On the Strong Approximation of Jump-Diffusion Processes

Nicola Bruti-Liberati[1] and **Eckhard Platen**[2]

April 15, 2005

Abstract. In financial modelling, filtering and other areas the underlying dynamics are often specified via stochastic differential equations (SDEs) of jump-diffusion type. The class of jump-diffusion SDEs that admits explicit solutions is rather limited. Consequently, there is a need for the systematic use of discrete time approximations in corresponding simulations. This paper presents a survey and new results on strong numerical schemes for SDEs of jump-diffusion type. These are relevant for scenario analysis, filtering and hedge simulation in finance. It provides a convergence theorem for the construction of strong approximations of any given order of convergence for SDEs driven by Wiener processes and Poisson random measures. The paper covers also derivative free, drift-implicit and jump adapted strong approximations. For the commutative case particular schemes are obtained. Finally, a numerical study on the accuracy of several strong schemes is presented.

200 *Mathematics Subject Classification:* primary 60H10; secondary 65C05.
JEL Classification: G10, G13.

Key words and phrases: jump-diffusion processes, stochastic Taylor expansion, discrete time approximation, simulation, strong convergence.

[1]University of Technology Sydney, School of Finance & Economics, PO Box 123, Broadway, NSW, 2007, Australia

[2]University of Technology Sydney, School of Finance & Economics and Department of Mathematical Sciences, PO Box 123, Broadway, NSW, 2007, Australia

1 Introduction

There is compelling evidence that the dynamics of prices of financial instruments exhibit jumps that cannot be adequately captured solely by diffusion processes, see Merton (1976). Several empirical studies, including Jorion (1988), Bates (1996) and Pan (2002), demonstrate the existence of jumps in stock markets, the foreign exchange market and bond markets. Jumps constitute also a key feature in the description of credit risk sensitive instruments. Therefore, models that incorporate jumps have become increasingly popular in finance, see, for instance, Merton (1976), Björk, Kabanov & Runggaldier (1997), Duffie, Pan & Singleton (2000), Kou (2002), Schönbucher (2003) and Chiarella & Nikitopoulos-Sklibosios (2004). Beyond finance there are many areas of application, including electrical engineering and biotechnology, that use jump-diffusion models. Since only a limited class of jump-diffusion SDEs admits explicit solutions, there is a strong need for the development of numerical schemes. In the current paper we consider the construction of strong discrete time approximations of SDEs driven by Wiener processes and Poisson random measures.

A discrete time approximation Y^{Δ} converges strongly with order $\gamma > 0$ at time T to the solution X of a given SDE if there exists a positive constant C and a $\Delta_0 > 0$ such that

$$\varepsilon(\Delta) = \sqrt{E(|X_T - Y_T^{\Delta}|^2)} \leq C\Delta^{\gamma}, \tag{1.1}$$

for each maximum time step size $\Delta \in (0, \Delta_0)$. Since strong approximations indicate pathwise closeness, the above criterion is appropriate for the classification of schemes for scenario analysis, filtering and hedge simulation. When the focus is on approximating the expectation of a payoff function of the solution of the underlying SDE, such as moments, derivative prices and risk measures, then a weaker criterion is sufficient. We refer to Kloeden & Platen (1999) and Platen (1999) for extensive surveys on both strong and weak approximations of SDEs in the case of diffusion and jump-diffusion processes.

The literature on strong approximations of jump-diffusion SDEs driven by Wiener processes and Poisson random measures is rather limited. Platen (1982a) describes a convergence theorem for strong schemes of any given strong order $\gamma \in \{0.5, 1, 1.5, \ldots\}$ and also introduces jump adapted approximations. Maghsoodi (1996, 1998) presents an analysis of some approximations up to strong order $\gamma = 1.5$. Gardoǹ (2004) presents a convergence theorem for strong schemes of any given order $\gamma \in \{0.5, 1, 1.5, \ldots\}$ similar to the one presented in Platen (1982a), but limited to SDEs driven by Wiener processes and homogeneous Poisson processes and without considering jump adapted approximations. Higham & Kloeden (2004) propose a class of implicit schemes of strong order $\gamma = 0.5$ for SDEs driven by Wiener processes and homogeneous Poisson processes. Additionally, the Euler scheme for the approximation of SDEs driven by more general semimartingales has been studied by Jacod & Protter (1998).

For general SDEs higher order strong schemes often require multiple stochastic integrals that cannot be easily generated in an efficient manner. For special classes of SDEs some multiple stochastic integrals cancel out, for instance, under commutativity. Additionally, for important applications such as hidden Markov chain filtering, see Elliott, Aggoun & Moore (1995), one can construct the required multiple stochastic integrals directly from the data.

In this paper we present jump adapted approximations, as introduced by Platen (1982a), for which the required multiple stochastic integrals do not involve any Poisson measure. Therefore, these schemes are easier to implement and computationally efficient for SDEs with low intensity Poisson measures.

In this paper we also consider schemes that avoid the computation of derivatives of the coefficient functions, which enhances the computational tractability. Implicit schemes are derived that improve the numerical stability, as shown in Hofmann & Platen (1996), Milstein, Platen & Schurz (1998) and Higham & Kloeden (2004). Moreover, along the analysis of the order 1.0 strong schemes with mark independent jump size, we derive a commutativity condition that permits to identify a class of jump-diffusion SDEs for which the computational efficiency of the order 1.0 strong schemes is independent of the jump intensity level.

Finally, we present three convergence theorems that extend the ones in Platen (1982a) to cover schemes of any given strong order for SDEs driven by Wiener processes and Poisson random measures, including derivative free, drift-implicit and jump adapted schemes. A numerical study on the accuracy of these strong schemes on the Merton (1976) model will be presented.

2 Model Dynamics

Given a filtered probability space $(\Omega, \mathcal{A}_T, \underline{\mathcal{A}}, P)$ satisfying the usual conditions and a mark space $(\mathcal{E}, \mathsf{B}(\mathcal{E}))$ with $\mathcal{E} \subseteq \mathbb{R}^r \backslash \{0\}$, for $r \in \{1, 2, \ldots\}$, we define on $\mathcal{E} \times [0, T]$ an $\underline{\mathcal{A}}$-adapted Poisson random measure $p_\phi(dv \times dt)$, where dv denotes an n-dimensional vector, with intensity measure

$$q_\phi(dv \times dt) = \phi(dv)dt. \tag{2.1}$$

We assume that the intensity $\lambda = \phi(\mathcal{E}) < \infty$ is finite. Thus, $p_\phi = \{p_\phi(t) := p_\phi(\mathcal{E} \times [0, t]), t \in [0, T]\}$ is a stochastic process that counts the number of jumps occurring in the time interval $[0, T]$. The Poisson random measure $p_\phi(dv \times dt)$ generates a sequence of pairs $\{(\tau_i, \xi_i), i \in \{1, 2, \ldots, p_\phi(T)\}\}$, where $\{\tau_i : \Omega \to \mathbb{R}_+, i \in \{1, 2, \ldots, p_\phi(T)\}\}$ is a sequence of increasing nonnegative random variables representing the jump times of a standard Poisson process with intensity λ, and $\{\xi_i : \Omega \to \mathcal{E}, i \in \{1, 2, \ldots, p_\phi(T)\}\}$ is a sequence of i.i.d. random variables with $\xi_i \sim \frac{\phi(du)}{\phi(\mathcal{E})}$. We can interpret τ_i as the time of the i-th event and the mark ξ_i as its amplitude. For a more general presentation of random measures we refer to Elliott (1982).

For the dynamics of the underlying d-dimensional factors we consider the SDE

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t + \int_{\mathcal{E}} c(t-, X_{t-}, v)\, p_\phi(dv \times dt), \qquad (2.2)$$

for $t \in [0, T]$, with $X_0 \in \mathbb{R}^d$, and $W = \{W_t = (W_t^1, \ldots, W_t^m)^\top, t \in [0, T]\}$ an $\underline{\mathcal{A}}$-adapted m-dimensional Wiener process. Here $a(t, x)$ and $c(t, x, v)$ are d-dimensional vectors of real valued functions on $[0, T] \times \mathbb{R}^d$ and on $[0, T] \times \mathbb{R}^d \times \mathcal{E}$, respectively. Furthermore, $b(t, x)$ is a $d \times m$-matrix of real valued functions on $[0, T] \times \mathbb{R}^d$. Here and in the sequel for a given vector a we adopt the notation a^i to denote its i-th component. Similarly by $b^{i,j}$ we will denote the component of the i-th row and j-th column of a given matrix b.

Moreover, we assume the Lipschitz conditions

$$\begin{aligned} |a(t, x) - a(t, y)|^2 &\leq C_1|x - y|^2, \qquad |b(t, x) - b(t, y)|^2 \leq C_2|x - y|^2, \\ &\int_{\mathcal{E}} |c(t, x, v) - c(t, y, v)|^2 \phi(dv) \leq C_3|x - y|^2, \end{aligned} \qquad (2.3)$$

for every $t \in [0, T]$ and $x, y \in \mathbb{R}^d$, and the linear growth conditions

$$\begin{aligned} |a(t, x)|^2 \leq\ & K_1(1 + |x|^2), \qquad |b(t, x)|^2 \leq K_2(1 + |x|^2), \\ &\int_{\mathcal{E}} |c(t, x, v)|^2 \phi(dv) \leq K_3(1 + |x|^2), \end{aligned} \qquad (2.4)$$

for all $t \in [0, T]$ and $x \in \mathbb{R}^d$. As shown in Ikeda & Watanabe (1989), under conditions (2.3) and (2.4) the SDE (2.2) admits a unique strong solution.

In (2.2) we have defined the jump impact via a stochastic integral with respect to a Poisson random measure as

$$\int_0^t \int_{\mathcal{E}} c(s-, X_{s-}, v) p_\phi(dv \times ds). \qquad (2.5)$$

This choice allows us to model a rather general jump behaviour. The only real restriction we impose on the jump component is the finiteness of the intensity.

If we consider the special case $d = m = r = 1$ with the coefficient functions

$$a(t, x) = \mu x, \quad b(t, x) = \sigma x, \quad c(t, x, v) = x\,(v - 1), \qquad (2.6)$$

and a Poisson measure $p_\phi(dv \times dt)$ with intensity measure $\phi(dv)dt = \lambda f(v) dv dt$, where $f(\cdot)$ is the density function of a lognormal random variable, then the SDE (2.2) reduces to

$$dX_t = X_{t-} \left(\mu dt + \sigma dW_t + \int_{\mathcal{E}} (v - 1)\, p_\phi(dv \times dt) \right), \qquad (2.7)$$

for $t \in [0,T]$. The SDE (2.7) admits the explicit solution

$$X_t = X_0\, e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W(t)} \prod_{i=1}^{p_\phi(t)} \xi_i, \tag{2.8}$$

where $\xi_i = e^{\zeta_i}$ is the i-th lognormal realization of the mark, with $\zeta_i \sim \mathcal{N}(\varrho, \varsigma)$ denoting an independent Gaussian random variable with mean ϱ and variance ς. Equation (2.8) represents a specification of the jump-diffusion asset price process proposed in Merton (1976), known as *Merton model*. A simple case is obtained when the lognormal random variable becomes degenerate and equals a positive constant. If we assume a log-Laplace density $f(\cdot)$ instead of a lognormal one, then we recover the *Kou model* proposed by Kou (2002).

To demonstrate the flexibility of the jump integral representation (2.5), we illustrate in the following three typical examples. It is possible to specify a jump component with time dependent intensity by choosing

$$c(t,x,v) = I_{\{\eta_1(t) \le v \le \eta_2(t)\}} f(t,x), \tag{2.9}$$

where η_1 and η_2 are deterministic functions of time. Here we obtain a jump integral of the form

$$\int_0^t \int_{\eta_1(s)}^{\eta_2(s)} f(s-, X_{s-}) p_\phi(dv \times ds). \tag{2.10}$$

If in (2.9) we allow the functions $\eta_1(t)$ and $\eta_2(t)$ to depend also on the solution X_t, thus permitting a feedback in the intensity, then we obtain the jump integral with stochastic intensity

$$\int_0^t \int_{\eta_1(s,X_{s-})}^{\eta_2(s,X_{s-})} f(s-, X_{s-}) p_\phi(dv \times ds). \tag{2.11}$$

In the modelling and pricing of defaultable claims, models with jumps have been proposed with the intensity being a stochastic process that is driven by a source of risk independent of the one driving the asset price before default. We can easily accommodate this feature with the two-dimensional SDE

$$\begin{aligned} dX_t^1 &= a^1(t, X_t^1)dt + b^1(t, X_t^1)dW_t^1 + \int_{\eta_1(t,X_{t-}^2)}^{\eta_2(t,X_{t-}^2)} f(t-, X_{t-}^1) p_\phi(dv \times dt) \\ dX_t^2 &= a^2(t, X_t^2)dt + b^2(t, X_t^2)dW_t^2, \end{aligned} \tag{2.12}$$

where the first state variable X_t^1 represents the asset price and the second state variable X_t^2 influences its jump intensity at time t. Furthermore, advanced credit risk models with multiple obligors and correlated intensities, as presented in Schönbucher (2003), can be specified via the SDE (2.2).

3 Strong Approximations

In this section we consider, for simplicity, the autonomous one-dimensional SDE

$$dX_t = a(X_t)dt + b(X_t)dW_t + \int_{\mathcal{E}} c(X_{t-}, v)\, p_\phi(dv \times dt), \tag{3.1}$$

for $t \in [0, T]$, with $X_0 \in \mathbb{R}$ and $W = \{W_t, t \in [0, T]\}$ an $\underline{\mathcal{A}}$-adapted one-dimensional Wiener process. We assume an $\underline{\mathcal{A}}$-adapted Poisson measure $p_\phi(dv \times dt)$ with a one-dimensional mark space $\mathcal{E} \subseteq \mathbb{R}\backslash\{0\}$ and with intensity measure $\phi(dv)\, dt = \lambda\, f(v)\, dv\, dt$, where $f(\cdot)$ is a given probability density function. Therefore, the SDE (3.1) can be written in integral form as

$$X_t \;=\; X_0 + \int_0^t a(X_s)ds + \int_0^t b(X_s)dW_s + \sum_{i=1}^{p_\phi(t)} c(X_{\tau_i -}, \xi_i), \tag{3.2}$$

where $\{(\tau_i, \xi_i), i \in \{1, 2 \ldots, p_\phi(t)\}\}$ is the double sequence generated by the Poisson random measure p_ϕ with $p_\phi(t) = p_\phi(\mathcal{E} \times [0, t])$ for $t \in [0, T]$.

In the following we will present several strong discrete time approximations of X in terms of multiple stochastic integrals and coefficient functions. In general, it is not obvious, especially for higher order schemes, how to efficiently obtain the required multiple stochastic integrals. We point out that in filtering applications it is possible to construct the required multiple stochastic integrals from the data. Furthermore, hardware implementations for the generation of multiple stochastic integrals may become available in the near future. It is therefore useful, in view of this kind of applications and developments, to derive higher order strong schemes.

For most applications, such as scenario simulation, needed for instance to check the performance of a hedging strategy, a discrete time approximation is implementable only if one is able to efficiently generate the involved multiple stochastic integrals.

3.1 Strong Taylor Schemes

Let us consider an equidistant time discretisation with n-th discretisation time $t_n = n\Delta$, $n \in \{0, 1, \ldots, N\}$ and time step size $\Delta = \frac{T}{N}$, on which we construct a discrete time approximation $Y^\Delta = \{Y_n^\Delta, n \in \{0, 1, \ldots, N\}\}$ of the solution X of (3.2).

The simplest scheme is the well-known *Euler scheme*, given by

$$\begin{aligned} Y_{n+1} &= Y_n + a(Y_n)\Delta + b(Y_n)\Delta W_n + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} c(Y_n, v)\, p_\phi(dv \times ds) \\ &= Y_n + a(Y_n)\Delta + b(Y_n)\Delta W_n + \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} c(Y_n, \xi_i), \end{aligned} \tag{3.3}$$

for $n \in \{0, 1, \ldots, N-1\}$ with initial value $Y_0 = X_0$.

Here $\Delta W_n = W_{t_{n+1}} - W_{t_n} \sim \mathcal{N}(0, \Delta)$ is the n-th increment of the Wiener process W and $p_\phi(s) = p_\phi(\mathcal{E} \times [0, s])$, as defined in Section 2, is a Poisson distributed random variable with mean $\lambda\, s$ representing the number of jumps of the random measure up to time s. It will be shown later that the Euler scheme (3.3) achieves, in general, a strong order of convergence $\gamma = 0.5$.

When we have a mark independent jump size, that means $c(x, v) = c(x)$, we obtain the Euler scheme

$$Y_{n+1} = Y_n + a(Y_n)\Delta + b(Y_n)\Delta W_n + c(Y_n)\Delta p_n, \tag{3.4}$$

where $\Delta p_n = p_\phi(t_{n+1}) - p_\phi(t_n)$ can be sampled from a Poisson distribution with mean $\lambda\Delta$. For the Merton model SDE (2.7) with jump coefficient $c(t, x, v) = x\,\beta$, with $\beta \geq -1$, we obtain the Euler scheme

$$Y_{n+1} = Y_n + \mu Y_n \Delta + \sigma Y_n \Delta W_n + Y_n\, \beta \Delta p_n. \tag{3.5}$$

When accuracy and efficiency in a simulation are required, it is important to be able to construct numerical methods with higher strong order of convergence. This can be achieved by including more terms from the Wagner-Platen expansion, see Platen (1982b), as will be shown later. It is possible to derive in this way the *order 1.0 strong Taylor scheme*

$$\begin{aligned} Y_{n+1} \;=\; & Y_n + a(Y_n)\Delta + b(Y_n)\Delta W_n + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} c(Y_n, v)\, p_\phi(dv \times ds) \\ & + b(Y_n) b'(Y_n) \int_{t_n}^{t_{n+1}} \int_{t_n}^{s_2} dW(s_1) dW(s_2) \\ & + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{s_2} b(Y_n) c'(Y_n, v) dW(s_1) p_\phi(dv \times ds_2) \\ & + \int_{t_n}^{t_{n+1}} \int_{t_n}^{s_2} \int_{\mathcal{E}} \Big\{ b\big(Y_n + c(Y_n, v)\big) - b\big(Y_n\big) \Big\} p_\phi(dv \times ds_1) dW(s_2) \\ & + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{s_2} \int_{\mathcal{E}} \Big\{ c\big(Y_n + c(Y_n, v_1), v_2\big) - c\big(Y_n, v_2\big) \Big\} \\ & \times p_\phi(dv_1 \times ds_1)\, p_\phi(dv_2 \times ds_2), \end{aligned} \tag{3.6}$$

where

$$b'(x) := \frac{d\, b(x)}{d\, x} \quad \text{and} \quad c'(x, v) := \frac{\partial c(x, v)}{\partial x}. \tag{3.7}$$

This scheme will be shown to achieve, in general, strong order $\gamma = 1.0$. In the case of mark independent jump size we obtain

$$\begin{aligned}Y_{n+1} &= Y_n + a(Y_n)\Delta + b(Y_n)\Delta W_n + c(Y_n)\Delta p_n + b(Y_n)b'(Y_n)\, I_{(1,1)} \\ &\quad + b(Y_n)\, c'(Y_n)\, I_{(1,-1)} + \{b(Y_n +\, c(Y_n)) - b(Y_n)\}\, I_{(-1,1)} \\ &\quad + \{c\,(Y_n + c(Y_n)) - c(Y_n)\}\, I_{(-1,-1)},\end{aligned} \tag{3.8}$$

with multiple stochastic integrals

$$\begin{aligned}I_{(1,1)} &= \int_{t_n}^{t_{n+1}} \int_{t_n}^{s_2} dW(s_1)dW(s_2), \\ I_{(1,-1)} &= \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{s_2} dW(s_1)p_\phi(dv \times ds_2), \\ I_{(-1,1)} &= \int_{t_n}^{t_{n+1}} \int_{t_n}^{s_2} \int_{\mathcal{E}} p_\phi(dv \times ds_1)dW(s_2), \\ I_{(-1,-1)} &= \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{s_2} \int_{\mathcal{E}} p_\phi(dv_1 \times ds_1)p_\phi(dv_2 \times ds_2).\end{aligned} \tag{3.9}$$

The level of complexity of the scheme (3.6), even in the case (3.8) of mark independent jump size, is quite high when compared to the Euler scheme (3.4). Indeed, it requires not only function evaluations of the drift, diffusion and jump coefficients, but also of their derivatives. This problem can be overcome by constructing derivative free schemes that will be presented in Section 3.2.

In view of applications in scenario simulations, a main problem concerns the generation of the multiple stochastic integrals appearing in (3.8). By application of Itô's lemma for jump-diffusion processes and the integration by parts formula, we can express the four stochastic integrals appearing in (3.8) as

$$\begin{aligned}I_{(1,1)} &= \tfrac{1}{2}\Big\{(\Delta W_n)^2 - \Delta\Big\}, \qquad & I_{(1,-1)} &= \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} W_{\tau_i} - \Delta p_n\, W_{t_n}, \\ I_{(-1,1)} &= \Delta p_n\, \Delta W_n - I_{(1,-1)}, \qquad & I_{(-1,-1)} &= \frac{1}{2}\Big\{(\Delta p_n)^2 - \Delta p_n\Big\}.\end{aligned} \tag{3.10}$$

The generation of $I_{(1,1)}$ and $I_{(-1,-1)}$ is straightforward once we have generated the random variables ΔW_n and Δp_n. The generation of the mixed multiple stochastic integrals, $I_{(1,-1)}$ and $I_{(-1,1)}$, is more complex as it requires to keep track of the jump times between discretisation points for the evaluation of W_{τ_i}. Conditioned on the number of jump events realized on the time interval $(t_n, t_{n+1}]$, the jump times are independent and uniformly distributed on this interval. Therefore, once we have generated Δp_n, we can sample Δp_n independent random numbers from a uniform distribution on $(t_n, t_{n+1}]$ in order to obtain the exact location of the jump times. However, from a computational point of view, this makes the efficiency of the algorithm heavily dependent on the level of the intensity of the Poisson measure. Indeed, the number of operations involved in an algorithm as the Euler scheme (3.4) seems not to depend on the level of the intensity. We are here neglecting the additional time needed to sample from a Poisson distribution with higher intensity. On the other hand, for the scheme (3.8) the number of computations is proportional to the number of jumps, due to the generation of the two double stochastic integrals $I_{(1,-1)}$ and $I_{(-1,1)}$. Therefore, this algorithm is not very efficient for the simulation of jump-diffusion SDEs driven by high intensity Poisson measures.

It is, in principle, possible to derive strong Taylor schemes of any given order, as will be demonstrated in Section 6. However, the schemes become rather complex and, in applications such as scenario simulation, the generation of the multiple stochastic integrals is not straightforward. Moreover, as explained above, for SDEs driven by high intensity Poisson measures, these schemes become inefficient. For these reasons we will not present in this section any scheme with order of strong convergence higher than $\gamma = 1.0$. For the construction of higher order schemes we refer to Section 4, where we are going to present jump adapted approximations that avoid all multiple stochastic integrals involving the Poisson measure, making the schemes much easier to derive and implement.

3.1.1 Commutativity Condition

As discussed previously, higher order Taylor schemes, even with mark independent jump size, become computationally inefficient when the intensity of the Poisson measure is high. Here the number of operations involved is proportional to the intensity level. Also the jump adapted schemes, to be presented in Section 4, show a similar dependence on the intensity of the jumps.

Analyzing the multiple stochastic integrals required for the scheme (3.8), we observe that the dependence on the jump times only affects the mixed multiple stochastic integrals $I_{(1,-1)}$ and $I_{(-1,1)}$. However, since by (3.10) we have

$$I_{(-1,1)} = \Delta p_n \, \Delta W_n - I_{(1,-1)}, \tag{3.11}$$

the sum of these integrals is obtained as

$$I_{(1,-1)} + I_{(-1,1)} = \Delta p_n \, \Delta W_n, \tag{3.12}$$

and thus independent of the particular jump times. Therefore, in the case of mark independent jump size $c(t, x, v) = c(t, x)$, for SDEs satisfying the *commutativity condition*

$$b(t,x)\frac{\partial c(t,x)}{\partial x} = b\Big(t, x + c(t,x)\Big) - b\Big(t,x\Big), \tag{3.13}$$

for all $t \in [0, T]$ and $x \in \mathbb{R}$, the order 1.0 strong Taylor scheme (3.8) does not require to keep track of the exact location of the jump times. Hence, its computational complexity is independent of the intensity level. This is an important observation from the practical point of view. If a given SDE satisfies the commutativity condition, then considerable savings in computational time can be achieved.

When we have a linear diffusion coefficient of the form

$$b(t,x) = a_1(t) + a_2(t)\,x, \tag{3.14}$$

as it frequently occurs in finance, the commutativity condition (3.13) implies the following ordinary differential equation for the jump coefficient

$$\frac{\partial c(t,x)}{\partial x} = \frac{a_2(t)\,c(t,x)}{a_1(t) + a_2(t)\,x}, \tag{3.15}$$

for all $t \in [0,T]$. Therefore, for linear diffusion coefficients of the form (3.14) the class of SDEs satisfying the commutativity condition (3.13) is identified by mark independent jump coefficients of the form

$$c(t,x) = e^{K(t)}\Big(a_1(t) + a_2(t)\,x\Big), \tag{3.16}$$

where $K(t)$ is an arbitrary function of time.

For instance, the SDE (2.7) with mark independent jump size $c(t,x,v) = x\,\beta$, with $\beta \geq -1$, satisfies the commutativity condition (3.13) and the corresponding order 1.0 strong Taylor scheme is given by

$$\begin{aligned} Y_{n+1} &= Y_n + \mu Y_n \Delta + \sigma Y_n \Delta W_n + \beta\, Y_n\, \Delta p_n + \frac{1}{2}\sigma^2\, Y_n\, \{(\Delta W_n)^2 - \Delta\} \\ &\quad + \sigma\,\beta\, Y_n \Delta p_n \Delta W_n + \frac{1}{2}\,\beta^2\, Y_n \{(\Delta p_n)^2 - \Delta p_n\}. \end{aligned} \tag{3.17}$$

In Table 1 we present some diffusion coefficients from models proposed in the finance literature and corresponding jump coefficients that satisfy the commutative condition (3.13).

$b(t,x)$	$c(t,x)$
$a_1(t)$	$K(t)$
$a_1(t) + a_2(t)\,x$	$e^{K(t)}(a_1(t) + a_2(t)\,x)$
$a_3(t)\sqrt{a_1(t) + a_2(t)x}$	$a_2(t)\,e^{K(t)} + 2e^{\frac{K(t)}{2}}\sqrt{a_1(t) + a_2(t)x}$
$a_1(t)(1 - e^{-x})$	$\log\{1 + e^{K(t)} - e^{-x+K(t)}\}$
$a_1(t)\,x^{\frac{3}{2}}$	$\dfrac{-2e^{\frac{3K(t)}{2}}x^{\frac{3}{2}} + 3e^{K(t)}x^2 - x^3}{e^{2K(t)} - 2e^{K(t)}x + x^2}$

Table 1: Coefficients satisfying the commutativity condition.

3.2 Derivative Free Schemes

Higher order schemes, as the order 1.0 strong Taylor scheme presented in Section 3.1, are rather complex as they involve the evaluation of derivatives of the drift and diffusion coefficients at each time step. For the implementation of general numerical routines for the approximation of jump-diffusion SDEs, that means without assuming a particular form for the coefficients, this constitutes a serious limitation, as one is required to include a symbolic differentiation in a numerical algorithm. In this section we propose strong schemes that avoid the computation of derivatives.

By replacing the derivatives in the scheme (3.6) with the corresponding difference ratios it is possible to obtain a scheme, with the same strong order of convergence, that does not require the evaluation of derivatives. However, to construct the difference ratios we need supporting values of the coefficients at additional points.

The *explicit order 1.0 strong Taylor scheme*, which achieves a strong order $\gamma = 1.0$, is given by

$$\begin{aligned} Y_{n+1} &= Y_n + a(Y_n)\Delta + b(Y_n)\Delta W_n + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} c(Y_n, v)\, p_\phi(dv \times ds) \\ &+ \frac{1}{\sqrt{\Delta}} \Big\{ b(\overline{Y}_n) - b(Y_n) \Big\} \int_{t_n}^{t_{n+1}} \int_{t_n}^{s_2} dW(s_1)\, dW(s_2) \\ &+ \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{s_2} \frac{1}{\sqrt{\Delta}} \Big\{ c(\overline{Y}_n, v) - c(Y_n, v) \Big\} dW(s_1)\, p_\phi(dv \times ds_2) \\ &+ \int_{t_n}^{t_{n+1}} \int_{t_n}^{s_2} \int_{\mathcal{E}} \Big\{ b(Y_n + c(Y_n, v)) - b(Y_n) \Big\} p_\phi(dv \times ds_1)\, dW(s_2) \\ &+ \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{s_2} \int_{\mathcal{E}} \Big\{ c(Y_n + c(Y_n, v_2), v_1) - c(Y_n, v_1) \Big\} \\ &\times p_\phi(dv_1 \times ds_1)\, p_\phi(dv_2 \times ds_2), \end{aligned} \tag{3.18}$$

with the supporting value

$$\overline{Y}_n = Y_n + b(Y_n)\sqrt{\Delta}. \tag{3.19}$$

Even in the case of a mark independent jump size, the derivative free coefficient of the multiple stochastic integral $I_{(1,-1)}$, which is

$$\frac{c(\overline{Y}_n) - c(Y_n)}{\sqrt{\Delta}}, \tag{3.20}$$

depends on the time step size Δ, while the the coefficient of the multiple stochastic integral $I_{(-1,1)}$,

$$b\Big(Y_n + c(Y_n)\Big) - b\Big(Y_n\Big), \tag{3.21}$$

is independent of Δ. Therefore, it is not possible to derive commutativity conditions similar to (3.13) that permit to identify special classes of SDEs for which the computational efficiency is independent of the jump intensity level. For instance, for the SDE (2.7) with mark independent jump size $c(t,x,v) = x\,\beta$, with $\beta \geq -1$, the explicit order 1.0 strong Taylor scheme is given by

$$\begin{aligned} Y_{n+1} &= Y_n + \mu Y_n \Delta + \sigma Y_n \Delta W_n + \beta\, Y_n\, \Delta p_n + \frac{\sigma}{\sqrt{\Delta}} \{\bar{Y}_n - Y_n\} I_{(1,1)} \\ &\quad + \frac{\beta}{\sqrt{\Delta}} \{\bar{Y}_n - Y_n\} I_{(1,-1)} + \sigma\,\beta\, Y_n\, I_{(-1,1)} + \beta^2\, Y_n I_{(-1,-1)}, \end{aligned} \tag{3.22}$$

with the supporting value

$$\bar{Y}_n = Y_n + \sigma Y_n \sqrt{\Delta}. \tag{3.23}$$

Since the evaluation of the multiple stochastic integrals $I_{(1,-1)}$ and $I_{(-1,1)}$, as given in (3.10), depends on the number of jumps, the computational efficiency of the scheme (3.22)-(3.23) will depend on the intensity λ of the jump measure.

For the special class of mark independent SDEs characterized by the commutativity condition (3.13), using the relationship

$$I_{(1,-1)} + I_{(-1,1)} = \Delta p_n\, \Delta W_n, \tag{3.24}$$

and substituting it in (3.8), we first derive the order 1.0 strong Taylor scheme, given by

$$\begin{aligned} Y_{n+1} &= Y_n + a(Y_n)\Delta + b(Y_n)\Delta W_n + c(Y_n)\Delta p_n \\ &\quad + \frac{b(Y_n)b'(Y_n)}{2} \{(\Delta W_n)^2 - \Delta\} + \{b(Y_n + c(Y_n)) - b(Y_n)\} \Delta p_n\, \Delta W_n \\ &\quad + \frac{\{c(Y_n + c(Y_n)) - c(Y_n)\}}{2} \{(\Delta p_n)^2 - \Delta p_n\}. \end{aligned} \tag{3.25}$$

Then, by replacing the derivative b' with the corresponding difference ratio, we obtain an explicit order 1.0 strong Taylor scheme

$$\begin{aligned} Y_{n+1} &= Y_n + a(Y_n)\Delta + b(Y_n)\Delta W_n + c(Y_n)\Delta p_n \\ &\quad + \frac{\{b(\bar{Y}_n) - b(Y_n)\}}{2\sqrt{\Delta}} \{(\Delta W_n)^2 - \Delta\} + \{b(Y_n + c(Y_n)) - b(Y_n)\} \Delta p_n\, \Delta W_n \\ &\quad + \frac{\{c(Y_n + c(Y_n)) - c(Y_n)\}}{2} \{(\Delta p_n)^2 - \Delta p_n\}, \end{aligned} \tag{3.26}$$

with the supporting value

$$\bar{Y}_n = Y_n + b(Y_n)\sqrt{\Delta}, \tag{3.27}$$

whose computational efficiency is independent on the intensity level.

For instance, for Merton's SDE (2.7) with $c(t,x,v) = x\,\beta$, with $\beta \geq -1$, we can derive the explicit order 1.0 strong Taylor scheme, which, due to the linearity of the diffusion coefficient, is the same as the order 1.0 strong Taylor scheme (3.17).

3.3 Implicit Schemes

As shown in Hofmann & Platen (1996) for the case of SDEs driven only by Wiener processes, when one has multiplicative noise explicit methods show narrow regions of numerical stability. SDEs with multiplicative noise are often employed when modelling asset prices in finance. They also arise in other important applications such as hidden Markov chain filtering. In order to construct approximate filters, one needs a strong discrete time approximation of an SDE with multiplicative noise, the Zakai equation, see Elliott, Aggoun & Moore (1995). Moreover, in filtering problems for large systems it is often not possible to employ small time step sizes, as the computations may not be performed fast enough to keep pace with the arrival of data. Therefore, for this kind of applications, higher order schemes with wide regions of stability are crucial. To overcome some of these problems, implicit schemes have been constructed that have good numerical stability properties.

In general, given an explicit scheme of strong order γ it is possible to obtain a drift-implicit scheme of the same order. Since the reciprocal of a Gaussian random variable does not have finite absolute moments, it is usually not possible to introduce implicitness easily in the diffusion coefficient. Regions of stability of drift-implicit schemes are typically wider than those of corresponding explicit schemes. Therefore, the former are often more suitable to filtering problems than corresponding explicit schemes.

In Higham & Kloeden (2004), a class of drift-implicit methods of strong order $\gamma = 0.5$ for jump-diffusion SDEs has been proposed and analyzed. We will present drift-implicit schemes of higher strong order for the jump-diffusion SDE (3.1).

From the Euler scheme (3.3), by introducing implicitness in the drift, we obtain the *drift-implicit Euler scheme*,

$$\begin{aligned} Y_{n+1} &= Y_n + \{\theta\, a(Y_{n+1}) + (1-\theta)\, a(Y_n)\}\, \Delta + b(Y_n)\Delta W_n \\ &\quad + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} c(Y_n, v)\, p_\phi(dv \times ds), \end{aligned} \tag{3.28}$$

where the parameter $\theta \in [0,1]$ characterizes the degree of implicitness. This scheme achieves a strong order of convergence $\gamma = 0.5$.

By comparing the drift-implicit Euler scheme (3.28) with the Euler scheme (3.3), one notices that there is an additional computational effort required to solve an algebraic equation at each time step. This can be performed, for instance, by a Newton-Raphson method.

In a similar way as for the Euler scheme, by introducing implicitness in the drift of the order 1.0 strong Taylor scheme (3.6), we obtain the *drift-implicit order 1.0* *strong Taylor scheme*

$$
\begin{aligned}
Y_{n+1} &= Y_n + \{\theta\, a(Y_{n+1}) + (1-\theta)\, a(Y_n)\}\,\Delta + b(Y_n)\Delta W_n \\
&\quad + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} c(Y_n, v)\, p_\phi(dv \times ds) + b(Y_n) b'(Y_n) \int_{t_n}^{t_{n+1}} \int_{t_n}^{s_2} dW(s_1) dW(s_2) \\
&\quad + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{s_2} b(Y_n) c'(Y_n, v) dW(s_1) p_\phi(dv \times ds_2) \\
&\quad + \int_{t_n}^{t_{n+1}} \int_{t_n}^{s_2} \int_{\mathcal{E}} \Big\{ b(Y_n + c(Y_n, v)) - b(Y_n) \Big\} p_\phi(dv \times ds_1) dW(s_2) \\
&\quad + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{s_2} \int_{\mathcal{E}} \Big\{ c(Y_n + c(Y_n, v_1), v_2) - c(Y_n, v_2) \Big\} \\
&\quad \times p_\phi(dv_1 \times ds_1)\, p_\phi(dv_2 \times ds_2),
\end{aligned} \tag{3.29}
$$

where the parameter $\theta \in [0,1]$, characterizes again the degree of implicitness. This scheme achieves a strong order of convergence $\gamma = 1.0$.

The commutativity condition (3.13), presented in Section 3.1.1, also applies to drift-implicit schemes. Therefore, for the class of SDEs identified by the commutativity condition (3.13) the computational efficiency of drift-implicit schemes of order $\gamma = 1.0$ is not dependent on the intensity level of the Poisson measure. For instance, for Merton's SDE (2.7) with $c(t,x,v) = x\,\beta$ and $\beta \geq -1$ it is possible to derive drift-implicit schemes that are efficient also for a high intensity jump measure.

4 Jump Adapted Approximations

In principle, by including enough terms from the Wagner-Platen expansion, to be presented in Section 6, it is possible to derive schemes of any given strong order of convergence. However, as noticed in Section 3.1, even for a one-dimensional autonomous SDE, higher order schemes are quite complex in that they involve multiple stochastic integrals with respect to the Wiener process and the Poisson random measure. In particular, when we have a mark dependent jump size, the generation of the required multiple stochastic integrals involving the Poisson measure can be complicated. As noticed before, there are applications, such as filtering, in which we are able to construct the multiple stochastic integrals directly from data. In these cases the proposed strong schemes can be readily applied. However, for scenario simulation we need to generate artificially the multiple stochastic integrals. To avoid the generation of multiple stochastic integrals with respect to the Poisson jump measure, Platen (1982a) proposed the so-called *jump adapted approximations* that significantly reduce the complexity of higher order schemes. A jump adapted time discretisation makes these schemes easily implementable for scenario simulation also in the case of a mark dependent jump size. Indeed, between the jump times the evolution of the SDE (2.2) is that of a diffusion without jumps and can be approximated by standard schemes, as presented in Kloeden & Platen (1999). At the jump time the prescribed jump is performed. Therefore, as we will show in this section, it is possible to develop tractable jump adapted higher order strong schemes also in the general case of mark dependent jump sizes, as the required multiple stochastic integrals involve only time and Wiener process integrations.

4.1 Jump Adapted Strong Schemes

We consider now a *jump adapted time discretisation* $0 = t_0 < t_1 < \ldots < t_N = T$, constructed by a superposition of the jump times $\{\tau_1, \tau_2, \ldots\}$ of the Poisson measure p_ϕ to a deterministic equidistant grid with maximum step size $\Delta > 0$. This means that we add all the random jump times to an equidistant grid, as the one presented in Section 3.1. In this way the maximum time step size of the jump adapted discretisation is assured to be Δ.

Within this time grid we can separate the diffusive part from the jumps, because the jumps can arise only at discretisation times. Therefore, we can approximate between discretisation points the diffusive part with a strong Taylor scheme for diffusion processes. We add the effect of a jump to the evolution of the approximate solution when we encounter a jump time as discretisation time. We remark that with jump adapted schemes the approximation of SDEs with mark dependent jump size becomes a trivial task. Therefore, in this section we consider the general case of a jump-diffusion SDE with mark dependent jump size given in (3.1). At first, note that we set $Y_n = Y_{t_n}$ and we define

$$Y_{t_{n+1}-} = \lim_{s \to t_{n+1}-} Y_s,$$

where $s < t_{n+1}-$ in the almost sure limit.

We present the *jump adapted Euler scheme* given by

$$Y_{t_{n+1}-} = Y_{t_n} + a(Y_{t_n})\Delta_{t_n} + b(Y_{t_n})\Delta W_{t_n} \tag{4.1}$$

and

$$Y_{t_{n+1}} = Y_{t_{n+1}-} + \int_{\mathcal{E}} c(Y_{t_{n+1}-}, v)\, p_\phi(dv \times \{t_{n+1}\}), \tag{4.2}$$

where $\Delta_{t_n} = t_{n+1} - t_n$ and $\Delta W_{t_n} = W_{t_{n+1}} - W_{t_n} \sim \mathcal{N}(0, \Delta_{t_n})$. The impact of jumps is simulated by (4.2). If t_{n+1} is a jump time, then $\int_{\mathcal{E}} p_\phi(dv \times \{t_{n+1}\}) = 1$ and

$$\int_{\mathcal{E}} c(Y_{t_{n+1}-}, v)\, p_\phi(dv \times \{t_{n+1}\}) = c(Y_{t_{n+1}-}, \xi_{p_\phi(t_{n+1})}), \tag{4.3}$$

while if t_{n+1} is not a jump time one has $Y_{t_{n+1}} = Y_{t_{n+1}-}$, as $\int_{\mathcal{E}} p_\phi(dv \times \{t_{n+1}\}) = 0$. Therefore, the strong order of convergence of the jump adapted Euler scheme is

$\gamma = 0.5$, resulting from the strong order of the approximation (4.1) of the diffusive component.

As the order of convergence of jump adapted schemes is, in general, the one induced by the approximation of the diffusive part, we can derive the *jump adapted order 1.0 strong Taylor scheme* given by

$$Y_{t_{n+1}-} = Y_{t_n} + a(Y_{t_n})\Delta_{t_n} + b(Y_{t_n})\Delta W_{t_n} + \frac{b(Y_{t_n})b'(Y_{t_n})}{2}\Big((\Delta W_{t_n})^2 - \Delta_{t_n}\Big) \quad (4.4)$$

and

$$Y_{t_{n+1}} = Y_{t_{n+1}-} + \int_{\mathcal{E}} c(Y_{t_{n+1}-}, v)\, p_\phi(dv \times \{t_{n+1}\}), \quad (4.5)$$

which achieves strong order $\gamma = 1.0$.

The comparison of the jump adapted order 1.0 strong scheme (4.4)-(4.5) with the order 1.0 strong Taylor scheme (3.6), shows that jump adapted schemes are much simpler. These avoid the problem of the generation of multiple stochastic integrals with respect to the Poisson measure. If we approximate the diffusive part of the SDE (3.1) with the order 1.5 strong Taylor scheme, see Kloeden & Platen (1999), we obtain the *jump adapted order 1.5 strong Taylor scheme* given by

$$\begin{aligned} Y_{t_{n+1}-} \;=\;& Y_{t_n} + a(Y_{t_n})\Delta_{t_n} + b(Y_{t_n})\Delta W_{t_n} + \frac{b(Y_{t_n})b'(Y_{t_n})}{2}\left((\Delta W_{t_n})^2 - \Delta_{t_n}\right) \\ & + a'(Y_{t_n})b(Y_{t_n})\Delta Z_{t_n} + \frac{1}{2}\left(a(Y_{t_n})a'(Y_{t_n}) + \frac{1}{2}\left(b(Y_{t_n})\right)^2 a''(Y_{t_n})\right)(\Delta_{t_n})^2 \\ & + \left(a(Y_{t_n})b'(Y_{t_n}) + \frac{1}{2}b(Y_{t_n})^2 b''(Y_{t_n})\right)(\Delta W_{t_n}\Delta_{t_n} - \Delta Z_{t_n}) \\ & + \frac{1}{2}b(Y_{t_n})\left(b(Y_{t_n})b''(Y_{t_n}) + (b'(Y_{t_n}))^2\right) \\ & \times \left\{\frac{1}{3}(\Delta W_{t_n})^2 - \Delta_{t_n}\right\}\Delta W_{t_n}, \end{aligned} \quad (4.6)$$

and

$$Y_{t_{n+1}} = Y_{t_{n+1}-} + \int_{\mathcal{E}} c(Y_{t_{n+1}-}, v)\, p_\phi(dv \times \{t_{n+1}\}), \quad (4.7)$$

where

$$\Delta Z_{t_n} = \int_{t_n}^{t_{n+1}} \int_{t_n}^{s_2} dW_{s_1}\, ds_2. \quad (4.8)$$

One can show that ΔZ_{t_n} has a Gaussian distribution with mean $E(\Delta Z_{t_n}) = 0$, variance $E((\Delta Z_{t_n})^2) = \frac{1}{3}(\Delta_{t_n})^3$ and covariance $E(\Delta Z_{t_n}\, \Delta W_{t_n}) = \frac{1}{2}(\Delta_{t_n})^2$. Therefore, with two independent $\mathcal{N}(0,1)$ distributed standard Gaussian random variables U_1 and U_2, we can obtain the required correlated random variables ΔZ_{t_n}

and ΔW_{t_n} by setting:

$$\Delta W_{t_n} = U_1 \sqrt{\Delta_{t_n}} \quad \text{and} \quad \Delta Z_{t_n} = \frac{1}{2} \left(\Delta_{t_n}\right)^{\frac{3}{2}} \left(U_1 + \frac{1}{\sqrt{3}} U_2 \right). \tag{4.9}$$

For the SDE (2.7) the terms involving the random variable ΔZ_{t_n} cancel out, thus yielding a rather simple jump adapted order 1.5 strong Taylor scheme.

Constructing strong schemes of higher order is, in principle, not difficult. However, as they involve multiple stochastic integrals of higher multiplicity, they can become quite complex. Therefore, we will not present here any scheme of strong order higher than $\gamma = 1.5$. Instead we refer to the convergence theorem to be presented in Section 8 that provides the methodology for the construction of jump adapted schemes of any given strong order.

4.2 Jump Adapted Derivative Free Schemes

As noticed in Section 3.2, it is convenient to develop higher order numerical approximations that do not require the evaluation of derivatives of the coefficient functions. With jump adapted schemes it is sufficient to replace the numerical scheme of the diffusive part with an equivalent derivative free scheme. We refer to Kloeden & Platen (1999) for derivative free schemes for diffusion processes.

The *jump adapted explicit order 1.0 strong Taylor scheme*, which achieves a strong order $\gamma = 1.0$, is given by

$$\begin{aligned} Y_{t_{n+1}-} &= Y_{t_n} + a(Y_{t_n})\Delta_{t_n} + b(Y_{t_n})\Delta W_{t_n} \\ &\quad + \frac{1}{2\sqrt{\Delta_{t_n}}} \{ b(\overline{Y}_{t_n}) - b(Y_{t_n}) \} \left((\Delta W_{t_n})^2 - \Delta_{t_n} \right), \end{aligned} \tag{4.10}$$

and

$$Y_{t_{n+1}} = Y_{t_{n+1}-} + \int_{\mathcal{E}} c(Y_{t_{n+1}-}, v)\, p_\phi(dv \times \{t_{n+1}\}), \tag{4.11}$$

with the supporting value

$$\overline{Y}_{t_n} = Y_{t_n} + b(Y_{t_n})\sqrt{\Delta_{t_n}}. \tag{4.12}$$

The *jump adapted explicit order 1.5 strong Taylor scheme* is given by

$$
\begin{aligned}
Y_{t_{n+1}-} \;=\; & Y_{t_n} + b(Y_{t_n})\Delta W_{t_n} + \frac{1}{2\sqrt{\Delta_{t_n}}}\left\{a(\overline{Y}^{+}_{t_n}) - a(\overline{Y}^{-}_{t_n})\right\}\Delta Z_{t_n} \\
& + \frac{1}{4}\left\{a(\overline{Y}^{+}_{t_n}) + 2a(Y_{t_n}) + a(\overline{Y}^{-}_{t_n})\right\}\Delta_{t_n} \\
& + \frac{1}{4\sqrt{\Delta_{t_n}}}\left\{b(\overline{Y}^{+}_{t_n}) - b(\overline{Y}^{-}_{t_n})\right\}\left((\Delta W_{t_n})^2 - \Delta_{t_n}\right) \\
& + \frac{1}{2\sqrt{\Delta_{t_n}}}\left\{b(\overline{Y}^{+}_{t_n}) + 2b(Y_{t_n}) + b(\overline{Y}^{-}_{t_n})\right\}\left(\Delta W_{t_n}\Delta_{t_n} - \Delta Z_{t_n}\right) \\
& + \frac{1}{4\sqrt{\Delta_{t_n}}}\left[b(\overline{\Phi}^{+}_{t_n}) - b(\overline{\Phi}^{-}_{t_n}) - b(\overline{Y}^{+}_{t_n}) + b(\overline{Y}^{-}_{t_n})\right] \\
& \times \left\{\frac{1}{3}(\Delta W_{t_n})^2 - \Delta_{t_n}\right\}\Delta W_{t_n}, \qquad (4.13)
\end{aligned}
$$

and

$$Y_{t_{n+1}} = Y_{t_{n+1}-} + \int_{\mathcal{E}} c(Y_{t_{n+1}-}, v)\, p_\phi(dv \times \{t_{n+1}\}), \qquad (4.14)$$

with

$$\overline{Y}^{\pm}_{t_n} = Y_{t_n} + a(Y_{t_n})\Delta_{t_n} \pm b(Y_{t_n})\Delta W_{t_n}, \qquad (4.15)$$

and

$$\overline{\Phi}^{\pm}_{t_n} = \overline{Y}^{\pm}_{t_n} \pm b(\overline{Y}^{+}_{t_n})\sqrt{\Delta_{t_n}}. \qquad (4.16)$$

4.3 Jump Adapted Implicit Schemes

As discussed previously, for applications such as filtering it is crucial to construct higher order schemes with wide regions of numerical stability. To achieve this one needs to introduce implicitness into the schemes. For deriving jump adapted drift-implicit schemes, it is sufficient to replace the explicit scheme for the diffusive part by a drift-implicit one. We refer to Kloeden & Platen (1999) for drift-implicit methods for SDEs driven by Wiener processes.

For the SDE (3.1) the *jump adapted drift-implicit Euler scheme* is given by:

$$Y_{t_{n+1}-} = Y_{t_n} + \left\{\theta\, a(Y_{t_{n+1}}) + (1-\theta)\, a(Y_{t_n})\right\}\Delta_{t_n} + b(Y_{t_n})\Delta W_{t_n}, \qquad (4.17)$$

and

$$Y_{t_{n+1}} = Y_{t_{n+1}-} + \int_{\mathcal{E}} c(Y_{t_{n+1}-}, v) p_\phi(dv \times \{t_{n+1}\}), \qquad (4.18)$$

where the parameter $\theta \in [0,1]$ characterizes the degree of implicitness.

Using a drift-implicit order 1.0 strong Taylor scheme for the diffusive part, we obtain the *jump adapted drift-implicit order 1.0 strong Taylor scheme*

$$
\begin{aligned}
Y_{t_{n+1}-} &= Y_{t_n} + \left\{\theta\, a(Y_{t_{n+1}}) + (1-\theta)\, a(Y_{t_n})\right\} \Delta_{t_n} + b(Y_{t_n})\Delta W_{t_n} \\
&\quad + \frac{b(Y_{t_n})b'(Y_{t_n})}{2}\left((\Delta W_{t_n})^2 - \Delta_{t_n}\right) \qquad (4.19)
\end{aligned}
$$

and

$$
Y_{t_{n+1}} = Y_{t_{n+1}-} + \int_{\mathcal{E}} c(Y_{t_{n+1}-}, v)\, p_\phi(dv \times \{t_{n+1}\}), \qquad (4.20)
$$

which achieves strong order $\gamma = 1.0$.

Finally, we present a *jump adapted drift-implicit order 1.5 strong Taylor scheme* given by

$$
\begin{aligned}
Y_{t_{n+1}-} &= Y_{t_n} + \frac{1}{2}\left\{a(Y_{t_{n+1}}) + a(Y_{t_n})\right\} \Delta_{t_n} + b(Y_{t_n})\, \Delta W_{t_n} \\
&\quad + \frac{b(Y_{t_n})b'(Y_{t_n})}{2}\left((\Delta W_{t_n})^2 - \Delta_{t_n}\right) \\
&\quad + \left(a(Y_{t_n})b'(Y_{t_n}) + \frac{1}{2}b(Y_{t_n})^2 b''(Y_{t_n})\right)\left(\Delta W_{t_n}\Delta_{t_n} - \Delta Z_{t_n}\right) \\
&\quad + a'(Y_{t_n})b(Y_{t_n})\left\{\Delta Z_{t_n} - \frac{1}{2}\Delta W_{t_n}\Delta_{t_n}\right\} \\
&\quad + \frac{1}{2}b(Y_{t_n})\left(b(Y_{t_n})b''(Y_{t_n}) + (b'(Y_{t_n}))^2\right) \\
&\quad \times \left\{\frac{1}{3}\left(\Delta W_{t_n}\right)^2 - \Delta_{t_n}\right\}\Delta W_{t_n}, \qquad (4.21)
\end{aligned}
$$

and

$$
Y_{t_{n+1}} = Y_{t_{n+1}-} + \int_{\mathcal{E}} c(Y_{t_{n+1}-}, v) p_\phi(dv \times \{t_{n+1}\}). \qquad (4.22)
$$

5 Numerical Results

In this section we present numerical results for the strong schemes presented for the SDE (2.7) describing the Merton model. We select the following default parameters: $\mu = -0.05$, $\sigma = 0.1$, $\lambda = 1$, $X_0 = 1$, $T = 0.5$. At first we consider the case of a mark independent jump size. We consider the SDE (2.7) with jump coefficient $c(t, x, v) = x\,\beta$ and set $\beta = 0.1$. In the following we report the strong error $\varepsilon(\Delta)$, as defined in (1.1), when comparing the results of the strong schemes with the closed form solution (2.8). In the corresponding plots we show the logarithm $\log_2(\varepsilon(\Delta))$ of the strong error versus the logarithm $\log_2(\Delta)$ of the time step size. The number of simulations depends on the scheme implemented.

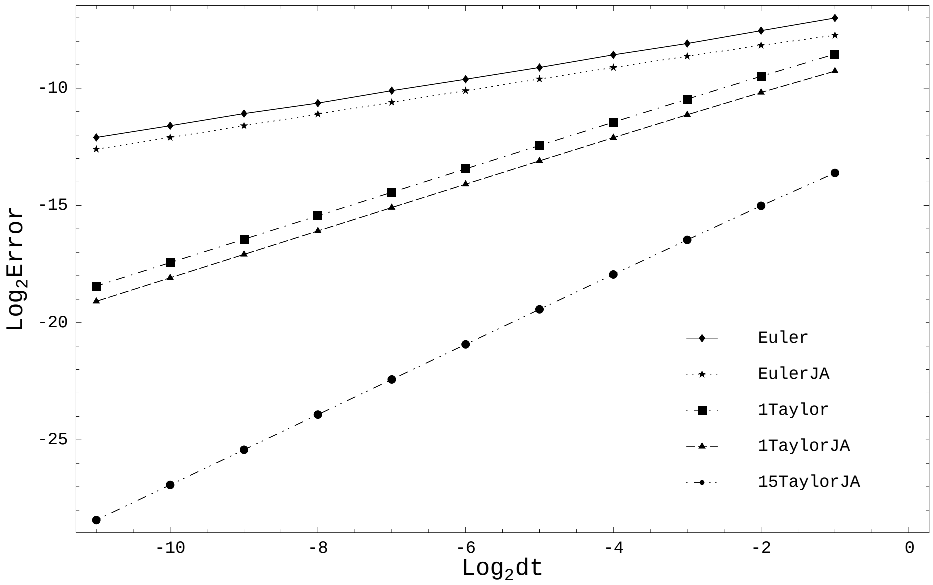

Figure 5.1: Log-log plot of strong error versus time step size.

It will always be chosen such that the statistical errors become negligible when compared to the systematic errors caused by the time discretisation.

In Figure 5.1, we report the results obtained from the Euler, jump adapted Euler, 1.0 Taylor, jump adapted 1.0 Taylor and jump adapted 1.5 Taylor schemes, presented in Sections 3 and 4. We confirm that the two Euler schemes achieve an order of strong convergence of about 0.5. The two 1.0 Taylor schemes achieve an order close to 1.0 and the jump adapted 1.5 Taylor scheme shows an order of strong convergence of about 1.5. These experimental results are consistent with the previously described strong orders, as will be stated in the convergence theorems to be presented in Section 6 and in Section 8. We also notice that when comparing a strong Taylor scheme with the jump adapted scheme of the same order, the jump adapted one is more accurate. This effect is due to the more accurate simulation of the jump impact at the correct jump time within the jump adapted schemes. However, as explained before, for higher intensity jump adapted schemes may not be computationally efficient.

We consider now the mark dependent jump coefficient $c(t, x, v) = x\,(v - 1)$, with the marks drawn from a lognormal distribution with a mean of 1.1 and a standard deviation of 0.02. As explained in Section 3, jump-diffusion SDEs with mark dependent jump size can be handled efficiently by resorting to jump adapted schemes. Therefore, in Figure 5.2 we compare the following jump adapted schemes: Euler, implicit Euler, 1.0 Taylor, implicit 1.0 Taylor, 1.5 Taylor and implicit 1.5 Taylor. Again, the orders of strong convergence obtained from our numerical experiments are the ones predicted by the theory. Comparing explicit with implicit schemes, we report that for this choice of parameters the implicit schemes are more accurate. Since the jump impact is simulated without creating extra errors, these differences are due to the approximation of the diffusive part. We remark that implicit schemes, which offer wider regions of stability, are more suitable for problems in which stability constitutes an important issue. This applies in the areas of filtering and finance, where SDEs with multiplicative noise naturally arise.

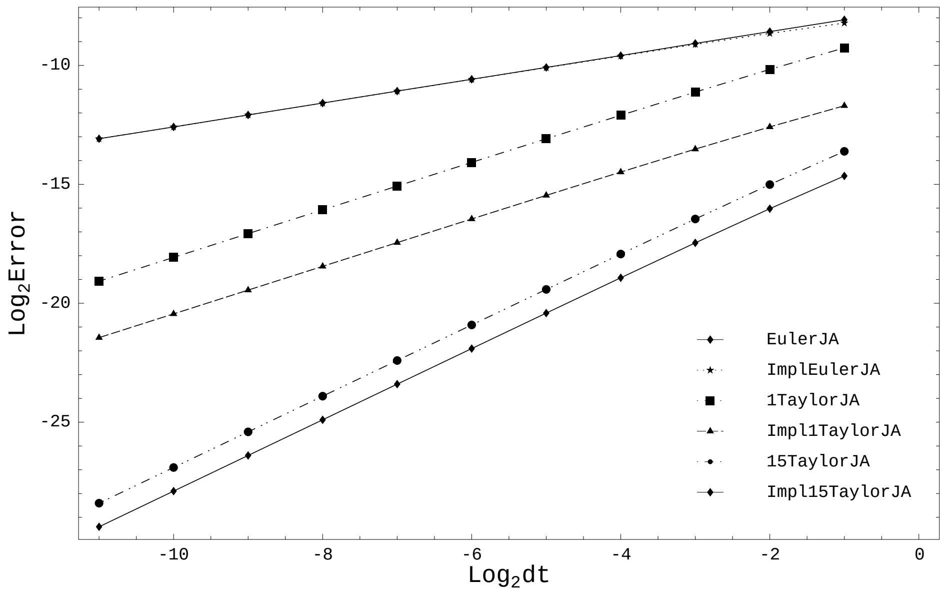

Figure 5.2: Log-log plot of strong error versus time step size.

6 Convergence Theorems

To analyze the order of strong convergence of the proposed numerical schemes we will exploit the Wagner-Platen expansion of the solution of the SDE (2.2), see Platen (1982a, 1982b, 1999). We rewrite the SDE (2.2) in a way such that the jump part will be expressed as a stochastic integral with respect to the *compensated Poisson measure*

$$\widetilde{p}_\phi(dv \times dt) := p_\phi(dv \times dt) - \phi(dv)dt. \tag{6.1}$$

By compensating the Poisson measure in the SDE (2.2) we obtain

$$dX_t = \widetilde{a}(t, X_t)dt + b(t, X_t)dW_t + \int_{\mathcal{E}} c(t-, X_{t-}, v)\widetilde{p}_\phi(dv \times dt), \tag{6.2}$$

where

$$\widetilde{a}(t, x) := a(t, x) + \int_{\mathcal{E}} c(t, x, v)\phi(dv), \tag{6.3}$$

for $t \in [0, T]$ and $x \in \mathbb{R}^d$.

Theorems 6.1 and 7.1, to be presented, analyze the order of convergence of strong approximations constructed with jump integrals with respect to the compensated Poisson measure $\widetilde{p}_\phi$. For computational convenience, we presented in Section 3 strong schemes with jump integrals with respect to the Poisson measure p_ϕ. When directly using the compensated Poisson measure $\widetilde{p}_\phi$, as employed in the strong Taylor schemes in Theorem 6.1 and in the strong Itô schemes in Theorem 7.1, some differences between the resulting schemes may arise. These can be shown to relate to terms of higher order, which do not affect the prescribed strong order of convergence. Therefore the strong order of convergence of the schemes presented in Section 3 can be derived from Theorems 6.1 and 7.1.

We now introduce a compact notation to express multiple stochastic integrals and the corresponding stochastic expansions. We call a row vector $\alpha = (j_1, j_2, \ldots, j_l)$, where $j_i \in \{-1, 0, 1, \ldots, m\}$ for $i \in \{1, 2, \ldots, l\}$, a *multi-index* of length $l := l(\alpha) \in \{1, 2, \ldots\}$. Here m represents the number of Wiener processes considered in the SDE (2.2). Then for $m \in \mathcal{N}$ the set of all multi-indices α is denoted by

$$\mathcal{M}_m = \{(j_1, \ldots, j_l) : j_i \in \{-1, 0, 1, 2, \ldots, l\}, i \in \{1, 2, \ldots, l\} \text{ for } l \in \mathcal{N}\} \cup \{v\}, \tag{6.4}$$

where v is the multi-index of length zero.

We write $n(\alpha)$ for the number of components of a multi-index α that are equal to 0 and $s(\alpha)$ for the number of components of a multi-index α that equal -1. Moreover, we write $\alpha-$ for the multi-index obtained by deleting the last component of α and $-\alpha$ for the multi-index obtained by deleting the first component of α. For instance, assuming $m = 2$,

$$l((0,-1,1)) = 3 \qquad l((0,1,-1,0,2)) = 5$$

$$n((0,-1,1)) = 1 \qquad n((0,1,-1,0,2)) = 2$$

$$s((0,-1,1)) = 1 \qquad s((0,1,-1,0,2)) = 1$$

$$(0,-1,1)- = (0,-1) \qquad (0,1,-1,0,2)- = (0,1,-1,0)$$

$$-(0,-1,1) = (-1,1) \qquad -(0,1,-1,0,2) = (1,-1,0,2).$$

We shall define some functional spaces of predictable stochastic processes $g = \{g(t), t \in [0,T]\}$ that are allowed to appear as integrands of the multiple stochastic integrals in the stochastic expansions to be presented. We define

$$\begin{aligned}
\mathcal{H}_v = & \quad \{g : \sup_{t\in[0,T]} E\left(|g(t,\omega)|\right) < \infty\} \\
\mathcal{H}_{(0)} = & \quad \{g : E\left(\int_0^T |g(s,\omega)|ds\right) < \infty\} \\
\mathcal{H}_{(-1)} = & \quad \{g : E\left(\int_0^T \int_{\mathcal{E}} |g(s,v,\omega)|^2 \phi(dv)ds\right) < \infty\} \\
\mathcal{H}_{(j)} = & \quad \{g : E\left(\int_0^T |g(s,\omega)|^2 ds\right) < \infty\},
\end{aligned} \tag{6.5}$$

for $j \in \{1, 2, \ldots, m\}$. The set $\mathcal{H}_\alpha$ for a multi-index $\alpha \in \mathcal{M}_m$ with $l(\alpha) > 1$ will be defined below.

Let ρ and τ be two stopping times with $0 \le \rho \le \tau \le T$ a.s. For a multi-index $\alpha \in \mathcal{M}_m$ and a predictable process $g(\cdot) \in \mathcal{H}_\alpha$ we define the multiple stochastic integral $I_\alpha[g(\cdot)]_{\rho,\tau}$ recursively by

$$I_\alpha[g(\cdot)]_{\rho,\tau} := \begin{cases} g(\tau) & \text{when}\, l = 0\, \text{and}\, \alpha = v \\ \int_\rho^\tau I_{\alpha-}[g(\cdot)]_{\rho,z}\, dz & \text{when}\, l \ge 1\, \text{and}\, j_l = 0 \\ \int_\rho^\tau I_{\alpha-}[g(\cdot)]_{\rho,z}\, dW_z^{j_l} & \text{when}\, l \ge 1\, \text{and}\, j_l \in \{1, \ldots, m\} \\ \int_\rho^\tau \int_{\mathcal{E}} I_{\alpha-}[g(\cdot)]_{\rho,z-}\, \widetilde{p}_\phi(dv \times dz) & \text{when}\, l \ge 1\, \text{and}\, j_l = -1, \end{cases} \tag{6.6}$$

where $g(\cdot) = g(\cdot, v)$, with $v \in \mathcal{E}^{s(\alpha)}$, and with $z-$ we denote the left hand limit of z. For simplicity, when it is not strictly necessary, here and in the sequel, we will omit the dependence of the integrand process g on one or more of the components $v^1, \ldots, v^{s(\alpha)}$ of the vector v expressing the marks of the Poisson jump measure.

The sets $\mathcal{H}_\alpha$, for every multi-index $\alpha = (j_1, \ldots, j_l) \in \mathcal{M}_m$ with $l(\alpha) > 1$, are defined recursively as the sets of predictable stochastic processes $g = \{g(t), t \ge 0\}$ such that the integral process $\{I_{\alpha-}[g(\cdot)]_{\rho,t}, t \in [0,T]\}$ satisfies

$$I_{\alpha-}[g(\cdot)]_{\rho,\cdot} \in \mathcal{H}_{(j_l)}. \tag{6.7}$$

As defined in (6.6), in a multi-index α the components that equal 0 refer to an integration with respect to time, the components that equal $j \in \{1, \ldots, m\}$ refer to an integration with respect to the j-th component of the Wiener process, while the components that equal -1 refer to an integration with respect to the Poisson martingale measure $\widetilde{p}_\phi(dv \times dt)$. For instance,

$$I_{(0,-1,1)}[g(\cdot)]_{\rho,\tau} = \int_\rho^\tau \int_\rho^{z_3} \int_{\mathcal{E}} \int_\rho^{z_2-} g(z_1, v_1)\, dz_1\, \widetilde{p}_\phi(dv_1 \times dz_2)\, dW_{z_3}^1 \tag{6.8}$$

and

$$I_{(2,0)}[g(\cdot)]_{\rho,\tau} = \int_\rho^\tau \int_\rho^{z_2} g(z_1)\, dW_{z_1}^2\, dz_2. \tag{6.9}$$

We need to define some sets of sufficiently smooth and integrable functions. $\mathcal{L}^0$ is the set of functions $f(t,x,u):[0,T]\times\mathbb{R}^d\times\mathcal{E}^{s(\alpha)}\longrightarrow\mathbb{R}^d$ for which

$$f\big(t,x+c(t,x,v),u\big)-f(t,x,u)-\sum_{i=1}^{d}c^i(t,x,v)\frac{\partial}{\partial x^i}f(t,x,u) \tag{6.10}$$

is $\phi(dv)$-integrable for all $t\in[0,T]$, $x\in\mathbb{R}^d$, $u\in\mathcal{E}^{s(\alpha)}$ and $f(\cdot,\cdot,u)\in\mathcal{C}^{1,2}$. Note that, according to the notation defined in Section 2, c^i denotes the i-th component of the jump coefficient vector c. With $\mathcal{L}^k$, $k\in\{1,\dots,m\}$, we denote the set of functions $f(t,x,u)$ with partial derivatives $\frac{\partial}{\partial x^i}f(t,x,u)$, $i\in\{1,\dots,d\}$. With $\mathcal{L}^{-1}$ we denote the set of functions for which

$$\big\{f\big(t,x+c\,(t,x,v)\,,u\big)-f\big(t,x,u\big)\big\}^2 \tag{6.11}$$

is $\phi(dv)$-integrable for all $t\in[0,T]$, $x\in\mathbb{R}^d$ and $u\in\mathcal{E}^{s(\alpha)}$.

Let us now define the following operators for a function $f(t,x,u)\in\mathcal{L}^k$, with $k\in\{-1,0,1,\dots,m\}$:

$$\begin{aligned}
L^{(0)}f(t,x,u) \;&:=\; \frac{\partial}{\partial t}f(t,x,u)+\sum_{i=1}^{d}a^i(t,x)\frac{\partial}{\partial x^i}f(t,x,u)\\
&\quad+\frac{1}{2}\sum_{i,r=1}^{d}\sum_{j=1}^{m}b^{i,j}(t,x)b^{r,j}(t,x)\frac{\partial^2}{\partial x^i\partial x^j}f(t,x,u)\\
&\quad+\int_{\mathcal{E}}\Big\{f\Big(t,x+c\big(t,x,v\big),u\Big)-f\Big(t,x,u\Big)\Big\}\;\phi(dv)\\
&=\; \frac{\partial}{\partial t}f(t,x,u)+\sum_{i=1}^{d}\widetilde{a}^i(t,x)\frac{\partial}{\partial x^i}f(t,x,u)\\
&\quad+\frac{1}{2}\sum_{i,r=1}^{d}\sum_{j=1}^{m}b^{i,j}(t,x)b^{r,j}(t,x)\frac{\partial^2}{\partial x^i\partial x^j}f(t,x,u)\\
&\quad+\int_{\mathcal{E}}\Big\{f\Big(t,x+c\big(t,x,v\big),u\Big)-f\Big(t,x,u\Big)\\
&\quad-\sum_{i=1}^{d}c^i(t,x,v)\frac{\partial}{\partial x^i}f(t,x,u)\Big\}\;\phi(dv),
\end{aligned} \tag{6.12}$$

$$L^{(k)}f(t,x,u):=\sum_{i=1}^{d}b^{i,k}(t,x)\frac{\partial}{\partial x^i}f(t,x,u),\quad \text{for}\, k\in\{1,\dots,m\} \tag{6.13}$$

and

$$L^{(-1)}f(t,x,u):=f\big(t,x+c\,(t,x,v)\,,u\big)-f(t,x,u), \tag{6.14}$$

for all $t \in [0,T]$, $x \in \mathbb{R}^d$ and $u \in \mathcal{E}^{s(\alpha)}$. Here the operator in (6.14) adds a new dependence on the component $v \in \mathcal{E}$, which we do not explicitly express in our notation to simplify the presentation.

For all $\alpha = (j_1, \ldots, j_{l(\alpha)}) \in \mathcal{M}_m$ and a function $f : [0,T] \times \mathbb{R}^d \longrightarrow \mathbb{R}^d$, we define recursively the *Itô coefficient functions*

$$f_\alpha(t,x,u) := \begin{cases} f\big(t,x\big) & \text{for} \quad l(\alpha) = 0, \\ f\big(t,\widetilde{a}(t,x)\big) & \text{for} \quad l(\alpha) = 1,\ j_1 = 0, \\ f\big(t,b^{j_1}(t,x)\big) & \text{for} \quad l(\alpha) = 1,\ j_1 \in \{1,\ldots,m\}, \\ f\big(t,c(t,x,u)\big) & \text{for} \quad l(\alpha) = 1,\ j_1 = -1, \\ L^{(j_1)} f_{-\alpha}(t,x,u^1,\ldots,u^{s(-\alpha)}) & \text{for} \quad l(\alpha) \geq 2,\ j_1 \in \{-1,0,\ldots,m\}. \end{cases} \tag{6.15}$$

Here by $b^{j_1}(t,x)$ we denote the d-dimensional vector of real valued functions on $[0,T] \times \mathbb{R}^d$ obtained by extracting the j_1-th column from the matrix $b(t,x)$ of coefficient functions. With $u^1, \ldots, u^{s(-\alpha)}$ we denote the components of the vector $u \in \mathcal{E}^{s(-\alpha)}$. We assume that the coefficients of the SDE (2.2) and the function f satisfy the smoothness and integrability conditions needed for the operators in (6.15) to be well defined. For instance, with $d = m = 1$, if we choose the identity function $f(t,x) = x$ we get

$$f_{(-1,0)}(t,x,u) = L^{(-1)}\widetilde{a}(t,x) = \widetilde{a}(t,x+c(t,x,u)) - \widetilde{a}(t,x), \tag{6.16}$$

$$\begin{aligned} f_{(0,1)}(t,x) &= L^{(0)}b(t,x) \\ &= \frac{\partial}{\partial t}b(t,x) + a(t,x)\frac{\partial}{\partial x}b(t,x) + \frac{1}{2}\big(b(t,x)\big)^2\frac{\partial^2}{\partial x^2}b(t,x) \\ &\quad + \int_{\mathcal{E}} \Big\{ b\big(t,x+c(t,x,v)\big) - b\big(t,x\big) \Big\} \phi(dv) \\ &= \frac{\partial}{\partial t}b(t,x) + \widetilde{a}(t,x)\frac{\partial}{\partial x}b(t,x) + \frac{1}{2}\big(b(t,x)\big)^2\frac{\partial^2}{\partial x^2}b(t,x) \\ &\quad + \int_{\mathcal{E}} \Big\{ b\big(t,x+c(t,x,v)\big) - b\big(t,x\big) - c(t,x,v)\frac{\partial}{\partial x}b(t,x) \Big\} \phi(dv) \end{aligned} \tag{6.17}$$

and

$$\begin{aligned} f_{(-1,-1)}(t,x,u) &= L^{(-1)}c(t,x,u^1) \\ &= c(t,x+c(t,x,u^2),u^1) - c(t,x,u^1). \end{aligned} \tag{6.18}$$

To define a stochastic Taylor expansion we finally need to specify some particular sets of multi-indices. A subset $\mathcal{A} \in \mathcal{M}_m$ is a *hierarchical* set if $\mathcal{A}$ is non-empty, the multi-indices in $\mathcal{A}$ are uniformly bounded in length, that means $\sup_{\alpha\in\mathcal{A}} l(\alpha) < \infty$, and if $-\alpha \in \mathcal{A}$ for each $\alpha \in \mathcal{A}\backslash\{v\}$. We also define the *remainder* set $\mathcal{B}(\mathcal{A})$ of $\mathcal{A}$ by

$$\mathcal{B}(\mathcal{A}) = \{\alpha \in \mathcal{M}_m\backslash\mathcal{A} : -\alpha \in \mathcal{A}\}. \tag{6.19}$$

Then the remainder set consists of all the next following multi-indices with respect to the given hierarchical set.

Given two stopping times ρ and τ with $0 \leq \rho \leq \tau \leq T$ a. s., a hierarchical set $\mathcal{A} \in \mathcal{M}_m$, and a function $f : [0,T] \times \mathbb{R}^d \longrightarrow \mathbb{R}^d$, we obtain the Wagner-Platen expansion

$$f(\tau, X_\tau) = \sum_{\alpha\in\mathcal{A}} I_\alpha[f_\alpha(\rho, X_\rho)]_{\rho,\tau} + \sum_{\alpha\in\mathcal{B}(\mathcal{A})} I_\alpha[f_\alpha(\cdot, X_\cdot)]_{\rho,\tau}, \tag{6.20}$$

where we have assumed that the function f and the coefficients of the SDE (2.2) are sufficiently smooth and integrable such that the coefficient functions f_α are well defined and all the multiple stochastic integrals exist.

By choosing as function f the identity functions $f(t,x) = x$ we can represent the process $X = \{X_t, t \in [0,T]\}$ as solution of the SDE (2.2) by the Wagner-Platen expansion

$$X_\tau = \sum_{\alpha\in\mathcal{A}} I_\alpha[f_\alpha(\rho, X_\rho)]_{\rho,\tau} + \sum_{\alpha\in\mathcal{B}(\mathcal{A})} I_\alpha[f_\alpha(\cdot, X_\cdot)]_{\rho,\tau}. \tag{6.21}$$

Note that in (6.21) we have suppressed in the notation the dependence of f_α on $u \in \mathcal{E}^{s(\alpha)}$ and we will do so also in the following where no misunderstanding is possible.

The proof of the Wagner-Platen expansion for jump-diffusion processes, which is based on an iterative application of the Itô formula, can be found in Platen (1982a, 1982b).

6.1 Strong Taylor Schemes

Let us consider a time discretization $0 \leq t_0 < t_1 < \ldots < t_{n_T} \leq T$ on which we will construct a discrete time approximation of the solution X of (2.2). We also introduce for all $t \in [0,T]$ the index

$$n_t = \max\{n \in \{0,1,\ldots\} : t_n \leq t\} \tag{6.22}$$

of the last discretization point before t. In the following we will assume a maximum step size $\Delta \in (0,1)$, that means for every $n \in \{0,1,2,\ldots,n_T - 1\}$ the discretisation time t_{n+1} is $\mathcal{A}_{t_n}$-measurable and $P(t_{n+1} - t_n \leq \Delta) = 1$. We also require to have a finite number of time discretisation points, that means $n_t < \infty$ almost surely for $t \in [0,T]$. We abbreviate a time discretisation of the above type by $(t)_\Delta$.

Moreover, for every $\gamma \in \{0.5, 1, 1.5, 2, \ldots\}$ we define the hierarchical set

$$\mathcal{A}_\gamma = \{\alpha \in \mathcal{M} : l(\alpha) + n(\alpha) \leq 2\gamma \quad \text{or} \quad l(\alpha) = n(\alpha) = \gamma + \frac{1}{2}\}. \tag{6.23}$$

For a time discretization with maximum step size $\Delta \in (0, 1)$, we define the *order γ strong Taylor scheme* by the vector equation

$$Y_{n+1}^\Delta = Y_n^\Delta + \sum_{\alpha \in \mathcal{A}_\gamma \setminus \{v\}} I_\alpha \left[f_\alpha(t_n, Y_n^\Delta) \right]_{t_n, t_{n+1}} = \sum_{\alpha \in \mathcal{A}_\gamma} I_\alpha \left[f_\alpha(t_n, Y_n^\Delta) \right]_{t_n, t_{n+1}}, \tag{6.24}$$

for $n \in \{0, 1, \ldots, n_T - 1\}$. Equation (6.24) gives us a numerical routine to generate approximate values of the solution of the SDE (2.2) at the discretization points.

In order to asses the strong order of convergence of these schemes we define, through a specific interpolation, the *order γ strong Taylor approximation $Y^\Delta = \{Y_t^\Delta,\ t \in [0, T]\}$*, by

$$Y_t^\Delta = \sum_{\alpha \in \mathcal{A}_\gamma} I_\alpha [f_\alpha(t_{n_t}, Y_{t_{n_t}}^\Delta)]_{t_{n_t}, t} \tag{6.25}$$

for $t \in [0, T]$, starting from a given $\mathcal{A}_0$-measurable random variable Y_0. This approximation defines a stochastic process $Y^\Delta = \{Y_t^\Delta,\ t \in [0, T]\}$, whose values coincide with the ones of the *order γ strong Taylor scheme* (6.24) on the discretisation points. Between the discretisation points the multiple stochastic integrals have constant coefficient functions but evolve randomly as a function of time, see (6.25).

We can now formulate a convergence theorem that will enable us to construct a strong Taylor approximation $Y^\Delta = \{Y_t^\Delta,\ t \in [0, T]\}$ of any given strong order $\gamma = \{0.5, 1, 1.5, 2, \ldots\}$.

Theorem 6.1 *For a given $\gamma \in \{0.5, 1, 1.5, 2, \ldots\}$, let $Y^\Delta = \{Y_t^\Delta,\ t \in [0, T]\}$ be the order γ strong Taylor approximation defined in* (6.25) *corresponding to a time discretisation with maximum step size $\Delta \in (0, 1)$.*

We assume that

$$E(|X_0|^2) < \infty \quad \text{and} \quad E(|X_0 - Y_0^\Delta|^2) \leq K_1 \Delta^{2\gamma}. \tag{6.26}$$

Moreover, suppose that the coefficient functions f_α satisfy the following conditions:

For $\alpha \in \mathcal{A}_\gamma$, $t \in [0, T]$, $u \in \mathcal{E}^{s(\alpha)}$ and $x, y \in \mathbb{R}^d$ the coefficient function f_α satisfies the Lipschitz type condition

$$|f_\alpha(t, x, u) - f_\alpha(t, y, u)| \leq K_1(u)|x - y|, \tag{6.27}$$

where $K_1(u)^2$ is $\phi(du)$-integrable.

For all $\alpha \in \mathcal{A}_\gamma \bigcup B(\mathcal{A}_\gamma)$ we assume

$$f_{-\alpha} \in \mathcal{C}^{1,2} \qquad \text{and} \qquad f_\alpha \in \mathcal{H}_\alpha, \tag{6.28}$$

and for $\alpha \in \mathcal{A}_\gamma \bigcup \mathcal{B}(\mathcal{A}_\gamma)$, $t \in [0,T]$, $u \in \mathcal{E}^{s(\alpha)}$ and $x \in \mathbb{R}^d$, we require

$$|f_\alpha(t,x,u)|^2 \leq K_2(u)(1+|x|^2), \tag{6.29}$$

where $K_2(u)$ is $\phi(du)$-integrable.

Then the estimate

$$\sqrt{E(\sup_{0\leq s\leq T} |X_s - Y_s^\Delta|^2 \, |\mathcal{A}_0)} \leq K_3 \Delta^\gamma \tag{6.30}$$

holds, where the constant K_3 does not depend on Δ.

The proof of the theorem will be given in Section 6.4.

We now present several results that are needed for the proof of the convergence Theorem 6.1.

6.2 Moments of Multiple Stochastic Integrals

The following two lemmas provide estimates of multiples stochastic integrals that will constitute the core of the proof of Theorem 6.1.

Lemma 6.2 *Let $\alpha \in \mathcal{M}_m \backslash \{v\}$, $g \in \mathcal{H}_\alpha$, $\Delta > 0$ and ρ and τ denote two stopping times with τ $\mathcal{A}_\rho$-measurable and $t_0 \leq \rho \leq \tau \leq \rho + \Delta \leq T$ almost surely. Then*

$$F_\tau^\alpha := E\left(\sup_{\rho\leq s\leq\tau} |\, I_\alpha[g(\cdot)]_{\rho,s}|^2 \middle| \mathcal{A}_\rho\right) \leq 4^{l(\alpha)-n(\alpha)} \Delta^{l(\alpha)+n(\alpha)-1} \int_\rho^\tau V_{\rho,z,s(\alpha)}\, dz, \tag{6.31}$$

where

$$V_{\rho,z,s(\alpha)} := \int_{\mathcal{E}} \ldots \int_{\mathcal{E}} E\left(\sup_{\rho\leq t\leq z} |g(t,v^1,\ldots,v^{s(\alpha)})|^2 \middle| \mathcal{A}_\rho\right) \phi(dv^1)\ldots\phi(dv^{s(\alpha)}) < \infty \tag{6.32}$$

for $z \in [\rho, \tau]$.

Proof: We will prove the assertion (6.31) by induction on $l(\alpha)$.

1. Let us assume that $l(\alpha) = 1$ and $\alpha = (0)$. By the Cauchy-Schwarz inequality we have the estimate

$$\left|\int_\rho^s g(z)\ dz\right|^2 \leq (s-\rho) \int_\rho^s |g(z)|^2\ dz. \tag{6.33}$$

Therefore, we obtain

$$
\begin{aligned}
F_{\tau}^{(0)} &= E\left(\sup_{\rho\le s\le\tau}\left|\int_{\rho}^{s} g(z)\ dz\right|^2 \middle| \mathcal{A}_{\rho}\right) \\
&\le E\left(\sup_{\rho\le s\le\tau}(s-\rho)\int_{\rho}^{s}|g(z)|^2\ dz \middle| \mathcal{A}_{\rho}\right) \\
&= E\left((\tau-\rho)\int_{\rho}^{\tau}|g(z)|^2\ dz \middle| \mathcal{A}_{\rho}\right) \\
&\le \Delta\ E\left(\int_{\rho}^{\tau}|g(z)|^2\ dz \middle| \mathcal{A}_{\rho}\right) \\
&= \Delta\int_{\rho}^{\tau} E\left(|g(z)|^2|\mathcal{A}_{\rho}\right)\ dz \\
&\le \Delta\int_{\rho}^{\tau} E\left(\sup_{\rho\le t\le z}|g(t)|^2 \middle| \mathcal{A}_{\rho}\right)\ dz \\
&= 4^{l(\alpha)-n(\alpha)}\,\Delta^{l(\alpha)+n(\alpha)-1}\int_{\rho}^{\tau} V_{\rho,z,s(\alpha)}\,dz,
\end{aligned}
\tag{6.34}
$$

where the interchange between expectation and integral holds by $\mathcal{A}_{\rho}$-measurability of τ and Fubini's theorem.

2. When $l(\alpha)=1$ and $\alpha=(j)$ with $j\in\{1,2,\dots,m\}$, we first observe that the process

$$
\{I_{\alpha}[g(\cdot)]_{\rho,t}, t\in[\rho,T]\} = \{\int_{\rho}^{t} g(s)\ dW_s^j, t\in[\rho,T]\} \tag{6.35}
$$

is a martingale. Therefore, applying Doob's inequality and Itô's isometry we have

$$
\begin{aligned}
F_\tau^{(j)} &= E\left(\sup_{\rho\le s\le\tau}\left|\int_\rho^s g(z)\ dW_z^j\right|^2\bigg|\mathcal{A}_\rho\right)\\
&\le 4\ E\left(\left|\int_\rho^\tau g(z)\ dW_z^j\right|^2\bigg|\mathcal{A}_\rho\right)\\
&= 4\ E\left(\int_\rho^\tau |g(z)|^2 dz\bigg|\mathcal{A}_\rho\right)\\
&= 4\ \int_\rho^\tau E\left(|g(z)|^2|\mathcal{A}_\rho\right)\, dz\\
&\le 4\ \int_\rho^\tau E\left(\sup_{\rho\le t\le z}|g(t)|^2\bigg|\mathcal{A}_\rho\right)\, dz\\
&= 4^{l(\alpha)-n(\alpha)}\,\Delta^{l(\alpha)+n(\alpha)-1}\int_\rho^\tau V_{\rho,z,s(\alpha)}\, dz,
\end{aligned}
\tag{6.36}
$$

where again the interchange between expectation and integral holds by $\mathcal{A}_\rho$ measurability of τ and Fubini's theorem.

3. Let us now consider the case with $l(\alpha)=1$ and $\alpha=(-1)$. The process

$$
\{I_\alpha[g(\cdot)]_{\rho,t}, t\in[\rho,T]\} = \{\int_\rho^t\int_{\mathcal{E}} g(s,v)\ \widetilde{p}_\phi(dv\times ds), t\in[\rho,T]\}
\tag{6.37}
$$

is a martingale. Then, by Doob's inequality and the isometry for jump processes, we obtain

$$
\begin{aligned}
F_\tau^{(-1)} &= E\left(\sup_{\rho\le s\le\tau}\left|\int_\rho^s\int_{\mathcal{E}} g(z,v)\ \widetilde{p}_\phi(dv\times dz)\right|^2\bigg|\mathcal{A}_\rho\right)\\
&\le 4\ E\left(\left|\int_\rho^\tau\int_{\mathcal{E}} g(z,v)\ \widetilde{p}_\phi(dv\times dz)\right|^2\bigg|\mathcal{A}_\rho\right)\\
&= 4\ E\left(\int_\rho^\tau\int_{\mathcal{E}} |g(z,v)|^2\,\phi(dv)\, dz\bigg|\mathcal{A}_\rho\right)\\
&= 4\ \int_\rho^\tau\int_{\mathcal{E}} E\left(|g(z,v)|^2|\mathcal{A}_\rho\right)\,\phi(dv)\, dz\\
&\le 4\ \int_\rho^\tau\int_{\mathcal{E}} E\left(\sup_{\rho\le t\le z}|g(t,v)|^2\bigg|\mathcal{A}_\rho\right)\phi(dv)\, dz\\
&= 4^{l(\alpha)-n(\alpha)}\,\Delta^{l(\alpha)+n(\alpha)-1}\int_\rho^\tau V_{\rho,z,s(\alpha)}\, dz,
\end{aligned}
\tag{6.38}
$$

since $s(\alpha) = 1$. This shows that the result of Lemma 6.2 holds for $l(\alpha) = 1$.

4. Now, let $l(\alpha) = n + 1$, where $\alpha = (j_1, \ldots, j_{n+1})$ and $j_{n+1} = 0$. Then, by applying the Cauchy-Schwarz inequality we obtain

$$\begin{aligned}
F_\tau^\alpha &= E\left(\sup_{\rho \le s \le \tau} \left| \int_\rho^s I_{\alpha-}[g(\cdot)]_{\rho,z}\; dz \right|^2 \middle| \mathcal{A}_\rho \right) \\
&\le E\left(\sup_{\rho \le s \le \tau} (s-\rho) \int_\rho^s |I_{\alpha-}[g(\cdot)]_{\rho,z}|^2 \; dz \middle| \mathcal{A}_\rho \right) \\
&= E\left((\tau-\rho) \int_\rho^\tau |I_{\alpha-}[g(\cdot)]_{\rho,z}|^2 \; dz \middle| \mathcal{A}_\rho \right) \\
&\le \Delta\, E\left(\int_\rho^\tau |I_{\alpha-}[g(\cdot)]_{\rho,z}|^2 \; dz \middle| \mathcal{A}_\rho \right) \\
&\le \Delta\, E\left(\int_\rho^\tau \sup_{\rho \le s \le \tau} |I_{\alpha-}[g(\cdot)]_{\rho,s}|^2 \; dz \middle| \mathcal{A}_\rho \right) \\
&= \Delta\, E\left(\int_\rho^\tau dz \times \sup_{\rho \le s \le \tau} |I_{\alpha-}[g(\cdot)]_{\rho,s}|^2 \middle| \mathcal{A}_\rho \right) \\
&\le \Delta^2\, E\left(\sup_{\rho \le s \le \tau} |I_{\alpha-}[g(\cdot)]_{\rho,s}|^2 \middle| \mathcal{A}_\rho \right). \qquad (6.39)
\end{aligned}$$

Then, by the inductive hypothesis it follows that

$$\begin{aligned}
F_\tau^\alpha &\le \Delta^2\, 4^{l(\alpha-)-n(\alpha-)}\, \Delta^{l(\alpha-)+n(\alpha-)-1} \int_\rho^\tau V_{\rho,z,s(\alpha-)}\, dz \\
&= 4^{l(\alpha)-n(\alpha)}\, \Delta^{l(\alpha)+n(\alpha)-1} \int_\rho^\tau V_{\rho,z,s(\alpha)}\, dz, \qquad (6.40)
\end{aligned}$$

where the last line holds considering that $l(\alpha) = l(\alpha-)+1$, $n(\alpha) = n(\alpha-)+1$ and $s(\alpha) = s(\alpha-)$.

5. Let us now consider the case when $l(\alpha) = n + 1$, where $\alpha = (j_1, \ldots, j_{n+1})$ and $j_{n+1} \in \{1, 2, \ldots, m\}$. The process

$$\{I_\alpha[g(\cdot)]_{\rho,t}, t \in [\rho, T]\} \qquad (6.41)$$

is a martingale. Therefore, by Doob's inequality and Itô's isometry we obtain

$$
\begin{aligned}
F_\tau^\alpha &= E\left(\sup_{\rho\le s\le\tau} \left| \int_\rho^s I_{\alpha-}[g(\cdot)]_{\rho,z} \; dW_z^{j_{n+1}} \right|^2 \middle| \mathcal{A}_\rho \right) \\
&\le 4 \; E\left(\left| \int_\rho^\tau I_{\alpha-}[g(\cdot)]_{\rho,z} \; dW_z^{j_{n+1}} \right|^2 \middle| \mathcal{A}_\rho \right) \\
&= 4 \; E\left(\int_\rho^\tau |I_{\alpha-}[g(\cdot)]_{\rho,z}|^2 \; dz \middle| \mathcal{A}_\rho \right) \\
&\le 4 \; E\left(\int_\rho^\tau \sup_{\rho\le s\le\tau} |I_{\alpha-}[g(\cdot)]_{\rho,s}|^2 \; dz \middle| \mathcal{A}_\rho \right) \\
&= 4 \; E\left((\tau-\rho) \sup_{\rho\le s\le\tau} |I_{\alpha-}[g(\cdot)]_{\rho,s}|^2 \middle| \mathcal{A}_\rho \right) \\
&\le 4\,\Delta\, E\left(\sup_{\rho\le s\le\tau} |I_{\alpha-}[g(\cdot)]_{\rho,s}|^2 \middle| \mathcal{A}_\rho \right).
\end{aligned}
\tag{6.42}
$$

By the inductive hypothesis we have

$$
\begin{aligned}
F_\tau^\alpha &\le 4\,\Delta\, 4^{l(\alpha-)-n(\alpha-)}\, \Delta^{l(\alpha-)+n(\alpha-)-1} \int_\rho^\tau V_{\rho,z,s(\alpha-)}\, dz \\
&= \Delta^{l(\alpha)+n(\alpha)-1}\, 4^{l(\alpha)-n(\alpha)} \int_\rho^\tau V_{\rho,z,s(\alpha)}\, dz,
\end{aligned}
\tag{6.43}
$$

since $l(\alpha) = l(\alpha-) + 1$, $n(\alpha) = n(\alpha-)$ and $s(\alpha) = s(\alpha-)$.

6. Finally, let us suppose that $l(\alpha) = n + 1$, where $\alpha = (j_1, \ldots, j_{n+1})$ and $j_{n+1} = -1$. The process

$$\{I_\alpha[g(\cdot)]_{\rho,t}, t \in [\rho, T]\} \tag{6.44}$$

is again a martingale. Therefore, by applying Doob's inequality and the isometry for jump processes, we obtain

$$\begin{aligned}
F_\tau^\alpha &= E\left(\sup_{\rho\le s\le\tau}\left|\int_\rho^s\int_{\mathcal{E}} I_{\alpha-}[g(\cdot,v^{s(\alpha)})]_{\rho,z}\ \widetilde{p}_\phi(dv^{s(\alpha)}\times dz)\right|^2\Bigg|\mathcal{A}_\rho\right)\\
&\le 4\ E\left(\left|\int_\rho^\tau\int_{\mathcal{E}} I_{\alpha-}[g(\cdot,v^{s(\alpha)})]_{\rho,z}\ \widetilde{p}_\phi(dv^{s(\alpha)}\times dz)\right|^2\Bigg|\mathcal{A}_\rho\right)\\
&= 4\ E\left(\int_\rho^\tau\int_{\mathcal{E}} |I_{\alpha-}[g(\cdot,v^{s(\alpha)})]_{\rho,z}|^2\,\phi(dv^{s(\alpha)})\,dz\Bigg|\mathcal{A}_\rho\right)\\
&\le 4\,E\left(\int_\rho^\tau\int_{\mathcal{E}}\sup_{\rho\le s\le\tau} |I_{\alpha-}[g(\cdot,v^{s(\alpha)})]_{\rho,s}|^2\,\phi(dv^{s(\alpha)})\,dz\Bigg|\mathcal{A}_\rho\right)\\
&= 4\,E\left((\tau-\rho)\int_{\mathcal{E}}\sup_{\rho\le s\le\tau} |I_{\alpha-}[g(\cdot,v^{s(\alpha)})]_{\rho,s}|^2\,\phi(dv^{s(\alpha)})\Bigg|\mathcal{A}_\rho\right)\\
&\le 4\,\Delta\int_{\mathcal{E}} E\left(\sup_{\rho\le s\le\tau} |I_{\alpha-}[g(\cdot,v^{s(\alpha)})]_{\rho,s}|^2\ \Bigg|\mathcal{A}_\rho\right)\phi(dv^{s(\alpha)}). \qquad (6.45)
\end{aligned}$$

By the inductive hypothesis we have

$$\begin{aligned}
F_\tau^\alpha &\le 4\,\Delta\,4^{l(\alpha-)-n(\alpha-)}\,\Delta^{l(\alpha-)+n(\alpha-)-1}\int_{\mathcal{E}}\int_\rho^\tau V_{\rho,z,s(\alpha-)}\,dz\,\phi(dv^{s(\alpha)})\\
&= 4^{l(\alpha)-n(\alpha)}\,\Delta^{l(\alpha)+n(\alpha)-1}\int_\rho^\tau V_{\rho,z,s(\alpha)}\,dz, \qquad (6.46)
\end{aligned}$$

since $l(\alpha) = l(\alpha-)+1$, $n(\alpha) = n(\alpha-)$ and $s(\alpha) = s(\alpha-)+1$, which completes the proof of Lemma 6.2. $\square$

Lemma 6.3 *For a given multi-index $\alpha\in\mathcal{M}_m\backslash\{v\}$, a time discretisation $(t)_\Delta$ with $\Delta\in(0,1)$ and $g\in\mathcal{H}_\alpha$ let*

$$V_{t_0,u,s(\alpha)} := \int_{\mathcal{E}}\dots\int_{\mathcal{E}} E\left(\sup_{t_0\le z\le u}|g(z,v^1,\dots,v^{s(\alpha)})|^2\Bigg|\mathcal{A}_{t_0}\right)\phi(dv^1)\dots\phi(dv^{s(\alpha)})<\infty \qquad (6.47)$$

and

$$F_t^\alpha := E\left(\sup_{t_0\le z\le t}\left|\sum_{n=0}^{n_z-1} I_\alpha[g(\cdot)]_{t_n,t_{n+1}} + I_\alpha[g(\cdot)]_{t_{n_z},z}\right|^2\Bigg|\mathcal{A}_{t_0}\right). \qquad (6.48)$$

Then

$$F_t^\alpha \le \begin{cases} (t-t_0)\ \Delta^{2(l(\alpha)-1)}\int_{t_0}^t V_{t_0,u,s(\alpha)}\,du & \text{when}: l(\alpha)=n(\alpha)\\ 4^{l(\alpha)-n(\alpha)+2}\ \Delta^{l(\alpha)+n(\alpha)-1}\int_{t_0}^t V_{t_0,u,s(\alpha)}\,du & \text{when}: l(\alpha)\ne n(\alpha)\end{cases}$$

almost surely, for every $t\in[t_0,T]$.

Proof:

1. By definition (6.22) of n_z we get, for $z \in [t_n, t_{n+1})$, the relation $t_{n_z} = t_n$. Then, for a multi-index $\alpha = (j_1, \ldots, j_n)$ with $j_n = 0$, we have

$$\begin{aligned}
&\sum_{n=0}^{n_z-1} I_\alpha[g(\cdot)]_{t_n,t_{n+1}} + I_\alpha[g(\cdot)]_{t_{n_z},z} \\
&= \sum_{n=0}^{n_z-1} \int_{t_n}^{t_{n+1}} I_{\alpha-}[g(\cdot)]_{t_n,s} ds + \int_{t_{n_z}}^{z} I_{\alpha-}[g(\cdot)]_{t_{n_z},s} ds \\
&= \sum_{n=0}^{n_z-1} \int_{t_n}^{t_{n+1}} I_{\alpha-}[g(\cdot)]_{t_{n_s},s} ds + \int_{t_{n_z}}^{z} I_{\alpha-}[g(\cdot)]_{t_{n_s},s} ds \\
&= \int_{t_0}^{z} I_{\alpha-}[g(\cdot)]_{t_{n_s},s} ds. \qquad (6.49)
\end{aligned}$$

The same type of equality holds analogously for every $j_n \in \{-1, 0, 1, \ldots, m\}$.

2. Let us first consider the case with $l(\alpha) = n(\alpha)$. By the Cauchy-Schwarz inequality we have

$$\begin{aligned}
F_t^\alpha &= E\left(\sup_{t_0 \le z \le t} \left| \int_{t_0}^{z} I_{\alpha-}[g(\cdot)]_{t_{n_u},u} \, du \right|^2 \middle| \mathcal{A}_{t_0} \right) \\
&\le E\left(\sup_{t_0 \le z \le t} (z - t_0) \int_{t_0}^{z} |I_{\alpha-}[g(\cdot)]_{t_{n_u},u}|^2 \, du \middle| \mathcal{A}_{t_0} \right) \\
&\le (t - t_0) \, E\left(\int_{t_0}^{t} |I_{\alpha-}[g(\cdot)]_{t_{n_u},u}|^2 \, du \middle| \mathcal{A}_{t_0} \right) \\
&= (t - t_0) \int_{t_0}^{t} E\left(|I_{\alpha-}[g(\cdot)]_{t_{n_u},u}|^2 \, |\mathcal{A}_{t_0} \right) du \\
&\le (t - t_0) \int_{t_0}^{t} E\left(\sup_{t_{n_u} \le z \le u} |I_{\alpha-}[g(\cdot)]_{t_{n_u},z}|^2 \middle| \mathcal{A}_{t_0} \right) du \\
&= (t - t_0) \int_{t_0}^{t} E\left(E\Big(\sup_{t_{n_u} \le z \le u} |I_{\alpha-}[g(\cdot)]_{t_{n_u},z}|^2 \Big| \mathcal{A}_{t_{n_u}} \Big) \middle| \mathcal{A}_{t_0} \right) du,
\end{aligned} \qquad (6.50)$$

where the last line holds because $t_0 \le t_{n_u}$ a.s. and then $\mathcal{A}_{t_0} \subseteq \mathcal{A}_{t_{n_u}}$ for $u \in [t_0, t]$. Therefore, applying Lemma 6.2 to

$$E\left(\sup_{t_{n_u} \le z \le u} |I_{\alpha-}[g(\cdot)]_{t_{n_u},z}|^2 \middle| \mathcal{A}_{t_{n_u}} \right) \qquad (6.51)$$

yields

$$\begin{aligned}
F_t^\alpha &\le (t-t_0)\ 4^{l(\alpha-)-n(\alpha-)} \\
&\quad \times \int_{t_0}^{t} E\left((u-t_{n_u})^{l(\alpha-)+n(\alpha-)-1} \int_{t_{n_u}}^{u} V_{t_{n_u},z,s(\alpha-)}\, dz \,\Big|\, \mathcal{A}_{t_0} \right) du \\
&\le (t-t_0)\, 4^{l(\alpha-)-n(\alpha-)} \int_{t_0}^{t} E\left((u-t_{n_u})^{l(\alpha-)+n(\alpha-)}\, V_{t_{n_u},u,s(\alpha-)} \,\Big|\, \mathcal{A}_{t_0} \right) du \\
&\le (t-t_0)\, 4^{l(\alpha-)-n(\alpha-)} \Delta^{l(\alpha-)+n(\alpha-)} \int_{t_0}^{t} E\left(V_{t_{n_u},u,s(\alpha-)} \,\Big|\, \mathcal{A}_{t_0} \right) du, \qquad (6.52)
\end{aligned}$$

where the last line holds as $(u-t_{n_u}) \le \Delta$ for $u \in [t_0,t]$ and $t \in [t_0,T]$. Since $\mathcal{A}_{t_0} \subseteq \mathcal{A}_{t_{n_u}}$, we notice that for $u \in [t_0,t]$

$$\begin{aligned}
& E\left(V_{t_{n_u},u,s(\alpha-)} | \mathcal{A}_{t_0}\right) \\
=\ & E\Bigg(\int_{\mathcal{E}} \cdots \int_{\mathcal{E}} E\Big(\sup_{t_{n_u} \le s \le u} |g(s,v^1,\ldots,v^{s(\alpha-)})|^2 \,\Big|\, \mathcal{A}_{t_{n_u}} \Big) \\
& \times \phi(dv^1)\ldots\phi(dv^{s(\alpha-)}) \,\Bigg|\, \mathcal{A}_{t_0} \Bigg) \\
=\ & \int_{\mathcal{E}} \cdots \int_{\mathcal{E}} E\Bigg(E\Big(\sup_{t_{n_u} \le s \le u} |g(s,v^1,\ldots,v^{s(\alpha-)})|^2 \,\Big|\, \mathcal{A}_{t_{n_u}} \Big) \,\Bigg|\, \mathcal{A}_{t_0} \Bigg) \\
& \times \phi(dv^1)\ldots\phi(dv^{s(\alpha-)}) \\
=\ & \int_{\mathcal{E}} \cdots \int_{\mathcal{E}} E\left(\sup_{t_{n_u} \le s \le u} |g(s,v^1,\ldots,v^{s(\alpha-)})|^2 \,\Big|\, \mathcal{A}_{t_0} \right) \phi(dv^1)\ldots\phi(dv^{s(\alpha-)}) \\
\le\ & \int_{\mathcal{E}} \cdots \int_{\mathcal{E}} E\left(\sup_{t_0 \le s \le u} |g(s,v^1,\ldots,v^{s(\alpha-)})|^2 \,\Big|\, \mathcal{A}_{t_0} \right) \phi(dv^1)\ldots\phi(dv^{s(\alpha-)}) \\
=\ & V_{t_0,u,s(\alpha-)}. \qquad (6.53)
\end{aligned}$$

It then follows

$$\begin{aligned}
F_t^\alpha &\le (t-t_0)\ 4^{l(\alpha-)-n(\alpha-)}\ \Delta^{l(\alpha-)+n(\alpha-)} \int_{t_0}^{t} V_{t_0,u,s(\alpha-)}\, du \\
&= (t-t_0)\ \Delta^{2(l(\alpha)-1)} \int_{t_0}^{t} V_{t_0,u,s(\alpha)}\, du, \qquad (6.54)
\end{aligned}$$

since $l(\alpha-) = n(\alpha-)$, $s(\alpha) = s(\alpha-)$ and this completes the proof for the case $l(\alpha) = n(\alpha)$.

3. Let us now consider the case with a multi-index $\alpha = (j_1,\ldots,j_l)$ with $l(\alpha) \neq n(\alpha)$ and $j_l \in \{1,\ldots,m\}$. In this case the multiple stochastic integral is a

martingale. Hence, by Doob's inequality, Itô's isometry and Lemma 6.2 we obtain

$$
\begin{aligned}
F_t^{\alpha} &= E\left(\sup_{t_0\le z\le t}\left|\int_{t_0}^{z} I_{\alpha-}[g(\cdot)]_{t_{n_u},u}\,dW_u^{j_l}\right|^2\Bigg|\mathcal{A}_{t_0}\right)\\
&\le 4\,E\left(\left|\int_{t_0}^{t} I_{\alpha-}[g(\cdot)]_{t_{n_u},u}\,dW_u^{j_l}\right|^2\Bigg|\mathcal{A}_{t_0}\right)\\
&\le 4\int_{t_0}^{t} E\left(|I_{\alpha-}[g(\cdot)]_{t_{n_u},u}|^2\Big|\mathcal{A}_{t_0}\right)du\\
&= 4\int_{t_0}^{t} E\left(E\left(|I_{\alpha-}[g(\cdot)]_{t_{n_u},u}|^2\Big|\mathcal{A}_{t_{n_u}}\right)\Bigg|\mathcal{A}_{t_0}\right)du\\
&\le 4\int_{t_0}^{t} E\left(E\left(\sup_{t_{n_u}\le z\le u}|I_{\alpha-}[g(\cdot)]_{t_{n_u},z}|^2\Big|\mathcal{A}_{t_{n_u}}\right)\Bigg|\mathcal{A}_{t_0}\right)du\\
&\le 4\;4^{l(\alpha-)-n(\alpha-)}\\
&\quad\times\int_{t_0}^{t} E\left((u-t_{n_u})^{l(\alpha-)+n(\alpha-)-1}\int_{t_{n_u}}^{u} V_{t_{n_u},z,s(\alpha-)}\,dz\Bigg|\mathcal{A}_{t_0}\right)du\\
&\le 4\;4^{l(\alpha-)-n(\alpha-)}\;\Delta^{l(\alpha-)+n(\alpha-)}\int_{t_0}^{t} V_{t_0,u,s(\alpha-)}\,du\\
&= 4^{l(\alpha)-n(\alpha)}\;\Delta^{l(\alpha)+n(\alpha)-1}\int_{t_0}^{t} V_{t_0,u,s(\alpha-)}\,du\\
&\le 4^{l(\alpha)-n(\alpha)+2}\;\Delta^{l(\alpha)+n(\alpha)-1}\int_{t_0}^{t} V_{t_0,u,s(\alpha)}\,du, \qquad (6.55)
\end{aligned}
$$

where the last passage holds since $s(\alpha) = s(\alpha-)$ and this completes the proof in this case.

4. Let us now consider the case with a multi-index $\alpha = (j_1, \ldots, j_l)$ with $l(\alpha) \neq n(\alpha)$ and $j_l = -1$. The multiple stochastic integral is again a martingale. Therefore, by Doob's inequality, Lemma 6.2 and steps similar

to the previous case we obtain

$$\begin{aligned}
F_t^\alpha &= E\left(\sup_{t_0\le z\le t}\left|\int_{t_0}^{z}\int_{\mathcal{E}} I_{\alpha-}[g(\cdot,v^{s(\alpha)})]_{t_{n_u},u}\,\widetilde{p}_\phi(dv^{s(\alpha)}\times du)\right|^2\middle|\mathcal{A}_{t_0}\right)\\
&\le 4E\left(\left|\int_{t_0}^{t}\int_{\mathcal{E}} I_{\alpha-}[g(\cdot,v^{s(\alpha)})]_{t_{n_u},u}\,\widetilde{p}_\phi(dv^{s(\alpha)}\times du)\right|^2\middle|\mathcal{A}_{t_0}\right)\\
&= 4\int_{t_0}^{t}\int_{\mathcal{E}} E\left(|I_{\alpha-}[g(\cdot,v^{s(\alpha)})]_{t_{n_u},u}|^2\middle|\mathcal{A}_{t_0}\right)\phi(dv^{s(\alpha)})\,du\\
&= 4\int_{t_0}^{t}\int_{\mathcal{E}} E\left(E\left(|I_{\alpha-}[g(\cdot,v^{s(\alpha)})]_{t_{n_u},u}|^2\middle|\mathcal{A}_{t_{n_u}}\right)\middle|\mathcal{A}_{t_0}\right)\phi(dv^{s(\alpha)})\,du\\
&\le 4\int_{t_0}^{t}\int_{\mathcal{E}} E\left(E\left(\sup_{t_{n_u}\le z\le u}|I_{\alpha-}[g(\cdot,v^{s(\alpha)})]_{t_{n_u},z}|^2\middle|\mathcal{A}_{t_{n_u}}\right)\middle|\mathcal{A}_{t_0}\right)\phi(dv^{s(\alpha)})\,du\\
&\le 4^{l(\alpha-)-n(\alpha-)+1}\\
&\quad\times\int_{t_0}^{t}\int_{\mathcal{E}} E\left((u-t_{n_u})^{l(\alpha-)+n(\alpha-)-1}\int_{t_{n_u}}^{u} V_{t_{n_u},z,s(\alpha-)}\,dz\middle|\mathcal{A}_{t_0}\right)\phi(dv^{s(\alpha)})\,du\\
&\le 4^{l(\alpha-)-n(\alpha-)+1}\,\Delta^{l(\alpha-)+n(\alpha-)}\int_{t_0}^{t}\int_{\mathcal{E}} E\left(V_{t_{n_u},u,s(\alpha-)}\middle|\mathcal{A}_{t_0}\right)\phi(dv^{s(\alpha)})\,du.
\end{aligned}\tag{6.56}$$

Hence, using (6.53) we have

$$\begin{aligned}
F_t^\alpha &\le 4^{l(\alpha-)-n(\alpha-)+1}\,\Delta^{l(\alpha-)+n(\alpha-)}\int_{t_0}^{t}\int_{\mathcal{E}} V_{t_0,u,s(\alpha-)}\,\phi(dv^{s(\alpha)})\,du\\
&= 4^{l(\alpha)-n(\alpha)}\,\Delta^{l(\alpha)+n(\alpha)-1}\int_{t_0}^{t} V_{t_0,u,s(\alpha)}\,du\\
&\le 4^{l(\alpha)-n(\alpha)+2}\,\Delta^{l(\alpha)+n(\alpha)-1}\int_{t_0}^{t} V_{t_0,u,s(\alpha)}\,du
\end{aligned}\tag{6.57}$$

since $l(\alpha)=l(\alpha-)+1$, $n(\alpha)=n(\alpha-)$, $s(\alpha)=s(\alpha-)+1$ and this completes the proof in this case.

5. Finally, we assume that $\alpha=(j_1,\dots,j_l)$ with $l(\alpha)\ne n(\alpha)$ and $j_l=0$. It can be shown that the discrete time process

$$\{\sum_{n=0}^{k} I_\alpha[g(\cdot)]_{t_n,t_{n+1}}\,,k\in\{0,1\dots,n_T-1\}\}\tag{6.58}$$

is a discrete time martingale. See Lemma 5.7.1 in (Kloeden & Platen 1999) for the diffusion case.

Using Cauchy's inequality we obtain

$$\begin{aligned} F_t^\alpha &= E\left(\sup_{t_0\le z\le t}\left|\sum_{n=0}^{n_z-1} I_\alpha[g(\cdot)]_{t_n,t_{n+1}} + I_\alpha[g(\cdot)]_{t_{n_z},z}\right|^2 \middle| \mathcal{A}_{t_0}\right) \\ &\le 2\,E\left(\sup_{t_0\le z\le t}\left|\sum_{n=0}^{n_z-1} I_\alpha[g(\cdot)]_{t_n,t_{n+1}}\right|^2 \middle| \mathcal{A}_{t_0}\right) \\ &\quad +2\,E\left(\sup_{t_0\le z\le t}\left|I_\alpha[g(\cdot)]_{t_{n_z},z}\right|^2 \middle| \mathcal{A}_{t_0}\right). \end{aligned} \tag{6.59}$$

Applying Doob's inequality to the first term of the equation (6.59) we get

$$\begin{aligned} & E\left(\sup_{t_0\le z\le t}\left|\sum_{n=0}^{n_z-1} I_\alpha[g(\cdot)]_{t_n,t_{n+1}}\right|^2 \middle| \mathcal{A}_{t_0}\right) \\ \le\;& 4\,E\left(\left|\sum_{n=0}^{n_t-1} I_\alpha[g(\cdot)]_{t_n,t_{n+1}}\right|^2 \middle| \mathcal{A}_{t_0}\right) \\ \le\;& 4\,E\Bigg(\Bigg[\Big|\sum_{n=0}^{n_t-2} I_\alpha[g(\cdot)]_{t_n,t_{n+1}}\Big|^2 \\ & +2\,\Big|\sum_{n=0}^{n_t-2} I_\alpha[g(\cdot)]_{t_n,t_{n+1}}\Big|\,E\Big(|I_\alpha[g(\cdot)]_{t_{n_t-1},t_{n_t}}|\,\Big|\mathcal{A}_{t_{n_t-1}}\Big) \\ & +E\left(|I_\alpha[g(\cdot)]_{t_{n_t-1},t_{n_t}}|^2\,\middle|\mathcal{A}_{t_{n_t-1}}\right)\Bigg]\Bigg|\mathcal{A}_{t_0}\Bigg) \\ \le\;& 4\,E\Bigg(\Bigg[\Big|\sum_{n=0}^{n_t-2} I_\alpha[g(\cdot)]_{t_n,t_{n+1}}\Big|^2 \\ & +E\left(|I_\alpha[g(\cdot)]_{t_{n_t-1},t_{n_t}}|^2\middle|\mathcal{A}_{t_{n_t-1}}\right)\Bigg]\Bigg|\mathcal{A}_{t_0}\Bigg), \end{aligned} \tag{6.60}$$

where the last line holds because, by the discrete time martingale property of the involved stochastic integrals, $E(I_\alpha[g(\cdot)]_{t_{n_t-1},t_{n_t}}|\mathcal{A}_{t_{n_t-1}}) = 0$.

Then by applying Lemma 6.2 we obtain

$$
\begin{aligned}
& E\left(\sup_{t_0\le z\le t}\left|\sum_{n=0}^{n_z-1} I_\alpha[g(\cdot)]_{t_n,t_{n+1}}\right|^2 \middle| \mathcal{A}_{t_0}\right) \\
\le\ & 4\,E\Bigg(\Big[|\sum_{n=0}^{n_t-2} I_\alpha[g(\cdot)]_{t_n,t_{n+1}}|^2 \\
& +4^{l(\alpha)-n(\alpha)}\ \Delta^{l(\alpha)+n(\alpha)-1}\int_{t_{n_t-1}}^{t_{n_t}} V_{t_{n_t-1},u,s(\alpha)}\,du\Big]\Bigg|\mathcal{A}_{t_0}\Bigg) \\
\le\ & 4\,E\Bigg(\Big[|\sum_{n=0}^{n_t-3} I_\alpha[g(\cdot)]_{t_n,t_{n+1}}|^2 \\
& +4^{l(\alpha)-n(\alpha)}\ \Delta^{l(\alpha)+n(\alpha)-1}\int_{t_{n_t-2}}^{t_{n_t-1}} V_{t_{n_t-2},u,s(\alpha)}\,du \\
& +4^{l(\alpha)-n(\alpha)}\ \Delta^{l(\alpha)+n(\alpha)-1}\int_{t_{n_t-1}}^{t_{n_t}} V_{t_{n_t-1},u,s(\alpha)}\,du\Big]\Bigg|\mathcal{A}_{t_0}\Bigg) \\
\le\ & 4\,E\Bigg(\Big[|\sum_{n=0}^{n_t-3} I_\alpha[g(\cdot)]_{t_n,t_{n+1}}|^2 \\
& +4^{l(\alpha)-n(\alpha)}\ \Delta^{l(\alpha)+n(\alpha)-1}\int_{t_{n_t-2}}^{t_{n_t-1}} V_{t_{n_t-2},u,s(\alpha)}\,du \\
& +4^{l(\alpha)-n(\alpha)}\ \Delta^{l(\alpha)+n(\alpha)-1}\int_{t_{n_t-1}}^{t_{n_t}} V_{t_{n_t-2},u,s(\alpha)}\,du\Big]\Bigg|\mathcal{A}_{t_0}\Bigg), \qquad (6.61)
\end{aligned}
$$

where the last passage holds since $V_{t_{n_t-1},u,s(\alpha)} \le V_{t_{n_t-2},u,s(\alpha)}$. Applying this procedure repetitively and using (6.53) we finally obtain

$$
\begin{aligned}
& E\left(\sup_{t_0\le z\le t}\left|\sum_{n=0}^{n_z-1} I_\alpha[g(\cdot)]_{t_n,t_{n+1}}\right|^2 \middle| \mathcal{A}_{t_0}\right) \\
\le\ & 4^{l(\alpha)-n(\alpha)+1}\ \Delta^{l(\alpha)+n(\alpha)-1}\, E\left(\int_{t_0}^{t} V_{t_0,u,s(\alpha)}\,du \middle| \mathcal{A}_{t_0}\right) \\
=\ & 4^{l(\alpha)-n(\alpha)+1}\ \Delta^{l(\alpha)+n(\alpha)-1}\int_{t_0}^{t} V_{t_0,u,s(\alpha)}\,du. \qquad (6.62)
\end{aligned}
$$

For the second term of equation (6.59), by applying the Cauchy-Schwarz inequality, similar steps as the ones used previously and Lemma 6.2, we get

$$E\left(\sup_{t_0\le z\le t}\left|I_\alpha[g(\cdot)]_{t_{n_z},z}\right|^2\middle|\mathcal{A}_{t_0}\right)$$

$$= E\left(\sup_{t_0\le z\le t}\left|\int_{t_{n_z}}^{z} I_{\alpha-}[g(\cdot)]_{t_{n_z},u}\,du\right|^2\middle|\mathcal{A}_{t_0}\right)$$

$$\le E\left(\sup_{t_0\le z\le t}(z-t_{n_z})\int_{t_{n_z}}^{z}|I_{\alpha-}[g(\cdot)]_{t_{n_z},u}|^2\,du\middle|\mathcal{A}_{t_0}\right)$$

$$\le \Delta\int_{t_0}^{t} E\left(E\Big(\sup_{t_{n_u}\le z\le u}|I_{\alpha-}[g(\cdot)]_{t_{n_u},z}|^2\Big|\mathcal{A}_{t_{n_u}}\Big)\middle|\mathcal{A}_{t_0}\right)du$$

$$\le \Delta\ 4^{l(\alpha-)-n(\alpha-)}\ \Delta^{l(\alpha-)+n(\alpha-)-1}\int_{t_0}^{t} E\left(\int_{t_{n_u}}^{u} V_{t_{n_u},z,s(\alpha-)}\,dz\middle|\mathcal{A}_{t_0}\right)du$$

$$\le \Delta\ 4^{l(\alpha-)-n(\alpha-)}\ \Delta^{l(\alpha-)+n(\alpha-)-1}\ \Delta\int_{t_0}^{t} V_{t_0,u,s(\alpha-)}\,du$$

$$= 4^{l(\alpha)-n(\alpha)}\ \Delta^{l(\alpha)+n(\alpha)-1}\int_{t_0}^{t} V_{t_0,u,s(\alpha)}\,du, \tag{6.63}$$

where the last passage holds since $l(\alpha)=l(\alpha-)+1$, $n(\alpha)=n(\alpha-)+1$ and $s(\alpha)=s(\alpha-)$.

Therefore, combining equations (6.62) and (6.63) we finally obtain

$$\begin{aligned} F_t^\alpha &\le 2\Bigg(4^{l(\alpha)-n(\alpha)+1}\ \Delta^{l(\alpha)+n(\alpha)-1}\int_{t_0}^{t} V_{t_0,u,s(\alpha)}\,du \\ &\qquad + 4^{l(\alpha)-n(\alpha)}\ \Delta^{l(\alpha)+n(\alpha)-1}\int_{t_0}^{t} V_{t_0,u,s(\alpha)}\,du\Bigg) \\ &\le 4^{l(\alpha)-n(\alpha)+2}\ \Delta^{l(\alpha)+n(\alpha)-1}\int_{t_0}^{t} V_{t_0,u,s(\alpha)}du, \end{aligned} \tag{6.64}$$

which completes the proof of Lemma 6.3. □

6.3 Moment Estimates for the SDE

We finally need an estimate of the moments of the solution of the SDE (2.2) that we shall use in the proof of Theorem 6.1.

Theorem 6.4 *Suppose that the coefficient functions* $a(\cdot), b(\cdot)$ *and* $c(\cdot)$ *of the SDE* (2.2) *satisfy the Lipschitz conditions* (2.3) *and the linear growth conditions* (2.4).

Moreover, let

$$E(|X_{t_0}|^2) < \infty. \tag{6.65}$$

Then the solution X_t *of* (2.2) *satisfies*

$$E\left(\sup_{t_0 \le s \le T} |X_s|^2 \middle| \mathcal{A}_{t_0}\right) \le C\Big(1 + E\left(|X_{t_0}|^2\right)\Big) \tag{6.66}$$

for $t \in [t_0, T]$ *with* $T < \infty$*, where* C *is a positive constant depending only on* $(T - t_0)$ *and the linear growth bound.*

A proof of this result, for the more general case of SDEs driven by semimartingales, can be found in Protter (2003).

6.4 Proof of Theorem 6.1

We can now present the proof of our main result, the Theorem 6.1.

Proof:

1. With the Wagner-Platen expansion (6.21) we can represent the solution of the SDE (2.2) as

$$X_\tau = \sum_{\alpha \in \mathcal{A}_\gamma} I_\alpha[f_\alpha(\rho, X_\rho)]_{\rho,\tau} + \sum_{\alpha \in \mathcal{B}(\mathcal{A}_\gamma)} I_\alpha[f_\alpha(\cdot, X_\cdot)]_{\rho,\tau}, \tag{6.67}$$

for any two stopping times ρ and τ with $0 \le \rho \le \tau \le T$ a.s. Therefore, we can express the solution of the SDE (2.2) at time $t \in [0, T]$ as

$$\begin{aligned} X_t &= X_0 + \sum_{\alpha \in \mathcal{A}_\gamma \setminus \{v\}} \left\{ \sum_{n=0}^{n_t - 1} I_\alpha[f_\alpha(t_n, X_{t_n})]_{t_n, t_{n+1}} + I_\alpha[f_\alpha(t_{n_t}, X_{t_{n_t}})]_{t_{n_t}, t} \right\} \\ &\quad + \sum_{\alpha \in \mathcal{B}(\mathcal{A}_\gamma)} \left\{ \sum_{n=0}^{n_t - 1} I_\alpha[f_\alpha(\cdot, X_\cdot)]_{t_n, t_{n+1}} + I_\alpha[f_\alpha(\cdot, X_\cdot)]_{t_{n_t}, t} \right\}, \end{aligned} \tag{6.68}$$

where n_t is defined as in equation (6.22).

We recall that the order γ strong Taylor approximation Y^Δ at time $t \in [0, T]$ is given by

$$Y_t^\Delta = Y_0^\Delta + \sum_{\alpha \in \mathcal{A}_\gamma \setminus \{v\}} \left\{ \sum_{n=0}^{n_t - 1} I_\alpha[f_\alpha(t_n, Y_{t_n}^\Delta)]_{t_n, t_{n+1}} + I_\alpha[f_\alpha(t_{n_t}, Y_{t_{n_t}}^\Delta)]_{t_{n_t}, t} \right\}. \tag{6.69}$$

From the estimate of Theorem 6.4 we have

$$E\left(\sup_{0 \le s \le T} |X_s|^2 \middle| \mathcal{A}_0\right) \le C\Big(1 + E\left(|X_0|^2\right)\Big). \tag{6.70}$$

2. We can also show a similar bound for the approximation Y^Δ. By definition (6.69) we have

$$\begin{aligned}
E\left(\sup_{0\le s\le T}|Y_s^\Delta|^2\Big|\mathcal{A}_0\right) &\le E\left(\sup_{0\le s\le T}(1+|Y_s^\Delta|^2)\Big|\mathcal{A}_0\right)\\
&\le E\Bigg(\sup_{0\le s\le T}\Big(1+\Big|Y_0^\Delta+\sum_{\alpha\in\mathcal{A}_\gamma\setminus\{v\}}\Big\{\sum_{n=0}^{n_s-1}I_\alpha[f_\alpha(t_n,Y_{t_n}^\Delta)]_{t_n,t_{n+1}}\\
&\quad +I_\alpha[f_\alpha(t_{n_s},Y_{t_{n_s}}^\Delta)]_{t_{n_s},s}\Big\}\Big|^2\Big)\Big|\mathcal{A}_0\Bigg)\\
&\le E\Bigg(\sup_{0\le s\le T}\Big(1+2\,|Y_0^\Delta|^2+2\Big|\sum_{\alpha\in\mathcal{A}_\gamma\setminus\{v\}}\\
&\quad \Big\{\sum_{n=0}^{n_s-1}I_\alpha[f_\alpha(t_n,Y_{t_n}^\Delta)]_{t_n,t_{n+1}}+I_\alpha[f_\alpha(t_{n_s},Y_{t_{n_s}}^\Delta)]_{t_{n_s},s}\Big\}\Big|^2\Big)\Big|\mathcal{A}_0\Bigg)\\
&\le C_1\left(1+\left|Y_0^\Delta\right|^2\right)+2\,K\sum_{\alpha\in\mathcal{A}_\gamma\setminus\{v\}}E\Bigg(\sup_{0\le s\le T}\\
&\quad \Big|\sum_{n=0}^{n_s-1}I_\alpha[f_\alpha(t_n,Y_{t_n}^\Delta)]_{t_n,t_{n+1}}+I_\alpha[f_\alpha(t_{n_s},Y_{t_{n_s}}^\Delta)]_{t_{n_s},s}\Big|^2\Big|\mathcal{A}_0\Bigg),
\end{aligned} \tag{6.71}$$

where K is a positive constant depending only on the strong order γ of the approximation. By Lemma 6.3 and the linear growth condition (6.29) we obtain

$$
\begin{aligned}
E\left(\sup_{0\le s\le T}|Y_s^\Delta|^2\Big|\mathcal{A}_0\right) &\le C_1(1+|Y_0^\Delta|^2)\\
&\quad +2\,K_1\sum_{\alpha\in\mathcal{A}_\gamma\setminus\{v\}}\left\{\int_0^T\int_{\mathcal{E}}\dots\int_{\mathcal{E}}\right.\\
&\quad \left.\times E\left(\sup_{0\le s\le u}\left|f_\alpha(s,Y_s^\Delta)\right|^2\Big|\mathcal{A}_0\right)\,\phi(dv^1)\dots\phi(dv^{s(\alpha)})\,du\right\}\\
&\le C_1\left(1+|Y_0^\Delta|^2\right)+2\,K_1\\
&\quad \times\sum_{\alpha\in\mathcal{A}_\gamma\setminus\{v\}}\left\{\int_{\mathcal{E}}\dots\int_{\mathcal{E}}K_2(v^1,\dots,v^{s(\alpha)})\phi(dv^1)\dots\phi(dv^{s(\alpha)})\right.\\
&\quad \left.\times\int_0^T E\left(\sup_{0\le s\le u}\left(1+|Y_s^\Delta|^2\right)\Big|\mathcal{A}_0\right)du\right\}\\
&\le C_1\left(1+|Y^\Delta(0)|^2\right)+C_2\int_0^T E\left(\sup_{0\le s\le u}\left(1+|Y_s^\Delta|^2\right)\Big|\mathcal{A}_0\right)du.
\end{aligned}
\tag{6.72}
$$

Then by applying the Gronwall inequality we obtain

$$
E\Big(\sup_{0\le s\le T}|Y_s^\Delta|^2\Big|\mathcal{A}_0\Big)\le C\,(1+|Y_0^\Delta|^2), \tag{6.73}
$$

where C is a positive finite constant.

3. Let us now analyze the mean square error of the order γ strong Taylor

approximation Y^Δ. By (6.68), (6.69) and Cauchy's inequality we obtain

$$\begin{aligned}
Z(t) &:= E\left(\sup_{0\le s\le t} |X_s - Y_s^\Delta|^2 \middle| \mathcal{A}_0\right) \\
&= E\Bigg(\sup_{0\le s\le t} \Bigg| X_0 - Y_0^\Delta \\
&\quad + \sum_{\alpha\in\mathcal{A}_\gamma\setminus\{v\}} \Big\{ \sum_{n=0}^{n_s-1} I_\alpha[f_\alpha(t_n, X_{t_n}) - f_\alpha(t_n, Y_{t_n}^\Delta)]_{t_n,t_{n+1}} \\
&\quad + I_\alpha[f_\alpha(t_{n_s}, X_{t_{n_s}}) - f_\alpha(t_{n_s}, Y_{t_{n_s}}^\Delta)]_{t_{n_s},s} \Big\} \\
&\quad + \sum_{\alpha\in\mathcal{B}(\mathcal{A}_\gamma)} \Big\{ \sum_{n=0}^{n_s-1} I_\alpha[f_\alpha(\cdot, X_\cdot)]_{t_n,t_{n+1}} + I_\alpha[f_\alpha(\cdot, X_\cdot)]_{t_{n_s},s} \Big\} \Bigg|^2 \Bigg| \mathcal{A}_0 \Bigg) \\
&\le C_3 \left\{ \left|X_0 - Y_0^\Delta\right|^2 + \sum_{\alpha\in\mathcal{A}_\gamma\setminus\{v\}} S_t^\alpha + \sum_{\alpha\in\mathcal{B}(\mathcal{A}_\gamma)} U_t^\alpha \right\} \qquad (6.74)
\end{aligned}$$

for all $t \in [0,T]$, where S_t^α and U_t^α are defined as

$$\begin{aligned}
S_t^\alpha &:= E\Bigg(\sup_{0\le s\le t} \Bigg| \sum_{n=0}^{n_s-1} I_\alpha\left[f_\alpha(t_n, X_{t_n}) - f_\alpha(t_n, Y_{t_n}^\Delta)\right]_{t_n,t_{n+1}} \\
&\quad + I_\alpha\left[f_\alpha(t_{n_s}, X_{t_{n_s}}) - f_\alpha(t_{n_s}, Y_{t_{n_s}}^\Delta)\right]_{t_{n_s},s} \Bigg|^2 \Bigg| \mathcal{A}_0 \Bigg), \qquad (6.75)
\end{aligned}$$

$$U_t^\alpha := E\left(\sup_{0\le s\le t} \left| \sum_{n=0}^{n_s-1} I_\alpha[f_\alpha(\cdot, X_\cdot)]_{t_n,t_{n+1}} + I_\alpha[f_\alpha(\cdot, X_\cdot)]_{t_{n_s},s} \right|^2 \middle| \mathcal{A}_0\right). \qquad (6.76)$$

4. By using again Lemma 6.3 and Lipschitz condition (6.27) we obtain

$$\begin{aligned}
S_t^\alpha &= E\Bigg(\sup_{0\le s\le t} \Big| \sum_{n=0}^{n_s-1} I_\alpha\left[f_\alpha(t_n, X_{t_n}) - f_\alpha(t_n, Y_{t_n}^\Delta)\right]_{t_n,t_{n+1}} \\
&\quad + I_\alpha\left[f_\alpha(t_{n_s}, X_{t_{n_s}}) - f_\alpha(t_{n_s}, Y_{t_{n_s}}^\Delta)\right]_{t_{n_s},s} \Big|^2 \Bigg| \mathcal{A}_0 \Bigg) \\
&\le C_4 \int_0^t \int_{\mathcal{E}} \cdots \int_{\mathcal{E}} E\left(\sup_{0\le s\le u} |f_\alpha(t_{n_s}, X_{t_{n_s}}) - f_\alpha(t_{n_s}, Y_{t_{n_s}}^\Delta)|^2 \Big| \mathcal{A}_0 \right) \\
&\quad \times \phi(dv^1)\ldots\phi(dv^{s(\alpha)})du \\
&\le C_4 \int_{\mathcal{E}} \cdots \int_{\mathcal{E}} \left(K_1(v^1,\ldots,v^{s(\alpha)})\right)^2 \phi(dv^1)\ldots\phi(dv^{s(\alpha)}) \\
&\quad \times \int_0^t E\left(\sup_{0\le s\le u} |X_{t_{n_s}} - Y_{t_{n_s}}^\Delta|^2 \Big| \mathcal{A}_0 \right) du \\
&\le C_5 \int_0^t Z(u)\, du. \qquad (6.77)
\end{aligned}$$

Applying again Lemma 6.3 and the linear growth condition (6.29) we obtain

$$\begin{aligned}
U_t^\alpha &= E\left(\sup_{0\le s\le t} \Big| \sum_{n=0}^{n_s-1} I_\alpha[f_\alpha(\cdot, X_\cdot)]_{t_n,t_{n+1}} + I_\alpha[f_\alpha(\cdot, X_\cdot)]_{t_{n_s},s} \Big|^2 \Bigg| \mathcal{A}_0 \right) \\
&\le C_5\, \Delta^{\psi(\alpha)} \int_0^t \int_{\mathcal{E}} \cdots \int_{\mathcal{E}} E\left(\sup_{0\le s\le u} |f_\alpha(s, X_s)|^2 \Big| \mathcal{A}_0 \right) \phi(dv^1)\ldots\phi(dv^{s(\alpha)})du \\
&\le C_5\, \Delta^{\psi(\alpha)} \int_{\mathcal{E}} \cdots \int_{\mathcal{E}} K_2(v^1,\ldots,v^{s(\alpha)}) \phi(dv^1)\ldots\phi(dv^{s(\alpha)}) \\
&\quad \times \int_0^t E\left(\sup_{0\le s\le u} (1 + |X_s|^2) \Big| \mathcal{A}_0 \right) du \\
&\le C_6\, \Delta^{\psi(\alpha)} \left(t + \int_0^t E\Big(\sup_{0\le s\le u} |X_s|^2 \Big| \mathcal{A}_0 \Big) du \right), \qquad (6.78)
\end{aligned}$$

where

$$\psi(\alpha) = \begin{cases} 2l(\alpha) - 2 & : l(\alpha) = n(\alpha) \\ l(\alpha) + n(\alpha) - 1 & : l(\alpha) \neq n(\alpha). \end{cases}$$

Since we are now considering $\alpha \in \mathcal{B}(\mathcal{A}_\gamma)$, we have that $l(\alpha) \ge \gamma + 1$ when $l(\alpha) = n(\alpha)$ and $l(\alpha) + n(\alpha) \ge 2\gamma + 1$ when $l(\alpha) \neq n(\alpha)$, so that $\psi(\alpha) \ge 2\gamma$.

Therefore, applying estimate (6.66) of Theorem 6.4 we obtain

$$\begin{aligned} U_t^\alpha &\leq C_6\,\Delta^{2\gamma}\left(t+\int_0^t C_1(1+|X_0|^2)du\right) \\ &\leq C_7\,\Delta^{2\gamma}\,(1+|X_0|^2). \end{aligned} \tag{6.79}$$

5. Combining equations (6.74), (6.77) and (6.79) we obtain

$$Z(t) \leq C_8 \left\{ \; |X_0 - Y_0^\Delta|^2 + C_9\,\Delta^{2\gamma}\left(1+|X_0|^2\right) + C_{10}\int_0^t Z(u)du \right\}. \tag{6.80}$$

By equations (6.70) and (6.73) $Z(t)$ is bounded. Therefore, by the Gronwall inequality we obtain

$$Z(T) \leq K_4\,\left(1+|X_0|^2\right)\,\Delta^{2\gamma} + K_5\left(|X_0 - Y_0^\Delta|^2\right). \tag{6.81}$$

Finally, by equation (6.26), we obtain

$$\sqrt{E(\sup_{0\leq s\leq T}|X_s - Y_s^\Delta|^2\,|\mathcal{A}_0)} = \sqrt{Z(T)} \leq K_3\;\Delta^\gamma, \tag{6.82}$$

which completes the proof of Theorem 6.1. □

7 General Strong Schemes

Now we consider more general strong schemes, the strong Itô schemes, constructed with the same multiple stochastic integrals underlying the scheme (6.24), but with different coefficients. Under particular conditions on these coefficients, the strong Itô schemes converge to the solution X of the SDE (2.2) with the same strong order γ of the corresponding strong Taylor schemes. Therefore, we can construct more general strong approximations of any given order, in particular, derivative free and implicit schemes.

For a time discretization with maximum step size $\Delta \in (0,1)$, as the one introduced in Section 6.1, we define the *order γ strong Itô scheme* by the vector equation

$$Y_{n+1}^\Delta = Y_n^\Delta + \sum_{\alpha\in\mathcal{A}_\gamma\setminus\{v\}} I_\alpha\left[h_{\alpha,n}\right]_{t_n,t_{n+1}} + R_n\,, \tag{7.1}$$

with $n \in \{0,1,\ldots,n_T-1\}$, if the coefficients $h_{\alpha,n}$ are $\mathcal{A}_{t_n}$-measurable and satisfy the estimate

$$E\left(\max_{0\leq n\leq n_T-1}|h_{\alpha,n} - f_\alpha(t_n,Y_n)|^2\right) \leq C(u)\,\Delta^{2\gamma-\psi(\alpha)}, \tag{7.2}$$

for all $\alpha \in \mathcal{A}_\gamma \backslash \{v\}$, where $C : \mathcal{E}^{s(\alpha)} \to \mathbb{R}$ is a $\phi(du)$- integrable function. Here

$$\psi(\alpha) = \begin{cases} 2l(\alpha) - 2 & : l(\alpha) = n(\alpha) \\ l(\alpha) + n(\alpha) - 1 & : l(\alpha) \neq n(\alpha), \end{cases}$$

and the R_n satisfy

$$E\left(\max_{1 \leq n \leq n_T} \left| \sum_{0 \leq k \leq n-1} R_k \right|^2 \right) \leq K\, \Delta^{2\gamma}. \tag{7.3}$$

We can now formulate a convergence theorem that will enable us to construct strong Itô approximations of any given strong order, including derivative free and drift-implicit schemes.

Theorem 7.1 *Let* $Y^\Delta = \{Y_n^\Delta, n \in \{0, 1, \ldots, n_T\}\}$ *be a discrete time approximation generated via the strong Itô scheme* (7.1)*, for a given time discretisation with maximum time step size* $\Delta \in (0, 1)$*, and for* $\gamma \in \{0.5, 1, 1.5, 2, \ldots\}$*. If the conditions of Theorem 6.1 are satisfied, then*

$$\sqrt{E\left(\max_{0 \leq n \leq n_T} |X_{t_n} - Y_n^\Delta|^2 \right)} \leq K\, \Delta^\gamma. \tag{7.4}$$

Proof: Since we have already proved in Theorem 6.1 that the strong Taylor scheme (6.24) converges with strong order γ, here it will be sufficient to show that the Itô scheme (7.1) converges with strong order γ to the Taylor scheme. With $\widetilde{Y}^\Delta$ we denote here the strong Taylor scheme (6.24). Let us also assume, for simplicity, that $\widetilde{Y}_0 = Y_0$. Then applying Jensen's inequality and Cauchy's inequality, we obtain for all $t \in [0,T]$ the estimate

$$
\begin{aligned}
H_t &:= E\left(\max_{1\le n\le n_t}\left|\widetilde{Y}_n^{\Delta} - Y_n^{\Delta}\right|^2\right) \\
&= E\left(\max_{1\le n\le n_t}\left|\sum_{k=0}^{n-1}\sum_{\alpha\in\mathcal{A}_\gamma\setminus\{v\}} I_\alpha\left[f_\alpha(t_k,\widetilde{Y}_k^{\Delta})\right]_{t_k,t_{k+1}}\right.\right. \\
&\qquad \left.\left. - \sum_{k=0}^{n-1}\left(\sum_{\alpha\in\mathcal{A}_\gamma\setminus\{v\}} I_\alpha\left[h_{\alpha,k}\right]_{t_k,t_{k+1}} + R_n\right)\right|^2\right) \\
&\le K_1 \sum_{\alpha\in\mathcal{A}_\gamma\setminus\{v\}}\left\{E\left(\max_{1\le n\le n_t}\left|\sum_{k=0}^{n-1} I_\alpha\left[f_\alpha(t_k,\widetilde{Y}_k^{\Delta}) - f_\alpha(t_k,Y_k^{\Delta})\right]_{t_k,t_{k+1}}\right|^2\right)\right. \\
&\qquad \left. +E\left(\max_{1\le n\le n_t}\left|\sum_{k=0}^{n-1} I_\alpha\left[f_\alpha(t_k,Y_k^{\Delta}) - h_{\alpha,k}\right]_{t_k,t_{k+1}}\right|^2\right)\right\} \\
&\qquad +K_1\, E\left(\max_{1\le n\le n_t}\left|\sum_{k=0}^{n-1} R_n\right|^2\right). \qquad (7.5)
\end{aligned}
$$

Applying Lemma 6.3, condition (7.3), the Lipschitz condition (6.27) and condition (7.2) yields

$$
\begin{aligned}
H_t &\le K_2 \sum_{\alpha\in\mathcal{A}_\gamma\setminus\{v\}}\left\{\int_0^t\int_{\mathcal{E}}\cdots\int_{\mathcal{E}}\Big(E\big(\max_{1\le n\le n_u}|f_\alpha(t_k,\widetilde{Y}_k^{\Delta}) - f_\alpha(t_k,Y_k^{\Delta})|^2\big)\right. \\
&\qquad \left. +E\big(\max_{1\le n\le n_u}|f_\alpha(t_k,Y_k^{\Delta}) - h_{\alpha,n}|^2\big)\Big)\phi(dv^1)\ldots\phi(dv^{s(\alpha)})\,du\right\}\Delta^{\psi(\alpha)} + K_3\Delta^{2\gamma} \\
&\le K_5\int_0^t E\big(\max_{1\le n\le n_u}|\widetilde{Y}_k^{\Delta} - Y_k^{\Delta}|^2\big)du \times \sum_{\alpha\in\mathcal{A}_\gamma\setminus\{v\}}\Delta^{\psi(\alpha)} \;+\; K_6\,\Delta^{2\gamma} \\
&\le K_7\int_0^t H_u\,du + K_6\,\Delta^{2\gamma}, \qquad (7.6)
\end{aligned}
$$

where the last inequality holds since $\Delta \in (0,1)$. With the estimate of Theorem 6.4 and a similar estimate on the numerical solution Y^{Δ}, one can show that H_t is bounded. Therefore, by applying the Gronwall inequality to (7.6), we obtain

$$H_t \le K_5\,\Delta^{2\gamma}\,e^{K_7 t}. \qquad (7.7)$$

Since we assume, for simplicity, $\widetilde{Y}_0^{\Delta} = Y_0^{\Delta}$, we have

$$E\big(\max_{0\le n\le n_T}|\widetilde{Y}_n^{\Delta} - Y_n^{\Delta}|^2\,\big) \le K\,\Delta^{2\gamma}. \qquad (7.8)$$

Finally, by the estimate of Theorem 6.1 we obtain

$$\sqrt{E(\max_{0\le n\le n_T} |X_{t_n} - Y_n^\Delta|^2)} = \sqrt{E(\max_{0\le n\le n_T} |X_{t_n} - \widetilde{Y}_n^\Delta + \widetilde{Y}_n^\Delta - Y_n^\Delta|^2)} \le K\,\Delta^\gamma, \tag{7.9}$$

which finalises the proof of Theorem 7.1. □

7.1 Derivative Free Schemes

The strong Itô scheme (7.1) and the related convergence Theorem 7.1 allow us to asses the strong order of convergence of general approximations. In this section we show how it is possible to rewrite derivative free schemes, as the ones presented in Section 3.2, as strong Itô schemes.

We recall here that the explicit order 1.0 strong Taylor scheme, presented in Section 3.2, is given as

$$\begin{aligned} Y_{n+1} \;=\;& Y_n + a(Y_n)\Delta + b(Y_n)\Delta W_n + \int_{t_n}^{t_{n+1}}\int_{\mathcal{E}} c(Y_n,v)\,p_\phi(dv\times ds) \\ &+\frac{1}{\sqrt{\Delta}}\Big\{b(\overline{Y}_n) - b(Y_n)\Big\}\int_{t_n}^{t_{n+1}}\int_{t_n}^{s_2} dW(s_1)\,dW(s_2) \\ &+\int_{t_n}^{t_{n+1}}\int_{\mathcal{E}}\int_{t_n}^{s_2}\frac{1}{\sqrt{\Delta}}\Big\{c(\overline{Y}_n,v) - c(Y_n,v)\Big\}\,dW(s_1)\,p_\phi(dv\times ds_2) \\ &+\int_{t_n}^{t_{n+1}}\int_{t_n}^{s_2}\int_{\mathcal{E}}\Big\{b(Y_n + c(Y_n,v)) - b(Y_n)\Big\}\,p_\phi(dv\times ds_1)\,dW(s_2) \\ &+\int_{t_n}^{t_{n+1}}\int_{\mathcal{E}}\int_{t_n}^{s_2}\int_{\mathcal{E}}\Big\{c(Y_n + c(Y_n,v_2),v_1) - c(Y_n,v_1)\Big\} \\ &\times p_\phi(dv_1\times ds_1)\,p_\phi(dv_2\times ds_2), \end{aligned} \tag{7.10}$$

with the supporting value

$$\overline{Y}_n = Y_n + b(Y_n)\sqrt{\Delta}. \tag{7.11}$$

From the deterministic Taylor expansion, we obtain

$$b(\overline{Y}_n) = b(Y_n) + b'(Y_n)\Big\{\overline{Y}_n - Y_n\Big\} + \frac{b''\Big(Y_n + \theta(\overline{Y}_n - Y_n)\Big)}{2}\Big\{\overline{Y}_n - Y_n\Big\}^2, \tag{7.12}$$

and

$$c(\overline{Y}_n,v) = c(Y_n,v) + c'(Y_n,v)\Big\{\overline{Y}_n - Y_n\Big\} + \frac{c''\Big(Y_n + \theta(\overline{Y}_n - Y_n),v\Big)}{2}\Big\{\overline{Y}_n - Y_n\Big\}^2, \tag{7.13}$$

with

$$b'(x) := \frac{db(x)}{dx} \quad \text{and} \quad b''(x) := \frac{d^2b(x)}{dx^2}, \tag{7.14}$$

$$c'(x,v) := \frac{\partial c(x,v)}{\partial x} \quad \text{and} \quad c''(x,v) := \frac{\partial^2 c(x,v)}{\partial x^2} \tag{7.15}$$

for every $v \in \mathcal{E}$, where $\theta \in (0,1)$.

Therefore, we can rewrite the scheme (7.10) as

$$\begin{aligned} Y_{n+1} &= Y_n + I_{(0)}[h_{(0),n}]_{t_n,t_{n+1}} + I_{(1)}[h_{(1),n}]_{t_n,t_{n+1}} + I_{(-1)}[h_{(-1),n}]_{t_n,t_{n+1}} \\ &\quad + I_{(1,1)}[h_{(1,1),n}]_{t_n,t_{n+1}} + I_{(1,-1)}[h_{(1,-1),n}]_{t_n,t_{n+1}} \\ &\quad + I_{(-1,1),n}[h_{(-1,1),n}]_{t_n,t_{n+1}} + I_{(-1,-1)}[h_{(-1,-1),n}]_{t_n,t_{n+1}}, \end{aligned} \tag{7.16}$$

with

$$\begin{aligned} &h_{(0),n} = a(Y_n), \qquad h_{(1),n} = b(Y_n), \qquad h_{(-1),n} = c(Y_n, v), \\ &h_{(1,-1),n} = \frac{\Big\{c(\overline{Y}_n, v) - c(Y_n, v)\Big\}}{\sqrt{\Delta}}, \quad h_{(-1,1),n} = b\Big(Y_n + c(Y_n,v)\Big) - b\Big(Y_n\Big), \\ &h_{(-1,-1),n} = c\Big(Y_n + c(Y_n, v_2), v_1\Big) - c\Big(Y_n, v_1\Big), \quad h_{(1,1),n} = \frac{\Big\{b(\overline{Y}_n) - b(Y_n)\Big\}}{\sqrt{\Delta}}. \end{aligned} \tag{7.17}$$

The coefficients $h_{\alpha,n}$ are different from the coefficients $f_{\alpha,n}$ of the order 1.0 strong Taylor scheme (3.6), only for $\alpha = (1,1)$ and $\alpha = (1,-1)$. Therefore, to prove that the scheme (7.10) is an order 1.0 strong Itô scheme, it remains to check condition (7.2) for these two coefficients.

By the linear growth condition (6.29) of Theorem 6.1, we have

$$\begin{aligned} \left| b(Y_n)^2 b''\left(Y_n + \theta\, b(Y_n)\sqrt{\Delta}\right)\right|^2 &\leq K_1(1+|Y_n|^4)\, K_2\left(1+|Y_n|^2\right) \\ &= C_1\left(1+|Y_n|^2+|Y_n|^4+|Y_n|^6\right). \end{aligned} \tag{7.18}$$

In a similar way we also obtain

$$\left| b(Y_n)^2 c''\left(Y_n + \theta\, b(Y_n)\sqrt{\Delta}, v\right)\right|^2 \leq C_2(v)\left(1+|Y_n|^2+|Y_n|^4+|Y_n|^6\right), \tag{7.19}$$

where $C_2(v) : \mathcal{E} \to \mathbb{R}$ is a $\phi(dv)$- integrable function.

Following similar steps as the ones used in the first part of the proof of Theorem 6.1, we can show that

$$E\left(\max_{0\leq n\leq n_T-1} |Y_n|^{2q}\right) \leq K\Big(1 + E(|Y_0|^{2q})\Big), \tag{7.20}$$

for $q \in \mathrm{N}$. Therefore, assuming $E(|Y_0|^6) < \infty$, by conditions (7.18), (7.19), and (7.20), we obtain

$$\begin{aligned} E\left(\max_{0 \le n \le n_T - 1} \left| h_{(1,1),n} - f_{(1,1)}(t_n, Y_n) \right|^2 \right) &\le E\left(\max_{0 \le n \le n_T - 1} \left| \frac{b(Y_n)^2 b''(Y_n)}{2} \sqrt{\Delta} \right|^2 \right) \\ &\le K\Delta \Big(1 + E(|Y_0|^6) \Big) \\ &\le C\, \Delta^{2\gamma - \psi(\alpha)}. \qquad (7.21) \end{aligned}$$

We also have

$$E\left(\max_{0 \le n \le n_T - 1} \left| h_{(1,-1),n} - f_{(1,1)}(t_n, Y_n) \right|^2 \right) \le C(v)\, \Delta^{2\gamma - \psi(\alpha)}, \qquad (7.22)$$

where $C(v) : \mathcal{E} \to \mathbb{R}$ is a $\phi(dv)$- integrable function, which shows that the scheme (7.10) is a strong Itô scheme of order $\gamma = 1.0$.

7.2 Implicit Schemes

As explained in Section 3.3, for any strong Taylor scheme of order γ it is possible to obtain a drift-implicit scheme of the same strong order of convergence. To avoid problems due to the reciprocal of Gaussian random variables, one can, in general, introduce implicitness only in the drift terms. Drift-implicit schemes of order γ can be derived by an application of the Wagner-Platen expansion to the drift terms of a correspondent strong Taylor scheme of order γ. If we apply the Wagner-Platen expansion to the drift term $a(x)$ we can write

$$\begin{aligned} a(X_t) &= a(X_{t+\Delta}) - L^0 a(X_t)\Delta - L^1 a(X_t) \left(W(t+\Delta) - W(t) \right) \\ &\quad - L^{-1} a(X_t) \left(p_\phi(t+\Delta) - p_\phi(t) \right) - R_1(t), \qquad (7.23) \end{aligned}$$

where

$$\begin{aligned} R_1(t) &= \int_t^{t+\Delta} \left\{ \int_t^s L^0 L^0 a(X_u) du + \int_t^s L^1 L^0 a(X_u) dW_u \right. \\ &\quad \left. + \int_t^s \int_{\mathcal{E}} L^{-1} L^0 a(X_u) p_\phi(dv \times du) \right\} ds \\ &\quad + \int_t^{t+\Delta} \left\{ \int_t^s L^0 L^1 a(X_u) du + \int_t^s L^1 L^1 a(X_u) dW_u \right. \\ &\quad \left. + \int_t^s \int_{\mathcal{E}} L^{-1} L^1 a(X_u) p_\phi(dv_1 \times du) \right\} dW_s \\ &\quad + \int_t^{t+\Delta} \left\{ \int_t^s L^0 L^{-1} a(X_u) du + \int_t^s L^1 L^{-1} a(X_u) dW_u \right. \\ &\quad \left. + \int_t^s \int_{\mathcal{E}} L^{-1} L^{-1} a(X_u) p_\phi(dv_1 \times du) \right\} p_\phi(dv_2 \times ds), \qquad (7.24) \end{aligned}$$

and the operators L^0, L^1 and L^{-1} are defined in (6.12), (6.13) and (6.14), respectively.

For any $\theta \in [0,1]$, we can rewrite the Euler scheme (3.3) as

$$\begin{aligned} Y_{n+1} &= Y_n + \{\theta\, a(Y_n) + (1-\theta)\, a(Y_n)\}\, \Delta + b(Y_n)\Delta W_n \\ &\quad + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} c(Y_n, v)\, p_\phi(dv \times ds), \end{aligned} \tag{7.25}$$

and by replacing the first drift coefficient $a(Y_n)$ with its implicit expansion (7.23), we obtain

$$\begin{aligned} Y_{n+1} &= Y_n + \Big\{\theta\, a(Y_{n+1}) + (1-\theta)\, a(Y_n)\Big\}\Delta + b(Y_n)\Delta W_n \\ &\quad + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} c(Y_n, v) p_\phi(dv \times ds) \\ &\quad - \Big\{L^0 a(Y_n)\Delta + L^1 a(Y_n)\Delta W_n + L^{-1} a(Y_n)\Delta p_n + R_1(t)\Big\}\theta\, \Delta. \end{aligned} \tag{7.26}$$

However, the terms in the last line of equation (7.26) are not necessary for a scheme with strong order $\gamma = 0.5$. Therefore, they can be discarded when deriving the implicit Euler scheme (3.28).

By applying the same procedure to every time integral appearing in a higher order strong Taylor scheme it is possible to derive higher order implicit schemes as, for instance, the drift-implicit order 1.0 strong scheme (3.29).

To prove the strong order of convergence of drift-implicit schemes it is sufficient to show that one can rewrite these as strong Itô schemes. The drift-implicit Euler scheme, for instance, can be written as an order 0.5 strong Itô scheme given by

$$Y_{n+1} = Y_n + I_{(0)}[h_{(0),n}]_{t_n,t_{n+1}} + I_{(1)}[h_{(1),n}]_{t_n,t_{n+1}} + I_{(-1)}[h_{(-1),n}]_{t_n,t_{n+1}} + R_n \tag{7.27}$$

with

$$h_{(0),n} = a(Y_n), \qquad h_{(1),n} = b(Y_n), \qquad h_{(-1),n} = c(Y_n, v), \tag{7.28}$$

and

$$R_n = \theta\, \Delta\, \left(a(Y_{n+1}) - a(Y_n)\right). \tag{7.29}$$

Since the coefficients $h_{\alpha,n}$ are the same as the ones employed in the Euler scheme (3.3), we have only to check condition (7.3) for the remainder term R_n. Following similar steps as the ones used in the first part of the proof of Theorem 6.1, we can show that

$$E\left(\max_{0 \le n \le n_T - 1} |Y_n|^2\right) \le K\Big(1 + E(|Y_0|^2)\Big). \tag{7.30}$$

By applying Jensen's inequality, the Chauchy-Schwarz inequality, the linear growth condition (2.4) and the estimate (7.30), we obtain

$$\begin{aligned} E\left(\max_{1\leq n\leq n_T} \left| \sum_{0\leq k\leq n-1} R_k \right|^2 \right) &= E\left(\max_{1\leq n\leq n_T} \left| \sum_{0\leq k\leq n-1} \theta\,\Delta\left(a(Y_{k+1}) - a(Y_k)\right) \right|^2 \right) \\ &\leq K\,\Delta^2\, E\left(\max_{1\leq n\leq n_T} \sum_{0\leq k\leq n-1} \left(2\,|a(Y_{k+1})|^2 + 2\,|a(Y_k)|^2 \right) \right) \\ &\leq K\,\Delta^2 \left\{ K_1 + E\left(\sum_{0\leq k\leq n_T-1} \left(2\,|Y_{k+1}|^2 + 2\,|Y_k|^2 \right) \right) \right\} \\ &\leq K\Delta^2 \Big(1 + E(|Y_0|^2) \Big) \\ &\leq C\,\Delta^{2\gamma}. \end{aligned} \tag{7.31}$$

Therefore, the convergence of the drift-implicit Euler scheme follows from Theorem 7.1 since we have shown that it can be rewritten as a strong Itô scheme of order 0.5. In a similar way it is possible to show that the drift-implicit order $\gamma = 1.0$ strong Taylor scheme (3.29) can be rewritten as an order $\gamma = 1.0$ strong Itô scheme.

8 Jump Adapted Schemes

In this section we present a convergence theorem for jump adapted approximations that allows us to asses the strong order of convergence of the schemes presented in Section 4.

We consider here a *jump adapted time discretisation* $0 = t_0 < t_1 < \ldots < t_N = T$ with maximum step size $\Delta \in (0,1)$ as suggested in Platen (1982a). The term "jump adapted" means that the time discretisation includes all the jump times $\{\tau_1, \tau_2, \ldots\}$ of the Poisson measure p_ϕ. A maximum step size $\Delta \in (0,1)$, means that for every $n \in \{0, 1, 2, \ldots, n_T - 1\}$ $P(t_{n+1} - t_n \leq \Delta) = 1$ and, if the discretisation time t_{n+1} is not a jump time, then t_{n+1} is $\mathcal{A}_{t_n}$-measurable. We also require to have a finite number of time discretisation points, that means $n_t < \infty$ a.s. for $t \in [0, T]$, where n_t is defined in (6.22). For instance, the superposition of the jump times to an equidistant time discretisation, as presented in Section 4, satisfies these assumptions.

As explained in Section 4, by construction the jumps arise only at discretisation points. Therefore, between discretisation points we can approximate the stochastic process X with a strong Taylor scheme for diffusions. For this reason we use here a slightly modified notation from the one introduced in Section 6, as will be outlined below.

For $m \in \mathcal{N}$ the set of all multi-indices α that do not include components equal to -1 is now denoted by

$$\overline{\mathcal{M}}_m = \{(j_1, \ldots, j_l) : j_i \in \{0, 1, 2, \ldots, l\}, i \in \{1, 2, \ldots, l\} \text{ for } l \in \mathcal{N}\} \cup \{v\}, \quad (8.1)$$

where v is the multi-index of length zero.

Let $\overline{\mathcal{L}}^0$ be the set of functions $f(t,x) : [0,T] \times \mathbb{R}^d \longrightarrow \mathbb{R}^d$ from $\mathcal{C}^{1,2}$ and $\overline{\mathcal{L}}^k$, for $k \in \{1, \ldots, m\}$, the set of functions $f(t,x)$ with partial derivatives $\frac{\partial}{\partial x^i} f(t,x)$, $i \in \{1, \ldots, d\}$. We also introduce the following operators for a function $f(t,x) \in \overline{\mathcal{L}}^k$, with $k \in \{0, 1, \ldots, m\}$:

$$\begin{aligned} \overline{L}^{(0)} f(t,x) \;&:=\; \frac{\partial}{\partial t} f(t,x) + \sum_{i=1}^{d} a^i(t,x) \frac{\partial}{\partial x^i} f(t,x) \\ &\quad + \frac{1}{2} \sum_{i,r=1}^{d} \sum_{j=1}^{m} b^{i,j}(t,x) b^{r,j}(t,x) \frac{\partial^2}{\partial x^i \partial x^j} f(t,x) \end{aligned} \quad (8.2)$$

and

$$\overline{L}^{(k)} f(t,x) := \sum_{i=1}^{d} b^{i,k}(t,x) \frac{\partial}{\partial x^i} f(t,x), \qquad \text{for } k \in \{1, \ldots, m\} \quad (8.3)$$

for all $t \in [0,T]$ and $x \in \mathbb{R}^d$.

For all $\alpha = (j_1, \ldots, j_{l(\alpha)}) \in \overline{\mathcal{M}}_m$ and a function $f : [0,T] \times \mathbb{R}^d \longrightarrow \mathbb{R}^d$, we define recursively the *Itô coefficient functions* $\overline{f}_\alpha$

$$\overline{f}_\alpha(t,x) := \begin{cases} f\big(t,x\big) & \text{for} \quad l(\alpha) = 0, \\ f\big(t, a(t,x)\big) & \text{for} \quad l(\alpha) = 1,\ j_1 = 0, \\ f\big(t, b^{j_1}(t,x)\big) & \text{for} \quad l(\alpha) = 1,\ j_1 \in \{1, \ldots, m\}, \\ \overline{L}^{(j_1)} \overline{f}_{-\alpha}(t,x) & \text{for} \quad l(\alpha) \geq 2,\ j_1 \in \{0, \ldots, m\}, \end{cases} \quad (8.4)$$

assuming that the coefficients of the SDE (2.2) satisfy the conditions of smoothness and integrability needed for the operators in (8.4) to be well defined.

Given a set $\mathcal{A} \subset \overline{\mathcal{M}}_m$, we also define the *remainder* set $\overline{\mathcal{B}}(\mathcal{A})$ of $\mathcal{A}$ by

$$\overline{\mathcal{B}}(\mathcal{A}) = \{\alpha \in \overline{\mathcal{M}}_m \backslash \mathcal{A} : -\alpha \in \mathcal{A}\}. \quad (8.5)$$

Moreover, for every $\gamma \in \{0.5, 1, 1.5, 2, \ldots\}$ we define the hierarchical set

$$\overline{\mathcal{A}}_\gamma = \{\alpha \in \overline{\mathcal{M}} : l(\alpha) + n(\alpha) \leq 2\gamma \quad \text{or} \quad l(\alpha) = n(\alpha) = \gamma + \frac{1}{2}\}. \quad (8.6)$$

Then for a jump adapted time discretisation, with maximum time step size $\Delta \in (0,1)$, we define the *jump adapted order γ strong Taylor scheme* by

$$\overline{Y}_{t_{n+1}-} = \overline{Y}_{t_n} + \sum_{\alpha \in \overline{\mathcal{A}}_\gamma \backslash \{v\}} \overline{f}_\alpha(t_n, \overline{Y}_{t_n}) I_\alpha \quad (8.7)$$

and

$$\overline{Y}_{t_{n+1}} = \overline{Y}_{t_{n+1}-} + \int_{\mathcal{E}} c(t_n-, \overline{Y}_{t_{n+1}-}, v)\, p_\phi(dv \times \{t_{n+1}\}), \tag{8.8}$$

where I_α is the multiple stochastic integral of the multi-index α over the time period $(t_n, t_{n+1}]$ and $n \in \{0, 1, \ldots, n_T - 1\}$.

As in Section 6.1, to asses the order of strong convergence of these schemes, we define through a specific interpolation the *jump adapted order γ strong Taylor approximation* by

$$\overline{Y}_t = \sum_{\alpha \in \overline{\mathcal{A}}_\gamma \setminus \{v\}} I_\alpha[\overline{f}_\alpha(t_{n_t}, \overline{Y}_{t_{n_t}})]_{t_{n_t}, t} \tag{8.9}$$

since there are no jumps between grid points. We can now formulate a convergence theorem for jump adapted schemes.

Theorem 8.1 *For a given $\gamma \in \{0.5, 1, 1.5, 2, \ldots\}$, let $\overline{Y}^\Delta = \{\overline{Y}^\Delta_t,\ t \in [0,T]\}$ be the order γ jump adapted strong Taylor approximation corresponding to a jump adapted time discretisation with maximum step size $\Delta \in (0,1)$. We assume that*

$$E(|X_0|^2) < \infty \quad \textit{and} \quad E(|X_0 - \overline{Y}^\Delta_0|^2) \leq C\, \Delta^{2\gamma}. \tag{8.10}$$

Moreover, suppose that the coefficient functions $\overline{f}_\alpha$ satisfy the following conditions:

For $\alpha \in \overline{\mathcal{A}}_\gamma$, $t \in [0,T]$ and $x, y \in \mathbb{R}^d$, the coefficient $\overline{f}_\alpha$ satisfies the Lipschitz type condition

$$\left|\overline{f}_\alpha(t,x) - \overline{f}_\alpha(t,y)\right| \leq K_1\, |x - y|. \tag{8.11}$$

For all $\alpha \in \overline{\mathcal{A}}_\gamma \bigcup \overline{B}(\overline{\mathcal{A}}_\gamma)$ we assume

$$\overline{f}_{-\alpha} \in \mathcal{C}^{1,2} \qquad \textit{and} \qquad \overline{f}_\alpha \in \mathcal{H}_\alpha, \tag{8.12}$$

and for $\alpha \in \overline{\mathcal{A}}_\gamma \bigcup \overline{\mathcal{B}}(\overline{\mathcal{A}}_\gamma)$, $t \in [0,T]$ and $x \in \mathbb{R}^d$, we require

$$\left|\overline{f}_\alpha(t,x)\right|^2 \leq K_2\,(1 + |x|^2). \tag{8.13}$$

Then the estimate

$$\sqrt{E(\sup_{0 \leq s \leq T} |X_s - \overline{Y}^\Delta_s|^2\, |\mathcal{A}_0)} \leq K_3\, \Delta^\gamma \tag{8.14}$$

holds, where the constant K_3 does not depend on Δ.

Proof: Since the jump adapted time discretisation contains all jump points of the solution X of the SDE (2.2), with the aid of the Wagner-Platen expansion for diffusion processes we can write

$$\begin{aligned} X_t &= X_0 + \sum_{\alpha \in \overline{\mathcal{A}}_\gamma \setminus \{v\}} \Big\{ \sum_{n=0}^{n_t - 1} I_\alpha[\overline{f}_\alpha(t_n, X_{t_n})]_{t_n, t_{n+1}} + I_\alpha[\overline{f}_\alpha(t_{n_t}, X_{t_{n_t}})]_{t_{n_t}, t} \Big\} \\ &+ \sum_{\alpha \in \overline{\mathcal{B}}(\overline{\mathcal{A}}_\gamma)} \Big\{ \sum_{n=0}^{n_t - 1} I_\alpha[\overline{f}_\alpha(\cdot, X_\cdot)]_{t_n, t_{n+1}} + I_\alpha[\overline{f}_\alpha(\cdot, X_\cdot)]_{t_{n_t}, t} \Big\} \\ &+ \int_0^t \int_{\mathcal{E}} c(t_{n_s}-, X_{t_{n_s}-}, v)\, p_\phi(dv \times ds), \end{aligned} \tag{8.15}$$

for $t \in [0, T]$.

The jump adapted order γ strong Taylor scheme can be written as

$$\begin{aligned} \overline{Y}_t &= \overline{Y}_0 + \sum_{\alpha \in \overline{\mathcal{A}}_\gamma \setminus \{v\}} \Big\{ \sum_{n=0}^{n_t - 1} I_\alpha[\overline{f}_\alpha(t_n, \overline{Y}_{t_n})]_{t_n, t_{n+1}} + I_\alpha[\overline{f}_\alpha(t_{n_t}, \overline{Y}_{t_{n_t}})]_{t_{n_t}, t} \Big\} \\ &+ \int_0^t \int_{\mathcal{E}} c(t_{n_s}-, \overline{Y}_{t_{n_s}-}, v)\, p_\phi(dv \times ds), \end{aligned} \tag{8.16}$$

for every $t \in [0, T]$.

From the estimate of Theorem 6.4 we have

$$E\left(\sup_{0 \le s \le T} |X_s|^2 \Big| \mathcal{A}_0 \right) \le C\Big(1 + E\left(|X_0|^2\right)\Big). \tag{8.17}$$

Moreover, with similar steps as the ones used in the first part of the proof of Theorem 6.1, we can show the following estimate

$$E\left(\sup_{0 \le s \le T} |\overline{Y}_s^\Delta|^2 \Big| \mathcal{A}_0 \right) \le C\Big(1 + E\Big(|\overline{Y}_0^\Delta|^2\Big)\Big). \tag{8.18}$$

The mean square error is given by

$$\begin{aligned}
Z(t) &:= E\left(\sup_{0\le s\le t} |X_s - \overline{Y}^{\Delta}_s|^2 \middle| \mathcal{A}_0\right) \\
&= E\Bigg(\sup_{0\le s\le t} \Bigg| X_0 - \overline{Y}^{\Delta}_0 \\
&\quad + \sum_{\alpha\in\overline{\mathcal{A}}_\gamma\setminus\{v\}} \Bigg\{ \sum_{n=0}^{n_s-1} I_\alpha[\overline{f}_\alpha(t_n, X_{t_n}) - \overline{f}_\alpha(t_n, \overline{Y}^{\Delta}_{t_n})]_{t_n,t_{n+1}} \\
&\quad + I_\alpha[\overline{f}_\alpha(t_{n_s}, X_{t_{n_s}}) - \overline{f}_\alpha(t_{n_s}, \overline{Y}^{\Delta}_{t_{n_s}})]_{t_{n_s},s} \Bigg\} \\
&\quad + \sum_{\alpha\in\overline{\mathcal{B}}(\overline{\mathcal{A}}_\gamma)} \Bigg\{ \sum_{n=0}^{n_s-1} I_\alpha[\overline{f}_\alpha(\cdot, X_\cdot)]_{t_n,t_{n+1}} + I_\alpha[\overline{f}_\alpha(\cdot, X_\cdot)]_{t_{n_s},s} \Bigg\} \\
&\quad + \int_0^s \int_{\mathcal{E}} \Big\{ c(t_{n_u}-, X_{t_{n_u}-}, v) - c(t_{n_u}-, \overline{Y}_{t_{n_u}-}, v) \Big\} p_\phi(dv\times du) \Bigg|^2 \Bigg| \mathcal{A}_0 \Bigg) \\
&\le C_3 \Bigg\{ \left| X_0 - \overline{Y}^{\Delta}_0 \right|^2 + \sum_{\alpha\in\overline{\mathcal{A}}_\gamma\setminus\{v\}} S^\alpha_t + \sum_{\alpha\in\overline{\mathcal{B}}(\overline{\mathcal{A}}_\gamma)} U^\alpha_t + P_t \Bigg\} \qquad (8.19)
\end{aligned}$$

for all $t \in [0,T]$, where S^α_t, U^α_t and P_t are defined by

$$\begin{aligned}
S^\alpha_t &:= E\Bigg(\sup_{0\le s\le t} \Bigg| \sum_{n=0}^{n_s-1} I_\alpha\left[\overline{f}_\alpha(t_n, X_{t_n}) - \overline{f}_\alpha(t_n, \overline{Y}^{\Delta}_{t_n})\right]_{t_n,t_{n+1}} \\
&\quad + I_\alpha\left[\overline{f}_\alpha(t_{n_s}, X_{t_{n_s}}) - \overline{f}_\alpha(t_{n_s}, \overline{Y}^{\Delta}_{t_{n_s}})\right]_{t_{n_s},s} \Bigg|^2 \Bigg| \mathcal{A}_0 \Bigg), \qquad (8.20)
\end{aligned}$$

$$U^\alpha_t := E\left(\sup_{0\le s\le t} \left| \sum_{n=0}^{n_s-1} I_\alpha[f_\alpha(\cdot, X_\cdot)]_{t_n,t_{n+1}} + I_\alpha[f_\alpha(\cdot, X_\cdot)]_{t_{n_s},s} \right|^2 \middle| \mathcal{A}_0 \right), \qquad (8.21)$$

and

$$P_t := E\left(\sup_{0\le s\le t} \left| \int_0^s \int_{\mathcal{E}} \Big\{ c(t_{n_u}-, X_{t_{n_u}-}, v) - c(t_{n_u}-, \overline{Y}_{t_{n_u}-}, v) \Big\} p_\phi(dv\times du) \right|^2 \middle| \mathcal{A}_0 \right). \qquad (8.22)$$

Therefore, the terms S^α_t and U^α_t can be estimated as in the proof of Theorem 6.1, while for P_t, applying Jensen's and Doob's inequalities, Itô's isometry for jump processes, the Chauchy-Schwarz inequality and the Lipschitz condition (2.3), we obtain

$$\begin{aligned}
P_t &= E\left(\sup_{0\le s\le t}\left|\int_0^s\int_{\mathcal{E}}\left\{c(t_{n_u}-,X_{t_{n_u}-},v)-c(t_{n_u}-,\overline{Y}_{t_{n_u}-},v)\right\}\widetilde{p}_\phi(dv\times du)\right.\right.\\
&\quad\left.\left.+\int_0^s\int_{\mathcal{E}}\left\{c(t_{n_u},X_{t_{n_u}-},v)-c(t_{n_u},\overline{Y}_{t_{n_u}-},v)\right\}\phi(dv)du\right|^2\middle|\mathcal{A}_0\right)\\
&\le 8\,E\left(\left|\int_0^t\int_{\mathcal{E}}\left\{c(t_{n_u}-,X_{t_{n_u}-},v)-c(t_{n_u}-,\overline{Y}_{t_{n_u}-},v)\right\}\widetilde{p}_\phi(dv\times du)\right|^2\middle|\mathcal{A}_0\right)\\
&\quad+2\,E\left(\sup_{0\le s\le t}\left|\int_0^s\int_{\mathcal{E}}\left\{c(t_{n_u},X_{t_{n_u}-},v)-c(t_{n_u},\overline{Y}_{t_{n_u}-},v)\right\}\phi(dv)du\right|^2\middle|\mathcal{A}_0\right)\\
&\le 8\,E\left(\int_0^t\int_{\mathcal{E}}\left|c(t_{n_u},X_{t_{n_u}-},v)-c(t_{n_u},\overline{Y}_{t_{n_u}-},v)\right|^2\phi(dv)\,du\middle|\mathcal{A}_0\right)\\
&\quad+2\,\lambda\,t\,E\left(\int_0^t\int_{\mathcal{E}}\left|c(t_{n_u},X_{t_{n_u}-},v)-c(t_{n_u},\overline{Y}_{t_{n_u}-},v)\right|^2\phi(dv)\,du\middle|\mathcal{A}_0\right)\\
&\le K\,E\left(\int_0^t\left|X_{t_{n_u}-}-\overline{Y}_{t_{n_u}-}\right|^2du\middle|\mathcal{A}_0\right)\\
&\le C\int_0^t Z(u)du. \qquad (8.23)
\end{aligned}$$

Therefore, since by (8.17) and (8.18) *Z(t)* is bounded, applying the Gronwall inequality to (8.19) we can complete the proof of Theorem 8.1. □

Acknowledgement

The authors wish to express their thanks to Christina Nikitopoulos-Sklibosios for valuable comments and suggestions in preparing this paper.

References

Bates, D. (1996). Jumps and stochastic volatility: exchange rate processes implicit in deutschemark options. *Rev. Financial Studies* **9**(1), 96–107.

Björk, T., Y. Kabanov, & W. Runggaldier (1997). Bond market structure in the presence of marked point processes. *Math. Finance* **7**(2), 211–239.

Chiarella, C. & C. Nikitopoulos-Sklibosios (2004). A class of jump-diffusion bond pricing models within the Heath-Jarrow-Morton framework. *Asia-Pacific Financial Markets*. Forthcoming.

Duffie, D., J. Pan, & K. Singleton (2000). Transform analysis and option pricing for affine jump diffusions. *Econometrica* **68**, 1343–1376.

Elliott, R. J. (1982). *Stochastic Calculus and Applications*. Springer.

Elliott, R. J., L. Aggoun, & J. B. Moore (1995). *Hidden Markov Models: Estimation and Control*, Volume 29 of *Appl. Math.* Springer.

Gardoǹ, A. (2004). The order of approximations for solutions of Itô-type stochastic differential equations with jumps. *Stochastic Analysis and Applications* **22**(3), 679–699.

Higham, D. & P. Kloeden (2004). Convergence and stability of implicit methods for jump-diffusion systems. *International Journal of Numerical Analysis & Modeling*. (to appear).

Hofmann, N. & E. Platen (1996). Stability of superimplicit numerical methods for stochastic differential equations. *Fields Inst. Commun.* **9**, 93–104.

Ikeda, N. & S. Watanabe (1989). *Stochastic Differential Equations and Diffusion Processes* (2nd ed.). North-Holland. (first edition (1981)).

Jacod, J. & P. Protter (1998). Asymptotic error distribution for the Euler method for stochastic differential equations. *Ann. Probab.* **26**(1), 267–307.

Jorion, P. (1988). On jump processes in the foreign exchange and stock markets. *Rev. Financial Studies* **1**(4), 427–445.

Kloeden, P. E. & E. Platen (1999). *Numerical Solution of Stochastic Differential Equations*, Volume 23 of *Appl. Math.* Springer. Third corrected printing.

Kou, S. G. (2002). A jump diffusion model for option pricing. *Management Science* **48**, 1086–1101.

Maghsoodi, Y. (1996). Mean-square efficient numerical solution of jump-diffusion stochastic differential equations. *SANKHYA A* **58**(1), 25–47.

Maghsoodi, Y. (1998). Exact solutions and doubly efficient approximations of jump-diffusion Itô equations. *Stochastic Anal. Appl.* **16**(6), 1049–1072.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *J. Financial Economics* **2**, 125–144.

Milstein, G. N., E. Platen, & H. Schurz (1998). Balanced implicit methods for stiff stochastic systems. *SIAM J. Numer. Anal.* **35**(3), 1010–1019.

Pan, D. (2002). The jump risk-premia implicit in options: an integrated time-series study. *J. Financial Economics* **63**, 3–50.

Platen, E. (1982a). An approximation method for a class of Itô processes with jump component. *Liet. Mat. Rink.* **22**(2), 124–136.

Platen, E. (1982b). A generalized Taylor formula for solutions of stochastic differential equations. *SANKHYA A* **44**(2), 163–172.

Platen, E. (1999). An introduction to numerical methods for stochastic differential equations. *Acta Numer.* **8**, 197–246.

Protter, P. (2003). *Stochastic Integration and Differential Equations*, Volume 21 of *Appl. Math.* Springer. Second edition.

Schönbucher, P. J. (2003). *Credit Derivatives Pricing Models: Models, Pricing and Implementation*. Wiley, Chichester.