $Y_i \stackrel{iid}{\sim} U(-\frac{1}{2}, \frac{1}{2})$ ,  $i = 1, \dots, N_T$ . We use the Euler scheme (4.7), the strong order 1.0 Taylor scheme (4.8) and the weak order 2.0 Taylor scheme (4.14) to solve (7.2), respectively. For simplicity, we set  $\dot{\lambda} = 1.0$ . In Tables 4 and 5, we have listed the errors and convergence rates of the schemes for different initial values of  $x_0$  and  $X_0$ .

Table 4: Errors and convergence rates of the schemes with  $x_0 = X_0 = 0.1$

$N$	16	32	64	128	256	CR
$\mathbb{E} X(T) - X_N $						
Euler	3.404E-01	1.956E-01	1.050E-01	5.330E-02	3.003E-02	0.888
S-1.0	2.968E-01	1.697E-01	9.092E-02	4.684E-02	2.324E-02	0.921
W-2.0	3.207E-02	9.929E-03	3.726E-03	1.848E-03	1.110E-03	1.213
$\mathbb{E}[X(T) - X_N]$						
Euler	2.802E-01	1.605E-01	8.578E-02	4.349E-02	2.110E-02	0.935
S-1.0	2.965E-01	1.696E-01	9.086E-02	4.682E-02	2.324E-02	0.920
W-2.0	2.417E-02	6.095E-03	1.432E-03	2.852E-04	7.891E-05	2.094

From the numerical results in Tables 4 and 5, we come to the conclusion that the Euler scheme (4.7), the strong order 1.0 Taylor scheme (4.8) and the weak order 2.0 Taylor scheme (4.14) are stable and accurate for solving the nonlinear MSDEJ (7.2) with different initial values of  $x_0$  and  $X_0$ . Tables 4 and 5 also show that the orders of strong convergence of the schemes (4.7), (4.8) and (4.14) are 0.5, 1.0 and 1.0, respectively, and the orders of weak convergence are 1.0, 1.0 and 2.0, respectively, which verify again our theoretical conclusions.

Table 5: Errors and convergence rates of the schemes with  $x_0 = 0.15$  and  $X_0 = 0.05$ 

$N$	16	32	64	128	256	CR
$\mathbb{E} X(T) - X_N $						
Euler	7.643E-01	4.595E-01	2.545E-01	1.315E-01	7.290E-02	0.859
S-1.0	5.989E-01	3.559E-01	1.945E-01	9.997E-02	4.898E-02	0.906
W-2.0	8.133E-02	3.081E-02	1.054E-02	4.521E-03	2.761E-03	1.253
$\mathbb{E}[X(T) - X_N]$						
Euler	5.947E-01	3.527E-01	1.923E-01	9.924E-02	4.858E-02	0.906
S-1.0	5.946E-01	3.527E-01	1.923E-01	9.926E-02	4.861E-02	0.905
W-2.0	7.884E-02	2.870E-02	7.801E-03	1.535E-03	2.047E-04	2.140

Note that the efficiency of the proposed schemes, especially the high order ones depends on the level of the intensity of the Poisson measure. This is mainly due to the existence of the double integrals involving the Poisson measure  $I_{\{-1,-1\}}$ ,  $I_{\{1,-1\}}$ ,  $I_{\{-1,1\}}$ ,  $I_{\{0,-1\}}$  and  $I_{\{-1,0\}}$ , the computation complexity of which is dependent on the number of jumps. Hence, to construct more efficient high order schemes for MSDEJs, we will focus on the jump-adapted methods in our future work, which avoid the integrals involving the Poisson measure.

**8. Conclusions.** In this paper, we developed the Itô formula and the Itô-Taylor expansion for MSDEJs, then based on which we proposed the strong order  $\gamma$  and the weak order  $\eta$  Itô-Taylor schemes for solving MSDEJs. We rigorously proved the error estimates of the proposed schemes, which show that the order of strong convergence of the strong order  $\gamma$  Taylor scheme and the order of weak convergence of the weak order  $\eta$  Taylor scheme are  $\gamma$  and  $\eta$ , respectively. Numerical experiments verify our theoretical conclusions and indicate that the efficiency of the schemes depends on the level of the intensity of the Poisson measure. In the future work, we shall consider the jump-adapted methods for MSDEJs.

## REFERENCES

- [1] A. BENSOUSSAN, J. FREHSE, AND P. YAM, *Mean Field Games and Mean Field Type Control Theory*, Springer Briefs in Mathematics, Springer, New York, 2013.
- [2] M. BOSSY AND D. TALAY, *A stochastic particle method for the McKean-Vlasov and the Burgers equation*, Math. Comput., 66 (1997), pp. 157–192.
- [3] R. BUCKDAHN, B. DJEHICHE, J. LI AND S. PENG, *Mean-field backward stochastic differential equations: a limit approach*, Ann. Probab., 37 (2009), pp. 1524–1565.
- [4] R. BUCKDAHN, J. LI AND S. PENG, *Mean-field backward stochastic differential equations and related partial differential equations*, Stochastic Process. Appl., 119 (2009), pp. 3133–3154.
- [5] R. BUCKDAHN, J. LI, AND S. PENG, *Mean-field stochastic differential equations and associated PDEs*, Ann. Probab., 45 (2017), pp. 824–878.
- [6] A. BENSOUSSAN, P. YAM, AND Z. ZHANG, *Well-posedness of mean-field type forward-backward stochastic differential equations*, Stochastic Process. Appl., 125 (2015), pp. 3327–3354.
- [7] D. BELOMESTNY AND J. SCHOENMAKERS, *Projected particle methods for solving McKean-Vlasov stochastic differential equations*, SIAM J. Numer. Anal., 56 (2018), pp. 3169–3195.
- [8] R. CARMONA AND F. DELARUE, *Probabilistic analysis of mean-field games*, SIAM J. Control Optim., 51 (2012), pp. 2705–2734.
- [9] P. CARDALIAGUET, F. DELARUE, J. M. LASRY, AND P. L. LIONS, *The master equation and the convergence problem in mean field games*, volume 201 of Annals of Mathematics Studies, Princeton University Press, Princeton, NJ, 2019.
- [10] R. CONT AND P. TANKOV, *Financial Modelling with Jump Processes*, Chapman and Hall/CRC

- Press, London, 2004.
- [11] Y. FU, W. ZHAO, AND T. ZHOU, *Multistep schemes for forward backward stochastic differential equations with jumps*, J. Sci. Comput., 69 (2016), pp. 1–22.
  - [12] O. GUÉANT, J. LASRY AND P. LIONS, *Mean-field games and applications*, Paris-Princeton lectures on mathematical finance 2010, Springer-Verlag Berlin, 2003, pp. 205–266.
  - [13] M. HAFAYED, S. ABBAS, AND A. ABBA, *On mean-field partial information maximum principle of optimal control for stochastic systems with levy processes*, J. Optimiz. Theory App., 167 (2015), pp. 1051–1069.
  - [14] M. HAFAYED, M. TABET, AND S. BOUKAF, *Mean-field maximum principle for optimal control of forward-backward stochastic systems with jumps and its application to mean-variance portfolio problem*, Commun. Math. Stat., 3 (2015), pp. 163–186.
  - [15] D. HIGHAM AND P. KLOEDEN, *Numerical methods for nonlinear stochastic differential equations with jumps*, Numerische Mathematik, 101 (2005), pp. 101–119.
  - [16] T. HAO AND J. LI, *Mean-field SDEs with jumps and nonlocal integral-PDEs*, Nonlinear Differential Equations Appl., 23 (2016), pp. 1–51.
  - [17] P. KLOEDEN AND E. PLATEN, *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, Berlin, 1992.
  - [18] P. KLOEDEN AND T. SHARDLOW, *Gauss-quadrature method for one-dimensional mean-field SDEs*, SIAM J. Sci. Comput., 39 (2017), pp. A2784–A2807.
  - [19] P. KOTELENEZ, *A class of quasilinear stochastic partial differential equations of McKean-Vlasov type with mass conservation*, Probab. Theory Related Fields, 102 (1995), pp. 159–188.
  - [20] J. LASRY AND P. LIONS, *Mean-field games*, Japan J. Math., 2 (2007), pp. 229–260.
  - [21] X. LIU AND W. LI, *Weak approximations and extrapolations of stochastic differential equations with jumps*, SIAM J. Numer. Anal., 37 (2000), pp. 1747–1767.
  - [22] J. LI, *Stochastic maximum principle in the mean-field controls*, Automatica, 48 (2012), pp. 366–373.
  - [23] J. LI, *Mean-field forward and backward SDEs with jumps and associated nonlocal quasi-linear integral-PDEs*, Stochastic Process. Appl., 128 (2018), pp. 3118–3180.
  - [24] J. LI AND H. MIN, *Controlled mean-field backward stochastic differential equations with jumps involving the value function*, J. Syst. Sci. Complex, 29 (2016), pp. 1238–1268.
  - [25] H. MCKEAN, *A class of markov processes associated with nonlinear parabolic equations*, PNAS, 55 (1966), pp. 1907–1911.
  - [26] S. MÉLÉARD, *Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models in probabilistic models for nonlinear partial differential equations*, Springer-Verlag, Berlin, 1996, pp. 42–95.
  - [27] M. MENDOZA, M. AGUILAR AND F. VALLE, *A mean-field approach that combines quantum mechanics and molecular dynamics simulation: the water molecule in liquid water*, J. Mol. Struct., 426 (1998), pp. 181–190.
  - [28] G. MILSTEIN AND M. TRETYAKOV, *Stochastic Numerics for Mathematical Physics*, Springer-Verlag, Berlin, 2004.
  - [29] Y. NI, X. LI AND J. ZHANG, *Mean-field stochastic linear-quadratic optimal control with Markov jump parameters*, Syst. Control Lett., 93 (2016), pp. 69–76.
  - [30] M. KAC, *Foundations of kinetic theory*, Proceedings of the 3rd Berkeley Symposium on Mathematical Statistics and Probability, 3 (1956), pp. 171–197.
  - [31] M. KAC, *Probability and Related Topics in the Physical Sciences*, Interscience Publishers, New York, 1958.
  - [32] B. ØKSENDAL, *Stochastic Differential Equations*, Springer-Verlag, Berlin, 2003.
  - [33] E. PLATEN AND N. BRUTI-LIBERATI, *Numerical Solution of Stochastic Differential Equations with Jumps in Finance*, Stochastic Modelling and Applied Probability, Springer, Berlin, 2010.
  - [34] P. STEVENSON, J. STONE AND M. STRAYER, *Hartree-Fock mean-field models using separable interactions*, Office of Scientific & Technical Information Technical Reports, 217 (1999), U8.
  - [35] Y. SUN, J. YANG, AND W. ZHAO, *Itô-Taylor schemes for solving mean-field stochastic differential equations*, Numer. Math. Theor. Meth. Appl., 10 (2017), pp. 798–828.
  - [36] Y. SUN AND W. ZHAO, *An explicit second-order numerical scheme for mean-field forward backward stochastic differential equations*, Numer. Algorithms, 2019, accepted, DOI: 10.1007/s11075-019-00754-2.
  - [37] Y. SUN, W. ZHAO, AND T. ZHOU, *Explicit  $\theta$ -scheme for solving mean-field backward stochastic differential equations*, SIAM J. Numer. Anal., 56 (2018), pp. 2672–2697.
  - [38] D. TALAY AND O. VAILLANT, *A stochastic particle method with random weights for the computation of statistical solutions of McKean-Vlasov equations*, Ann. Appl. Probab., 13 (2003),

- pp. 140–180.
- [39] B. WANG AND J. ZHANG, *Mean-field games for large population multi-agent systems with Markov jump parameters*, SIAM J. Control Optim., 50 (2012), pp. 2308–2334.
  - [40] G. ZHANG, W. ZHAO, C. WEBSTER AND M. GUNZBURGER, *Numerical methods for a class of nonlocal diffusion problems with the use of backward SDEs*, Comput. Math. Appl., 71 (2015), pp. 2479–2496.
  - [41] W. ZHAO, W. ZHANG, AND G. ZHANG, *Second-order numerical schemes for decoupled forward-backward stochastic differential equations with jumps*, J. Comput. Math., 35 (2017), pp. 213–244.