Numerical Solution of Stochastic Differential Equations with Jumps in Finance

A Thesis Submitted for the Degree of
Doctor of Philosophy

by

Nicola Bruti-Liberati
B.Sc.(Bocconi University, Milan)
M.A. (Columbia University, New York)
Nicola.BrutiLiberati@student.uts.edu.au

in

School of Finance and Economics
University of Technology, Sydney
PO Box 123 Broadway
NSW 2007, Australia

July 31, 2007

Certificate

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirement for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

Signed Production Note:
Signature removed prior to publication.

Date 3/8/2007

Acknowledgements

I would like to thank my supervisor, Professor Eckhard Platen, for his guidance throughout my PhD studies. I am very grateful to him for sharing his deep knowledge in stochastic numerical methods and finance, for teaching me the principles of research and academic status, and for giving me the opportunity of several academic visits abroad. I would also like to thank my co-supervisor, Associate Professor Erik Schlögl, for his additional supervision.

I have also benefited from discussions with Giovanni Barone-Adesi, Lorenzo Casavecchia, Carl Chiarella, Ernst Eberlein, Paul Embrechts, Damir Filipovic, David Heath, Christina Nikitopoulos-Sklibosios, Massimo Piccardi, Andrea Roncoroni and Wolfgang Runggaldier. A special thank goes to Hardy Hulley, who has given me many useful advices on several technical aspects of this thesis.

I would like to thank Sandro Salsa and Emilio Barucci for their hospitality during several visits at Politecnico di Milano.

The financial assistance received with the scholarships of the Quantitative Finance Research Centre and of Bocconi University is greatly appreciated. Additionally, I acknowledge the financial support that I received from the School of Finance and Economics, from the Financial Integrity Research Network, from the Faculty of Business and from the Vice-Chancellor to attend various conferences.

Finally, I would like to thank my family for their continual support and visits from another continent, and Ellie for being always supportive and for her love.

Contents

Basic Notation

$x^\top$	transpose of a vector or matrix x;
$x = (x^1, \dots, x^d)^\top$	column vector $x \in \mathbb{R}^d$ with ith component x^i;
$\lvert x \rvert$	absolute value of x or Euclidean norm;
$A = [a^{i,j}]_{i,j=1}^{k,d}$	$(k \times d)$-matrix A with ijth component $a^{i,j}$;
$\mathbb{N} = \{1, 2, \dots\}$	set of natural numbers;
$\mathbb{R} = (-\infty, \infty)$	set of real numbers;
$\mathbb{R}^+ = [0, \infty)$	set of nonnegative real numbers;
$\mathbb{R}^d$	d-dimensional Euclidean space;
(a, b)	open interval $a < x < b$ in $\mathbb{R}$;
$[a, b]$	closed interval $a \leq x \leq b$ in $\mathbb{R}$;
Ω	sample space;
$\emptyset$	empty set;
Δ	time step size of a time discretization;
$n! = 1 \cdot 2 \cdot \ldots \cdot n$	factorial of n;
$\binom{i}{l} = \frac{i!}{l!(i-l)!}$	combinatorial coefficient;
$[a]$	largest integer not exceeding $a \in \mathbb{R}$;
$(\bmod c)$	modulo c;
$(a)^+ = \max(a, 0)$	maximum of a and 0;

$\ln(a)$	natural logarithm of a;
i.i.d.	independent identically distributed;
a.s.	almost surely;
$f : Q_1 \to Q_2$	function f from Q_1 into Q_2;
f'	first derivative of $f : \mathbb{R} \to \mathbb{R}$;
f''	second derivative of $f : \mathbb{R} \to \mathbb{R}$;
$\frac{\partial u}{\partial x^i}$	ith partial derivative of $u : \mathbb{R}^d \to \mathbb{R}$;
$\partial^k_{x^i} u$ or $\left(\frac{\partial}{\partial x^i}\right)^k u$	kth order partial derivative of u with respect to x^i;
$C^k(\mathbb{R}^d, \mathbb{R})$	set of k times continuously differentiable functions;
$C^k_P(\mathbb{R}^d, \mathbb{R})$	set of k times continuously differentiable functions which, together with their partial derivatives of order up to k, have polynomial growth;
$\mathbf{1}_A$	indicator function for event A to be true;
$\mathcal{N}(\cdot)$	Gaussian distribution function;
$\mathcal{A}$	collection of events, sigma-algebra;
$\underline{\mathcal{A}}$	filtration;
$E(X)$	expectation of X;
$E(X \mid \mathcal{A})$	conditional expectation of X under $\mathcal{A}$;
$P(A)$	probability of A;
$\mathcal{B}(U)$	smallest sigma-algebra on U;
SDE	stochastic differential equation;

Letters such as K, $\widetilde{K}$, C, $\widetilde{C}$, $\ldots$ represent finite positive real constants that can vary from line to line. All these constants are assumed to be independent of the time step size Δ. The remaining notation is either standard or will be introduced when used.

Abstract

This thesis concerns the design and analysis of new discrete time approximations for stochastic differential equations (SDEs) driven by Wiener processes and Poisson random measures. In financial modelling, SDEs with jumps are often used to describe the dynamics of state variables such as credit ratings, stock indices, interest rates, exchange rates and electricity prices. The jump component can capture event-driven uncertainties, such as corporate defaults, operational failures or central bank announcements. The thesis proposes new, efficient, and numerically stable strong and weak approximations. Strong approximations provide efficient tools for problems such as filtering, scenario analysis and hedge simulation, while weak approximations are useful for handling problems such as derivative pricing, the evaluation of moments, and the computation of risk measures and expected utilities. The discrete time approximations proposed are divided into regular and jump-adapted schemes. Regular schemes employ time discretizations that do not include the jump times of the Poisson measure. Jump-adapted time discretizations, on the other hand, include these jump times.

The first part of the thesis introduces stochastic expansions for jump diffusions and proves new, powerful lemmas providing moment estimates of multiple stochastic integrals. The second part presents strong approximations with a new strong convergence theorem for higher order general approximations. Innovative strong derivative-free and predictor-corrector schemes are derived. Furthermore, the strong convergence of higher order schemes for pure jump SDEs is established under conditions weaker than those required for jump diffusions. The final part of the thesis presents a weak convergence theorem for jump-adapted higher order general approximations. These approximations include new derivative-free, predictor-corrector, and simplified schemes. Finally, highly efficient implementations of simplified weak schemes based on random bit generators and hardware accelerators are developed and tested.

Chapter 1

Introduction

1.1 Brief Survey of Results

Before we discuss the motivation of this research and provide an overview of the existing literature, we give in this section a brief survey of the results presented in this thesis. The topic of this thesis is the numerical solution of stochastic differential equations (SDEs) with jumps. The thesis is divided into three parts. The first part, covering Chapters 2 and 3, discusses the class of SDEs with jumps and introduces the Wagner-Platen expansion along with some new results on estimates of multiple stochastic integrals. The second part, comprising Chapters 4, 5, 6, 7 and 8, considers strong approximations of jump-diffusion and pure jump SDEs. Finally, the third part, which is composed of Chapters 9, 10, 11 and 12, introduces weak approximations and discusses efficient implementations of weak schemes.

Some of the new results of this thesis have already appeared or have been accepted in eight refereed publications. Below, these papers are put into the context of the thesis. The rigorous derivation and description of the various new results did not allow any significantly shortened presentation of the thesis. However, a substantial effort has been made to present these results systematically, which should allow an efficient reading.

A main result of the first part of the thesis is the derivation of new moment estimates (Lemmas 3.6.1 and 3.6.2) of multiple stochastic integrals with respect to Wiener processes, Poisson measures and compensated Poisson measures. These estimates constitute the core of the proof of new strong convergence theorems. We also present results (Lemmas 3.6.3 and 3.6.4) that will be used in the proofs of new weak convergence theorems. These moment estimates build on some results presented in Liu & Li (2000) and Li & Liu (2000), who consider only multiple stochastic integrals with respect to the Poisson random measure. Here we also consider multiple stochastic integrals with respect to the compensated Poisson measure, which are necessary to prove the convergence of certain strong and weak approximations.

There are three main results of the second part of the thesis. First, the order of convergence of regular strong Taylor approximations (Theorem 4.5.1) is established. A detailed proof is provided, which extends the results presented in Platen (1982a) and Gardoǹ (2004). The new result considers the more general case of a driving Poisson measure and covers also the strong convergence of regular approximations based on a compensated Poisson measure.

Second, the strong convergence of the fairly general class of regular strong Itô approximations (Theorem 5.5.1) is derived. The corresponding strong convergence theorem allows the construction of new derivative-free, drift-implicit and predictor-corrector schemes (see Bruti-Liberati, Nikitopoulos-Sklibosios & Platen (2006) and Bruti-Liberati & Platen (2007a)). The proposed innovative strong predictor-corrector schemes appear to be new, even for the case of pure diffusion SDEs. These new schemes are of particular importance, since they combine the efficiency of explicit schemes and the enhanced numerical stability properties inherited from corresponding implicit schemes. Most importantly, such strong predictor-corrector schemes introduce quasi-implicitness also in the diffusion terms of the algorithm. In the context of strong approximations, this is a particularly difficult task that, so far, only balanced implicit schemes were able to achieve, see Milstein, Platen & Schurz (1998), Kahl & Schurz (2006) and Alcock & Burrage (2006). The new predictor-corrector methods avoid the key drawback of balanced implicit methods, which is the requirement for solving an algebraic equation at each time step.

Third, the strong order of convergence of strong Taylor approximations for pure jump SDEs (Theorem 8.8.4) has been established under weaker conditions than usually required (see Bruti-Liberati & Platen (2007c)). We show that the differentiability conditions on the jump coefficient, typical of strong convergence theorems for jump diffusions, are here not needed. In the special case of pure jump dynamics, Lipschitz and linear growth conditions on the jump coefficient are sufficient for strong schemes with any order of strong convergence.

The last part of the thesis contains two main results. First, the order of weak convergence of general jump-adapted weak Taylor approximations (Theorem 10.7.1) has been established. The corresponding new theorem extends the results in Mikulevicius & Platen (1988). It covers the weak convergence of new jump-adapted weak derivative-free and predictor-corrector schemes. Furthermore, new jump-adapted simplified schemes, which use multi-point distributed random variables between jump times, have been proposed.

Second, new efficient implementations of simplified weak schemes, based on multi-point distributed random variables, have been proposed (Chapter 12). These results are discussed in the case of pure diffusion SDEs and can be readily applied to the diffusive part of a jump-adapted scheme when approximating SDEs with jumps. The main idea underlying these implementations (see Bruti-Liberati & Platen (2004)) is the systematic use of efficient random bit generators for the multi-point distributed random numbers needed in simplified weak schemes. Software implementations are tested and the obtained speedups of up to 29 times on the entire simulation, when compared to an implementation based on Gaussian random variables, are reported. Finally, new hardware accelerators, based on field programmable gate arrays, are proposed, which provide further speedups of up to three times (see Martini, Piccardi, Bruti-Liberati & Platen (2005), Bruti-Liberati, Platen, Martini & Piccardi (2005), and Bruti-Liberati, Martini, Piccardi & Platen (2007)).

1.2 Motivation

Key features of advanced financial models are event-driven uncertainties such as corporate defaults, operational failures or central bank announcements. By analyzing time series properties of historical prices and other financial quantities, many authors have argued for the presence of jumps, see Jorion (1988) for foreign exchange and stock markets, and Johannes (2004) for short-term interest rates. Jumps are also used to generate the short-term smile effect observed in implied volatilities of option prices, see Cont & Tankov (2004). Furthermore, jumps are needed to properly model credit events like defaults and credit rating changes, see for instance Jarrow, Lando & Turnbull (1997). The short rate, typically set by a central bank, jumps up or down, usually by a quarter of a percent, after some random waiting times, see Babbs & Webber (1995). Therefore, models for the dynamics of financial quantities specified by SDEs with jumps have become increasingly popular. Models of this kind can be found, for instance, in Merton (1976), Björk, Kabanov & Runggaldier (1997), Duffie, Pan & Singleton (2000), Kou (2002), Schönbucher (2003), Glasserman & Kou (2003), Cont & Tankov (2004) and Geman

& Roncoroni (2006). The areas of application of SDEs with jumps go far beyond finance. Other areas of application include economics, insurance, population dynamics, epidemiology, structural mechanics, physics, chemistry and biotechnology. In chemistry, for instance, the reactions of single molecules or coupled reactions yield stochastic models with jumps, see Turner, Schnell & Burrage (2004).

Since only a small class of jump-diffusion SDEs admits explicit solutions, it is important to construct discrete time approximations. The topic of this thesis is the numerical solution of SDEs with jumps via simulation. We consider pathwise simulation, for which strong schemes are used, and Monte Carlo simulation, for which weak schemes are employed. Note that there exist alternative methods to Monte Carlo simulation that we do not consider in this thesis. These include Markov chain, tree, and partial differential equation methods. The class of SDEs here considered is that driven by Wiener processes and finite intensity Poisson random measures. Some authors consider the smaller class of SDEs driven by Wiener processes and homogeneous Poisson processes, while other authors analyze the larger class of SDEs driven by fairly general semimartingales. The class of SDEs driven by Wiener processes and Poisson random measures with finite intensity appears to be just large enough for a realistic modelling of the underlying dynamics in finance. Here continuous trading noise and few single events are typical sources of uncertainty. Furthermore, stochastic jump sizes and stochastic intensities, can be conveniently covered by a Poisson random measure. At present the development of a rich theory on simulation methods for SDEs with jumps, similar to that established for pure diffusion SDEs in Kloeden & Platen (1999), is under way. This thesis aims to contribute to some aspects of this new theory.

We consider discrete time approximations of solutions of SDEs constructed on time discretizations $(t)_\Delta$, with maximum step size $\Delta \in (0, \Delta_0)$, with $\Delta_0 \in (0, 1)$. We call a time discretization *regular* if the jump times, generated by the Poisson measure, are not discretization times. On the other hand, if the jump times are included in the time discretization, then a *jump-adapted* time discretization is obtained. Accordingly, discrete time approximations constructed on regular time discretizations are called *regular schemes*, while approximations constructed on jump-adapted time discretizations are called *jump-adapted schemes*.

Discrete time approximations can be divided into two major classes: *strong* approximations and *weak* approximations, see Kloeden & Platen (1999). We say that a discrete time approximation Y^Δ, constructed on a time discretization $(t)_\Delta$, with maximum step size $\Delta > 0$, converges with *strong* order γ at time T to the solution X of a given SDE, if there exists a positive constant C, independent of Δ, and a finite number $\Delta_0 \in (0, 1)$, such that

$$\varepsilon_s(\Delta) := \sqrt{E(|X_T - Y_T^\Delta|^2)} \leq C\Delta^\gamma, \tag{1.2.1}$$

for all $\Delta \in (0, \Delta_0)$. From the definition of the strong error $\varepsilon_s(\Delta)$, in (1.2.1), one notices that strong schemes provide pathwise approximations of the original solution X of the given SDE. Therefore, these methods are suitable for problems such as filtering, scenario and hedge simulation, and the testing of statistical and other quantitative methods.

We say that a discrete time approximation Y^Δ converges *weakly* with order β to X at time T, if for each $g \in \mathcal{C}_P^{2(\beta+1)}(\mathbb{R}^d, \mathbb{R})$ there exists a positive constant C, independent of Δ, and a finite number, $\Delta_0 \in (0, 1)$, such that

$$\varepsilon_w(\Delta) := |E(g(X_T)) - E(g(Y_T^\Delta))| \leq C\Delta^\beta, \tag{1.2.2}$$

for each $\Delta \in (0, \Delta_0)$. Here $\mathcal{C}_P^{2(\beta+1)}(\mathbb{R}^d, \mathbb{R})$ denotes the set of $2(\beta + 1)$ continuously differentiable functions which, together with their partial derivatives of order up to $2(\beta + 1)$, have polynomial growth. This means that for $g \in \mathcal{C}_P^{2(\beta+1)}(\mathbb{R}^d, \mathbb{R})$ there exist constants $K > 0$ and $r \in \mathbb{N}$, depending on g, such that

$$|\partial_y^j g(y)| \leq K(1 + |y|^{2r}), \tag{1.2.3}$$

for all $y \in \mathbb{R}^d$ and any partial derivative $\partial_y^j g(y)$ of order $j \leq 2(\beta+1)$. Weak schemes provide approximations of the probability measure generated by the solution of a given SDE. These schemes are appropriate for problems such as derivative pricing, the evaluation of moments and the computation of risk measures and expected utilities.

Let us briefly discuss some relationships between strong and weak approximations. The following remark, easily obtained by Jensen's inequality, see (1.2.10) in Appendix A, provides some insights.

Remark 1.2.1 *Let Y^Δ be a discrete time approximation, constructed on a time discretization $(t)_\Delta$, with strong order of convergence γ, see* (1.2.1). *Consider a* *function $g : \mathbb{R}^d \longrightarrow \mathbb{R}$ satisfying the Lipschitz condition*

$$|g(x) - g(y)| \le K|x - y|, \tag{1.2.4}$$

for every $x, y \in \mathbb{R}^d$, where K is a positive constant. Then there exists a positive constant C, independent of Δ, and a finite number, $\Delta_0 \in (0,1)$, such that

$$|E(g(X_T)) - E(g(Y_T^\Delta))| \le C\Delta^\gamma, \tag{1.2.5}$$

for each $\Delta \in (0, \Delta_0)$.

Since the set of Lipschitz continuous functions includes the set $C_P^{2(\beta+1)}(\mathbb{R}^d, \mathbb{R})$, Remark 1.2.1 implies that if a discrete time approximation Y^Δ achieves an order γ of strong convergence, then it also achieves *at least* an order $\beta = \gamma$ of weak convergence. We emphasize that the weak order obtained in Remark 1.2.1 is usually not sharp and, thus, the order of weak convergence could actually be higher than that of strong convergence. For instance, it is well-known and we will later show that the Euler scheme typically achieves strong order $\gamma = 0.5$ and weak order $\beta = 1.0$.

In light of Remark 1.2.1, one could think that the design of strong approximations is sufficient for any type of application, since these approximations can be also applied to weak problems, achieving at least the same order of convergence. This is of course true, but the resulting schemes might be not optimal in terms of computational efficiency. Let us consider as an example, the strong Milstein scheme for pure diffusion SDEs, see Milstein (1974). By adding the double Wiener integrals to the Euler scheme one obtains the Milstein scheme, thus enhancing the order of strong convergence from $\gamma = 0.5$ to $\gamma = 1.0$. Nonetheless, the order of weak convergence of the Milstein scheme equals $\beta = 1.0$, which is not an improvement over the order of weak convergence of the Euler scheme. Therefore, to price a European call option, for example, the Euler scheme is computationally more efficient than the Milstein scheme, since it has fewer terms and the same order of weak convergence. This simple example indicates that to construct efficient higher order weak approximations, one should not take the naive approach of just using higher order strong approximations. Furthermore, as it will be discussed in Chapter 12, when designing weak schemes one has the freedom of using simple multi-point distributed random variables to approximate the underlying multiple stochastic integrals. These multi-point distributed random variables lead to highly efficient implementations of weak schemes.

1.3 Literature Review

An extensive literature on the numerical solution of pure diffusion SDEs has been developed over several decades. We refer the reader to the monographs Kloeden & Platen (1999), Kloeden, Platen & Schurz (2003), Milstein (1995), Milstein & Tretyakov (2004), and to the surveys Platen (1999), Higham (2001), Kloeden (2002) and Burrage, Burrage & Tian (2004). Implementations of these methods with mathematical software, such as MAPLE and MATLAB, have been discussed in Cyganowski, Kloeden & Ombach (2001) and Higham & Kloeden (2002). Also relevant to our work are the monographs by Jäckel (2002), Glasserman (2004), McLeish (2005) and Chan & Wong (2006) on Monte Carlo simulation in finance. Specific algorithms for the approximation of certain pure diffusion SDEs commonly arising in finance are described in Hunter, Jäckel & Joshi (2001), Higham & Mao (2005), Kahl & Jäckel (2005), Lord, Koekkoek & van Dijk (2006), Joshi & Stacey (2007), Broadie & Kaya (2006) and Andersen (2007). More general monographs on computational finance, covering also tree and partial differential equations methods, include Higham (2004) and Seydel (2006).

The literature on discrete time approximations of SDEs with jumps is still quite limited. In the sequel we aim to give an overview on this literature, first with focus on strong approximations and later on weak approximations.

1.3.1 Strong Approximations

The early paper by Wright (1980) considers the Euler scheme and some Runge-Kutta type schemes for the approximation of pure jump SDEs. Some numerical experiments are reported, but no proofs of convergence results are provided. Platen (1982a) described a convergence theorem for regular strong Taylor schemes of any given strong order $\gamma \in \{0.5, 1, 1.5, \ldots\}$. This paper also introduced jump-adapted strong Taylor approximations, which are constructed on time discretizations that include the jump times of the Poisson measure. Maghsoodi (1996) presented an analysis of some regular and jump-adapted strong discrete time approximations with strong order up to $\gamma = 1.5$. Later Maghsoodi (1998) proposed certain regular and jump-adapted schemes that are doubly efficient, meaning that they achieve strong order $\gamma = 1.0$ and weak order $\beta = 2.0$. A sequence of papers by Li & Liu (1997), Li, Wu & Liu (1998) and Liu & Li (1999) discussed the approximation of multiple stochastic integrals arising in higher order regular strong schemes. Cyganowski, Grüne & Kloeden (2002) described MAPLE implementations of the Euler scheme and of the strong schemes presented in Maghsoodi (1996). Gardoň (2004) proved a convergence theorem for regular strong Taylor schemes of any given order $\gamma \in \{0.5, 1, 1.5, \ldots\}$, similar to that presented in Platen (1982a). However, this result is limited to SDEs driven by Wiener processes and homogeneous Poisson processes. Turner, Schnell & Burrage (2004) introduced a class of Poisson Runge-Kutta methods for the approximation of pure jump SDEs. In a series of papers Higham & Kloeden (2005, 2006, 2007) proposed a fairly general class of implicit schemes of strong order $\gamma = 0.5$ for SDEs that are driven by Wiener processes and homogeneous Poisson processes. They also provided a detailed analysis of numerical stability properties. Jimenez & Carbonell (2006) proposed strong approximations based on some local linearization method.

The following papers consider discrete time approximations for SDEs with jumps, constructed such that they satisfy criteria slightly different from that introduced in (1.2.1) for the classification of strong convergence. The work in Maghsoodi & Harris (1987) analyzed the so-called in-probability approximations for jump-diffusion SDEs. The papers by Li (1995) and Li & Liu (2000) analyzed the almost sure convergence of jump-diffusion approximations.

In the case of SDEs driven by fairly general semimartingales, the convergence of the Euler scheme, in terms of strong error or similar criteria for pathwise approximations, has been studied in Protter (1985), Kurtz & Protter (1991), Kohatsu-Higa & Protter (1994), Jacod & Protter (1998), Rubenthaler (2003), Rubenthaler & Wiktorsson (2003) and Jacod (2004).

As discussed in Section 1.1, the results in the second part of this thesis extend the literature of strong approximations for SDEs with jumps. The results in this direction that are already published or accepted for publications shall be briefly mentioned in the following. The article Bruti-Liberati, Nikitopoulos-Sklibosios & Platen (2006) proposes fairly general strong schemes with order of convergence $\gamma = 1.0$, including derivative-free and drift-implicit schemes. In Bruti-Liberati & Platen (2007c) strong schemes for pure jump SDEs are proposed and analyzed. In this paper a strong convergence result is obtained under weaker conditions on the jump coefficient than usually required for jump-diffusion SDEs. Finally, the work Bruti-Liberati & Platen (2007a) provides an introduction and a survey of strong and weak discrete time approximations for SDEs with jumps, with a discussion on applications in finance and economics.

1.3.2 Weak Approximations

In Mikulevicius & Platen (1988) a theorem for the weak convergence of jump-adapted weak Taylor schemes of any weak order $\beta \in \{1.0, 2.0, \ldots\}$ was derived. The paper Liu & Li (2000) analyzed regular weak Taylor schemes of any weak order, which are based on regular time discretizations that do not include jump times. In this paper a weak convergence theorem and the leading coefficients of the global error are derived for the Euler method and the order 2.0 weak Taylor scheme. Extrapolation methods are also presented in this context. In Kubilius & Platen (2002) the weak convergence of the jump-adapted Euler scheme in the case of Hölder continuous coefficients is treated. The paper by Glasserman & Merener (2003b) considered the weak convergence of the jump-adapted Euler and the order 2.0 weak Taylor schemes under weakened conditions on the jump coefficient. This is important for the numerical approximation of jump-diffusion SDEs with state-dependent intensities. Mordecki, Szepessy, Tempone & Zouraris (2006) analyzed the Euler scheme with adaptive time stepping. Broadie & Kaya (2006) developed exact approximations for certain specific pure diffusion and jump-diffusion SDEs arising in financial applications.

The weak convergence of the Euler scheme for the approximation of SDEs driven by fairly general semimartingales is studied in Platen & Rebolledo (1985), Protter & Talay (1997), Hausenblas (2002), and Jacod, Kurtz, Méléard & Protter (2005). We also refer to the monograph by Janicki (1996), which considered the numerical approximation of SDEs driven by α-stable Lévy measures. In this book a corresponding Euler scheme is proposed and its weak convergence in the Skorokhod topology is proved. The numerical solution of stochastic partial differential equations driven by Poisson random measures is considered, for instance, in Hausenblas (2003, 2006).

As we shall see later, when using weak schemes one can approximate the underlying multiple stochastic integrals by simple multi-point distributed random variables. The resulting approximations are called simplified weak schemes. The last chapter of this thesis will consider efficient implementations of simplified weak schemes by using random bit generators and hardware accelerators. The related literature in this direction is quite limited. Random bit generators are at the heart of random number generation, see Knuth (1981). Nonetheless, the use of random bit generators to obtain multi-point distributed random numbers for simplified weak schemes appeared to be new when published in Bruti-Liberati & Platen (2004). We subsequently became aware that Milstein & Tretyakov (2004) had independently suggested a similar method without any actual implementation nor efficiency analysis. Also the idea of using hardware accelerators, based on field programmable gate arrays (FPGAs), for Monte Carlo simulation in finance, first published in Martini, Piccardi, Bruti-Liberati & Platen (2005), appears to be innovative. To our knowledge the only other article on a related topic seems to be that of Zhang, Leong, Ho, Tsoi, Cheung, Lee, Cheung & Luk (2005). Finally, it should be mentioned that random number generators on FPGAs have been proposed in Martin (2002), Tsoi, Leung & Leong (2003), Lee, Luk, Villasenor & Cheung (2004) and Lee, Villasenor, Luk & Leong (2006).

As mentioned in Section 1.1, the third part of this thesis contributes to the literature of weak approximations in several directions. Among the results already published or accepted for publications, we mention the following articles. Bruti-Liberati & Platen (2007a) and Bruti-Liberati & Platen (2007b) consider fairly general weak approximations including new weak predictor-corrector schemes. The article Bruti-Liberati & Platen (2004) proposes efficient software implementations of simplified schemes based on random bit generators. Finally, hardware accelerators for weak simplified schemes are presented in Martini, Piccardi, Bruti-Liberati & Platen (2005), Bruti-Liberati, Platen, Martini & Piccardi (2005), and Bruti-Liberati, Martini, Piccardi & Platen (2007).

Chapter 2

Stochastic Differential Equations with Jumps

This chapter introduces the class of SDEs driven by Wiener processes and finite intensity Poisson random measures. We also discuss financial applications and results on the existence and uniqueness of strong solutions.

2.1 Introduction

Let be given a filtered probability space $(\Omega, \mathcal{A}_T, \underline{\mathcal{A}}, P)$, with $\underline{\mathcal{A}} = (\mathcal{A}_t)_{t\in[0,T]}$, $T \in [0,\infty)$, satisfying the usual conditions, and a mark space $(\mathcal{E}, \mathsf{B}(\mathcal{E}))$, with $\mathcal{E} \subseteq \mathbb{R}\backslash\{0\}$. Note that one could also use a multi-dimensional mark space without any complication. We define on $\mathcal{E} \times [0,T]$ an $\underline{\mathcal{A}}$-adapted Poisson random measure $p_\phi(dv, dt)$, with intensity measure $q_\phi(dv, dt) = \phi(dv)dt$. We assume that the total intensity

$$\lambda = \phi(\mathcal{E}) < \infty \tag{2.1.1}$$

of p_ϕ is finite. Thus, $p_\phi = \{p_\phi(t) = p_\phi(\mathcal{E}, [0,t]), t \in [0,T]\}$ is a stochastic process that counts the number of jumps occurring in the time interval $[0,T]$. The Poisson random measure $p_\phi(dv, dt)$ generates a sequence of pairs $\{(\tau_i, \xi_i), i \in \{1, 2, \ldots, p_\phi(T)\}\}$, where $\{\tau_i \in [0,T], i \in \{1, 2, \ldots, p_\phi(T)\}\}$ is a sequence of increasing nonnegative random variables representing the jump times of a standard Poisson process with intensity λ, and $\{\xi_i \in \mathcal{E}, i \in \{1, 2, \ldots, p_\phi(T)\}\}$ is a sequence of independent, identically distributed random variables. Here ξ_i is distributed according to $\frac{\phi(dv)}{\lambda} = F(dv)$. Therefore, we call $F(\cdot)$ the distribution function of the marks. We can interpret τ_i as the time of the ith event and the mark ξ_i as its amplitude. For a more general presentation of random measures we refer to Elliott (1982).

In this thesis we consider the d-dimensional jump-diffusion SDE

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t + \int_{\mathcal{E}} c(t, X_{t-}, v)\, p_\phi(dv, dt), \tag{2.1.2}$$

for $t \in [0, T]$, with initial value $X_0 \in \mathbb{R}^d$, an $\underline{\mathcal{A}}$-adapted m-dimensional Wiener process $W = \{W_t = (W_t^1, \ldots, W_t^m), t \in [0, T]\}$ and the previously introduced Poisson random measure p_ϕ. Note that in (2.1.2) we denote by X_{t-} the almost sure left-hand limit of X at time t. A solution of an SDE of the type (2.1.2) is called *jump diffusion* or *Itô process*.

The coefficients $a(t, x)$ and $c(t, x, v)$ are d-dimensional vectors of Borel measurable real valued functions on $[0, T] \times \mathbb{R}^d$ and on $[0, T] \times \mathbb{R}^d \times \mathcal{E}$, respectively. Additionally, $b(t, x)$ is a $d \times m$-matrix of Borel measurable 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 ith component. Similarly with $b^{i,j}$ we will denote the component of the ith row and jth column of a given matrix b.

The SDE (2.1.2) can be also written in integral form as

$$X_t = X_0 + \int_0^t a(s, X_s)ds + \int_0^t b(s, X_s)dW_s + \sum_{i=1}^{p_\phi(t)} c(\tau_i, X_{\tau_i-}, \xi_i), \tag{2.1.3}$$

where $\{(\tau_i, \xi_i), i \in \{1, 2 \ldots, p_\phi(t)\}\}$ is the above described sequence of pairs of jump times and corresponding marks generated by the Poisson random measure p_ϕ. Note that in this thesis we adopt the convention that the summation $\sum_{i=j_1}^{j_2} c_i$ yields zero for all possible summands c_i if $j_1 > j_2$. Therefore, if there are no jumps up to time t, which means that $p_\phi(t) = 0$, the last term in (2.1.3) vanishes.

In the SDE (2.1.2) we have defined the impact of a jump via an Itô stochastic integral with respect to a Poisson random measure as

$$\int_0^t \int_{\mathcal{E}} c(s, X_{s-}, v) p_\phi(dv, ds). \tag{2.1.4}$$

This stochastic integral allows us to model a rather general jump behavior. The only restriction we impose on the jump component is the finiteness of the total intensity as requested in condition (2.1.1).

Let us consider the special case of the type of coefficient functions $a(t, x) = \mu x$,

$b(t,x) = \sigma x$ and $c(t,x,v) = x\,(v-1)$. Then the SDE (2.1.2) reduces to

$$dX_t = X_{t-}\left(\mu dt + \sigma dW_t + \int_{\mathcal{E}} (v-1)\, p_\phi(dv, dt)\right), \tag{2.1.5}$$

for $t \in [0,T]$ and admits the explicit solution

$$X_t = X_0\, e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t} \prod_{i=1}^{p_\phi(t)} \xi_i, \tag{2.1.6}$$

where the marks $\xi_i \geq 0$ are distributed according to a given probability measure $F(dv) = \frac{\phi(dv)}{\lambda}$. When we choose a lognormal probability measure, which means that $\zeta_i = \ln(\xi_i)$ is an independent Gaussian random variable, $\zeta_i \sim \mathcal{N}(\varrho, \varsigma)$, with mean ϱ and variance ς, then equation (2.1.6) represents a specification of the jump-diffusion asset price model proposed in Merton (1976), known as *Merton model*. A simple case is obtained when the lognormal random variable ξ_i becomes degenerate, which is for zero variance $\varsigma = 0$, where $\xi_i = e^{\varrho}$. If we assume a log-Laplace distribution, instead of a lognormal one, then we recover the *Kou model* proposed in Kou (2002).

The flexibility in the definition of the jump integral (2.1.4) is illustrated in the following examples.

First, one can easily construct several independent driving Poisson processes by splitting the mark space $\mathcal{E}$ into corresponding disjoint subsets $\mathcal{E}_1, \ldots, \mathcal{E}_r$, for $r \in \mathbb{N}$, and obtaining with $P^j = \{P_t^j := p_\phi(\mathcal{E}_j, [0,t]), t \in [0,T]\}$ the jth Poisson process having intensity $\lambda^j = \phi(\mathcal{E}_j)$.

It is also possible to specify a jump component with time-dependent intensity by choosing

$$c(t,x,v) = \mathbf{1}_{\{v \in [\eta_1(t), \eta_2(t)]\}} f(t,x,v), \tag{2.1.7}$$

where η_1 and η_2 are deterministic functions of time and $\mathbf{1}_{\{\cdot\}}$ is the indicator function defined by

$$\mathbf{1}_{\{v \in A\}} = \begin{cases} 1 & \text{for} \quad v \in A \\ 0 & \text{for} \quad v \notin A \end{cases} \tag{2.1.8}$$

for a given set $A \subseteq \mathcal{E}$. Then we obtain a jump integral of the form

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

Moreover, if we allow the functions η_1 and η_2 to be time and state-dependent, then we obtain a jump integral of the form

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

which allows the modelling of stochastic intensity. Note that in this case the coefficient function c is, in general, discontinuous in the state variable and, therefore, could be not Lipschitz. With the above specifications also advanced credit risk models with multiple obligors and correlated intensities, as discussed in Schönbucher (2003), can be modeled via the SDE (2.1.2). In the fast growing energy markets, quantities such as electricity prices are often described by rather complex jump-diffusion SDEs; see, for instance, Geman & Roncoroni (2006) for a class of jump-diffusion models that capture the typical jump reversion feature of electricity prices.

Other important examples of jump diffusions of the form (2.1.2) arise in the pricing and hedging of complex interest rate derivatives. We refer to Björk, Kabanov & Runggaldier (1997) for the HJM framework with jumps and to Glasserman & Kou (2003) for the LIBOR market model with jumps. For example, let us consider a specific LIBOR market model with jumps proposed in Samuelides & Nahum (2001) for pricing short-term interest rate derivatives. Given a set of equidistant tenor dates $T_1, \ldots, T_{d+1}$, with $T_{i+1} - T_i = \delta$ for $i \in \{1, \ldots, d\}$, the components of the vector $X_t = (X_t^1, \ldots, X_t^d)^\top$ represent discrete compounded forward rates at time t, maturing at tenor dates $T_1, \ldots, T_d$, respectively. In this case the model is driven by one Wiener process, $m = 1$, and two Poisson processes, $r = 2$. The diffusion coefficient is specified as $b(t, x) = \sigma x$, with σ a d-dimensional vector of positive numbers, and the jump coefficient is given by $c(t, x) = \beta x$, where β is a $d \times 2$-matrix with $\beta^{i,1} > 0$ and $\beta^{i,2} < 0$, for $i \in \{1, \ldots, d\}$. In this way the first jump process generates upward jumps, while the second one creates downward jumps. Moreover, the marks are set to $\xi_i = 1$ so that the two driving jump processes are standard Poisson processes. A no-arbitrage restriction on the evolution of forward rates under the T_{d+1}-forward measure, see Björk, Kabanov & Runggaldier (1997) and Glasserman & Kou (2003), imposes a particular non-linear form on the drift coefficient $a(t,x)$. Its ith component is given by

$$
\begin{aligned}
a^i(t,x^1,\dots,x^d) &= -\Big\{ \sum_{j=i+1}^{d} \frac{\delta x^j}{1+\delta x^j}\sigma^j + \lambda^1 \prod_{j=i+1}^{d} \big(1+\beta^{j,1}\frac{\delta x^j}{1+\delta x^j}\big) \\
&\quad + \lambda^2 \prod_{j=i+1}^{d} \big(1+\beta^{j,2}\frac{\delta x^j}{1+\delta x^j}\big)\Big\},
\end{aligned} \tag{2.1.9}
$$

for $i \in \{1,\dots,d\}$, where λ^1 and λ^2 denote the intensities of the two Poisson processes. A complex non-linear drift coefficient, as the one in (2.1.9), is a typical feature of LIBOR market models. This makes the application of numerical techniques essential for the pricing and hedging of complex interest rate derivatives. We refer to Glasserman & Merener (2003b) for numerical approximations of jump-diffusion LIBOR market models.

2.2 Existence and Uniqueness of Strong Solutions

Let us now state a theorem on the existence and uniqueness of strong solutions of SDEs with jumps. This ensures that the objects we aim to approximate are well defined. For the definition of strong solutions of jump-diffusion SDEs, we can refer, for instance, to Ikeda & Watanabe (1989).

We assume that the coefficient functions of the SDE (2.1.2) satisfy the Lipschitz conditions

$$
\begin{aligned}
|a(t,x)-a(t,y)|^2 &\le C_1|x-y|^2, \qquad |b(t,x)-b(t,y)|^2 \le C_2|x-y|^2, \\
&\int_{\mathcal{E}} |c(t,x,v)-c(t,y,v)|^2\phi(dv) \le C_3|x-y|^2,
\end{aligned} \tag{2.2.10}
$$

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

$$
\begin{aligned}
|a(t,x)|^2 \le\ & K_1(1+|x|^2), \qquad |b(t,x)|^2 \le K_2(1+|x|^2), \\
&\int_{\mathcal{E}} |c(t,x,v)|^2\phi(dv) \le K_3(1+|x|^2),
\end{aligned} \tag{2.2.11}
$$

for all $t \in [0,T]$ and $x \in \mathbb{R}^d$. Note that the linear growth conditions can be usually derived by the corresponding Lipschitz conditions.

Moreover, we assume that the initial condition X_0 is $\mathcal{A}_0$-measurable with

$$E(|X_0|^2) < \infty. \tag{2.2.12}$$

Theorem 2.2.1 *Suppose that the coefficient functions $a(\cdot), b(\cdot)$ and $c(\cdot)$ of the SDE* (2.1.2) *satisfy the Lipschitz conditions* (2.2.10) *and the linear growth conditions* (2.2.11). *Then the jump-diffusion SDE* (2.1.2) *admits a unique strong solution. Moreover, the solution X_t of the SDE* (2.1.2) *satisfies the estimate*

$$E\Big(\sup_{0\leq s\leq T} |X_s|^2|\Big) \leq C\left(1 + E(|X_0|^2)\right) \tag{2.2.13}$$

with $T < \infty$, where C is a finite positive constant.

Proofs of Theorem 2.2.1 can be found in Ikeda & Watanabe (1989) or Situ (2005).

For certain applications, the Lipschitz condition (2.2.10) on the jump coefficient c is too restrictive. For instance, for modelling state-dependent intensities, as discussed in Section 2.1, it is convenient to use jump coefficients that are not Lipschitz continuous. Athreya, Kliemann & Koch (1988) provide some results on the existence and uniqueness of the solution of the jump-diffusion SDE (2.1.2) when the jump coefficient c is not Lipschitz continuous. Glasserman & Merener (2003a) analyze the convergence of the jump-adapted Euler scheme, which we will discuss in Chapter 10, in such a case. The Yamada condition, see Ikeda & Watanabe (1989), provides another condition for SDEs that do not satisfy Lipschitz type conditions.

Chapter 3

Stochastic Expansions with Jumps

In this chapter we present the *Wagner-Platen expansion* for solutions of SDEs with jumps. This stochastic expansion generalizes the deterministic Taylor formula and the Wagner-Platen expansion for diffusions to the case of SDEs with jumps. It allows expanding the increments of smooth functions of Itô processes in terms of multiple stochastic integrals. As we will see, it is the key tool for the construction of higher order stochastic numerical methods.

3.1 Introduction

Let us first rewrite the SDE (2.1.2) in a way where the jump part will be expressed as a stochastic integral with respect to the *compensated Poisson measure*

$$\widetilde{p}_{\phi}(dv, dt) := p_{\phi}(dv, dt) - \phi(dv)dt. \tag{3.1.1}$$

By compensating the Poisson measure in the SDE (2.1.2), we obtain

$$dX_t = \widetilde{a}(t, X_t)dt + b(t, X_t)dW_t + \int_{\mathcal{E}} c(t, X_{t-}, v)\widetilde{p}_{\phi}(dv, dt), \tag{3.1.2}$$

where the *compensated drift* coefficient is given by

$$\widetilde{a}(t, x) := a(t, x) + \int_{\mathcal{E}} c(t, x, v)\phi(dv), \tag{3.1.3}$$

for $t \in [0, T]$, with initial value $X_0 \in \mathbb{R}^d$ and mark space $\mathcal{E} \subseteq \mathbb{R}\backslash\{0\}$. Note that by relation (3.1.3), the Cauchy-Schwarz inequality (see (1.2.1) in Appendix A), and conditions (2.2.10)-(2.2.11), the compensated drift coefficient $\widetilde{a}$ satisfies the Lipschitz condition

$$|\widetilde{a}(t, x) - \widetilde{a}(t, y)|^2 \leq K|x - y|^2, \tag{3.1.4}$$

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

$$|\widetilde{a}(t, x)|^2 \leq K(1 + |x|^2), \tag{3.1.5}$$

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

To construct discrete time approximations we will exploit Wagner-Platen expansions of solutions of the SDE (2.1.2). With these formulas we expand smooth functions of solutions of the SDE (2.1.2) in terms of multiple stochastic integrals. Note that, in a similar way as in (3.1.2), a multiple stochastic integral involving jumps can be defined by using as integrator either the Poisson measure p_ϕ or the compensated Poisson measure $\widetilde{p}_\phi$. Therefore, we will derive two different types of stochastic expansions. We will call the former *Wagner-Platen expansions* and the latter *compensated Wagner-Platen expansions*. To express these expansions in a compact form, we will introduce below certain notation.

3.2 Multiple Stochastic Integrals

In this section we introduce a compact notation for multiple stochastic integrals that will appear in corresponding stochastic expansions.

3.2.1 Multi-Indices

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 $l := l(\alpha) \in \mathbb{N}$. Here m represents the number of Wiener processes considered in the SDE (2.1.2). Then for $m \in \mathbb{N}$ the set of all multi-indices α is denoted by

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

where v is the multi-index of length zero. Later, by a component $j \in \{1, \ldots, m\}$ of a multi-index we will denote an integration with respect to the jth Wiener process. A component $j = 0$ will denote an integration with respect to time. Finally, a component $j = -1$ will denote an integration with respect to the compensated Poisson measure $\widetilde{p}_\phi$.

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 $m = 2$,

$$\begin{aligned}
&l((0,-1,1)) = 3 \qquad && l((0,1,-1,0,2)) = 5 \\
&n((0,-1,1)) = 1 \qquad && n((0,1,-1,0,2)) = 2 \\
&s((0,-1,1)) = 1 \qquad && s((0,1,-1,0,2)) = 1 \\
&(0,-1,1)- = (0,-1) \qquad && (0,1,-1,0,2)- = (0,1,-1,0) \\
&-(0,-1,1) = (-1,1) \qquad && -(0,1,-1,0,2) = (1,-1,0,2).
\end{aligned} \tag{3.2.1}$$

Additionally, given two multi-indices $\alpha_1 = (j_1, \ldots, j_k)$ and $\alpha_2 = (i_1, \ldots, i_l)$, we introduce the *concatenation operator* $\star$ on $\mathcal{M}_m$ defined by

$$\alpha_1 \star \alpha_2 = (j_1, \ldots, j_k, i_1, \ldots, i_l). \tag{3.2.2}$$

This operator allows us to combine two multi-indices.

3.2.2 Multiple Integrals

We shall define certain sets of adapted stochastic processes $g = \{g(t), t \in [0,T]\}$ that will be allowed to appear as integrands of multiple stochastic integrals in the stochastic expansions. We define

$$\begin{aligned}
\mathcal{H}_v = &\quad \{g : \sup_{t\in[0,T]} E\left(|g(t,\omega)|\right) < \infty\} \\
\mathcal{H}_{(0)} = &\quad \{g : E\left(\int_0^T |g(s,\omega)| ds\right) < \infty\} \\
\mathcal{H}_{(-1)} = &\quad \{g : E\left(\int_0^T \int_{\mathcal{E}} |g(s,v,\omega)|^2 \phi(dv) ds\right) < \infty\} \\
\mathcal{H}_{(j)} = &\quad \{g : E\left(\int_0^T |g(s,\omega)|^2 ds\right) < \infty\},
\end{aligned} \tag{3.2.3}$$

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

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

$$I_\alpha[g(\cdot)]_{\rho,\tau} := \begin{cases} g(\tau) & \text{when}\, l = 0 \,\text{and}\, \alpha = v \\ \int_\rho^\tau I_{\alpha-}[g(\cdot)]_{\rho,z}\, dz & \text{when}\, l \geq 1 \,\text{and}\, j_l = 0 \\ \int_\rho^\tau I_{\alpha-}[g(\cdot)]_{\rho,z}\, dW_z^{j_l} & \text{when}\, l \geq 1 \,\text{and}\, j_l \in \{1, \ldots, m\} \\ \int_\rho^\tau \int_{\mathcal{E}} I_{\alpha-}[g(\cdot)]_{\rho,z-}\, p_\phi(dv_{s(\alpha)}, dz) & \text{when}\, l \geq 1 \,\text{and}\, j_l = -1, \end{cases} \tag{3.2.4}$$

where $g(\cdot) = g(\cdot, v_1, \ldots, v_{s(\alpha)})$. As previously, by $z-$ we denote the left-hand limit of z. However, $\alpha-$ for a multi-index α has another meaning as described earlier. For simplicity, when it is not strictly necessary, here and in the sequel we may omit the dependence of the integrand process g on one or more of the variables $v_1, \ldots, v_{s(\alpha)}$ that express the dependence on the marks of the Poisson jump measure.

As defined in (3.2.4), 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 jth component of the Wiener process, while the components that equal -1 refer to an integration with respect to the Poisson measure p_ϕ. For instance, for $m = 2$ one has

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

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

and

$$I_{(-1,-1)}[g(\cdot)]_{\rho,\tau} = \int_\rho^\tau \int_{\mathcal{E}} \int_\rho^{z_2-} \int_{\mathcal{E}} g(z_1-, v_1, v_2)\, p_\phi(dv_1, dz_1) p_\phi(dv_2, dz_2).$$

Now for every multi-index $\alpha = (j_1, \ldots, j_l) \in \mathcal{M}_m$ with $l(\alpha) > 1$, we can recursively define the sets $\mathcal{H}_\alpha$ as the sets of adapted stochastic processes $g = \{g(t), t \geq 0\}$

such that the integral process $\{I_{\alpha-}[g(\cdot)]_{\rho,t}, t \in [0,T]\}$ satisfies

$$I_{\alpha-}[g(\cdot)]_{\rho,\cdot} \in \mathcal{H}_{(j_l)}. \tag{3.2.5}$$

As we shall see later, it is also useful to define multiple stochastic integrals where the integrations are defined with respect to the compensated Poisson measure $\widetilde{p}_\phi$ instead of the Poisson measure p_ϕ. Therefore, for a multi-index $\alpha \in \mathcal{M}_m$ and an adapted stochastic process $g \in \mathcal{H}_\alpha$, we define the *compensated multiple stochastic integral* $\widetilde{I}_\alpha[g(\cdot)]_{\rho,\tau}$ recursively by

$$\widetilde{I}_\alpha[g(\cdot)]_{\rho,\tau} := \begin{cases} g(\tau) & \text{when } l = 0 \text{ and } \alpha = v \\ \int_\rho^\tau \widetilde{I}_{\alpha-}[g(\cdot)]_{\rho,z}\, dz & \text{when } l \geq 1 \text{ and } j_l = 0 \\ \int_\rho^\tau \widetilde{I}_{\alpha-}[g(\cdot)]_{\rho,z}\, dW_z^{j_l} & \text{when } l \geq 1 \text{ and } j_l \in \{1,\ldots,m\} \\ \int_\rho^\tau \int_{\mathcal{E}} \widetilde{I}_{\alpha-}[g(\cdot)]_{\rho,z-}\, \widetilde{p}_\phi(dv_{s(\alpha)}, dz) & \text{when } l \geq 1 \text{ and } j_l = -1, \end{cases} \tag{3.2.6}$$

where $g(\cdot) = g(\cdot, v_1, \ldots, v_{s(\alpha)})$. Note that the multiple stochastic integral $I_\alpha[g(\cdot)]_{\rho,\tau}$ defined previously in (3.2.4) involves integrations with respect to the Poisson jump measure p_ϕ, while the compensated multiple stochastic integral $\widetilde{I}_\alpha[g(\cdot)]_{\rho,\tau}$ is defined in terms of integrations with respect to the compensated Poisson measure $\widetilde{p}_\phi$. For instance, for $m = 2$ one obtains

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

$$\widetilde{I}_{(2,0)}[g(\cdot)]_{\rho,\tau} = I_{(2,0)}[g(\cdot)]_{\rho,\tau}$$

and

$$\widetilde{I}_{(-1,-1)}[g(\cdot)]_{\rho,\tau} = \int_\rho^\tau \int_{\mathcal{E}} \int_\rho^{z_2-} \int_{\mathcal{E}} g(z_1-, v_1, v_2)\, \widetilde{p}_\phi(dv_1, dz_1) \widetilde{p}_\phi(dv_2, dz_2).$$

We now illustrate the obvious link between a multiple stochastic integral of the type (3.2.4) with respect to the Poisson jump measure p_ϕ and a compensated multiple stochastic integral of the type (3.2.6) with respect to the compensated Poisson measure $\widetilde{p}_\phi$. This link is governed by the relationship between the Poisson jump measure p_ϕ and its compensated version $\widetilde{p}_\phi$, see (3.1.1).

Remark 3.2.1 *Let $\alpha \in \mathcal{M}_m$, and ρ and τ denote two stopping times with τ being*

$\mathcal{A}_\rho$-measurable and $0 \leq \rho \leq \tau \leq T$ almost surely. Consider an adapted stochastic process $g \in \mathcal{H}_\alpha$. Then

$$\widetilde{I}_\alpha[g(\cdot)]_{\rho,\tau} = I_\alpha[g(\cdot)]_{\rho,\tau} + \sum_{i=1}^{2^{s(\alpha)}-1} H_{\alpha,i}. \tag{3.2.7}$$

The terms $H_{\alpha,i}$ denote multiple stochastic integrals of the adapted process $g(\cdot)$ that use as integrators the time, the Wiener processes, the Poisson jump measure p_ϕ and the intensity measure ϕ. A complete description of these terms could be given by defining recursively a new type of multiple stochastic integral. Then by using the relationship (3.1.1) together with (3.2.4) and (3.2.6), the terms $H_{\alpha,i}$ could be completely characterized. However, since this characterization requires a rather technical notation, for the sake of simplicity we omit this obvious but complex characterization. Instead, we provide the following two illustrative examples.

For $\alpha = (0, -1, 1)$ one has

$$\begin{aligned}
\widetilde{I}_{(0,-1,1)}[g(\cdot)]_{\rho,\tau} &= \int_\rho^\tau \int_\rho^{z_3} \int_{\mathcal{E}} \int_\rho^{z_2-} g(z_1, v_1)\, dz_1\, \widetilde{p}_\phi(dv_1, dz_2)\, dW^1_{z_3} \\
&= \int_\rho^\tau \int_\rho^{z_3} \int_{\mathcal{E}} \int_\rho^{z_2-} g(z_1, v_1)\, dz_1\, p_\phi(dv_1, dz_2)\, dW^1_{z_3} \\
&\quad - \int_\rho^\tau \int_\rho^{z_3} \int_{\mathcal{E}} \int_\rho^{z_2} g(z_1, v_1)\, dz_1\, \phi(dv_1) dz_2\, dW^1_{z_3} \\
&= I_{(0,-1,1)}[g(\cdot)]_{\rho,\tau} + H_{(0,-1,1),1},
\end{aligned}$$

so that

$$H_{(0,-1,1),1} = -\int_\rho^\tau \int_\rho^{z_3} \int_{\mathcal{E}} \int_\rho^{z_2} g(z_1, v_1)\, dz_1\, \phi(dv_1) dz_2\, dW^1_{z_3}.$$

For $\alpha = (-1, -1)$ one obtains

$$\begin{aligned}
\widetilde{I}_{(-1,-1)}[g(\cdot)]_{\rho,\tau} &= \int_\rho^\tau \int_{\mathcal{E}} \int_\rho^{z_2-} \int_{\mathcal{E}} g(z_1-, v_1, v_2)\, \widetilde{p}_\phi(dv_1, dz_1) \widetilde{p}_\phi(dv_2, dz_2) \\
&= \int_\rho^\tau \int_{\mathcal{E}} \int_\rho^{z_2-} \int_{\mathcal{E}} g(z_1-, v_1, v_2)\, p_\phi(dv_1, dz_1) p_\phi(dv_2, dz_2) \\
&\quad - \int_\rho^\tau \int_{\mathcal{E}} \int_\rho^{z_2} \int_{\mathcal{E}} g(z_1-, v_1, v_2)\, p_\phi(dv_1, dz_1) \phi(dv_2) dz_2
\end{aligned}$$

$$
\begin{aligned}
& - \int_\rho^\tau \int_{\mathcal{E}} \int_\rho^{z_2-} \int_{\mathcal{E}} g(z_1, v_1, v_2)\, \phi(dv_1) dz_1 p_\phi(dv_2, dz_2) \\
& + \int_\rho^\tau \int_{\mathcal{E}} \int_\rho^{z_2} \int_{\mathcal{E}} g(z_1, v_1, v_2)\, \phi(dv_1) dz_1 \phi(dv_2) dz_2 \\
= \; & I_{(-1,-1)}[g(\cdot)]_{\rho,\tau} + H_{(-1,-1),1} + H_{(-1,-1),2} + H_{(-1,-1),3},
\end{aligned}
$$

where each term in the sum expresses the corresponding double integral.

Similarly, we can express a multiple stochastic integral of the type (3.2.4), which uses as integrator for the jumps the Poisson jump measure p_ϕ, as a sum of terms involving multiple stochastic integrals that use as integrators for the jumps the compensated Poisson measure $\widetilde{p}_\phi$.

Remark 3.2.2 *Let $\alpha \in \mathcal{M}_m$, and ρ and τ denote two stopping times with τ being $\mathcal{A}_\rho$-measurable and $0 \leq \rho \leq \tau \leq T$ almost surely. Consider an adapted stochastic process $g \in \mathcal{H}_\alpha$. Then*

$$
I_\alpha[g(\cdot)]_{\rho,\tau} = \widetilde{I}_\alpha[g(\cdot)]_{\rho,\tau} + \sum_{i=1}^{2^{s(\alpha)}-1} \widetilde{H}_{\alpha,i}. \tag{3.2.8}
$$

Here the terms $\widetilde{H}_{\alpha,i}$ are multiple stochastic integrals of the process $g(\cdot)$ that use as integrators the time, the Wiener processes, the compensated Poisson measure $\widetilde{p}_\phi$ and the intensity measure ϕ. For instance, for $\alpha = (-1,-1)$ one obtains

$$
\begin{aligned}
I_{(-1,-1)}[g(\cdot)]_{\rho,\tau} &= \int_\rho^\tau \int_{\mathcal{E}} \int_\rho^{z_2-} \int_{\mathcal{E}} g(z_1-, v_1, v_2)\, p_\phi(dv_1, dz_1) p_\phi(dv_2, dz_2) \\
&= \int_\rho^\tau \int_{\mathcal{E}} \int_\rho^{z_2-} \int_{\mathcal{E}} g(z_1-, v_1, v_2)\, \widetilde{p}_\phi(dv_1, dz_1) \widetilde{p}_\phi(dv_2, dz_2) \\
&\quad + \int_\rho^\tau \int_{\mathcal{E}} \int_\rho^{z_2} \int_{\mathcal{E}} g(z_1-, v_1, v_2)\, \widetilde{p}_\phi(dv_1, dz_1) \phi(dv_2) dz_2 \\
&\quad + \int_\rho^\tau \int_{\mathcal{E}} \int_\rho^{z_2-} \int_{\mathcal{E}} g(z_1, v_1, v_2)\, \phi(dv_1) dz_1 \widetilde{p}_\phi(dv_2, dz_2) \\
&\quad + \int_\rho^\tau \int_{\mathcal{E}} \int_\rho^{z_2} \int_{\mathcal{E}} g(z_1, v_1, v_2)\, \phi(dv_1) dz_1 \phi(dv_2) dz_2 \\
&= \widetilde{I}_{(-1,-1)}[g(\cdot)]_{\rho,\tau} + \widetilde{H}_{(-1,-1),1} + \widetilde{H}_{(-1,-1),2} + \widetilde{H}_{(-1,-1),3}.
\end{aligned}
$$

3.3 Coefficient Functions

We need to define some sets of sufficiently smooth and integrable functions to be used as coefficient functions of the stochastic expansions. Let $s \in \{0, 1, \ldots\}$ denote some integer that will become clear from the context when the following definitions will be applied. By $\mathcal{L}^0$ we denote the set of functions $f(t, x, u) : [0, T] \times \mathbb{R}^d \times \mathcal{E}^s \longrightarrow \mathbb{R}^d$ for which

$$f\big(t, x + c(t, x, v), u\big) - f(t, x, u) \tag{3.3.1}$$

is $\phi(dv)$-integrable for all $t \in [0, T]$, $x \in \mathbb{R}^d$, $u \in \mathcal{E}^s$ and $f(\cdot, \cdot, u) \in \mathcal{C}^{1,2}$. Here $\mathcal{C}^{1,2}$ denotes the set of functions that are continuously differentiable with respect to time and twice continuously differentiable with respect to the spatial variables. Note that, according to the notation adopted in this thesis, c^i denotes the ith component of the jump coefficient vector c. Moreover, with $\widetilde{\mathcal{L}}^0$ we denote the set of functions $f(t, x, u) : [0, T] \times \mathbb{R}^d \times \mathcal{E}^s \longrightarrow \mathbb{R}^d$ for which

$$f\big(t, x + c(t, x, v), u\big) - f(t, x, u) - \sum_{i=1}^{d} c^i(t, x, v) \frac{\partial}{\partial x^i} f(t, x, u) \tag{3.3.2}$$

is $\phi(dv)$-integrable for all $t \in [0, T]$, $x \in \mathbb{R}^d$, $u \in \mathcal{E}^s$ and $f(\cdot, \cdot, u) \in \mathcal{C}^{1,2}$. With $\mathcal{L}^k$, $k \in \{1, \ldots, m\}$, we denote the set of functions $f(t, x, u)$ with partial derivatives $\frac{\partial}{\partial x^i} f(t, x, u)$, $i \in \{1, \ldots, d\}$. With $\mathcal{L}^{-1}$ we denote the set of functions for which

$$\big| f\big(t, x + c\,(t, x, v)\,, u\big) - f\big(t, x, u\big) \big|^2 \tag{3.3.3}$$

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

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

$$\begin{aligned} L^{(0)} f(t, x, u) \;&:=\; \frac{\partial}{\partial t} f(t, x, u) + \sum_{i=1}^{d} a^i(t, x) \frac{\partial}{\partial x^i} f(t, x, u) \\ &\quad + \frac{1}{2} \sum_{i,l=1}^{d} \sum_{j=1}^{m} b^{i,j}(t, x) b^{l,j}(t, x) \frac{\partial^2}{\partial x^i \partial x^l} f(t, x, u) \end{aligned} \tag{3.3.4}$$

$$L^{(k)} f(t, x, u) := \sum_{i=1}^{d} b^{i,k}(t, x) \frac{\partial}{\partial x^i} f(t, x, u), \tag{3.3.5}$$

for $k \in \{1, \ldots, m\}$, and

$$L_v^{(-1)} f(t,x,u) := f\big(t, x + c\,(t,x,v)\,, u\big) - f(t,x,u), \tag{3.3.6}$$

for all $t \in [0,T]$, $x \in \mathbb{R}^d$ and $u \in \mathcal{E}^s$. Note that the operator in (3.3.6) adds an extra dependence $v \in \mathcal{E}$ on the mark components. Let us also define the operator

$$\begin{aligned}
\widetilde{L}^{(0)} f(t,x,u) &:= L^{(0)} f(t,x,u) + \int_{\mathcal{E}} \left\{ f\Big(t, x + c\big(t,x,v\big), u\Big) - f\Big(t,x,u\Big) \right\} \phi(dv) \\
&= \frac{\partial}{\partial t} f(t,x,u) + \sum_{i=1}^{d} \widetilde{a}^i(t,x) \frac{\partial}{\partial x^i} f(t,x,u) \\
&\quad + \frac{1}{2} \sum_{i,l=1}^{d} \sum_{j=1}^{m} b^{i,j}(t,x) b^{l,j}(t,x) \frac{\partial^2}{\partial x^i \partial x^l} f(t,x,u) \\
&\quad + \int_{\mathcal{E}} \left\{ f\Big(t, x + c\big(t,x,v\big), u\Big) - f\Big(t,x,u\Big) \right. \\
&\quad \left. - \sum_{i=1}^{d} c^i(t,x,v) \frac{\partial}{\partial x^i} f(t,x,u) \right\} \phi(dv),
\end{aligned} \tag{3.3.7}$$

for all $t \in [0,T]$, $x \in \mathbb{R}^d$ and $u \in \mathcal{E}^s$, which allows us to describe conveniently the impact of the compensated Poisson measure.

By using the above definitions, 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*

$$f_\alpha(t,x,u) := \begin{cases} f\big(t,x\big) & \text{for} \quad l(\alpha) = 0, \\ L^{(j_1)} f_{-\alpha}(t,x,u_1,\ldots,u_{s(-\alpha)}) & \text{for} \quad l(\alpha) \geq 1,\ j_1 \in \{0,\ldots,m\} \\ L^{(-1)}_{u_{s(\alpha)}} f_{-\alpha}(t,x,u_1,\ldots,u_{s(-\alpha)}) & \text{for} \quad l(\alpha) \geq 1,\ j_1 = -1. \end{cases} \tag{3.3.8}$$

By $u = (u^1, \ldots, u^{s(\alpha)})$ we denote a vector $u \in \mathcal{E}^{s(\alpha)}$. Note that the dependence on u in (3.3.8) is introduced by the repeated application of the operator (3.3.6). Additionally, we assume that the coefficients of the SDE (2.1.2) and the function f satisfy the smoothness and integrability conditions needed for the operators in (3.3.8) to be well defined, see also Remark 3.5.2.

If we choose the identity function $f(t,x) = x$, then for the case $d = m = 1$ we can write the examples

$$f_{(-1,0)}(t,x,u) = L_u^{(-1)} a(t,x) = a(t, x + c(t,x,u)) - a(t,x),$$

$$\begin{aligned} f_{(0,1)}(t,x) &= L^{(0)} b(t,x) \\ &= \frac{\partial}{\partial t} b(t,x) + a(t,x) \frac{\partial}{\partial x} b(t,x) + \frac{1}{2} \big(b(t,x)\big)^2 \frac{\partial^2}{\partial x^2} b(t,x) \end{aligned}$$

and

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

with $u = (u^1, u^2) \in \mathcal{E}^2$.

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

$$\tilde{f}_\alpha(t,x,u) := \begin{cases} f(t,x) & \text{for} \quad l(\alpha) = 0, \\ \widetilde{L}^{(0)} \tilde{f}_{-\alpha}(t,x,u_1,\ldots,u_{s(-\alpha)}) & \text{for} \quad l(\alpha) \geq 1,\ j_1 = 0 \\ L^{(j_1)} \tilde{f}_{-\alpha}(t,x,u_1,\ldots,u_{s(-\alpha)}) & \text{for} \quad l(\alpha) \geq 1,\ j_1 \in \{1,\ldots,m\} \\ L_{u_{s(\alpha)}}^{(-1)} \tilde{f}_{-\alpha}(t,x,u_1,\ldots,u_{s(-\alpha)}) & \text{for} \quad l(\alpha) \geq 1,\ j_1 = -1. \end{cases} \tag{3.3.9}$$

Here we assume again that the coefficients of the SDE (2.1.2) and the function f satisfy the smoothness and integrability conditions needed for the operators in (3.3.9) to be well defined. For illustration, if we choose the identity function $f(t,x) = x$, then for the case $d = m = 1$ we can write the examples

$$\begin{aligned} \tilde{f}_{(-1,0)}(t,x,u) &= L_u^{(-1)} \tilde{a}(t,x) \\ &= \tilde{a}(t, x + c(t,x,u)) - \tilde{a}(t,x) \\ &= a(t, x + c(t,x,u)) - a(t,x) \\ &\quad + \int_{\mathcal{E}} \{c(t, x + c(t,x,u), v) - c(t,x,v)\} \, \phi(dv), \end{aligned}$$

$$\begin{aligned}
\tilde{f}_{(0,1)}(t,x) &= \widetilde{L}^{(0)} b(t,x) \\
&= \frac{\partial}{\partial t} b(t,x) + a(t,x) \frac{\partial}{\partial x} b(t,x) + \frac{1}{2} \big(b(t,x)\big)^2 \frac{\partial^2}{\partial x^2} b(t,x) \\
&\quad + \int_{\mathcal{E}} \big\{ b\big(t, x + c(t,x,v)\big) - b\big(t,x\big) \big\} \, \phi(dv) \\
&= \frac{\partial}{\partial t} b(t,x) + \tilde{a}(t,x) \frac{\partial}{\partial x} b(t,x) + \frac{1}{2} \big(b(t,x)\big)^2 \frac{\partial^2}{\partial x^2} b(t,x) \\
&\quad + \int_{\mathcal{E}} \left\{ b\big(t, x + c(t,x,v)\big) - b\big(t,x\big) - c(t,x,v) \frac{\partial}{\partial x} b(t,x) \right\} \phi(dv)
\end{aligned}$$

and

$$\tilde{f}_{(-1,-1)}(t,x,u) = f_{(-1,-1)}(t,x,u). \tag{3.3.10}$$

3.4 Hierarchical and Remainder Sets

To define a Wagner-Platen expansion we finally need to select some appropriate sets of multi-indices that characterize its expansion part. 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, which means $\sup_{\alpha \in \mathcal{A}} l(\alpha) < \infty$, and $-\alpha \in \mathcal{A}$ for each $\alpha \in \mathcal{A} \backslash \{v\}$.

We also define the *remainder* set $\mathcal{B}(\mathcal{A})$ of a hierarchical set $\mathcal{A}$ by

$$\mathcal{B}(\mathcal{A}) = \{\alpha \in \mathcal{M}_m \backslash \mathcal{A} : -\alpha \in \mathcal{A}\}. \tag{3.4.1}$$

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

In the following we give an example of a hierarchical set and the corresponding remainder set. When $m = 1$, we will later see that the hierarchical set $\mathcal{A}_E$ corresponding to the Euler scheme is given by

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

The corresponding remainder set $\mathcal{B}(\mathcal{A}_E)$ is then of the form

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

3.5 Wagner-Platen Expansions

We will now present the general form of Wagner-Platen expansions for the solution of the d-dimensional jump-diffusion SDE

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t + \int_{\mathcal{E}} c(t, X_{t-}, v)\, p_\phi(dv, dt), \tag{3.5.1}$$

for $t \in [0, T]$, with $X_0 \in \mathbb{R}^d$, as described in (2.1.2) and (3.1.2).

Theorem 3.5.1 *For two given 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 corresponding Wagner-Platen expansion*

$$f(\tau, X_\tau) = \sum_{\alpha \in \mathcal{A}} I_\alpha[f_\alpha(\rho, X_\rho)]_{\rho,\tau} + \sum_{\alpha \in \mathcal{B}(\mathcal{A})} I_\alpha[f_\alpha(\cdot, X_\cdot)]_{\rho,\tau} \tag{3.5.2}$$

and the compensated Wagner-Platen expansion

$$f(\tau, X_\tau) = \sum_{\alpha \in \mathcal{A}} \tilde{I}_\alpha[\tilde{f}_\alpha(\rho, X_\rho)]_{\rho,\tau} + \sum_{\alpha \in \mathcal{B}(\mathcal{A})} \tilde{I}_\alpha[\tilde{f}_\alpha(\cdot, X_\cdot)]_{\rho,\tau}, \tag{3.5.3}$$

assuming that the function f and the coefficients of the SDE (3.5.1) are sufficiently smooth and integrable such that the arising coefficient functions f_α and $\tilde{f}_\alpha$ are well defined and the corresponding multiple stochastic integrals exist.

Note that in (3.5.2) and (3.5.3) we have suppressed the dependence of f_α and $\tilde{f}_\alpha$ on $u \in \mathcal{E}^{s(\alpha)}$ in our notation and we will do so also in the following when no misunderstanding is possible.

Remark 3.5.2 *Sharp conditions to be satisfied by the function f and the coefficients of the SDE (3.5.1) such that the coefficient functions f_α and $\tilde{f}_\alpha$ are well defined, and the corresponding multiple stochastic integrals exist, can be obtained by the definitions of the operators (3.3.4)-(3.3.7) and of the sets in (3.2.3). Alternatively, one can provide simple sufficient conditions. For instance, if the function f and the coefficients a, b, and c of the SDE (3.5.1) are $2(l(\alpha) + 1)$ times continuously differentiable, uniformly bounded with uniformly bounded derivatives, then all coefficient functions and multiple stochastic integrals in the expansions (3.5.2) and (3.5.3) are well defined.*

The proof of the Wagner-Platen expansion is based on the iterated application of the Itô formula. Since it is formally the same as in Platen (1982a, 1982b), it is here omitted.

By choosing as function f the identity functions $f(t,x) = x$, we can represent the solution $X = \{X_t, t \in [0,T]\}$ of the SDE (3.5.1) by the *Wagner-Platen expansion*

$$X_\tau = \sum_{\alpha \in \mathcal{A}} I_\alpha[f_\alpha(\rho, X_\rho)]_{\rho,\tau} + \sum_{\alpha \in \mathcal{B}(\mathcal{A})} I_\alpha[f_\alpha(\cdot, X_\cdot)]_{\rho,\tau} \tag{3.5.4}$$

and also by the *compensated Wagner-Platen expansion*

$$X_\tau = \sum_{\alpha \in \mathcal{A}} \widetilde{I}_\alpha[\tilde{f}_\alpha(\rho, X_\rho)]_{\rho,\tau} + \sum_{\alpha \in \mathcal{B}(\mathcal{A})} \widetilde{I}_\alpha[\tilde{f}_\alpha(\cdot, X_\cdot)]_{\rho,\tau}, \tag{3.5.5}$$

where ρ and τ are two given stopping times with $0 \leq \rho \leq \tau \leq T$ a.s.

We remark that Wagner-Platen expansions are also generalizations of the deterministic Taylor formula. On the other hand, since they are obtained by an iterative application of the Itô formula, they generalize the Itô formula. As we shall see later, discrete time approximations of an Itô process will be constructed by using truncated Wagner-Platen expansions that neglect remainder terms.

For instance, for the hierarchical set $\mathcal{A}_E$ corresponding to the Euler scheme, with $m = 1$, we obtain from (3.5.4) the Wagner-Platen expansion

$$X_\tau = X_\rho + a(\rho, X_\rho) \int_\rho^\tau dz + b(\rho, X_\rho) \int_\rho^\tau dW_z + \int_\rho^\tau \int_{\mathcal{E}} c(\rho, X_\rho, v) p_\phi(dv, dz) + R$$

with remainder

$$\begin{aligned}
R \;=\;& \int_\rho^\tau \int_\rho^s L^{(0)} a(z, X_z) dz ds + \int_\rho^\tau \int_\rho^s L^{(1)} a(z, X_z) dW_z ds \\
&+ \int_\rho^\tau \int_\rho^s \int_{\mathcal{E}} L_v^{(-1)} a(z, X_z) p_\phi(dv, dz) ds \\
&+ \int_\rho^\tau \int_\rho^s L^{(0)} b(z, X_z) dz dW_s + \int_\rho^\tau \int_\rho^s L^{(1)} b(z, X_z) dW_z dW_s \\
&+ \int_\rho^\tau \int_\rho^s \int_{\mathcal{E}} L_v^{(-1)} b(z, X_z) p_\phi(dv, dz) dW_s \\
&+ \int_\rho^\tau \int_{\mathcal{E}} \int_\rho^s L^{(0)} c(z, X_z, v) dz p_\phi(dv, ds)
\end{aligned}$$

$$+\int_\rho^\tau \int_{\mathcal{E}} \int_\rho^s L^{(1)} c(z, X_z, v) dW_z p_\phi(dv, ds)$$

$$+\int_\rho^\tau \int_{\mathcal{E}} \int_\rho^s \int_{\mathcal{E}} L_u^{(-1)} c(z, X_z, v) p_\phi(dv, dz) p_\phi(du, ds). \tag{3.5.6}$$

There is an analogous Wagner-Platen expansion following from (3.5.5) where the Poisson measure is replaced by the compensated Poisson measure, the drift by the compensated drift and the Itô coefficient functions by their counterparts for the compensated case. Note that in this particular case (for the hierarchical set $\mathcal{A}_E$), one can show that these two expansions are equivalent. However, in general these two expansions are different.

3.6 Moments of Multiple Stochastic Integrals

The lemmas in this section provide estimates of multiple stochastic integrals. These constitute the basis of the proofs of convergence theorems that will be presented later. We will consider both the case of multiple stochastic integrals $I_\alpha[g(\cdot)]_{\rho,\tau}$ with respect to the Poisson jump measure and compensated multiple stochastic integrals $\widetilde{I}_\alpha[g(\cdot)]_{\rho,\tau}$ with respect to the compensated Poisson measure. The estimates to be derived in these two cases differ in the values of some finite constants, but show the same structural dependence on the length of the interval of integration $(\tau - \rho)$. Because of the martingale property of the compensated Poisson measure $\widetilde{p}_\phi$, the proofs flow more naturally in the case of the compensated multiple stochastic integral $\widetilde{I}_\alpha[g(\cdot)]_{\rho,\tau}$. Therefore, in the following proofs we will first consider this case in full detail. Then, by using the decomposition

$$p_\phi(dv, dt) = \widetilde{p}_\phi(dv, dt) + \phi(dv)dt, \tag{3.6.1}$$

see (3.1.1), we will prove corresponding estimates for the case of multiple stochastic integrals $I_\alpha[g(\cdot)]_{\rho,\tau}$.

The following lemma provides a uniform mean-square estimate of multiple stochastic integrals.

Lemma 3.6.1 *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 being $\mathcal{A}_\rho$-measurable and $0 \le \rho \le \tau \le \rho + \Delta \le T$ almost surely. Then*

$$\widetilde{F}_{\tau}^{\alpha} := E\left(\sup_{\rho \le s \le \tau} \left| \widetilde{I}_{\alpha}[g(\cdot)]_{\rho,s} \right|^{2} \middle| \mathcal{A}_{\rho} \right) \le 4^{l(\alpha)-n(\alpha)} \Delta^{l(\alpha)+n(\alpha)-1} \int_{\rho}^{\tau} V_{\rho,z,s(\alpha)}\, dz, \quad (3.6.2)$$

and

$$F_{\tau}^{\alpha} := E\left(\sup_{\rho \le s \le \tau} | I_{\alpha}[g(\cdot)]_{\rho,s}|^{2} \middle| \mathcal{A}_{\rho} \right) \le 4^{l(\alpha)-n(\alpha)} \Delta^{l(\alpha)+n(\alpha)-1} \hat{K}^{s(\alpha)} \int_{\rho}^{\tau} V_{\rho,z,s(\alpha)}\, dz, \quad (3.6.3)$$

where

$$V_{\rho,z,s(\alpha)} := \int_{\mathcal{E}} \cdots \int_{\mathcal{E}} E\left(\sup_{\rho \le t \le z} |g(t, v^{1}, \ldots, v^{s(\alpha)})|^{2} \middle| \mathcal{A}_{\rho} \right) \phi(dv^{1}) \ldots \phi(dv^{s(\alpha)}) < \infty \quad (3.6.4)$$

for $z \in [\rho, \tau]$, *and* $\hat{K} = \frac{1}{2}(4 + T\lambda)$.

Proof: We will first prove assertion (3.6.2) by induction with respect to $l(\alpha)$.

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

$$\left| \int_{\rho}^{s} g(z)\ dz \right|^{2} \le (s - \rho) \int_{\rho}^{s} |g(z)|^{2}\ dz. \quad (3.6.5)$$

Therefore, we obtain

$$\begin{aligned}
\widetilde{F}_{\tau}^{(0)} &= E\left(\sup_{\rho \le s \le \tau} \left| \int_{\rho}^{s} g(z)\ dz \right|^{2} \middle| \mathcal{A}_{\rho} \right) \\
&\le E\left(\sup_{\rho \le s \le \tau} (s - \rho) \int_{\rho}^{s} |g(z)|^{2}\ dz \middle| \mathcal{A}_{\rho} \right) \\
&= E\left((\tau - \rho) \int_{\rho}^{\tau} |g(z)|^{2}\ dz \middle| \mathcal{A}_{\rho} \right) \\
&\le \Delta\ E\left(\int_{\rho}^{\tau} |g(z)|^{2}\ dz \middle| \mathcal{A}_{\rho} \right) \\
&= \Delta \int_{\rho}^{\tau} E\left(|g(z)|^{2} | \mathcal{A}_{\rho} \right)\ dz
\end{aligned}$$

$$
\begin{aligned}
&\leq \Delta \int_\rho^\tau E\left(\sup_{\rho\leq t\leq z} |g(t)|^2 \middle| \mathcal{A}_\rho\right) dz \\
&= 4^{l(\alpha)-n(\alpha)}\,\Delta^{l(\alpha)+n(\alpha)-1} \int_\rho^\tau V_{\rho,z,s(\alpha)}\,dz, \qquad (3.6.6)
\end{aligned}
$$

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

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

$$
\{\widetilde{I}_\alpha[g(\cdot)]_{\rho,t}, t \in [\rho, T]\} = \{\int_\rho^t g(s)\ dW_s^j, t \in [\rho, T]\} \qquad (3.6.7)
$$

is a local martingale. Since $g \in \mathcal{H}_{(j)}$, by (3.2.3) it is also a martingale. Therefore, applying Doob's inequality, see Appendix A, and Itô's isometry we have

$$
\begin{aligned}
\widetilde{F}_\tau^{(j)} &= E\left(\sup_{\rho\leq s\leq\tau}\left|\int_\rho^s g(z)\ dW_z^j\right|^2 \middle| \mathcal{A}_\rho\right) \\
&\leq 4\ E\left(\left|\int_\rho^\tau g(z)\ dW_z^j\right|^2 \middle| \mathcal{A}_\rho\right) \\
&= 4\ E\left(\int_\rho^\tau |g(z)|^2 dz \middle| \mathcal{A}_\rho\right) \\
&= 4 \int_\rho^\tau E\left(|g(z)|^2 | \mathcal{A}_\rho\right) dz \\
&\leq 4 \int_\rho^\tau E\left(\sup_{\rho\leq t\leq z} |g(t)|^2 \middle| \mathcal{A}_\rho\right) dz \\
&= 4^{l(\alpha)-n(\alpha)}\,\Delta^{l(\alpha)+n(\alpha)-1} \int_\rho^\tau V_{\rho,z,s(\alpha)}\,dz. \qquad (3.6.8)
\end{aligned}
$$

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

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

$$\{\widetilde{I}_\alpha[g(\cdot)]_{\rho,t}, t \in [\rho, T]\} = \{\int_\rho^t \int_{\mathcal{E}} g(s-, v)\ \widetilde{p}_\phi(dv, ds), t \in [\rho, T]\} \tag{3.6.9}$$

is by (3.1.1) and (3.2.3) a martingale. Then, by Doob's inequality and the isometry for Poisson type jump martingales, we obtain

$$\begin{aligned}
\widetilde{F}_\tau^{(-1)} &= E\left(\sup_{\rho\le s\le\tau}\left|\int_\rho^s \int_{\mathcal{E}} g(z-, v)\ \widetilde{p}_\phi(dv, dz)\right|^2 \middle| \mathcal{A}_\rho\right) \\
&\le 4\ E\left(\left|\int_\rho^\tau \int_{\mathcal{E}} g(z-, v)\ \widetilde{p}_\phi(dv, dz)\right|^2 \middle| \mathcal{A}_\rho\right) \\
&= 4\ E\left(\int_\rho^\tau \int_{\mathcal{E}} |g(z, v)|^2\, \phi(dv)\, dz \middle| \mathcal{A}_\rho\right) \\
&= 4 \int_\rho^\tau \int_{\mathcal{E}} E\left(|g(z, v)|^2 | \mathcal{A}_\rho\right)\, \phi(dv)\, dz \\
&\le 4 \int_\rho^\tau \int_{\mathcal{E}} E\left(\sup_{\rho\le t\le z} |g(t, v)|^2 \middle| \mathcal{A}_\rho\right)\, \phi(dv)\, dz \\
&= 4^{l(\alpha)-n(\alpha)}\, \Delta^{l(\alpha)+n(\alpha)-1} \int_\rho^\tau V_{\rho,z,s(\alpha)}\, dz,
\end{aligned} \tag{3.6.10}$$

since $s(\alpha) = 1$. This shows that the result of Lemma 3.6.1 holds for $l(\alpha) = 1$.

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

$$\begin{aligned}
\widetilde{F}_\tau^\alpha &= E\left(\sup_{\rho\le s\le\tau}\left|\int_\rho^s \widetilde{I}_{\alpha-}[g(\cdot)]_{\rho,z}\ dz\right|^2 \middle| \mathcal{A}_\rho\right) \\
&\le E\left(\sup_{\rho\le s\le\tau} (s-\rho) \int_\rho^s |\widetilde{I}_{\alpha-}[g(\cdot)]_{\rho,z}|^2\ dz \middle| \mathcal{A}_\rho\right) \\
&= E\left((\tau-\rho) \int_\rho^\tau |\widetilde{I}_{\alpha-}[g(\cdot)]_{\rho,z}|^2\ dz \middle| \mathcal{A}_\rho\right) \\
&\le \Delta\, E\left(\int_\rho^\tau |\widetilde{I}_{\alpha-}[g(\cdot)]_{\rho,z}|^2\ dz \middle| \mathcal{A}_\rho\right)
\end{aligned}$$

$$
\begin{aligned}
&\leq \Delta\, E\left(\int_\rho^\tau \sup_{\rho\leq s\leq\tau} |\tilde{I}_{\alpha-}[g(\cdot)]_{\rho,s}|^2 \, dz \,\middle|\, \mathcal{A}_\rho \right) \\
&= \Delta\, E\left(\int_\rho^\tau dz \times \sup_{\rho\leq s\leq\tau} |\tilde{I}_{\alpha-}[g(\cdot)]_{\rho,s}|^2 \,\middle|\, \mathcal{A}_\rho \right) \\
&\leq \Delta^2\, E\left(\sup_{\rho\leq s\leq\tau} |\tilde{I}_{\alpha-}[g(\cdot)]_{\rho,s}|^2 \,\middle|\, \mathcal{A}_\rho \right). \qquad (3.6.11)
\end{aligned}
$$

Then, by the inductive hypothesis it follows that

$$
\begin{aligned}
\tilde{F}_\tau^\alpha &\leq \Delta^2\, 4^{l(\alpha-)-n(\alpha-)}\, \Delta^{l(\alpha-)+n(\alpha-)-1} \int_\rho^\tau V_{\rho,z,s(\alpha-)}\, dz \\
&= 4^{l(\alpha)-n(\alpha)}\, \Delta^{l(\alpha)+n(\alpha)-1} \int_\rho^\tau V_{\rho,z,s(\alpha)}\, dz, \qquad (3.6.12)
\end{aligned}
$$

where the last line holds since $l(\alpha) = l(\alpha-) + 1$, $n(\alpha) = n(\alpha-) + 1$ 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

$$
\{\tilde{I}_\alpha[g(\cdot)]_{\rho,t}, t \in [\rho, T]\} \qquad (3.6.13)
$$

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

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

$$
\begin{aligned}
&= 4\ E\left((\tau - \rho) \sup_{\rho \le s \le \tau} |\widetilde{I}_{\alpha-}[g(\cdot)]_{\rho,s}|^2 \,\middle|\, \mathcal{A}_\rho \right) \\
&\le 4\,\Delta\, E\left(\sup_{\rho \le s \le \tau} |\widetilde{I}_{\alpha-}[g(\cdot)]_{\rho,s}|^2 \,\middle|\, \mathcal{A}_\rho \right). \qquad (3.6.14)
\end{aligned}
$$

By the inductive hypothesis we have

$$
\begin{aligned}
\widetilde{F}^{\alpha}_{\tau} &\le 4\,\Delta\, 4^{l(\alpha-)-n(\alpha-)}\, \Delta^{l(\alpha-)+n(\alpha-)-1} \int_\rho^T V_{\rho,z,s(\alpha-)}\, dz \\
&= \Delta^{l(\alpha)+n(\alpha)-1}\, 4^{l(\alpha)-n(\alpha)} \int_\rho^T V_{\rho,z,s(\alpha)}\, dz, \qquad (3.6.15)
\end{aligned}
$$

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

$$
\{\widetilde{I}_\alpha[g(\cdot)]_{\rho,t}, t \in [\rho, T]\} \qquad (3.6.16)
$$

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

$$
\begin{aligned}
\widetilde{F}^{\alpha}_{\tau} &= E\left(\sup_{\rho \le s \le \tau} \left| \int_\rho^s \int_{\mathcal{E}} \widetilde{I}_{\alpha-}[g(\cdot, v^{s(\alpha)})]_{\rho,z-}\ \widetilde{p}_\phi(dv^{s(\alpha)}, dz) \right|^2 \,\middle|\, \mathcal{A}_\rho \right) \\
&\le 4\ E\left(\left| \int_\rho^T \int_{\mathcal{E}} \widetilde{I}_{\alpha-}[g(\cdot, v^{s(\alpha)})]_{\rho,z-}\ \widetilde{p}_\phi(dv^{s(\alpha)}, dz) \right|^2 \,\middle|\, \mathcal{A}_\rho \right) \\
&= 4\ E\left(\int_\rho^T \int_{\mathcal{E}} |\widetilde{I}_{\alpha-}[g(\cdot, v^{s(\alpha)})]_{\rho,z}|^2\, \phi(dv^{s(\alpha)})\, dz \,\middle|\, \mathcal{A}_\rho \right) \\
&\le 4\,E\left(\int_\rho^T \int_{\mathcal{E}} \sup_{\rho \le s \le \tau} |\widetilde{I}_{\alpha-}[g(\cdot, v^{s(\alpha)})]_{\rho,s}|^2\, \phi(dv^{s(\alpha)})\, dz \,\middle|\, \mathcal{A}_\rho \right) \\
&= 4\,E\left((\tau - \rho) \int_{\mathcal{E}} \sup_{\rho \le s \le \tau} |\widetilde{I}_{\alpha-}[g(\cdot, v^{s(\alpha)})]_{\rho,s}|^2\, \phi(dv^{s(\alpha)}) \,\middle|\, \mathcal{A}_\rho \right) \\
&\le 4\,\Delta \int_{\mathcal{E}} E\left(\sup_{\rho \le s \le \tau} |\widetilde{I}_{\alpha-}[g(\cdot, v^{s(\alpha)})]_{\rho,s}|^2 \,\middle|\, \mathcal{A}_\rho \right) \phi(dv^{s(\alpha)}). \qquad (3.6.17)
\end{aligned}
$$

By the inductive hypothesis we have

$$
\begin{aligned}
\widetilde{F}^{\alpha}_{\tau} &\le 4\,\Delta\, 4^{l(\alpha-)-n(\alpha-)}\,\Delta^{l(\alpha-)+n(\alpha-)-1} \int_{\mathcal{E}} \int_{\rho}^{\tau} V_{\rho,z,s(\alpha-)}\, dz\, \phi(dv^{s(\alpha)}) \\
&= 4^{l(\alpha)-n(\alpha)}\,\Delta^{l(\alpha)+n(\alpha)-1} \int_{\rho}^{\tau} V_{\rho,z,s(\alpha)}\, dz, \qquad (3.6.18)
\end{aligned}
$$

since $l(\alpha) = l(\alpha-)+1$, $n(\alpha) = n(\alpha-)$ and $s(\alpha) = s(\alpha-)+1$, which completes the proof of assertion (3.6.2).

To prove assertion (3.6.3) we can proceed in a similar way by induction with respect $l(\alpha)$. The idea, which will be used also in other lemmas, is to rewrite the Poisson jump measure as a sum of the compensated Poisson measure and a time integral as follows by (3.6.1).

Let us consider the case of $l(\alpha) = 1$ with $\alpha = (-1)$. By (3.6.1) and the Cauchy-Schwarz inequality, we have

$$
\begin{aligned}
F^{(-1)}_{\tau} &= E\left(\sup_{\rho\le s\le\tau} \left| \int_{\rho}^{s} \int_{\mathcal{E}} g(z-,v)\ p_{\phi}(dv,dz) \right|^2 \middle| \mathcal{A}_{\rho} \right) \\
&= E\left(\sup_{\rho\le s\le\tau} \left| \int_{\rho}^{s} \int_{\mathcal{E}} g(z-,v)\ \widetilde{p}_{\phi}(dv,dz) + \int_{\rho}^{s} \int_{\mathcal{E}} g(z,v)\ \phi(dv)dz \right|^2 \middle| \mathcal{A}_{\rho} \right) \\
&\le 2E\left(\sup_{\rho\le s\le\tau} \left| \widetilde{I}_{(-1)}[g(\cdot)]_{\rho,\tau} \right|^2 \middle| \mathcal{A}_{\rho} \right) \\
&\quad +2E\left(\sup_{\rho\le s\le\tau} \left| \widetilde{I}_{(0)}[\int_{\mathcal{E}} g(\cdot,v)\phi(dv)]_{\rho,\tau} \right|^2 \middle| \mathcal{A}_{\rho} \right). \qquad (3.6.19)
\end{aligned}
$$

By applying the estimate of the already obtained assertion (3.6.2) and the Cauchy-Schwarz inequality, we have

$$
\begin{aligned}
F^{(-1)}_{\tau} &\le 8 \int_{\rho}^{\tau} \int_{\mathcal{E}} E\left(\sup_{\rho\le t\le z} |g(t,v)|^2 \middle| \mathcal{A}_{\rho} \right) \phi(dv)dz \\
&\quad +2\Delta \int_{\rho}^{\tau} E\left(\sup_{\rho\le t\le z} \left| \int_{\mathcal{E}} g(t,v)\phi(dv) \right|^2 \middle| \mathcal{A}_{\rho} \right) dz
\end{aligned}
$$

$$\begin{aligned} &\leq 4\hat{K} \int_\rho^\tau V_{\rho,z,s(\alpha)}\, dz \\ &= 4^{l(\alpha)-n(\alpha)} \Delta^{l(\alpha)+n(\alpha)-1} \hat{K}^{s(\alpha)} \int_\rho^\tau V_{\rho,z,s(\alpha)}\, dz, \end{aligned} \tag{3.6.20}$$

with $\hat{K} = \frac{1}{2}(4 + T\lambda)$.

To finalize the proof we consider the case of $l(\alpha) = n + 1$ where $\alpha = (j_1, \dots, j_{n+1})$ with $j_{n+1} = -1$. By (3.6.1), the Cauchy-Schwarz inequality, Doob's inequality, the Itô isometry and similar steps as those used in (3.6.17) and (3.6.11), one obtains

$$\begin{aligned} F_\tau^\alpha &= E\left(\sup_{\rho \leq s \leq \tau} \left| \int_\rho^s \int_{\mathcal{E}} I_{\alpha-}[g(\cdot, v^{s(\alpha)})]_{\rho,z-}\ p_\phi(dv^{s(\alpha)}, dz) \right|^2 \middle| \mathcal{A}_\rho \right) \\ &\leq 2E\left(\sup_{\rho \leq s \leq \tau} \left| \int_\rho^s \int_{\mathcal{E}} I_{\alpha-}[g(\cdot, v^{s(\alpha)})]_{\rho,z-}\ \widetilde{p}_\phi(dv^{s(\alpha)}, dz) \right|^2 \middle| \mathcal{A}_\rho \right) \\ &\quad + 2E\left(\sup_{\rho \leq s \leq \tau} \left| \int_\rho^s \int_{\mathcal{E}} I_{\alpha-}[g(\cdot, v^{s(\alpha)})]_{\rho,z-}\ \phi(dv^{s(\alpha)}) dz \right|^2 \middle| \mathcal{A}_\rho \right) \\ &\leq 8\,\Delta \int_{\mathcal{E}} E\left(\sup_{\rho \leq s \leq \tau} |I_{\alpha-}[g(\cdot, v^{s(\alpha)})]_{\rho,s}|^2 \middle| \mathcal{A}_\rho \right) \phi(dv^{s(\alpha)}) \\ &\quad + 2\,\lambda\,\Delta^2 \int_{\mathcal{E}} E\left(\sup_{\rho \leq s \leq \tau} |I_{\alpha-}[g(\cdot)]_{\rho,s}|^2 \middle| \mathcal{A}_\rho \right) \phi(dv^{s(\alpha)}). \end{aligned} \tag{3.6.21}$$

By the induction hypothesis, we finally obtain

$$\begin{aligned} F_\tau^\alpha &\leq 8\,\Delta\, 4^{l(\alpha-)-n(\alpha-)} \Delta^{l(\alpha-)+n(\alpha-)-1} \hat{K}^{s(\alpha-)} \int_{\mathcal{E}} \int_\rho^\tau V_{\rho,z,s(\alpha-)}\, dz \phi(dv^{s(\alpha)}) \\ &\quad + 2\,\lambda\,\Delta\, T\, 4^{l(\alpha-)-n(\alpha-)} \Delta^{l(\alpha-)+n(\alpha-)-1} \hat{K}^{s(\alpha-)} \int_{\mathcal{E}} \int_\rho^\tau V_{\rho,z,s(\alpha-)}\, dz \phi(dv^{s(\alpha)}) \\ &\leq 4^{l(\alpha)-n(\alpha)} \Delta^{l(\alpha)+n(\alpha)-1} \hat{K}^{s(\alpha)} \int_\rho^\tau V_{\rho,z,s(\alpha)}\, dz, \end{aligned} \tag{3.6.22}$$

which completes the proof of Lemma 3.6.1. □

The following lemma provides an estimate for higher moments of multiple stochastic integrals. A similar result is presented in Li & Liu (2000).

Lemma 3.6.2 *Let $\alpha \in \mathcal{M}_m \backslash \{v\}$, $g \in \mathcal{H}_\alpha$, and ρ and τ denote two stopping times* *with τ being $\mathcal{A}_\rho$-measurable and $0 \leq \rho \leq \tau \leq T$ almost surely. Then for any $q \in \{1, 2 \ldots\}$ there exist positive constants C_1 and C_2 independent of $(\tau - \rho)$ such that*

$$\begin{aligned} \widetilde{F}_\tau^\alpha &:= E\left(\left| \widetilde{I}_\alpha[g(\cdot)]_{\rho,\tau} \right|^{2q} \middle| \mathcal{A}_\rho \right) \\ &\leq C_1 \, (\tau - \rho)^{q\left(l(\alpha)+n(\alpha)-s(\alpha)\right)+s(\alpha)-1} \int_\rho^\tau V_{\rho,z,s(\alpha)} \, dz, \qquad (3.6.23) \end{aligned}$$

and

$$\begin{aligned} F_\tau^\alpha &:= E\left(| \, I_\alpha[g(\cdot)]_{\rho,\tau} |^{2q} \middle| \mathcal{A}_\rho \right) \\ &\leq C_2 \, (\tau - \rho)^{q\left(l(\alpha)+n(\alpha)-s(\alpha)\right)+s(\alpha)-1} \int_\rho^\tau V_{\rho,z,s(\alpha)} \, dz, \qquad (3.6.24) \end{aligned}$$

where

$$V_{\rho,z,s(\alpha)} := \int_{\mathcal{E}} \ldots \int_{\mathcal{E}} E\left(|g(z, v_1, \ldots, v_{s(\alpha)})|^{2q} \middle| \mathcal{A}_\rho \right) \phi(dv_1) \ldots \phi(dv_{s(\alpha)}) < \infty$$

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

Proof: We will first prove the estimate (3.6.23) by induction on $l(\alpha)$.

1. Let us first prove the assertion for a multi-index α of length one, which means $l(\alpha) = 1$. When $l(\alpha) = 1$ and $\alpha = (0)$, by applying the Hölder inequality, see Appendix A, we obtain

$$\begin{aligned} \widetilde{F}_\tau^{(0)} &:= E\left(\left| \int_\rho^\tau g(z) dz \right|^{2q} \middle| \mathcal{A}_\rho \right) \\ &\leq (\tau - \rho)^{2q-1} \int_\rho^\tau E\left(|g(z)|^{2q} \middle| \mathcal{A}_\rho \right) dz \qquad (3.6.25) \\ &= C(\tau - \rho)^{q\left(l(\alpha)+n(\alpha)-s(\alpha)\right)+s(\alpha)-1} \int_\rho^\tau E\left(|g(z)|^{2q} | \mathcal{A}_\rho \right) dz. \end{aligned}$$

Note that here and in the following we denote by C any finite positive constant independent of $(\tau - \rho)$.

2. For $l(\alpha) = 1$ and $\alpha = (j)$, with $j \in \{1, 2 \ldots, m\}$, we obtain, see Krylov (1980), Corollary 3, page 80,

$$\begin{aligned}
\widetilde{F}_\tau^{(j)} &:= E\left(\left| \int_\rho^\tau g(z) dW_z^j \right|^{2q} \middle| \mathcal{A}_\rho \right) \\
&\leq 2^q \, (2q-1)^q (\tau - \rho)^{q-1} \int_\rho^\tau E\left(|g(z)|^{2q} \middle| \mathcal{A}_\rho \right) dz \qquad (3.6.26) \\
&= C(\tau - \rho)^{q\left(l(\alpha)+n(\alpha)-s(\alpha)\right)+s(\alpha)-1} \int_\rho^\tau E\left(|g(z)|^{2q} |\mathcal{A}_\rho \right) dz.
\end{aligned}$$

3. For $l(\alpha) = 1$ and $\alpha = (-1)$, let us define

$$x_\tau = \int_\rho^\tau \int_{\mathcal{E}} g(z-, v) p_\phi(dv, dz).$$

By applying Itô's formula to $|x_\tau|^{2q}$ together with the Hölder inequality, we obtain

$$\begin{aligned}
|x_\tau|^{2q} &= \int_\rho^\tau \int_{\mathcal{E}} \left(|x_{z-} + g(z-, v)|^{2q} - |x_{z-}|^{2q} \right) p_\phi(dv, dz) \\
&\leq (2^{2q-1} - 1) \int_\rho^\tau \int_{\mathcal{E}} |x_{z-}|^{2q} \, p_\phi(dv, dz) \\
&\quad + 2^{2q-1} \int_\rho^\tau \int_{\mathcal{E}} |g(z-, v)|^{2q} p_\phi(dv, dz). \qquad (3.6.27)
\end{aligned}$$

Therefore, by the properties of the Poisson jump measure, we have

$$\begin{aligned}
E\left(|x_\tau|^{2q} |\mathcal{A}_\rho \right) &\leq (2^{2q-1} - 1) \, \lambda \int_\rho^\tau E\left(|x_z|^{2q} |\mathcal{A}_\rho \right) dz \\
&\quad + 2^{2q-1} \int_\rho^\tau \int_{\mathcal{E}} E\left(|g(z, v)|^{2q} |\mathcal{A}_\rho \right) \phi(dv) \, dz, \quad (3.6.28)
\end{aligned}$$

where we recall that $\lambda = \phi(\mathcal{E})$ is the total intensity. Moreover, let us note that, by the first line of (3.6.27) and the properties of the Poisson jump measure, we obtain

$$E\left(|x_\tau|^{2q} |\mathcal{A}_\rho \right) = \int_\rho^\tau \int_{\mathcal{E}} E\left(\left(|x_z + g(z, v)|^{2q} - |x_z|^{2q} \right) \middle| \mathcal{A}_\rho \right) \phi(dv) \, dz,$$

which proves the continuity of $E\big(|x_\tau|^{2q}|\mathcal{A}_\rho\big)$ as a function of τ. Therefore, by applying the Gronwall inequality (1.2.9) in Appendix A to (3.6.28), we obtain

$$\begin{aligned}
E\big(|x_\tau|^{2q}|\mathcal{A}_\rho\big) &= E\left(\left|\int_\rho^\tau \int_{\mathcal{E}} g(z-,v)p_\phi(dv,dz)\right|^{2q}\middle|\mathcal{A}_\rho\right) \\
&\leq \exp\big\{(2^{2q-1}-1)\,\lambda\,(\tau-\rho)\big\}\,2^{2q-1} \\
&\quad\times \int_\rho^\tau \int_{\mathcal{E}} E\big(|g(z,v)|^{2q}|\mathcal{A}_\rho\big)\phi(dv)\,dz \\
&\leq \exp\big\{(2^{2q-1}-1)\,\lambda\,T\big\}\,2^{2q-1} \\
&\quad\times \int_\rho^\tau \int_{\mathcal{E}} E\big(|g(z,v)|^{2q}|\mathcal{A}_\rho\big)\phi(dv)\,dz \\
&= C\,(\tau-\rho)^{q\big(l(\alpha)+n(\alpha)-s(\alpha)\big)+s(\alpha)-1} \qquad (3.6.29)\\
&\quad\times \int_\rho^\tau \int_{\mathcal{E}} E\left(|g(z,v_{s(\alpha)})|^{2q}\middle|\mathcal{A}_\rho\right)\phi(dv_{s(\alpha)})\,dz.
\end{aligned}$$

Finally, by (3.1.1), the Hölder inequality, (3.6.29) and (3.6.25), we obtain

$$\begin{aligned}
\widetilde{F}_\tau^{(-1)} &= E\left(\left|\int_\rho^\tau \int_{\mathcal{E}} g(z-,v)p_\phi(dv,dz) - \int_\rho^\tau \int_{\mathcal{E}} g(z,v)\phi(dv)\,dz\right|^{2q}\middle|\mathcal{A}_\rho\right) \\
&\leq 2^{2q-1}\Bigg\{E\left(\left|\int_\rho^\tau \int_{\mathcal{E}} g(z-,v)p_\phi(dv,dz)\right|^{2q}\middle|\mathcal{A}_\rho\right) \\
&\quad + E\left(\left|\int_\rho^\tau \int_{\mathcal{E}} g(z,v)\phi(dv)\,dz\right|^{2q}\middle|\mathcal{A}_\rho\right)\Bigg\} \\
&\leq C\,(\tau-\rho)^{q\big(l(\alpha)+n(\alpha)-s(\alpha)\big)+s(\alpha)-1} \\
&\quad\times\Bigg\{\int_\rho^\tau \int_{\mathcal{E}} E\left(|g(z,v)|^{2q}\middle|\mathcal{A}_\rho\right)\phi(dv)\,dz \\
&\quad + \int_\rho^\tau E\left(\left|\int_{\mathcal{E}} g(z,v)\phi(dv)\right|^{2q}\middle|\mathcal{A}_\rho\right)dz\Bigg\} \\
&\leq C\,(\tau-\rho)^{q\big(l(\alpha)+n(\alpha)-s(\alpha)\big)+s(\alpha)-1} \\
&\quad\times \int_\rho^\tau \int_{\mathcal{E}} E\left(\left|g(z,v_{s(\alpha)})\right|^{2q}\middle|\mathcal{A}_\rho\right)\phi(dv_{s(\alpha)})\,dz.
\end{aligned}$$

4. Let us now consider the case $l(\alpha) = n+1$ and $\alpha = (j_1, \ldots, j_{n+1})$, with $j_{n+1} = 0$. By (3.6.25) and the inductive hypothesis, we get

$$\begin{aligned}
\widetilde{F}_\tau^\alpha &:= E\left(\left|\int_\rho^\tau \widetilde{I}_{\alpha-}[g(\cdot)]_{\rho,s} ds\right|^{2q} \middle| \mathcal{A}_\rho\right) \\
&\leq (\tau-\rho)^{2q-1} \int_\rho^\tau E\left(|\widetilde{I}_{\alpha-}[g(\cdot)]_{\rho,s}|^{2q} \middle| \mathcal{A}_\rho\right) ds \\
&\leq C\,(\tau-\rho)^{2q-1}\,(\tau-\rho)^{q\left(l(\alpha-)+n(\alpha-)-s(\alpha-)\right)+s(\alpha-)-1} \\
&\qquad \times \int_\rho^\tau \int_\rho^s V_{\rho,z,s(\alpha-)} dz\, ds \\
&= C\,(\tau-\rho)^{q\left(l(\alpha)+n(\alpha)-s(\alpha)\right)+s(\alpha)-1} \int_\rho^\tau V_{\rho,z,s(\alpha)} dz, \qquad (3.6.30)
\end{aligned}$$

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

5. When $l(\alpha) = n+1$ and $\alpha = (j_1, \ldots, j_{n+1})$, with $j_{n+1} \in \{1, 2, \ldots, m\}$, by (3.6.26) and the inductive hypothesis, we obtain

$$\begin{aligned}
\widetilde{F}_\tau^\alpha &:= E\left(\left|\int_\rho^\tau \widetilde{I}_{\alpha-}[g(\cdot)]_{\rho,s} dW_s^{j_{n+1}}\right|^{2q} \middle| \mathcal{A}_\rho\right) \\
&\leq 2^q\,(2q-1)^q(\tau-\rho)^{q-1} \int_\rho^\tau E\left(|\widetilde{I}_{\alpha-}[g(\cdot)]_{\rho,s}|^{2q} \middle| \mathcal{A}_\rho\right) ds \\
&\leq C\,(\tau-\rho)^{q-1}\,(\tau-\rho)^{q\left(l(\alpha-)+n(\alpha-)-s(\alpha-)\right)+s(\alpha-)-1} \\
&\qquad \times \int_\rho^\tau \int_\rho^s V_{\rho,z,s(\alpha-)} dz\, ds \\
&= C\,(\tau-\rho)^{q\left(l(\alpha)+n(\alpha)-s(\alpha)\right)+s(\alpha)-1} \int_\rho^\tau V_{\rho,z,s(\alpha)} dz,
\end{aligned}$$

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

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

-1. By (3.6.29) and the inductive hypothesis, we obtain

$$\begin{aligned}
\widetilde{F}_\tau^\alpha &:= E\left(\left|\int_\rho^\tau \int_{\mathcal{E}} \widetilde{I}_{\alpha-}[g(\cdot, v_{s(\alpha)})]_{\rho,s-} \widetilde{p}_\phi(dv_{s(\alpha)}, ds)\right|^{2q} \middle| \mathcal{A}_\rho\right) \\
&\leq C \int_\rho^\tau \int_{\mathcal{E}} E\left(|\widetilde{I}_{\alpha-}[g(\cdot, v_{s(\alpha)})]_{\rho,s}|^{2q} \middle| \mathcal{A}_\rho\right) \phi(dv_{s(\alpha)})\, ds \\
&\leq C\,(\tau-\rho)^{q\left(l(\alpha-)+n(\alpha-)-s(\alpha-)\right)+s(\alpha-)-1} \\
&\qquad \times \int_\rho^\tau \int_\rho^s \int_{\mathcal{E}} V_{\rho,z,s(\alpha-)} \phi(dv_{s(\alpha)}) dz\, ds \\
&\leq C\,(\tau-\rho)^{q\left(l(\alpha)+n(\alpha)-s(\alpha)\right)+s(\alpha)-1} \int_\rho^\tau V_{\rho,z,s(\alpha)} dz, \qquad (3.6.31)
\end{aligned}$$

where the last line holds since $l(\alpha) = l(\alpha-) + 1$, $n(\alpha) = n(\alpha-)$ and $s(\alpha) = s(\alpha-) + 1$, and this completes the proof of the assertion (3.6.23).

The assertion (3.6.24) can be proved similarly by induction on $l(\alpha)$. The case of $l(\alpha) = 1$ with $\alpha = (-1)$ has been already proved in (3.6.29). The case of $l(\alpha) = n+1$, with $\alpha = (j_1, \ldots, j_{n+1})$ and $j_{n+1} = -1$, can be proved by using (3.1.1), the Hölder inequality, the inductive hypothesis, (3.6.30) and (3.6.31). This completes the proof of Lemma 3.6.2. □

Let us introduce some notation needed for the following lemma. For any $z \in \mathbb{R}$, let us denote by $[z]$ the integer part of z. Moreover, for a given integer $p \in \mathbb{N}$ we will use the set $\mathcal{A}_p$ of multi-indices $\alpha = (j_1, \ldots, j_l)$ of length $l \leq p$ with components $j_i \in \{-1, 0\}$, for $i \in \{1, \ldots, l\}$. For a given function $h : [\rho, \tau] \times \mathbb{R}^d \to \mathbb{R}$, to be defined in the lemma below, and a multi-index $\alpha \in \mathcal{A}_p$, we consider the coefficient function $\tilde{f}_\alpha$ defined in (3.3.9), with $f(t,x) = h(t,x)$. For instance, if $\alpha = (-1, 0, 0)$, then $\tilde{f}_\alpha(t, X_t) = L_v^{(-1)} \widetilde{L}^{(0)} \widetilde{L}^{(0)} h(t, X_t)$.

Furthermore, for $p \in \mathbb{N}$, we denote by $\mathcal{C}^{p,2p}([\rho,\tau] \times \mathbb{R}^d, \mathbb{R})$ the set of functions that are p times continuously differentiable with respect to time and $2p$ times continuously differentiable with respect to the spatial variables.

By using Lemma 3.6.2 we prove the following result similar as in Liu & Li (2000).

Lemma 3.6.3 *Consider $\alpha \in \mathcal{M}_m$, $p = l(\alpha) - [\frac{l(\alpha)+n(\alpha)}{2}]$, and let ρ and τ be two stopping times with τ being $\mathcal{A}_\rho$-measurable and $0 \leq \rho \leq \tau \leq T$ almost surely.*

Moreover, define the process $Z = \{Z_t = h(t, X_t), t \in [\rho, \tau]\}$, *with* $h \in \mathcal{C}^{p,2p}([\rho, \tau] \times \mathbb{R}^d, \mathbb{R})$. *Assume that there exist a finite positive constant* K *such that for any* $\alpha \in \mathcal{A}_p$ *we have the estimate* $E\big(\tilde{f}_\alpha(t, X_t)^2 | \mathcal{A}_\rho\big) \leq K$ *a.s. for* $t \in [\rho, \tau]$. *Additionally, consider an adapted process* $g(\cdot) \in \mathcal{H}_\alpha$, *where* $g = g(\cdot, v)$ *with* $v \in \mathcal{E}^{s(\alpha)}$. *If there exists a positive,* $\phi(dv)$*-integrable function* $K(v)$ *such that* $E\big(g(t, v)^2 | \mathcal{A}_\rho\big) < K(v)$ *a.s for* $t \in [\rho, \tau]$, *then we obtain*

$$\left| E\Big(Z_\tau\, \tilde{I}_\alpha[g(\cdot)]_{\rho,\tau} | \mathcal{A}_\rho\Big)\right| \leq C_1\,(\tau - \rho)^{l(\alpha)}, \tag{3.6.32}$$

and

$$\left| E\Big(Z_\tau\, I_\alpha[g(\cdot)]_{\rho,\tau} | \mathcal{A}_\rho\Big)\right| \leq C_2\,(\tau - \rho)^{l(\alpha)}, \tag{3.6.33}$$

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

Proof: Let us first prove the assertion (3.6.32) by induction on $l(\alpha)$. For $l(\alpha) = 0$ we can prove (3.6.32) by applying the Cauchy-Schwarz inequality

$$\left| E\Big(Z_\tau\, g(\tau) | \mathcal{A}_\rho\Big)\right| \leq \sqrt{E\Big(Z_\tau^2 | \mathcal{A}_\rho\Big)}\sqrt{E\Big(g(\tau)^2 | \mathcal{A}_\rho\Big)} \leq C,$$

where we denote by C any constant that does not depend on $(\tau - \rho)$. However, it may depend on α, λ, T, h and g.

Now consider the case $l(\alpha) = n + 1$. By Itô's formula we have

$$\begin{aligned} Z_\tau &= Z_\rho + \int_\rho^\tau L^{(0)} h(z, X_z) dz + \sum_{i=1}^m \int_\rho^\tau L^{(i)} h(z, X_z) dW_z^i \\ &\quad + \int_\rho^\tau \int_{\mathcal{E}} L_v^{(-1)} h(z, X_{z-}) p_\phi(dv, dz), \end{aligned}$$

where the operators $L^{(i)}$, with $i \in \{0, 1, \ldots, m\}$, and $L_v^{(-1)}$ are defined in (3.3.4)–(3.3.6).

1. If $\alpha = (j_1, \ldots, j_{l+1})$ with $j_{l+1} = -1$, we can write

$$\tilde{I}_\alpha[g(\cdot)]_{\rho,\tau} = \int_\rho^\tau \int_{\mathcal{E}} \tilde{I}_{\alpha-}[g(\cdot)]_{\rho,z-} p_\phi(dv, dz) - \int_\rho^\tau \int_{\mathcal{E}} \tilde{I}_{\alpha-}[g(\cdot)]_{\rho,z} \phi(dv) dz.$$

By the product rule of stochastic calculus we obtain

$$
\begin{aligned}
Z_\tau \widetilde{I}_\alpha[g(\cdot)]_{\rho,\tau} &= \int_\rho^\tau \Big\{ \big(L^{(0)} h(z, X_z)\big) \widetilde{I}_\alpha[g(\cdot)]_{\rho,z} \\
&\qquad - h(z, X_z) \int_{\mathcal{E}} \widetilde{I}_{\alpha-}[g(\cdot)]_{\rho,z} \phi(dv) \Big\} dz \\
&\quad + \sum_{i=1}^{m} \int_\rho^\tau \big(L^{(i)} h(z, X_z)\big) \widetilde{I}_\alpha[g(\cdot)]_{\rho,z} dW_z^i \\
&\quad + \int_\rho^\tau \int_{\mathcal{E}} \Big\{ h(z, X_{z-}) \widetilde{I}_{\alpha-}[g(\cdot)]_{\rho,z-} \\
&\qquad + \big(L_v^{(-1)} h(z, X_{z-})\big) \widetilde{I}_\alpha[g(\cdot)]_{\rho,z-} \\
&\qquad + \big(L_v^{(-1)} h(z, X_{z-})\big) \widetilde{I}_{\alpha-}[g(\cdot)]_{\rho,z-} \Big\} p_\phi(dv, dz).
\end{aligned}
$$

Therefore, by the properties of Itô's integral and the induction hypothesis we have

$$
\begin{aligned}
\Big| E\Big(Z_\tau \widetilde{I}_\alpha[g(\cdot)]_{\rho,\tau} | \mathcal{A}_\rho \Big) \Big| &= \Big| - \int_\rho^\tau E\Big(h(z, X_z) \int_{\mathcal{E}} \widetilde{I}_{\alpha-}[g(\cdot)]_{\rho,z} \phi(dv) | \mathcal{A}_\rho \Big) dz \\
&\quad + \int_\rho^\tau \int_{\mathcal{E}} E\Big(\big\{ h(z, X_z) + L_v^{(-1)} h(z, X_z) \big\} \\
&\qquad \times \widetilde{I}_{\alpha-}[g(\cdot)]_{\rho,z} | \mathcal{A}_\rho \Big) \phi(dv) dz \\
&\quad + \int_\rho^\tau E\Big(\big(L^{(0)} h(z, X_z)\big) \widetilde{I}_\alpha[g(\cdot)]_{\rho,z} | \mathcal{A}_\rho \Big) dz \\
&\quad + \int_\rho^\tau \int_{\mathcal{E}} E\Big(\big(L_v^{(-1)} h(z, X_z)\big) \widetilde{I}_\alpha[g(\cdot)]_{\rho,z} | \mathcal{A}_\rho \Big) \phi(dv) dz \Big| \\
&\leq C\, (\tau - \rho)^{n+1} \\
&\quad + \int_\rho^\tau \Big| E\Big(\big(\widetilde{L}^{(0)} h(z, X_z)\big) \widetilde{I}_\alpha[g(\cdot)]_{\rho,z} | \mathcal{A}_\rho \Big) \Big| dz, \qquad (3.6.34)
\end{aligned}
$$

where we have used the equality (3.3.7).

Note that $\widetilde{L}^{(0)} h \in \mathcal{C}^{p-1, 2(p-1)}([\rho, \tau] \times \mathbb{R}^d, \mathbb{R})$ and we have the estimate

$$
E\big(\big(\widetilde{L}^{(0)}(h(t, X_t))\big)^2 | \mathcal{A}_\rho \big) \leq K
$$

for $t \in [\rho, \tau]$. Therefore, by the same steps used so far, we can show that

$$\Big|E\Big(\big(\widetilde{L}^{(0)}h(\tau, X_z)\big)\widetilde{I}_\alpha[g(\cdot)]_{\rho,z}|\mathcal{A}_\rho\Big)\Big| \le C\,(z-\rho)^{n+1} + \int_\rho^z \Big|E\Big(\big(\widetilde{L}^{(0)}\big(\widetilde{L}^{(0)}h(z_1, X_{z_1})\big)\big)\widetilde{I}_\alpha[g(\cdot)]_{\rho,z_1}|\mathcal{A}_\rho\Big)\Big|dz_1, \tag{3.6.35}$$

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

By using (3.6.35) in (3.6.34), we obtain

$$\begin{aligned}\Big|E\Big(Z_\tau\widetilde{I}_\alpha[g(\cdot)]_{\rho,\tau}|\mathcal{A}_\rho\Big)\Big| &\le C\,(\tau-\rho)^{n+1} \\ &+ \int_\rho^\tau\int_\rho^{z_2} \Big|E\Big(\big(\widetilde{L}^{(0)}\big(\widetilde{L}^{(0)}h(z_1, X_{z_1})\big)\big)\widetilde{I}_\alpha[g(\cdot)]_{\rho,z_1}|\mathcal{A}_\rho\Big)\Big|dz_1dz_2.\end{aligned} \tag{3.6.36}$$

By applying again this procedure for $p-2$ times and by using the Cauchy-Schwarz inequality we obtain

$$\begin{aligned}\Big|E\Big(Z_\tau\widetilde{I}_\alpha[g(\cdot)]_{\rho,\tau}|\mathcal{A}_\rho\Big)\Big| &\le C\,(\tau-\rho)^{n+1} \\ &+ \int_\rho^\tau \cdots \int_\rho^{z_2} \Big|E\Big(\tilde{f}_{\alpha_p}(z_1, X_{z_1})\widetilde{I}_\alpha[g(\cdot)]_{\rho,z_1}|\mathcal{A}_\rho\Big)\Big|dz_1\dots dz_p \\ &\le C\,(\tau-\rho)^{n+1} \\ &+ \int_\rho^\tau \cdots \int_\rho^{z_2} \Big[E\Big(\tilde{f}_{\alpha_p}(z_1, X_{z_1})^2|\mathcal{A}_\rho\Big)\Big]^{\frac{1}{2}} \\ &\times\Big[E\Big(\big(\widetilde{I}_\alpha[g(\cdot)]_{\rho,z_1}\big)^2|\mathcal{A}_\rho\Big)\Big]^{\frac{1}{2}}dz_1\dots dz_p,\end{aligned} \tag{3.6.37}$$

where α_p is defined as the multi-index with length p and all zeros, which means $l(\alpha_p) = n(\alpha_p) = p$. Therefore, $\tilde{f}_{\alpha_p}(t, X_t)$ denotes the stochastic process resulting from applying p times the operator $L^{(0)}$, see (3.3.7), to $h(t, X_t)$. Note that C denotes here a different constant from that in (3.6.34).

Finally, by applying the result (3.6.23) of Lemma 3.6.2, with $q = 1$, to the last term in (3.6.37) and considering that one has the estimates
$E\big(\tilde{f}_{\alpha_p}(t, X_t)^2|\mathcal{A}_\rho\big) \le K < \infty$ and $E\big(g(t,v)^2|\mathcal{A}_\rho\big) \le K(v)$ a.s. for $t \in [\rho, \tau]$,

with $K(v)$ $\phi(dv)$-integrable, we obtain

$$\begin{aligned}\left|E\Big(Z_\tau \widetilde{I}_\alpha[g(\cdot)]_{\rho,\tau}|\mathcal{A}_\rho\Big)\right| &\leq C\,(\tau-\rho)^{n+1} \\ &\quad + K\int_\rho^\tau \cdots \int_\rho^{z_2} (\tau-\rho)^{\frac{l(\alpha)+n(\alpha)}{2}} dz_1 \ldots dz_p \\ &\leq C\,(\tau-\rho)^{l(\alpha)},\end{aligned}$$

which proves the case of $\alpha = (j_1, \ldots, j_{l+1})$ with $j_{l+1} = -1$.

2. If $\alpha = (j_1, \ldots, j_{l+1})$ with $j_{l+1} = 0$, we can write

$$\widetilde{I}_\alpha[g(\cdot)]_{\rho,\tau} = \int_\rho^\tau \widetilde{I}_{\alpha-}[g(\cdot)]_{\rho,z} dz.$$

The product rule for stochastic integrals yields

$$\begin{aligned}Z_\tau \widetilde{I}_\alpha[g(\cdot)]_{\rho,\tau} &= \int_\rho^\tau \Big\{\big(L^{(0)}h(z,X_z)\big)\widetilde{I}_\alpha[g(\cdot)]_{\rho,z} + h(z,X_z)\widetilde{I}_{\alpha-}[g(\cdot)]_{\rho,z}\Big\}dz \\ &\quad + \sum_{i=1}^m \int_\rho^\tau \big(L^{(i)}h(z,X_z)\big)\widetilde{I}_\alpha[g(\cdot)]_{\rho,z} dW_z^i \\ &\quad + \int_\rho^\tau \int_{\mathcal{E}} \Big\{\big(L_v^{(-1)}h(z,X_{z-})\big)\widetilde{I}_\alpha[g(\cdot)]_{\rho,z-}\Big\} p_\phi(dv,dz).\end{aligned}$$

Therefore, similar to (3.6.34), by the properties of Itô's integral and the induction hypothesis one obtains

$$\begin{aligned}\left|E\Big(Z_\tau \widetilde{I}_\alpha[g(\cdot)]_{\rho,\tau}|\mathcal{A}_\rho\Big)\right| &\leq C\,(\tau-\rho)^{n+1} \\ &\quad + \int_\rho^\tau \left|E\Big(\big(L^{(0)}h(z,X_z)\big)\widetilde{I}_\alpha[g(\cdot)]_{\rho,z}|\mathcal{A}_\rho\Big)\right| dz.\end{aligned}$$

Then, by applying again this procedure for $p-1$ times, and by using the Cauchy-Schwarz inequality and the same estimates as before, we have

$$\begin{aligned}\left|E\Big(Z_\tau \widetilde{I}_\alpha[g(\cdot)]_{\rho,\tau}|\mathcal{A}_\rho\Big)\right| &\leq C\,(\tau-\rho)^{n+1} + \int_\rho^\tau \cdots \int_\rho^{z_2} \Big[E\Big(\tilde{f}_{\bar{\alpha}_p}(z_1,X_{z_1})^2|\mathcal{A}_\rho\Big)\Big]^{\frac{1}{2}} \\ &\quad \times \Big[E\Big(\big(\widetilde{I}_\alpha[g(\cdot)]_{\rho,z_1}\big)^2|\mathcal{A}_\rho\Big)\Big]^{\frac{1}{2}} dz_1 \ldots dz_p \\ &\leq C\,(\tau-\rho)^{l(\alpha)},\end{aligned}$$

where $\tilde{f}_{\bar{\alpha}_p}(t, X_t)$ denotes the stochastic process resulting from applying p times the operator $L^{(0)}$, see (3.3.4), to $h(t, X_t)$.

3. Let us finally consider the case $\alpha = (j_1, \ldots, j_{l+1})$ with $j_{l+1} = (j)$ and $j \in \{1, \ldots, m\}$. Here, we have

$$\widetilde{I}_\alpha[g(\cdot)]_{\rho,\tau} = \int_\rho^\tau \widetilde{I}_{\alpha-}[g(\cdot)]_{\rho,z} dW_z^j.$$

Therefore, by applying the product rule for stochastic integrals we obtain

$$\begin{aligned} Z_\tau \widetilde{I}_\alpha[g(\cdot)]_{\rho,\tau} &= \int_\rho^\tau \Big\{ \big(L^{(0)} h(z, X_z)\big) \widetilde{I}_\alpha[g(\cdot)]_{\rho,z} + \big(L^{(j)} h(z, X_z)\big) \widetilde{I}_{\alpha-}[g(\cdot)]_{\rho,z} \Big\} dz \\ &\quad + \sum_{i=1}^m \int_\rho^\tau \big(L^{(i)} h(z, X_z)\big) \widetilde{I}_\alpha[g(\cdot)]_{\rho,z} dW_z^i \\ &\quad + \int_\rho^\tau h(z, X_z) \widetilde{I}_{\alpha-}[g(\cdot)]_{\rho,z} dW_z^j \\ &\quad + \int_\rho^\tau \int_{\mathcal{E}} \Big\{ \big(L_v^{(-1)} h(z, X_{z-})\big) \widetilde{I}_\alpha[g(\cdot)]_{\rho,z-} \Big\} p_\phi(dv, dz). \end{aligned}$$

Again, by the properties of the Itô integral and the induction hypothesis we have

$$\begin{aligned} \Big| E\Big(Z_\tau \widetilde{I}_\alpha[g(\cdot)]_{\rho,\tau} \big| \mathcal{A}_\rho \Big) \Big| &\leq C\,(\tau - \rho)^{n+1} \\ &\quad + \int_\rho^\tau \Big| E\Big(\big(L^{(0)} h(z, X_z)\big) \widetilde{I}_\alpha[g(\cdot)]_{\rho,z} \big| \mathcal{A}_\rho \Big) \Big| dz, \end{aligned}$$

as in the previous case of $\alpha = (j_1, \ldots, j_{l+1})$ with $j_{l+1} = 0$. Therefore, in the same way, we can obtain the assertion (3.6.32).

To prove the estimate (3.6.33) we need only to check the case of $l(\alpha) = n+1$ with $\alpha = (j_1, \ldots, j_{n+1})$ and $j_{n+1} = -1$.

By using the product rule of stochastic calculus one obtains

$$
\begin{aligned}
Z_\tau I_\alpha[g(\cdot)]_{\rho,\tau} &= \int_\rho^\tau \left(L^{(0)}h(z,X_z)\right) I_\alpha[g(\cdot)]_{\rho,z} dz \\
&\quad + \sum_{i=1}^m \int_\rho^\tau \left(L^{(i)}h(z,X_z)\right) I_\alpha[g(\cdot)]_{\rho,z} dW_z^i \\
&\quad + \int_\rho^\tau \int_{\mathcal{E}} \Big\{ h(z,X_{z-}) I_{\alpha-}[g(\cdot)]_{\rho,z-} + \left(L_v^{(-1)}h(z,X_{z-})\right) I_\alpha[g(\cdot)]_{\rho,z-} \\
&\qquad + \left(L_v^{(-1)}h(z,X_{z-})\right) I_{\alpha-}[g(\cdot)]_{\rho,z-} \Big\} p_\phi(dv,dz).
\end{aligned}
$$

By the properties of Itô integrals and the induction hypothesis we obtain

$$
\begin{aligned}
\left| E\Big(Z_\tau I_\alpha[g(\cdot)]_{\rho,\tau} | \mathcal{A}_\rho\Big)\right| &= \Bigg| \int_\rho^\tau \int_{\mathcal{E}} E\Big(\left\{ h(z,X_z) + L_v^{(-1)}h(z,X_z)\right\} \\
&\qquad \times I_{\alpha-}[g(\cdot)]_{\rho,z} | \mathcal{A}_\rho\Big) \phi(dv) dz \\
&\quad + \int_\rho^\tau E\Big(\left(L^{(0)}h(z,X_z)\right) I_\alpha[g(\cdot)]_{\rho,z} | \mathcal{A}_\rho\Big) dz \\
&\quad + \int_\rho^\tau \int_{\mathcal{E}} E\Big(\left(L_v^{(-1)}h(z,X_z)\right) I_\alpha[g(\cdot)]_{\rho,z} | \mathcal{A}_\rho\Big) \phi(dv) dz \Bigg| \\
&\le C\,(\tau-\rho)^{n+1} \\
&\quad + \int_\rho^\tau \left| E\Big(\left(\widetilde{L}^{(0)}h(z,X_z)\right) I_\alpha[g(\cdot)]_{\rho,z} | \mathcal{A}_\rho\Big)\right| dz, \qquad (3.6.38)
\end{aligned}
$$

where we have used again the equality (3.3.7).

From this point we can proceed in the same way as in the proof of assertion (3.6.32). This completes the proof of Lemma 3.6.3. □

We also propose the following lemma, similar to a result in Liu & Li (2000).

Lemma 3.6.4 *Let $\alpha \in \mathcal{M}_m$, ρ and τ denote two stopping times with τ being $\mathcal{A}_\rho$-measurable and $0 \le \rho \le \tau \le T$ almost surely. Moreover, let $h = \{h(t), t \in [\rho,\tau]\}$ be an adapted process such that $E\big(h(t)^2 | \mathcal{A}_\rho\big) \le K < \infty$ for $t \in [\rho,\tau]$. Additionally, consider an adapted process $g(\cdot) \in \mathcal{H}_\alpha$, where $g = g(\cdot, v)$ with $v \in \mathcal{E}^{s(\alpha)}$. If for a given $q \in \{1, 2 \ldots\}$ there exists a positive, $\phi(dv)$-integrable function $K(v)$ such that*

$E\big(g(t,v)^{2^{s(\alpha)+3}q}|\mathcal{A}_\rho\big) < K(v)$ a.s. for $t \in [\rho,\tau]$, then we obtain

$$\widetilde{F}^\alpha_\tau := E\left(h(\tau)\left|\widetilde{I}_\alpha[g(\cdot)]_{\rho,\tau}\right|^{2q}\Big|\mathcal{A}_\rho\right) \leq C_1\,(\tau-\rho)^{q\left(l(\alpha)+n(\alpha)-s(\alpha)\right)+s(\alpha)}, \qquad (3.6.39)$$

and

$$F^\alpha_\tau := E\left(h(\tau)\,|I_\alpha[g(\cdot)]_{\rho,\tau}|^{2q}\Big|\mathcal{A}_\rho\right) \leq C_2\,(\tau-\rho)^{q\left(l(\alpha)+n(\alpha)-s(\alpha)\right)+s(\alpha)}, \qquad (3.6.40)$$

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

Proof: We first prove the estimate (3.6.39) by induction with respect to $s(\alpha)$.

1. If $s(\alpha) = 0$, then by the Cauchy-Schwarz inequality and Lemma 3.6.2 we obtain

$$\begin{aligned}\widetilde{F}^\alpha_\tau &= E\left(h(\tau)\left|\widetilde{I}_\alpha[g(\cdot)]_{\rho,\tau}\right|^{2q}\Big|\mathcal{A}_\rho\right)\\ &\leq \left[E\left(h(\tau)^2\Big|\mathcal{A}_\rho\right)\right]^{\frac{1}{2}}\left[E\left(\left|\widetilde{I}_\alpha[g(\cdot)]_{\rho,\tau}\right|^{4q}\Big|\mathcal{A}_\rho\right)\right]^{\frac{1}{2}}\\ &\leq C\,(\tau-\rho)^{q\left(l(\alpha)+n(\alpha)\right)}.\end{aligned}$$

   Note that here and in the following we denote again by C any positive constant that does not depend on $(\tau-\rho)$.

2. Let us consider the case $s(\alpha) = s \in \{1,2,\ldots\}$.

   By the relationship 3.2.7 and some straightforward estimates we obtain

$$\begin{aligned}\widetilde{F}^\alpha_\tau \leq\; & C\Big\{E\left(h(\tau)\,|I_\alpha[g(\cdot)]_{\rho,\tau}|^{2q}\Big|\mathcal{A}_\rho\right)\\ &+\sum_{i=1}^{2^{s(\alpha)}-1} E\left(h(\tau)\,|H_{\alpha,i}|^{2q}\Big|\mathcal{A}_\rho\right)\Big\},\end{aligned} \qquad (3.6.41)$$

   where terms $H_{\alpha,i}$ are described in Remark 3.2.1.

   Since $s(\alpha) = s \in \{1,2,\ldots\}$, we can write $\alpha = \alpha_1 \star (-1) \star \alpha_2 \star (-1) \ldots \star (-1) \star \alpha_{s+1}$, where $s(\alpha_j) = 0$ for $j \in \{1,2,\ldots,s+1\}$ and $\star$ denotes the concatenation operation on multi-indices defined in (3.2.2).

Thus,

$$I_{\alpha_1 \star (-1)}[g(\cdot)]_{\rho,\tau} = \sum_{i=p_\phi(\rho)+1}^{p_\phi(\tau)} I_{\alpha_1}[g(\cdot)]_{\rho,\tau_i},$$

where τ_i, with $i \in \{1, 2, \ldots, p_\phi(T)\}$, are the jump times generated by the Poisson measure. Similarly, one can show that

$$I_\alpha[g(\cdot)]_{\rho,\tau} = \sum_{i=p_\phi(\rho)+1}^{p_\phi(\tau)} I_{\alpha_1}[g(\cdot)]_{\rho,\tau_i}\, I_{\alpha_2;\tau_i,\tau_{i+1}} \ldots I_{\alpha_{s+1};\tau_{i+s-1},\tau},$$

where $I_{\alpha_j;\tau_i,\tau_{i+1}} = I_{\alpha_j}[1]_{\tau_i,\tau_{i+1}}$ denotes the multiple stochastic integral for the multi-index α_j over the time interval $[\tau_i, \tau_{i+1}]$. Let us note that $I_\alpha[g(\cdot)]_{\rho,\tau} = 0$ if $p_\phi([\rho,\tau)) < s$. Therefore, by the Cauchy-Schwarz inequality and Lemma 3.6.2, we obtain

$$\begin{aligned}
&E\left(h(\tau)\ |I_\alpha[g(\cdot)]_{\rho,\tau}|^{2q} \Big| \mathcal{A}_\rho\right) \\
=\ & \sum_{n\geq s} E\left(h(\tau)\left|\sum_{i=1}^{n-s+1} I_{\alpha_1}[g(\cdot)]_{\rho,\tau_i}\, I_{\alpha_2;\tau_i,\tau_{i+1}} \ldots I_{\alpha_{s+1};\tau_{i+s-1},\tau}\right|^{2q} \Big| \mathcal{A}_\rho\right) \\
& \quad \times P\big(p_\phi([\rho,\tau)) = n\big) \\
\leq\ & e^{-\lambda(\tau-\rho)} \sum_{n\geq s} (n-s+1)^{2q-1} \frac{\big(\lambda(\tau-\rho)\big)^n}{n!} \sum_{i=1}^{n-s+1} \left[E\left(h(\tau)^2 \Big| \mathcal{A}_\rho\right)\right]^{\frac{1}{2}} \\
& \times \left[E\left(\left| I_{\alpha_1}[g(\cdot)]_{\rho,\tau_i}\right|^{8q} \Big| \mathcal{A}_\rho\right)\right]^{\frac{1}{4}} \left[E\left(\left| I_{\alpha_2;\tau_i,\tau_{i+1}}\right|^{16q} \Big| \mathcal{A}_\rho\right)\right]^{\frac{1}{8}} \\
& \times \ldots \times \left[E\left(\left| I_{\alpha_{s+1};\tau_{i+s-1},\tau}\right|^{2^{s+3}q} \Big| \mathcal{A}_\rho\right)\right]^{\frac{1}{2^{s+2}}} \\
\leq\ & C e^{-\lambda(\tau-\rho)} \sum_{n\geq s} (n-s+1)^{2q-1} \frac{\big(\lambda(\tau-\rho)\big)^n}{n!} \\
& \quad \times \sum_{i=1}^{n-s+1} (\tau-\rho)^{q\sum_{i=1}^{s+1}\left(l(\alpha_i)+n(\alpha_i)\right)} \\
\leq\ & C\,(\tau-\rho)^{q\left(l(\alpha)+n(\alpha)-s(\alpha)\right)}\, e^{-\lambda(\tau-\rho)}\,(\tau-\rho)^s \\
& \times \sum_{n\geq s} (n-s+1)^{2q} \frac{\lambda^n(\tau-\rho)^{n-s}}{n!}
\end{aligned}$$

$$\leq \quad C\,(\tau-\rho)^{q\left(l(\alpha)+n(\alpha)-s(\alpha)\right)+s(\alpha)}. \tag{3.6.42}$$

Note that the last line of (3.6.42) follows by

$$\begin{aligned}
&\sum_{n\geq s}(n-s+1)^{2q}\frac{\lambda^n(\tau-\rho)^{n-s}}{n!}\\
&\quad\leq (\lambda)^s\sum_{j=0}^{\infty}(j+1)^{2q}(j+s)\ldots(j+1)\frac{\left(\lambda(\tau-\rho)\right)^j}{j!}\\
&\quad\leq K\sum_{j=0}^{\infty}(j+s)^{2q+s}\,\frac{\left(\lambda(\tau-\rho)\right)^j}{j!}\\
&\quad\leq K\left\{\sum_{j=0}^{\infty}\frac{\left(\lambda(\tau-\rho)\right)^j}{j!}+\sum_{j=0}^{\infty}j^{2q+s}\,\frac{\left(\lambda(\tau-\rho)\right)^j}{j!}\right\}\\
&\quad= K\,e^{\lambda(\tau-\rho)}\left(1+B_{2q+s}(\lambda(\tau-\rho))\right)\\
&\quad\leq K\,e^{\lambda(\tau-\rho)},
\end{aligned}$$

where we denote by $B_n(x)$ the Bell polynomial of degree n, see Bell (1934).

To complete the proof we have to show that the bound (3.6.42) also holds for the other terms in (3.6.41). Note that, as shown in Remark 3.2.1, the terms $H_{\alpha,i}$, for $i\in\{1,\ldots,2^{s(\alpha)}-1\}$, involve multiple stochastic integrals with the same multiplicity $l(\alpha)$ of the multiple stochastic integrals I_α and replace some of the integrations with respect to the Poisson jump measure p_ϕ with integrations with respect to time and to the intensity measure ϕ. Therefore, by following the steps above, one can easily establish the bound (3.6.42) for all terms in (3.6.41). This completes the proof of Lemma 3.6.4. $\square$

3.7 Weak Truncated Expansions

In this section we present weak truncated Wagner-Platen expansions that will be used in the construction of weak schemes for Monte Carlo simulation.

For $\beta \in \{1, 2, \ldots\}$ we define the hierarchical set

$$\Gamma_\beta = \{\alpha \in \mathcal{M}_m : l(\alpha) \leq \beta\}. \tag{3.7.1}$$

By (3.7.1) and (3.4.1) the corresponding remainder set is then

$$\mathcal{B}(\Gamma_\beta) = \{\alpha \in \mathcal{M}_m : l(\alpha) = \beta + 1\}. \tag{3.7.2}$$

We consider the truncated *weak Itô-Taylor expansion*

$$\eta_t = \sum_{\alpha \in \Gamma_\beta} I_\alpha[f_\alpha(0, X_0)]_{0,t}, \tag{3.7.3}$$

for $t \in [0, T]$, where f_α are the Itô coefficient functions defined in (3.3.8) corresponding to $f(t, x) = x$. Moreover, we assume that the coefficients of the SDE (2.1.2) are sufficiently smooth and integrable such that for any $\alpha \in \Gamma_\beta \bigcup \mathcal{B}(\Gamma_\beta)$ the coefficient functions f_α are well defined and all the multiple stochastic integrals exist. We also have the truncated *compensated weak Itô-Taylor expansion*

$$\widetilde{\eta}_t = \sum_{\alpha \in \Gamma_\beta} \widetilde{I}_\alpha[\tilde{f}_\alpha(0, X_0)]_{0,t}, \tag{3.7.4}$$

for $t \in [0, T]$, where $\tilde{f}_\alpha$ are the Itô coefficient functions defined in (3.3.9) corresponding to $f(t, x) = x$.

In what follows we shall fix a $\beta \in \{1, 2, \ldots\}$ and write

$$\eta_t = \sum_{\alpha \in \Gamma_\beta} I_\alpha[f_\alpha(X_0)]_t, \tag{3.7.5}$$

and

$$\widetilde{\eta}_t = \sum_{\alpha \in \Gamma_\beta} \widetilde{I}_\alpha[\tilde{f}_\alpha(X_0)]_t, \tag{3.7.6}$$

for $t \in [0, T]$. Let us denote by X_t^y the solution of the SDE (2.1.2) that starts from $y \in \mathbb{R}^d$ at time $t = 0$.

Lemma 3.7.1 *Let $\beta \in \{1, 2, \ldots\}$ and $T \in (0, \infty)$ be given and suppose that the drift, diffusion and jump coefficients of the SDE* (2.1.2) *are smooth enough to apply the Wagner-Platen expansions with hierarchical set Γ_β, see also Remark 3.5.2.*

Then we have

$$X_t^{X_0} - \eta_t = \sum_{\alpha \in \mathcal{B}(\Gamma_\beta)} I_\alpha[f_\alpha(X_\cdot^{X_0})]_t \tag{3.7.7}$$

and

$$X_t^{X_0} - \tilde{\eta}_t = \sum_{\alpha \in \mathcal{B}(\Gamma_\beta)} \tilde{I}_\alpha[\tilde{f}_\alpha(X_\cdot^{X_0})]_t, \tag{3.7.8}$$

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

Proof: By using (3.7.5) we obtain

$$\begin{aligned} X_t^{X_0} - \eta_t &= X_t^{X_0} - X_0 - (\eta_t - X_0) \\ &= X_t^{X_0} - X_0 - \sum_{\alpha \in \Gamma_\beta \setminus \{v\}} I_\alpha[f_\alpha(X_0)]_t. \end{aligned} \tag{3.7.9}$$

Moreover, by the Wagner-Platen expansion (3.5.4) and (3.7.1) we have

$$X_t^{X_0} - X_0 = \sum_{\alpha \in \Gamma_\beta \setminus \{v\}} I_\alpha[f_\alpha(X_0)]_t + \sum_{\alpha \in \mathcal{B}(\Gamma_\beta)} I_\alpha[f_\alpha(X_\cdot^{X_0})]_t.$$

Inserting the last equation in (3.7.9) we obtain (3.7.7). In the same way one can prove (3.7.8) $\square$

By similar arguments as those used in the standard proof of the existence and uniqueness of the solution of the SDE (2.1.2), see Ikeda & Watanabe (1989), one obtains the following lemma.

Lemma 3.7.2 *Let $T \in (0, \infty)$ be given and assume that the drift, diffusion and jump coefficients of the SDE (2.1.2) satisfy Lipschitz and linear growth conditions. Then, for each $p \in \{1, 2, \ldots\}$ there exists a constant $K \in (0, \infty)$ such that*

$$E\Big(\sup_{0 \le t \le T} |X_t^y|^{2q} \Big| \mathcal{A}_0\Big) \le K\left(1 + |y|^{2q}\right), \tag{3.7.10}$$

for all $y \in \Re^d$, $t \in [0, T]$ and $q \in \{1, \ldots, p\}$.

Chapter 4

Regular Strong Taylor Approximations

In this chapter we present regular strong approximations obtained directly from a truncated Wagner-Platen expansion. The desired strong order of convergence determines how many terms of the stochastic expansion one should include in the approximation. We call these schemes *regular* strong approximations as opposed to the *jump-adapted* strong approximations that will be presented in Chapter 6. The term *regular* refers to the time discretizations used to construct these approximations. These are called regular because they do not include the jump times of the Poisson random measure. A convergence theorem for approximations of any given strong order of convergence $\gamma \in \{0.5, 1, 1.5, 2, \ldots\}$ will be presented at the end of this chapter. Some of the results in this chapters have been published in Bruti-Liberati, Nikitopoulos-Sklibosios & Platen (2006) and Bruti-Liberati & Platen (2007c).

4.1 Introduction

First we consider, for simplicity, the one-dimensional SDE, $d = 1$, in the form

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t + \int_{\mathcal{E}} c(t, X_{t-}, v)\, p_\phi(dv, dt), \tag{4.1.1}$$

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. As previously, we assume an $\underline{\mathcal{A}}$-adapted Poisson measure $p_\phi(dv, dt)$ with mark space $\mathcal{E} \subseteq \mathbb{R}\backslash\{0\}$, and with intensity measure $\phi(dv)\, dt = \lambda\, F(dv)\, dt$, where $F(\cdot)$ is a given probability distribution function for the realizations of the marks.

The SDE (4.1.1) can be written in integral form as

$$\begin{aligned} X_t &= X_0 + \int_0^t a(s, X_s)ds + \int_0^t b(s, X_s)dW_s + \int_0^t \int_{\mathcal{E}} c(s, X_{s-}, v)\, p_\phi(dv, ds) \\ &= X_0 + \int_0^t a(s, X_s)ds + \int_0^t b(s, X_s)dW_s + \sum_{i=1}^{p_\phi(t)} c(\tau_i, X_{\tau_i -}, \xi_i), \end{aligned} \tag{4.1.2}$$

where $\{(\tau_i, \xi_i), i \in \{1, 2 \ldots, p_\phi(t)\}\}$ is the double sequence of jump times and corresponding marks generated by the Poisson random measure. We express the ith mark at time τ_i by $\xi_i \in \mathcal{E}$. For simplicity, we have assumed a one-dimensional mark space $\mathcal{E} \subseteq \mathbb{R} \backslash \{0\}$. Multi-dimensional mark spaces can be similarly considered.

The case of a mark-independent jump size, which means $c(t, x, v) = c(t, x)$, is of particular interest as it simplifies the derivation and the implementation of numerical schemes. Therefore, we also consider a one-dimensional SDE with mark-independent jump size, which, in integral form, is given by

$$X_t = X_0 + \int_0^t a(s, X_s)ds + \int_0^t b(s, X_s)dW_s + \sum_{i=1}^{p_\phi(t)} c(\tau_i, X_{\tau_i -}). \tag{4.1.3}$$

Later, we will present strong Taylor approximations for a d-dimensional SDE, as introduced in Section 2.1, given by

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t + \int_{\mathcal{E}} c(t, X_{t-}, v)\, p_\phi(dv, dt), \tag{4.1.4}$$

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. Moreover, p_ϕ is again an $\underline{\mathcal{A}}$-adapted Poisson measure. 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 $[0, T] \times \mathbb{R}^d$. We recall that in this thesis we use superscripts to denote vector components.

Moreover, we consider a *regular time discretization* $0 = t_0 < t_1 < \ldots < t_N = T$, on which we will construct a *discrete time approximation* $Y^\Delta = \{Y_t^\Delta, t \in [0, T]\}$ of the solution of the SDE (4.1.4). For a given maximum time step size $\Delta \in (0, 1)$ we require the regular time discretization

$$(t)_\Delta = \{0 = t_0 < t_1 < \ldots < t_N = T\}, \tag{4.1.5}$$

to satisfy the following conditions:

$$P(t_{n+1} - t_n \leq \Delta) = 1, \tag{4.1.6}$$

where

$$t_{n+1} \quad \text{is} \quad \mathcal{A}_{t_n} - \text{measurable}, \tag{4.1.7}$$

for $n \in \{0, 1, \ldots, N-1\}$ and

$$n_t := \max\{n \in \{0, 1, \ldots\} : t_n \leq t\} < \infty \quad \text{a.s.}, \tag{4.1.8}$$

denoting the largest integer n such that $t_n \leq t$, for all $t \in [0, T]$. Such a time discretization is called *regular*, as opposed to the jump-adapted one to be presented later in Chapter 6, because it does not include the jump times generated by the Poisson measure. For instance, we could consider an equidistant time discretization with nth discretization time $t_n = n\Delta$, $n \in \{0, 1, \ldots, N\}$, and time step size $\Delta = \frac{T}{N}$. However, the discretization times could also be random, as needed if one wants to employ a step size control. Conditions (4.1.6),(4.1.7) and (4.1.8) pose some restrictions on the choice of the random discretization times. Condition (4.1.6) requests that the maximum step size in the time discretization cannot be larger than Δ. Condition (4.1.7) ensures that the length $\Delta_n = t_{n+1} - t_n$ of the next time step depends only on the information available at the last discretization time t_n. Condition (4.1.8) guarantees a finite number of discretization points in any bounded interval $[0, t]$.

For simplicity, when describing the schemes we will use the abbreviation

$$f = f(t_n, Y_n) \tag{4.1.9}$$

for a function f when no misunderstanding is possible. For the jump coefficient we may also write

$$c(v) = c(t_n, Y_n, v) \qquad \text{and} \qquad c(\xi_i) = c(t_n, Y_n, \xi_i), \tag{4.1.10}$$

if convenient, where ξ_i is the ith mark of the Poisson measure. Similarly, we write

$$c'(v) = c'(t_n, Y_n, v) \qquad \text{and} \qquad c'(\xi_i) = c'(t_n, Y_n, \xi_i). \tag{4.1.11}$$

Note that here and in the sequel the prime $'$ in (4.1.11) denotes the derivative with respect to the second argument, which is the spatial variable. Moreover, we will omit mentioning the initial value Y_0 and the step numbers $n \in \{0, 1, \ldots, N\}$ if this is not confusing.

4.2 Euler Scheme

The simplest scheme is the well-known *Euler scheme*, which, in the one-dimensional case $d = m = 1$, is given by the algorithm

$$\begin{aligned} Y_{n+1} &= Y_n + a\Delta_n + b\Delta W_n + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} c(v)\, p_\phi(dv, dz) \\ &= Y_n + a\Delta_n + b\Delta W_n + \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} c(\xi_i), \end{aligned} \tag{4.2.1}$$

for $n \in \{0, 1, \ldots, N-1\}$ with initial value $Y_0 = X_0$. Note that $a = a(t_n, Y_n)$, $b = b(t_n, Y_n)$, $c(v) = c(t_n, Y_n, v)$ and $c(\xi_i) = c(t_n, Y_n, \xi_i)$, according to the abbreviation defined in (4.1.9)–(4.1.11). Here

$$\Delta_n = t_{n+1} - t_n = I_{0,n} \tag{4.2.2}$$

is the length of the time step size $[t_n, t_{n+1}]$ and

$$\Delta W_n = W_{t_{n+1}} - W_{t_n} \tag{4.2.3}$$

is the nth Gaussian $\mathcal{N}(0, \Delta_n)$ distributed increment of the Wiener process W, $n \in \{0, 1, \ldots, N-1\}$. Furthermore,

$$p_\phi(t) = p_\phi(\mathcal{E}, [0, t]) \tag{4.2.4}$$

represents the total number of jumps of the Poisson random measure up to time t, which is Poisson distributed with mean $\lambda\, t$. Finally,

$$\xi_i \in \mathcal{E} \tag{4.2.5}$$

is the ith mark of the Poisson random measure p_ϕ at the ith jump time τ_i, with distribution function $F(\cdot)$. The Euler scheme (4.2.1) generally achieves a strong order of convergence $\gamma = 0.5$, as will be shown at the end of this chapter.

When we have a mark-independent jump size, which means $c(t, x, v) = c(t, x)$, the Euler scheme reduces to

$$Y_{n+1} = Y_n + a\Delta_n + b\Delta W_n + c\Delta p_n, \tag{4.2.6}$$

where

$$\Delta p_n = p_\phi(t_{n+1}) - p_\phi(t_n) \tag{4.2.7}$$

follows a Poisson distribution with mean $\lambda\Delta_n$.

In the multi-dimensional case with scalar driving Wiener process, which means $m = 1$, and mark-dependent jump size, the kth component of the *Euler scheme* is given by

$$Y^k_{n+1} = Y^k_n + a^k\Delta_n + b^k\Delta W_n + \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} c^k(\xi_i), \tag{4.2.8}$$

for $k \in \{1, 2, \ldots, d\}$, where a^k, b^k, and c^k are the kth components of the drift, diffusion and jump coefficients, respectively.

In the multi-dimensional case with scalar Wiener process, $m = 1$, and mark-independent jump size, the kth component of the *Euler scheme* is given by

$$Y^k_{n+1} = Y^k_n + a^k\Delta_n + b^k\Delta W_n + c^k\Delta p_n, \tag{4.2.9}$$

for $k \in \{1, 2, \ldots, d\}$.

For the general multi-dimensional case with mark-dependent jump size the kth component of the *Euler scheme* is of the form

$$Y^k_{n+1} = Y^k_n + a^k\Delta_n + \sum_{j=1}^{m} b^{k,j}\Delta W^j_n + \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} c^k(\xi_i),$$

where a^k and c^k are the kth components of the drift and the jump coefficients, respectively, and $b^{k,j}$ is the component of the kth row and jth column of the diffusion matrix b, for $k \in \{1, 2, \ldots, d\}$, and $j \in \{1, 2, \ldots, m\}$. Moreover,

$$\Delta W^j_n = W^j_{t_{n+1}} - W^j_{t_n} \tag{4.2.10}$$

is the $\mathcal{N}(0, \Delta_n)$ distributed nth increment of the jth Wiener process. We recall that

$$\xi_i \in \mathcal{E} \tag{4.2.11}$$

is the ith mark generated by the Poisson random measure.

In the multi-dimensional case with mark-independent jump size we obtain the kth component of the *Euler scheme*

$$Y_{n+1}^k = Y_n^k + a^k \Delta_n + \sum_{j=1}^{m} b^{k,j} \Delta W_n^j + c^k \Delta p_n,$$

for $k \in \{1, 2, \ldots, d\}$.

The Euler scheme includes only the time integral and the stochastic integrals of multiplicity one from the Wagner-Platen expansions (3.5.4) and (3.5.5), which both give the same truncated expansion. As we will see later, it can be interpreted as the order 0.5 strong Taylor scheme. In the following we will use the Wagner-Platen expansion (3.5.4) for the construction of strong Taylor schemes. Similarly, one can start from the compensated Wagner-Platen expansion (3.5.5) obtaining slightly different compensated Taylor schemes, as will be discussed later.

4.3 Order 1.0 Taylor Scheme

When accuracy and efficiency are required, it is important to construct numerical methods with higher strong order of convergence. This can be achieved by adding more terms of the Wagner-Platen expansion (3.5.4) to the scheme. In this way, it is possible to derive the *order 1.0 strong Taylor scheme*, which, in the one-dimensional case, $d = m = 1$, is given by

$$\begin{aligned}
Y_{n+1} &= Y_n + a\Delta_n + b\Delta W_n + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} c(v)\, p_\phi(dv, dz) \\
&\quad + bb' \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} dW(z_1) dW(z_2) \\
&\quad + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} b\, c'(v) dW(z_1) p_\phi(dv, dz_2) \\
&\quad + \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} \int_{\mathcal{E}} \Big\{ b\big(t_n, Y_n + c(v)\big) - b \Big\} p_\phi(dv, dz_1) dW(z_2) \qquad (4.3.1) \\
&\quad + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} \int_{\mathcal{E}} \Big\{ c\big(t_n, Y_n + c(v_1), v_2\big) - c(v_2) \Big\} p_\phi(dv_1, dz_1) p_\phi(dv_2, dz_2),
\end{aligned}$$

where

$$b' := b'(t,x) = \frac{\partial\, b(t,x)}{\partial\, x} \quad \text{and} \quad c'(v) := c'(t,x,v) = \frac{\partial c(t,x,v)}{\partial x}. \qquad (4.3.2)$$

For simplicity, the abbreviation (4.1.9)–(4.1.11) were used in (4.3.1). The scheme (4.3.1) achieves strong order $\gamma = 1.0$, as we will see later. It represents a generalization of the Milstein scheme, see Milstein (1974), to the case of jump diffusions.

In view of applications to scenario simulations, a main problem concerns the generation of the multiple stochastic integrals appearing in (4.3.1). By application of Itô's formula for jump-diffusion processes, see Ikeda & Watanabe (1989), and the integration by parts formula, we can simplify the four double stochastic integrals appearing in (4.3.1) and rewrite the order 1.0 strong Taylor scheme (4.3.1) as

$$\begin{aligned}
Y_{n+1} &= Y_n + a\Delta_n + b\Delta W_n + \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} c(\xi_i) + \frac{b\,b'}{2}\big((\Delta W_n)^2 - \Delta_n\big) \\
&\quad + b \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} c'(\xi_i)\big(W(\tau_i) - W(t_n)\big) \\
&\quad + \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} \Big\{ b\big(Y_n + c(\xi_i)\big) - b \Big\} \big(W(t_{n+1}) - W(\tau_i)\big) \\
&\quad + \sum_{j=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} \sum_{i=p_\phi(t_n)+1}^{p_\phi(\tau_j)} \Big\{ c\big(Y_n + c(\xi_i), \xi_j\big) - c\big(\xi_j\big) \Big\}. \qquad (4.3.3)
\end{aligned}$$

This scheme is readily applicable for a scenario simulation.

In the special case of a mark-independent jump coefficient, $c(t,x,v) = c(t,x)$, the order 1.0 strong Taylor scheme reduces to

$$\begin{aligned} Y_{n+1} &= Y_n + a\Delta_n + b\Delta W_n + c\Delta p_n + b\,b'\,I_{(1,1)} + b\,c'\,I_{(1,-1)} \\ &\quad + \{b(t_n, Y_n + c) - b\}\,I_{(-1,1)} + \{c\,(t_n, Y_n + c) - c\}\,I_{(-1,-1)}, \quad (4.3.4) \end{aligned}$$

with multiple stochastic integrals

$$\begin{aligned} I_{(1,1)} &:= \int_{t_n}^{t_{n+1}} \int_{t_n}^{s_2} dW_{s_1} dW_{s_2} = \frac{1}{2}\Big\{(\Delta W_n)^2 - \Delta_n\Big\}, \\ I_{(1,-1)} &:= \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{s_2} dW_{s_1} p_\phi(dv, ds_2) = \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} W_{\tau_i} - \Delta p_n\, W_{t_n}, \\ I_{(-1,1)} &:= \int_{t_n}^{t_{n+1}} \int_{t_n}^{s_2} \int_{\mathcal{E}} p_\phi(dv, ds_1) dW_{s_2} = \Delta p_n\, \Delta W_n - I_{(1,-1)}, \quad (4.3.5) \\ I_{(-1,-1)} &:= \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{s_2} \int_{\mathcal{E}} p_\phi(dv_1, ds_1) p_\phi(dv_2, ds_2) = \frac{1}{2}\Big\{(\Delta p_n)^2 - \Delta p_n\Big\}. \end{aligned}$$

The level of complexity of the scheme (4.3.1), even in the simpler case (4.3.4) of mark-independent jump size, is quite substantial when compared to the Euler scheme (4.2.6). Indeed, it requires not only function evaluations of the drift, diffusion and jump coefficients, but also evaluations of their derivatives. The calculation of derivatives can be avoided by constructing derivative-free schemes that will be presented in the next chapter.

We point out that the simulation of some of the multiple stochastic integrals is computationally demanding. 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 since it requires to keep track of the jump times between discretization 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 the number of jumps Δp_n, we can sample Δp_n independent outcomes from a uniform distribution on $(t_n, t_{n+1}]$ in order to obtain the exact location of the jump times. However, from the computational point of view, this demonstrates that the efficiency of the algorithm is heavily dependent on the level of the intensity of the Poisson measure.

For the special case of a mark-independent jump coefficient, the number of calculations involved in an algorithm as the Euler scheme (4.2.6) does not depend on the level of the intensity. Note that we are here neglecting the additional time needed to sample from a Poisson distribution with higher intensity. On the other hand, even in this special case, for the scheme (4.3.4) the number of computations is directly related to the number of jumps because of the generation of the two double stochastic integrals $I_{(1,-1)}$ and $I_{(-1,1)}$. Therefore, this algorithm is not efficient for the simulation of jump-diffusion SDEs driven by a Poisson measure with high intensity.

It is, in principle, possible to derive strong Taylor schemes of any given order, as will be demonstrated later in Section 4.5. However, the schemes become rather complex. Moreover, as explained above, for SDEs driven by high intensity Poisson measures, these schemes are computationally 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 Chapter 6, where we will describe jump-adapted approximations that avoid multiple stochastic integrals involving the Poisson measure. This makes it much easier to derive and implement these schemes.

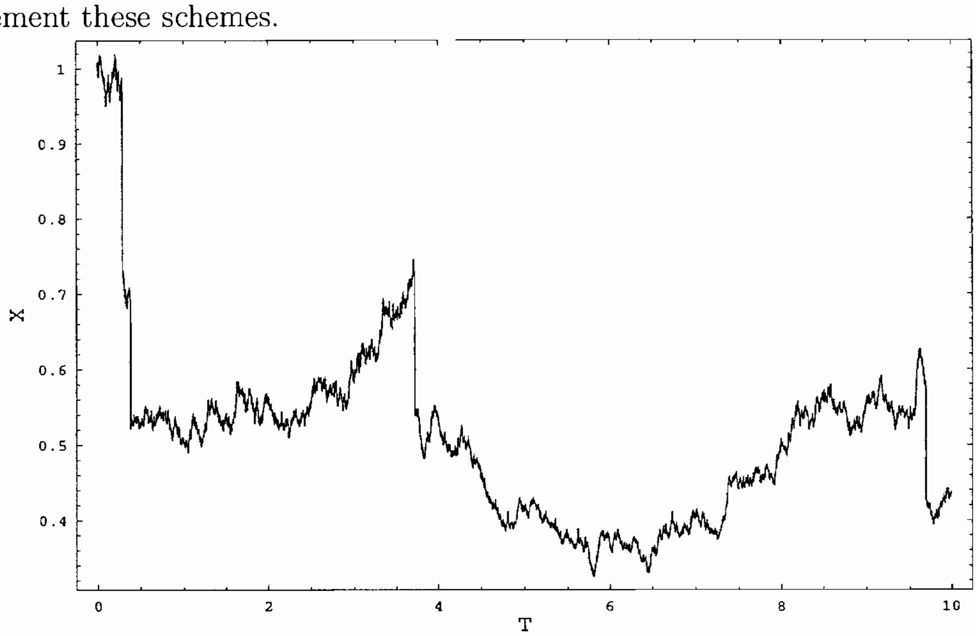

Figure 4.3.1: A path of the Merton model

Let us illustrate the higher accuracy achieved by the order 1.0 Taylor scheme on a scenario simulation. We consider the SDE (2.1.5) describing the Merton model with constant jump coefficient $c(t,x) = \psi x$. We use the following parameters: $X_0 = 1$, $\mu = 0.05$, $\sigma = 0.2$, $\psi = -0.25$, $T = 10$ and $\lambda = 0.3$. In Figure 4.3.1 we show a path $X = \{X_t, t \in [0,10]\}$ of the Merton model by using the explicit solution (2.1.6).

In Figure 4.3.2 we show the accuracy of the Euler scheme and of the order 1.0 Taylor scheme in approximating the true solution X on the same sample path shown in Figure 4.3.1, when a time step size $\Delta = 0.5$ is used. We plot the error $\left(|X_{t_n} - Y_{t_n}|^2\right)^{\frac{1}{2}}$, for $n \in \{0, 1, \ldots, n_T\}$, generated by these schemes. One clearly notices the higher accuracy of the order 1.0 Taylor scheme.

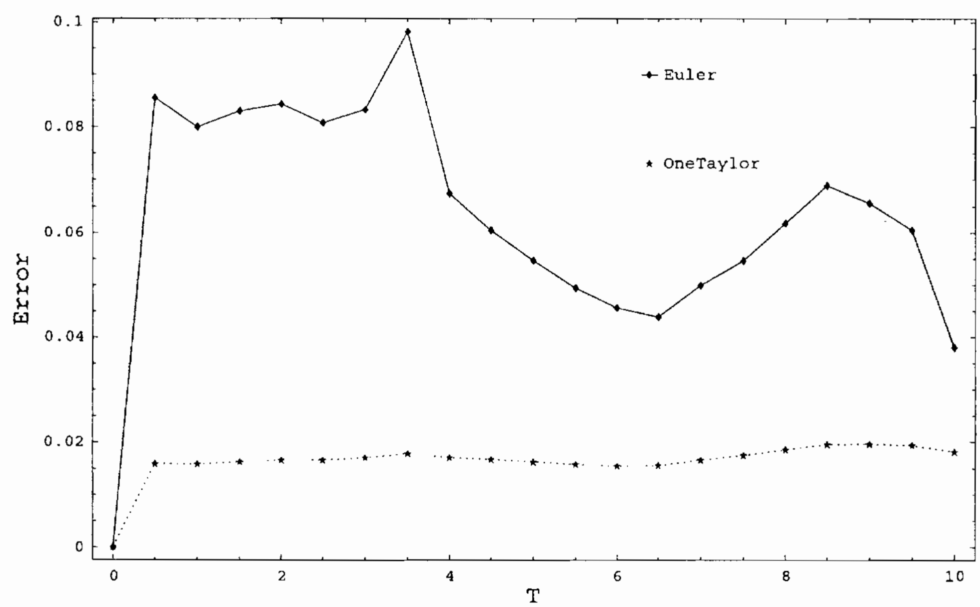

Figure 4.3.2: Strong error for the Euler and the order 1.0 strong Taylor schemes

In the multi-dimensional case with scalar Wiener process, $m = 1$, and mark-dependent jump size, the kth component of the *order 1.0 strong Taylor scheme* is given by the algorithm

$$\begin{aligned} Y_{n+1}^k &= Y_n^k + a^k \Delta_n + b^k \Delta W_n + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} c^k(v)\, p_\phi(dv, dz) \\ &+ \sum_{l=1}^{d} \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} b^l \frac{\partial b^k}{\partial x^l} dW(z_1) dW(z_2) \\ &+ \sum_{l=1}^{d} \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} b^l \frac{\partial c^k(v)}{\partial x^l}\, dW(z_1) p_\phi(dv, dz_2) \end{aligned}$$

$$
\begin{aligned}
&+\int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} \int_{\mathcal{E}} \Big\{ b^k\big(t_n, Y_n + c(v)\big) - b^k \Big\} p_\phi(dv, dz_1) dW(z_2) \\
&+\int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} \int_{\mathcal{E}} \Big\{ c^k\big(t_n, Y_n + c(v_1), v_2\big) - c^k\big(v_2\big) \Big\} \\
&\times p_\phi(dv_1, dz_1)\, p_\phi(dv_2, dz_2),
\end{aligned}
\tag{4.3.6}
$$

where a^k, b^k, and c^k are the kth components of the drift, diffusion and jump coefficients, respectively, for $k \in \{1, 2, \ldots, d\}$. Similar to (4.3.3) we can rewrite also this scheme in a form that is readily applicable for scenario simulation.

For the multi-dimensional case with one driving Wiener process and mark-independent jump size the kth component of the *order 1.0 strong Taylor scheme* simplifies to

$$
\begin{aligned}
Y_{n+1}^k &= Y_n^k + a^k \Delta_n + b^k \Delta W_n + c^k \Delta p_n + \sum_{l=1}^d b^l \frac{\partial b^k}{\partial x^l} I_{(1,1)} \\
&\quad + \sum_{l=1}^d b^l \frac{\partial c^k}{\partial x^l} I_{(1,-1)} + \Big\{ b^k\big(t_n, Y_n + c\big) - b^k \Big\} I_{(-1,1)} \\
&\quad + \Big\{ c^k\big(t_n, Y_n + c\big) - c^k \Big\} I_{(-1,-1)},
\end{aligned}
\tag{4.3.7}
$$

for $k \in \{1, 2, \ldots, d\}$. Here the four double stochastic integrals involved can be generated as described in (4.3.5).

In the general multi-dimensional case the kth component of the *order 1.0 strong Taylor scheme* is given by

$$
\begin{aligned}
Y_{n+1}^k &= Y_n^k + a^k \Delta_n + \sum_{j=1}^m b^{k,j} \Delta W_n^j + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} c^k(v)\, p_\phi(dv, dz) \\
&\quad + \sum_{j_1, j_2 = 1}^m \sum_{i=1}^d b^{i,j_1} \frac{\partial b^{k,j_2}}{\partial x^i} \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} dW^{j_1}(z_1) dW^{j_2}(z_2) \\
&\quad + \sum_{j_1=1}^m \sum_{i=1}^d \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} b^{i,j_1} \frac{\partial c^k(v)}{\partial x^i} dW^{j_1}(z_1) p_\phi(dv, dz_2) \\
&\quad + \sum_{j_1=1}^m \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} \Big\{ b^{k,j_1}\big(t_n, Y_n + c(v)\big) - b^{k,j_1} \Big\} p_\phi(dv, dz_2)\, dW^{j_1}(z_2)
\end{aligned}
$$

$$
\begin{aligned}
&+\int_{t_n}^{t_{n+1}}\int_{\mathcal{E}}\int_{t_n}^{z_2}\int_{\mathcal{E}}\Big\{c^k\big(t_n, Y_n + c(v_1), v_2\big) - c^k\big(v_2\big)\Big\}\\
&\times p_\phi(dv_1, dz_1)\, p_\phi(dv_2, dz_2), \qquad (4.3.8)
\end{aligned}
$$

for $k \in \{1, 2, \ldots, d\}$.

In the multi-dimensional case with mark-independent jump size, the kth component of the *order 1.0 strong Taylor scheme* is given by

$$
\begin{aligned}
Y_{n+1}^k &= Y_n^k + a^k \Delta_n + \sum_{j=1}^m b^{k,j} \Delta W_n^j + c^k \Delta p_n \\
&+ \sum_{j_1,j_2=1}^m \sum_{i=1}^d b^{i,j_1} \frac{\partial b^{k,j_2}}{\partial x^i} I_{(j_1,j_2)} + \sum_{j_1=1}^m \sum_{i=1}^d b^{i,j_1} \frac{\partial c^k}{\partial x^i} I_{(j_1,-1)} \qquad (4.3.9)\\
&+ \sum_{j_1=1}^m \Big\{b^{k,j_1}\big(t_n, Y_n + c\big) - b^{k,j_1}\Big\} I_{(-1,j_1)} + \Big\{c^k\big(t_n, Y_n + c\big) - c^k\Big\} I_{(-1,-1)},
\end{aligned}
$$

for $k \in \{1, 2, \ldots, d\}$.

The considerations on the generation of the mixed multiple stochastic integrals involving Wiener processes and the Poisson random measure, presented for the one-dimensional case, apply in a similar way to the implementation of the scheme (4.3.9). Indeed,

$$
\begin{aligned}
I_{(j_1,-1)} &:= \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{s_2} dW_{s_1}^{j_1} p_\phi(dv, ds_2) = \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} W_{\tau_i}^{j_1} - \Delta p_n\, W_{t_n}^{j_1},\\
I_{(-1,j_1)} &:= \int_{t_n}^{t_{n+1}} \int_{t_n}^{s_2} \int_{\mathcal{E}} p_\phi(dv, ds_1) dW_{s_2}^{j_1} = \Delta p_n\, \Delta W_n^{j_1} - I_{(j_1,-1)},
\end{aligned}
$$

for $j_1 \in \{1, \ldots, m\}$. Moreover, in (4.3.9) we also require the multiple stochastic integrals

$$
I_{(j_1,j_2)} = \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} dW^{j_1}(z_1) dW^{j_2}(z_2), \qquad (4.3.10)
$$

for $j_1, j_2 \in \{1, \ldots, m\}$. When $j_1 = j_2$, the corresponding double Wiener integral is given by

$$
I_{(j_1,j_1)} = \frac{1}{2}\{\big(\Delta W_n^{j_1}\big)^2 - \Delta_n\}. \qquad (4.3.11)
$$

However, when $j_1 \neq j_2$, this integral cannot be easily expressed in terms of the increments $\Delta W_n^{j_1}$ and $\Delta W_n^{j_2}$ of the corresponding Wiener processes. Nonetheless, employing the Karhunen-Loève expansion, see Kloeden & Platen (1999), it is possible to generate approximations of the required multiple stochastic integrals with the desired accuracy. An alternative and simpler way to approximate these multiple stochastic integrals is the following: we apply the Euler scheme to the SDE which describes the dynamics of the required multiple stochastic integrals with a time step size δ smaller than the original time step size Δ. Since the Euler scheme has strong order of convergence $\gamma = 0.5$, by choosing $\delta = \Delta^2$, the strong order $\gamma = 1.0$ of the scheme is preserved, see Milstein (1995) and Kloeden (2002).

The compensated strong Taylor schemes that arise from the compensated Wagner-Platen expansion (3.5.5) are generally different from that arising from the Wagner-Platen expansion (3.5.4) used in this section. For illustration, we describe now the *order* 1.0 *compensated strong Taylor scheme* in the special case $d = m = 1$. This is given by

$$\begin{aligned} Y_{n+1} &= Y_n + \widetilde{a}\Delta_n + b\Delta W_n + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} c(v)\, \widetilde{p}_\phi(dv, dz) \\ &\quad + bb' \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} dW(z_1) dW(z_2) \\ &\quad + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} b\, c'(v) dW(z_1) \widetilde{p}_\phi(dv, dz_2) \\ &\quad + \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} \int_{\mathcal{E}} \Big\{ b\big(t_n, Y_n + c(v)\big) - b \Big\} p_\phi(dv, dz_1) dW(z_2) \qquad (4.3.12) \\ &\quad + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} \int_{\mathcal{E}} \Big\{ c\big(t_n, Y_n + c(v_1), v_2\big) - c(v_2) \Big\} \widetilde{p}_\phi(dv_1, dz_1) \widetilde{p}_\phi(dv_2, dz_2). \end{aligned}$$

For application in scenario simulation, the generation of a stochastic integral involving the compensated Poisson measure $\widetilde{p}_\phi$ is implemented as follows: one should first generate a corresponding multiple stochastic integral with jump integrations with respect to the Poisson jump measure p_ϕ and then subtract its mean, which is given by the corresponding multiple stochastic integral with the integrations with respect to the Poisson jump measure p_ϕ replaced by integrations with respect to time and the intensity measure ϕ, see (3.1.1). For example, to generate the stochastic integral

$$\int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} c(v)\, \widetilde{p}_\phi(dv, dz),$$

we use the representation

$$\int_{t_n}^{t_{n+1}}\int_{\mathcal{E}} c(v)\,\widetilde{p}_\phi(dv,dz) = \int_{t_n}^{t_{n+1}}\int_{\mathcal{E}} c(v)\,p_\phi(dv,dz) - \int_{t_n}^{t_{n+1}}\int_{\mathcal{E}} c(v)\,\phi(dv)dz. \tag{4.3.13}$$

We recall, that, according to (4.1.10), we have used the abbreviation $c(v) = c(t_n, Y_n, v)$. Therefore, last term on the right-hand side of (4.3.13) is $\mathcal{A}_{t_n}$-measurable, but its implementation generally requires a numerical integration at each time step.

To compare the computational effort required by the order 1.0 compensated strong Taylor scheme (4.3.12) with that of the order 1.0 strong Taylor scheme (4.3.1), let us rewrite the former with integrations with respect to the Poisson jump measure p_ϕ. In this case order 1.0 compensated strong Taylor scheme is given by

$$\begin{aligned}
Y_{n+1} &= Y_n + a\Delta_n + b\Delta W_n + \int_{t_n}^{t_{n+1}}\int_{\mathcal{E}} c(v)\,p_\phi(dv,dz) \\
&\quad + bb'\int_{t_n}^{t_{n+1}}\int_{t_n}^{z_2} dW(z_1)dW(z_2) \\
&\quad + \int_{t_n}^{t_{n+1}}\int_{\mathcal{E}}\int_{t_n}^{z_2} b\,c'(v)dW(z_1)p_\phi(dv,dz_2) \\
&\quad + \int_{t_n}^{t_{n+1}}\int_{t_n}^{z_2}\int_{\mathcal{E}} \Big\{b\big(t_n, Y_n + c(v)\big) - b\Big\}p_\phi(dv,dz_1)dW(z_2) \\
&\quad + \int_{t_n}^{t_{n+1}}\int_{\mathcal{E}}\int_{t_n}^{z_2}\int_{\mathcal{E}} \Big\{c\big(t_n, Y_n + c(v_1), v_2\big) - c(v_2)\Big\}p_\phi(dv_1,dz_1)p_\phi(dv_2,dz_2) \\
&\quad - \int_{t_n}^{t_{n+1}}\int_{\mathcal{E}}\int_{t_n}^{z_2} b\,c'(v)dW(z_1)\phi(dv)dz_2 \\
&\quad - \int_{t_n}^{t_{n+1}}\int_{\mathcal{E}}\int_{t_n}^{z_2}\int_{\mathcal{E}} \Big\{c\big(t_n, Y_n + c(v_1), v_2\big) - c(v_2)\Big\}\phi(dv_1)dz_1p_\phi(dv_2,dz_2) \\
&\quad - \int_{t_n}^{t_{n+1}}\int_{\mathcal{E}}\int_{t_n}^{z_2}\int_{\mathcal{E}} \Big\{c\big(t_n, Y_n + c(v_1), v_2\big) - c(v_2)\Big\}p_\phi(dv_1,dz_1)\phi(dv_2)dz_2 \\
&\quad + \int_{t_n}^{t_{n+1}}\int_{\mathcal{E}}\int_{t_n}^{z_2}\int_{\mathcal{E}} \Big\{c\big(t_n, Y_n + c(v_1), v_2\big) - c(v_2)\Big\}\phi(dv_1)dz_1\phi(dv_2)dz_2.
\end{aligned}$$

Therefore, the implementation of the order 1.0 compensated strong Taylor scheme has the same computational effort of the order 1.0 strong Taylor scheme plus that required for the generation of the last four terms in (4.3.14). Note that both schemes have the same order of strong convergence $\gamma = 1.0$. Nonetheless, the compensated strong Taylor schemes might be more accurate when dealing with high intensity jump diffusions. However, we leave a comparison of the accuracy of these schemes for further research.

4.4 Commutativity Conditions

As previously discussed, higher order Taylor schemes, even with mark-independent jump size, become computationally inefficient when the intensity of the Poisson measure is high. In this case the number of operations involved is almost proportional to the intensity level. Also the jump-adapted schemes to be presented in Chapter 6 show a strong dependency in their efficiency on the intensity of the jumps.

By analyzing the multiple stochastic integrals required for the scheme (4.3.4), we observe that the dependence on the jump times affects only the mixed multiple stochastic integrals $I_{(1,-1)}$ and $I_{(-1,1)}$. However, since by (4.3.5) we have

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

the sum of these integrals is obtained as

$$I_{(1,-1)} + I_{(-1,1)} = \Delta p_n \, \Delta W_n, \tag{4.4.1}$$

which is independent of the particular jump times. Let us consider a one-dimensional SDE with mark-independent jump size, $c(t, x, v) = c(t, x)$, satisfying the *jump commutativity condition*

$$b(t,x)\frac{\partial c(t,x)}{\partial x} = b\Big(t, x + c(t,x)\Big) - b\Big(t, x\Big), \tag{4.4.2}$$

for all $t \in [0, T]$ and $x \in \mathbb{R}$. In this case the order 1.0 strong Taylor scheme (4.3.4) depends only on the sum $I_{(1,-1)} + I_{(-1,1)}$ expressed in (4.4.1). One does not need 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 jump commutativity condition (4.4.2), then considerable savings in computational time can be achieved.

When we have a linear diffusion coefficient of the form

$$b(t,x) = b_1(t) + b_2(t)\,x, \tag{4.4.3}$$

with $b(t,x) > 0$, as it frequently occurs in finance, the jump commutativity condition (4.4.2) implies the following ordinary differential equation (ODE) for the jump coefficient:

$$\frac{\partial c(t,x)}{\partial x} = \frac{b_2(t)\,c(t,x)}{b_1(t) + b_2(t)\,x}, \tag{4.4.4}$$

for all $t \in [0,T]$ and $x \in \mathbb{R}$. Therefore, for linear diffusion coefficients of the form (4.4.3) the class of SDEs satisfying the jump commutativity condition (4.4.2) is identified by mark-independent jump coefficients of the form

$$c(t,x) = e^{K(t)}\Big(b_1(t) + b_2(t)\,x\Big), \tag{4.4.5}$$

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

For instance, the SDE (2.1.5) with mark-independent, multiplicative jump size $c(t,x,v) = x\,\beta$, for $\beta \geq -1$, satisfies the jump commutativity condition (4.4.2). The corresponding order 1.0 strong Taylor scheme is given by

$$\begin{aligned} Y_{n+1} &= Y_n + \mu Y_n \Delta_n + \sigma Y_n \Delta W_n + \beta\, Y_n\, \Delta p_n + \frac{1}{2}\sigma^2\, Y_n\,\{(\Delta W_n)^2 - \Delta_n\} \\ &\quad + \sigma\,\beta\, Y_n \Delta p_n \Delta W_n + \frac{1}{2}\beta^2\, Y_n\{(\Delta p_n)^2 - \Delta p_n\}. \end{aligned} \tag{4.4.6}$$

Another interesting example that arises in the financial literature is the square root diffusion coefficient

$$b(t,x) = b_1(t)\sqrt{x}, \tag{4.4.7}$$

see Cox, Ingersoll & Ross (1985), Duffie & Kan (1994) and Platen (2001). If $b_1(t) \neq 0$ for $t \in [0,T]$, we have the ODE

$$\frac{\partial c(t,x)}{\partial x} = \frac{\sqrt{x + c(x)} - \sqrt{x}}{\sqrt{x}}, \tag{4.4.8}$$

and, thus, for the jump commutative case (4.4.2), the jump coefficient is of the form

$$c(t,x) = e^{K(t)} + 2\,e^{\frac{K(t)}{2}}\sqrt{x}. \tag{4.4.9}$$

$b(t,x)$	$c(t,x)$
$b_1(t)$	$K(t)$
$b_1(t)+b_2(t)\,x$	$e^{K(t)}(b_1(t)+b_2(t)\,x)$
$b_3(t)\sqrt{b_1(t)+b_2(t)x}$	$b_2(t)\,e^{K(t)}+2e^{\frac{K(t)}{2}}\sqrt{b_1(t)+b_2(t)x}$
$b_1(t)(1-e^{-x})$	$\log\{1+e^{K(t)}-e^{-x+K(t)}\}$
$b_1(t)\,x^{\frac{3}{2}}$	$\dfrac{-2e^{\frac{3K(t)}{2}}x^{\frac{3}{2}}+3e^{K(t)}x^2-x^3}{e^{2K(t)}-2e^{K(t)}x+x^2}$

Table 4.1: Coefficients satisfying the jump commutativity condition.

In Table 4.4 we present some diffusion coefficients from models proposed in the finance literature together with the corresponding jump coefficients that satisfy the jump commutative condition (4.4.2).

In the multi-dimensional case with scalar Wiener process and mark-independent jump size we obtain the *jump commutativity condition*

$$\sum_{l=1}^{d} b^l(t,x)\,\frac{\partial c^k(,t,x)}{\partial x^l}=b^k\Big(t,x+c(t,x)\Big)-b^k\Big(t,x\Big) \tag{4.4.10}$$

for $k\in\{1,2,\dots,d\}$, $t\in[0,T]$ and $x\in\mathbb{R}^d$.

For the general multi-dimensional case, one can show that the sum of two multiple stochastic integrals with respect to the j_1th component of the Wiener process and the Poisson measure, given by

$$I_{(j_1,-1)}+I_{(-1,j_1)}=\Delta p_n\,\Delta W_n^{j_1}, \tag{4.4.11}$$

is independent of the particular jump times. Therefore, for a general multi-dimensional SDE with mark-independent jump size, we obtain the *jump commutativity condition*

$$\sum_{l=1}^{d} b^{l,j_1}(t,x)\,\frac{\partial c^k(t,x)}{\partial x^l}=b^{k,j_1}\Big(t,x+c(t,x)\Big)-b^{k,j_1}\Big(t,x\Big) \tag{4.4.12}$$

for $j_1\in\{1,2,\dots,m\}$, $k\in\{1,2,\dots,d\}$, $t\in[0,T]$ and $x\in\mathbb{R}^d$. Thus, once the diffusion coefficients are specified, a solution of the first order semilinear $d\times m$-dimensional system of partial differential equations (PDEs) (4.4.12) provides a d-dimensional commutative jump coefficient. Note that the system (4.4.12) has $d\times m$ equations and only d unknown functions. Therefore, even for simple diffusion coefficients, there may not exist any jump coefficient satisfying (4.4.12).

Consider, for instance, the multi-dimensional case with *additive diffusion coefficient* $b(t,x) = b(t)$. The jump commutativity condition (4.4.12) reduces to the $m \times d$ system of first order homogeneous linear PDEs

$$\sum_{l=1}^{d} b^{l,j}(t) \frac{\partial c^k(t,x)}{\partial x^l} = 0 \tag{4.4.13}$$

for $j \in \{1,2,\ldots,m\}$, $k \in \{1,2,\ldots,d\}$, $t \in [0,T]$ and $x = (x^1, x^2, \ldots, x^d)^\top \in \mathbb{R}^d$. An *additive jump coefficient* $c(t,x) = c(t)$ satisfies the condition (4.4.13). In the scalar case, $m = 1$, or in the trivial case where $b^{i,j_1}(t) = b^{i,j_2}(t)$ for all $i \in \{1,2,\ldots,d\}$, $j_1, j_2 \in \{1,2\ldots,m\}$ and $t \in [0,T]$, we obtain the solution

$$c^k(t,x) = f(t,y) \tag{4.4.14}$$

for $k \in \{1,\ldots,d\}$. Here $y = (y^1,\ldots,y^{d-1})^\top \in \mathbb{R}^{d-1}$ has components

$$y^i = \frac{-b^{i+1}(t)x^1 + b^1(t)x^{i+1}}{b^1(t)}$$

and $f : [0,T] \times \mathbb{R}^{d-1}$ is an arbitrary function, differentiable with respect to the second argument y.

Let us consider the multi-dimensional, *multiplicative diffusion coefficient* $b(t,x)$, where the element in the ith row and jth column is given by $b^{i,j}(t,x) = \sigma^{i,j}(t)x^i$, with $\sigma^{i,j}(t) \in \mathbb{R}$ for $i \in \{1,2,\ldots,d\}$ and $j \in \{1,2\ldots,m\}$. In this case the jump commutativity condition (4.4.12) reduces to the $m \times d$ system of first order linear PDEs

$$\sum_{l=1}^{d} \sigma^{l,j}(t) x^i \frac{\partial c^k(t,x)}{\partial x^l} = \sigma^{k,j}(t)\, c(t,x) \tag{4.4.15}$$

for $j \in \{1,2,\ldots,m\}$, $k \in \{1,2,\ldots,d\}$, $t \in [0,T]$ and $x = (x^1,x^2,\ldots,x^d)^\top \in \mathbb{R}^d$. In the scalar case, with $m = 1$, or in the trivial case where $\sigma^{i,j_1}(t) = \sigma^{i,j_2}(t)$ for $i \in \{1,2,\ldots,d\}$, $j_1, j_2 \in \{1,2\ldots,m\}$ and $t \in [0,T]$, we obtain the solution

$$c^k(t,x) = (x^1)^{\frac{\sigma^k(t)}{\sigma^1(t)}} f(t,y) \tag{4.4.16}$$

for $k \in \{1, \ldots, d\}$. Here $y = (y^1, \ldots, y^{d-1})^\top \in \Re^{d-1}$ has components

$$y^i = (x^1)^{\frac{\sigma^{i+1}(t)}{\sigma^1(t)}} x^{i+1}$$

and $f : [0, T] \times \Re^{d-1}$ is an arbitrary function, differentiable with respect to the components of the second argument y.

We now discuss some commutativity conditions involving only the diffusion coefficients. These can be found also in Kloeden & Platen (1999), in the context of the approximation of pure diffusion SDEs. As noticed in Section 4.3, when we have a multi-dimensional driving Wiener process, the corresponding order 1.0 strong Taylor scheme requires multiple stochastic integrals with respect to the different components of the Wiener process. In general, these can be generated only resorting to approximations, such as the Karhunen-Loève expansion, see Kloeden & Platen (1999). However, in the special case of a *diffusion commutativity condition*, where

$$L^{j_1} b^{k,j_2}(t, x) = L^{j_2} b^{k,j_1}(t, x) \tag{4.4.17}$$

for $j_1, j_2 \in \{1, 2, \ldots, m\}$, $k \in \{1, 2, \ldots, d\}$, $t \in [0, T]$ and $x \in \Re^d$, it is possible to express all the double Wiener integrals in terms of the increments $\Delta W_n^{j_1}$ and $\Delta W_n^{j_2}$ of the Wiener processes. Therefore, for a multi-dimensional SDE satisfying the diffusion commutativity condition (4.4.17), the jump commutativity condition (4.4.12) and with mark-independent jump size, we obtain a computationally efficient order 1.0 strong Taylor scheme, whose kth component is given by

$$\begin{aligned} Y_{n+1}^k &= Y_n^k + a^k \Delta_n + \sum_{j=1}^m b^{k,j} \Delta W_n^j + c^k \Delta p_n \\ &+ \frac{1}{2} \sum_{j_1, j_2 = 1}^m \sum_{i=1}^d b^{i,j_1} \frac{\partial b^{k,j_2}}{\partial x^i} \left\{ \Delta W_n^{j_1} \Delta W_n^{j_2} - \Delta_n \right\} \\ &+ \sum_{j_1=1}^m \{ b^{k,j_1}(t_n, Y_n + c) - b^{k,j_1} \} \Big(\Delta p_n \Delta W_n^{j_1} \Big) \\ &+ \frac{1}{2} \{ c^k(t_n, Y_n + c) - c^k \} \Big(\big(\Delta p_n\big)^2 - \Delta p_n \Big), \end{aligned} \tag{4.4.18}$$

for $k \in \{1, 2, \ldots, d\}$. For instance, the special case of additive diffusion and jump coefficients, which means $b(t, x) = b(t)$ and $c(t, x) = c(t)$, satisfies all the required commutativity conditions and therefore leads to an efficient order 1.0 strong Taylor scheme. Note that in this particular case the last two lines of (4.4.18) equal zero and, thus, only double stochastic integrals with respect to Wiener processes are needed.

From this analysis it becomes clear that when selecting a suitable numerical scheme for a specific model it is important to check for particular commutativity properties of the SDE under consideration to potentially save computational time.

4.5 Convergence Results

In this section we introduce strong Taylor approximations and compensated strong Taylor approximations of any given strong order $\gamma \in \{0.5, 1, 1.5, 2, \ldots\}$. The key results underlying the construction of these strong approximations are the Wagner-Platen expansions (3.5.4) and (3.5.5) presented in Chapter 3. Including in a scheme enough terms from these expansions, we can obtain approximations with the desired order of strong convergence. More precisely, for an order $\gamma \in \{0.5, 1, 1.5, 2, \ldots\}$ strong Taylor scheme we need to use the hierarchical set

$$\mathcal{A}_\gamma = \{\alpha \in \mathcal{M} : l(\alpha) + n(\alpha) \leq 2\gamma \quad \text{or} \quad l(\alpha) = n(\alpha) = \gamma + \frac{1}{2}\}, \tag{4.5.1}$$

where $l(\alpha)$ denotes again the length of the multi-index α, and $n(\alpha)$ the number of its components equal to zero. Note that in Chapter 3 we have derived two types of Wagner-Platen expansions; the first uses the Poisson measure as integrator of jump type and the second employs the compensated Poisson measure as integrator involving jumps. Therefore, we will obtain two different types of strong approximations; the strong Taylor approximations and the compensated strong Taylor approximations.

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

$$Y_{n+1}^\Delta = Y_n^\Delta + \sum_{\alpha \in \mathcal{A}_\gamma \setminus \{v\}} I_\alpha \left[f_\alpha(t_n, Y_n^\Delta) \right]_{t_n, t_{n+1}} = \sum_{\alpha \in \mathcal{A}_\gamma} I_\alpha \left[f_\alpha(t_n, Y_n^\Delta) \right]_{t_n, t_{n+1}} \tag{4.5.2}$$

for $n \in \{0, 1, \ldots, n_T - 1\}$. Similarly, we define the *order γ compensated strong*

Taylor scheme by the vector equation

$$Y_{n+1}^{\Delta} = Y_n^{\Delta} + \sum_{\alpha \in \mathcal{A}_\gamma \setminus \{v\}} \widetilde{I}_\alpha \left[\tilde{f}_\alpha(t_n, Y_n^{\Delta})\right]_{t_n, t_{n+1}} = \sum_{\alpha \in \mathcal{A}_\gamma} \widetilde{I}_\alpha \left[\tilde{f}_\alpha(t_n, Y_n^{\Delta})\right]_{t_n, t_{n+1}} \tag{4.5.3}$$

for $n \in \{0, 1, \ldots, n_T - 1\}$.

Equations (4.5.2) and (4.5.3) provide recursive numerical routines generating approximate values of the solution of the SDE (4.1.1) at the time 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

$$Y_t^{\Delta} = \sum_{\alpha \in \mathcal{A}_\gamma} I_\alpha[f_\alpha(t_{n_t}, Y_{t_{n_t}}^{\Delta})]_{t_{n_t}, t} \tag{4.5.4}$$

and the *order γ compensated strong Taylor approximation* $Y^{\Delta} = \{Y_t^{\Delta},\, t \in [0, T]\}$, by

$$Y_t^{\Delta} = \sum_{\alpha \in \mathcal{A}_\gamma} \widetilde{I}_\alpha[\tilde{f}_\alpha(t_{n_t}, Y_{t_{n_t}}^{\Delta})]_{t_{n_t}, t} \tag{4.5.5}$$

for $t \in [0, T]$, starting from a given $\mathcal{A}_0$-measurable random variable Y_0, where n_t was defined in (4.1.8).

These two approximations define stochastic processes $Y^{\Delta} = \{Y_t^{\Delta},\, t \in [0, T]\}$, whose values coincide with the ones of the *order γ strong Taylor scheme* (4.5.2) and of the *order γ compensated strong Taylor scheme* (4.5.3), respectively, at the time discretization points. Between the discretization points the multiple stochastic integrals have constant coefficient functions but evolve randomly as a stochastic process.

The strong order of convergence of the compensated strong Taylor schemes presented above can be derived from the following theorem. This convergence theorem will enable us to construct a compensated strong Taylor approximation $Y^{\Delta} = \{Y_t^{\Delta},\, t \in [0, T]\}$ of any given strong order $\gamma \in \{0.5, 1, 1.5, 2, \ldots\}$.

Theorem 4.5.1 *For given $\gamma \in \{0.5, 1, 1.5, 2, \ldots\}$, let $Y^{\Delta} = \{Y_t^{\Delta},\, t \in [0, T]\}$ be the order γ compensated strong Taylor approximation defined in* (4.5.5)*, corresponding to a time discretization with maximum step size $\Delta \in (0, 1)$. We assume that*

$$E(|X_0|^2) < \infty \quad \textit{and} \quad E(|X_0 - Y_0^{\Delta}|^2) \leq K_1 \Delta^{2\gamma}. \tag{4.5.6}$$

Moreover, suppose that the coefficient functions $\tilde{f}_\alpha$ 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 $\tilde{f}_\alpha$ satisfies the Lipschitz type condition

$$\left|\tilde{f}_\alpha(t,x,u) - \tilde{f}_\alpha(t,y,u)\right| \leq K_1(u)|x-y|, \tag{4.5.7}$$

where $\left(K_1(u)\right)^2$ is $\phi(du^1) \times \ldots \times \phi(du^{s(\alpha)})$-integrable.

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

$$\tilde{f}_{-\alpha} \in \mathcal{C}^{1,2} \qquad \text{and} \qquad \tilde{f}_\alpha \in \mathcal{H}_\alpha, \tag{4.5.8}$$

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

$$\left|\tilde{f}_\alpha(t,x,u)\right|^2 \leq K_2(u)(1+|x|^2), \tag{4.5.9}$$

where $K_2(u)$ is $\phi(du^1) \times \ldots \times \phi(du^{s(\alpha)})$-integrable.

Then the estimate

$$\sqrt{E\big(\sup_{0\leq z\leq T} |X_z - Y_z^\Delta|^2 \,|\mathcal{A}_0\big)} \leq K_3 \Delta^\gamma \tag{4.5.10}$$

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

Remark 4.5.2 *By using the definitions of the operators* (3.3.4)–(3.3.7) *and of the sets in* (3.2.3), *it is possible to obtain conditions on the coefficients a, b and c of the SDE* (3.5.1) *which imply conditions* (4.5.7)–(4.5.9) *on the coefficient functions $\tilde{f}_\alpha$. For instance, if the drift, diffusion and jump coefficients of the SDE* (2.1.2) *have $2(\gamma+2)$ times continuously differentiable components $a^k, b^{k,j}, c^k$, for all $k \in \{1,2\ldots,d\}$ and $j \in \{1,2\ldots,m\}$, that are uniformly bounded with uniformly bounded derivatives, then conditions* (4.5.7)–(4.5.9) *are fulfilled.*

Theorem 4.5.1 generalizes a similar result for pure diffusions described in Kloeden & Platen (1999). The proof will be given at the end of this chapter in Section 4.7. A related result, with slightly different conditions, was published without proof in Platen (1982a).

Theorem 4.5.1 establishes the order of strong convergence of the order γ compensated strong Taylor approximation defined in (4.5.5). The following corollary permits us to obtain the order of strong convergence of the strong Taylor schemes presented in this chapter.

Corollary 4.5.3 *Let $Y^\Delta = \{Y_t^\Delta, t \in [0,T]\}$ be the order γ strong Taylor approximation* (4.5.4). *Assume that the Itô coefficient functions f_α satisfy the conditions of Theorem 4.5.1. Then, if also the conditions on the initial data* (4.5.6) *hold, we obtain the estimate*

$$\sqrt{E\left(\sup_{0\le z\le T}|X_z - Y_z^\Delta|^2\right)} \le K\,\Delta^\gamma, \tag{4.5.11}$$

where K is a finite positive constant independent of Δ.

A similar result, limited to SDEs driven by Wiener processes and homogeneous Poisson processes, is presented in Gardoǹ (2004).

4.6 Lemma on Multiple Itô Integrals

We present here a lemma on multiple stochastic integrals that we will need in the proof of Theorem 4.5.1. This lemma can be also used for other approximations.

Lemma 4.6.1 *For a given multi-index $\alpha \in \mathcal{M}_m \backslash \{v\}$, a time discretization $(t)_\Delta$ with $\Delta \in (0,1)$ and $g \in \mathcal{H}_\alpha$ let*

$$V_{t_0,u,s(\alpha)} := \int_{\mathcal{E}} \cdots \int_{\mathcal{E}} E\left(\sup_{t_0\le z\le u}|g(z,v^1,\ldots,v^{s(\alpha)})|^2 \middle| \mathcal{A}_{t_0}\right)\phi(dv^1)\ldots\phi(dv^{s(\alpha)}) < \infty, \tag{4.6.1}$$

$$\widetilde{F}_t^\alpha := E\left(\sup_{t_0\le z\le t}\left|\sum_{n=0}^{n_z-1}\widetilde{I}_\alpha[g(\cdot)]_{t_n,t_{n+1}} + \widetilde{I}_\alpha[g(\cdot)]_{t_{n_z},z}\right|^2 \middle| \mathcal{A}_{t_0}\right) \tag{4.6.2}$$

and

$$F_t^\alpha := E\left(\sup_{t_0\le z\le t}\left|\sum_{n=0}^{n_z-1} I_\alpha[g(\cdot)]_{t_n,t_{n+1}} + I_\alpha[g(\cdot)]_{t_{n_z},z}\right|^2 \middle| \mathcal{A}_{t_0}\right). \tag{4.6.3}$$

Then

$$\widetilde{F}_t^\alpha \leq \begin{cases} (t-t_0)\ \Delta^{2(l(\alpha)-1)} \int_{t_0}^t V_{t_0,u,s(\alpha)}\, du & when: l(\alpha) = n(\alpha) \\ 4^{l(\alpha)-n(\alpha)+2}\ \Delta^{l(\alpha)+n(\alpha)-1} \int_{t_0}^t V_{t_0,u,s(\alpha)}\, du & when: l(\alpha) \neq n(\alpha) \end{cases}$$

and

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

almost surely, for every $t \in [t_0, T]$. Here $\hat{C} = 4 + \lambda(T - t_0)$.

Proof: Let us first prove the assertion of the lemma for $\widetilde{F}_t^\alpha$.

1. By definition (4.1.8) 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 obtain

$$\begin{aligned} &\sum_{n=0}^{n_z-1} \widetilde{I}_\alpha[g(\cdot)]_{t_n,t_{n+1}} + \widetilde{I}_\alpha[g(\cdot)]_{t_{n_z},z} \\ &= \sum_{n=0}^{n_z-1} \int_{t_n}^{t_{n+1}} \widetilde{I}_{\alpha-}[g(\cdot)]_{t_n,s} ds + \int_{t_{n_z}}^{z} \widetilde{I}_{\alpha-}[g(\cdot)]_{t_{n_z},s} ds \\ &= \sum_{n=0}^{n_z-1} \int_{t_n}^{t_{n+1}} \widetilde{I}_{\alpha-}[g(\cdot)]_{t_{n_s},s} ds + \int_{t_{n_z}}^{z} \widetilde{I}_{\alpha-}[g(\cdot)]_{t_{n_s},s} ds \\ &= \int_{t_0}^{z} \widetilde{I}_{\alpha-}[g(\cdot)]_{t_{n_s},s} ds. \qquad (4.6.4) \end{aligned}$$

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

$$
\begin{aligned}
\widetilde{F}_t^\alpha &= E\left(\sup_{t_0\le z\le t}\left|\int_{t_0}^{z}\widetilde{I}_{\alpha-}[g(\cdot)]_{t_{n_u},u}\,du\right|^2\Bigg|\mathcal{A}_{t_0}\right)\\
&\le E\left(\sup_{t_0\le z\le t}(z-t_0)\int_{t_0}^{z}|\widetilde{I}_{\alpha-}[g(\cdot)]_{t_{n_u},u}|^2\,du\Bigg|\mathcal{A}_{t_0}\right)\\
&\le (t-t_0)\,E\left(\int_{t_0}^{t}|\widetilde{I}_{\alpha-}[g(\cdot)]_{t_{n_u},u}|^2\,du\Bigg|\mathcal{A}_{t_0}\right)\\
&= (t-t_0)\int_{t_0}^{t}E\left(|\widetilde{I}_{\alpha-}[g(\cdot)]_{t_{n_u},u}|^2\,|\mathcal{A}_{t_0}\right)du\\
&\le (t-t_0)\int_{t_0}^{t}E\left(\sup_{t_{n_u}\le z\le u}|\widetilde{I}_{\alpha-}[g(\cdot)]_{t_{n_u},z}|^2\Bigg|\mathcal{A}_{t_0}\right)du\\
&= (t-t_0)\int_{t_0}^{t}E\left(E\Big(\sup_{t_{n_u}\le z\le u}|\widetilde{I}_{\alpha-}[g(\cdot)]_{t_{n_u},z}|^2\Big|\mathcal{A}_{t_{n_u}}\Big)\Bigg|\mathcal{A}_{t_0}\right)du,
\end{aligned}
\tag{4.6.5}
$$

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 3.6.1 to

$$
E\left(\sup_{t_{n_u}\le z\le u}|\widetilde{I}_{\alpha-}[g(\cdot)]_{t_{n_u},z}|^2\Bigg|\mathcal{A}_{t_{n_u}}\right)
\tag{4.6.6}
$$

yields

$$
\begin{aligned}
\widetilde{F}_t^\alpha &\le (t-t_0)\;4^{l(\alpha-)-n(\alpha-)}\\
&\quad\times\int_{t_0}^{t}E\left((u-t_{n_u})^{l(\alpha-)+n(\alpha-)-1}\int_{t_{n_u}}^{u}V_{t_{n_u},z,s(\alpha-)}\,dz\Bigg|\mathcal{A}_{t_0}\right)du\\
&\le (t-t_0)\,4^{l(\alpha-)-n(\alpha-)}\int_{t_0}^{t}E\left((u-t_{n_u})^{l(\alpha-)+n(\alpha-)}\,V_{t_{n_u},u,s(\alpha-)}\Big|\mathcal{A}_{t_0}\right)du\\
&\le (t-t_0)\,4^{l(\alpha-)-n(\alpha-)}\Delta^{l(\alpha-)+n(\alpha-)}\int_{t_0}^{t}E\left(V_{t_{n_u},u,s(\alpha-)}\Big|\mathcal{A}_{t_0}\right)du,
\end{aligned}
\tag{4.6.7}
$$

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 $u \in [t_0, t]$

$$\begin{aligned}
& E\left(V_{t_{n_u},u,s(\alpha-)} \middle| \mathcal{A}_{t_0}\right) \\
&= E\Bigg(\int_{\mathcal{E}} \cdots \int_{\mathcal{E}} E\Big(\sup_{t_{n_u} \le z \le u} |g(z, v^1, \ldots, v^{s(\alpha-)})|^2 \,\Big|\, \mathcal{A}_{t_{n_u}}\Big) \\
&\qquad \times \phi(dv^1) \ldots \phi(dv^{s(\alpha-)}) \,\Bigg|\, \mathcal{A}_{t_0}\Bigg) \\
&= \int_{\mathcal{E}} \cdots \int_{\mathcal{E}} E\Bigg(E\Big(\sup_{t_{n_u} \le z \le u} |g(z, v^1, \ldots, v^{s(\alpha-)})|^2 \,\Big|\, \mathcal{A}_{t_{n_u}}\Big) \,\Bigg|\, \mathcal{A}_{t_0}\Bigg) \\
&\qquad \times \phi(dv^1) \ldots \phi(dv^{s(\alpha-)}) \\
&= \int_{\mathcal{E}} \cdots \int_{\mathcal{E}} E\left(\sup_{t_{n_u} \le z \le u} |g(z, v^1, \ldots, v^{s(\alpha-)})|^2 \,\middle|\, \mathcal{A}_{t_0}\right) \phi(dv^1) \ldots \phi(dv^{s(\alpha-)}) \\
&\le \int_{\mathcal{E}} \cdots \int_{\mathcal{E}} E\left(\sup_{t_0 \le z \le u} |g(z, v^1, \ldots, v^{s(\alpha-)})|^2 \,\middle|\, \mathcal{A}_{t_0}\right) \phi(dv^1) \ldots \phi(dv^{s(\alpha-)}) \\
&= V_{t_0,u,s(\alpha-)}. \qquad (4.6.8)
\end{aligned}$$

It then follows

$$\begin{aligned}
\tilde{F}_t^{\alpha} &\le (t - t_0)\; 4^{l(\alpha-)-n(\alpha-)}\; \Delta^{l(\alpha-)+n(\alpha-)} \int_{t_0}^{t} V_{t_0,u,s(\alpha-)}\, du \\
&= (t - t_0)\; \Delta^{2(l(\alpha)-1)} \int_{t_0}^{t} V_{t_0,u,s(\alpha)}\, du, \qquad (4.6.9)
\end{aligned}$$

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 3.6.1 we obtain

$$\begin{aligned}
\tilde{F}_t^{\alpha} &= E\left(\sup_{t_0 \le z \le t} \left| \int_{t_0}^{z} \tilde{I}_{\alpha-}[g(\cdot)]_{t_{n_u},u}\, dW_u^{j_l} \right|^2 \,\middle|\, \mathcal{A}_{t_0}\right) \\
&\le 4\, E\left(\left| \int_{t_0}^{t} \tilde{I}_{\alpha-}[g(\cdot)]_{t_{n_u},u}\, dW_u^{j_l} \right|^2 \,\middle|\, \mathcal{A}_{t_0}\right)
\end{aligned}$$

$$
\begin{aligned}
&\leq 4 \int_{t_0}^{t} E\left(|\widetilde{I}_{\alpha-}[g(\cdot)]_{t_{n_u},u}|^2 \Big| \mathcal{A}_{t_0} \right) du \\
&= 4 \int_{t_0}^{t} E\left(E\left(|\widetilde{I}_{\alpha-}[g(\cdot)]_{t_{n_u},u}|^2 \Big| \mathcal{A}_{t_{n_u}} \right) \Big| \mathcal{A}_{t_0} \right) du \\
&\leq 4 \int_{t_0}^{t} E\left(E\left(\sup_{t_{n_u} \leq z \leq u} |\widetilde{I}_{\alpha-}[g(\cdot)]_{t_{n_u},z}|^2 \Big| \mathcal{A}_{t_{n_u}} \right) \Big| \mathcal{A}_{t_0} \right) du \\
&\leq 4 \; 4^{l(\alpha-)-n(\alpha-)} \\
&\quad \times \int_{t_0}^{t} E\left((u - t_{n_u})^{l(\alpha-)+n(\alpha-)-1} \int_{t_{n_u}}^{u} V_{t_{n_u},z,s(\alpha-)} \, dz \Big| \mathcal{A}_{t_0} \right) du \\
&\leq 4 \; 4^{l(\alpha-)-n(\alpha-)} \; \Delta^{l(\alpha-)+n(\alpha-)} \int_{t_0}^{t} V_{t_0,u,s(\alpha-)} \, du \\
&= 4^{l(\alpha)-n(\alpha)} \; \Delta^{l(\alpha)+n(\alpha)-1} \int_{t_0}^{t} V_{t_0,u,s(\alpha-)} \, du \\
&\leq 4^{l(\alpha)-n(\alpha)+2} \; \Delta^{l(\alpha)+n(\alpha)-1} \int_{t_0}^{t} V_{t_0,u,s(\alpha)} \, du, \qquad (4.6.10)
\end{aligned}
$$

where the last passage holds since $s(\alpha) = s(\alpha-)$. 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 3.6.1 and steps similar to the previous case we obtain

$$
\begin{aligned}
\widetilde{F}_t^{\alpha} &= E\left(\sup_{t_0 \leq z \leq t} \left| \int_{t_0}^{z} \int_{\mathcal{E}} \widetilde{I}_{\alpha-}[g(\cdot, v^{s(\alpha)})]_{t_{n_u},u-} \, \widetilde{p}_{\phi}(dv^{s(\alpha)}, du) \right|^2 \Bigg| \mathcal{A}_{t_0} \right) \\
&\leq 4E\left(\left| \int_{t_0}^{t} \int_{\mathcal{E}} \widetilde{I}_{\alpha-}[g(\cdot, v^{s(\alpha)})]_{t_{n_u},u-} \, \widetilde{p}_{\phi}(dv^{s(\alpha)}, du) \right|^2 \Bigg| \mathcal{A}_{t_0} \right) \\
&= 4 \int_{t_0}^{t} \int_{\mathcal{E}} E\left(|\widetilde{I}_{\alpha-}[g(\cdot, v^{s(\alpha)})]_{t_{n_u},u}|^2 \Big| \mathcal{A}_{t_0} \right) \phi(dv^{s(\alpha)}) \, du \\
&= 4 \int_{t_0}^{t} \int_{\mathcal{E}} E\left(E\left(|\widetilde{I}_{\alpha-}[g(\cdot, v^{s(\alpha)})]_{t_{n_u},u}|^2 \Big| \mathcal{A}_{t_{n_u}} \right) \Big| \mathcal{A}_{t_0} \right) \phi(dv^{s(\alpha)}) \, du \\
&\leq 4 \int_{t_0}^{t} \int_{\mathcal{E}} E\left(E\left(\sup_{t_{n_u} \leq z \leq u} |\widetilde{I}_{\alpha-}[g(\cdot, v^{s(\alpha)})]_{t_{n_u},z}|^2 \Big| \mathcal{A}_{t_{n_u}} \right) \Big| \mathcal{A}_{t_0} \right) \phi(dv^{s(\alpha)}) \, du
\end{aligned}
$$

$$
\begin{aligned}
&\leq 4^{l(\alpha-)-n(\alpha-)+1} \\
&\quad \times \int_{t_0}^{t} \int_{\mathcal{E}} E\left((u - t_{n_u})^{l(\alpha-)+n(\alpha-)-1} \int_{t_{n_u}}^{u} V_{t_{n_u},z,s(\alpha-)}\, dz \,\middle|\, \mathcal{A}_{t_0} \right) \phi(dv^{s(\alpha)})\, du \\
&\leq 4^{l(\alpha-)-n(\alpha-)+1}\ \Delta^{l(\alpha-)+n(\alpha-)} \int_{t_0}^{t} \int_{\mathcal{E}} E\left(V_{t_{n_u},u,s(\alpha-)} \,\middle|\, \mathcal{A}_{t_0} \right) \phi(dv^{s(\alpha)})\, du.
\end{aligned}
\tag{4.6.11}
$$

Hence, using (4.6.8) we have

$$
\begin{aligned}
\widetilde{F}_t^{\alpha} &\leq 4^{l(\alpha-)-n(\alpha-)+1}\ \Delta^{l(\alpha-)+n(\alpha-)} \int_{t_0}^{t} \int_{\mathcal{E}} V_{t_0,u,s(\alpha-)}\, \phi(dv^{s(\alpha)})\, du \\
&= 4^{l(\alpha)-n(\alpha)}\ \Delta^{l(\alpha)+n(\alpha)-1} \int_{t_0}^{t} V_{t_0,u,s(\alpha)}\, du \\
&\leq 4^{l(\alpha)-n(\alpha)+2}\ \Delta^{l(\alpha)+n(\alpha)-1} \int_{t_0}^{t} V_{t_0,u,s(\alpha)}\, du
\end{aligned}
\tag{4.6.12}
$$

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 $\alpha = (j_1, \ldots, j_l)$ with $l(\alpha) \neq n(\alpha)$ and $j_l = 0$. It can be shown that the discrete time process

$$
\{\sum_{n=0}^{k} \widetilde{I}_{\alpha}[g(\cdot)]_{t_n,t_{n+1}}, k \in \{0, 1 \ldots, n_T - 1\}\}
\tag{4.6.13}
$$

is a discrete time martingale.

Using Cauchy-Schwarz inequality we obtain

$$
\begin{aligned}
\widetilde{F}_t^{\alpha} &= E\left(\sup_{t_0 \leq z \leq t} \left| \sum_{n=0}^{n_z-1} \widetilde{I}_{\alpha}[g(\cdot)]_{t_n,t_{n+1}} + \widetilde{I}_{\alpha}[g(\cdot)]_{t_{n_z},z} \right|^2 \,\middle|\, \mathcal{A}_{t_0} \right) \\
&\leq 2\, E\left(\sup_{t_0 \leq z \leq t} \left| \sum_{n=0}^{n_z-1} \widetilde{I}_{\alpha}[g(\cdot)]_{t_n,t_{n+1}} \right|^2 \,\middle|\, \mathcal{A}_{t_0} \right) \\
&\quad + 2\, E\left(\sup_{t_0 \leq z \leq t} \left| \widetilde{I}_{\alpha}[g(\cdot)]_{t_{n_z},z} \right|^2 \,\middle|\, \mathcal{A}_{t_0} \right).
\end{aligned}
\tag{4.6.14}
$$

Applying Doob's inequality to the first term of the equation (4.6.14) we have

$$
\begin{aligned}
& E\left(\sup_{t_0 \le z \le t} \left| \sum_{n=0}^{n_z - 1} \widetilde{I}_\alpha[g(\cdot)]_{t_n,t_{n+1}} \right|^2 \Bigg| \mathcal{A}_{t_0}\right) \\
\le\ & 4\,E\left(\left| \sum_{n=0}^{n_t - 1} \widetilde{I}_\alpha[g(\cdot)]_{t_n,t_{n+1}} \right|^2 \Bigg| \mathcal{A}_{t_0}\right) \\
\le\ & 4\,E\Bigg(\Bigg[\left| \sum_{n=0}^{n_t - 2} \widetilde{I}_\alpha[g(\cdot)]_{t_n,t_{n+1}} \right|^2 \\
& + 2 \left| \sum_{n=0}^{n_t - 2} \widetilde{I}_\alpha[g(\cdot)]_{t_n,t_{n+1}} \right| E\left(|\widetilde{I}_\alpha[g(\cdot)]_{t_{n_t - 1},t_{n_t}}| \Big| \mathcal{A}_{t_{n_t - 1}}\right) \\
& + E\left(|\widetilde{I}_\alpha[g(\cdot)]_{t_{n_t - 1},t_{n_t}}|^2 \Big| \mathcal{A}_{t_{n_t - 1}}\right) \Bigg] \Bigg| \mathcal{A}_{t_0}\Bigg) \\
\le\ & 4\,E\Bigg(\Bigg[\left| \sum_{n=0}^{n_t - 2} \widetilde{I}_\alpha[g(\cdot)]_{t_n,t_{n+1}} \right|^2 \\
& + E\left(|\widetilde{I}_\alpha[g(\cdot)]_{t_{n_t - 1},t_{n_t}}|^2 \Big| \mathcal{A}_{t_{n_t - 1}}\right) \Bigg] \Bigg| \mathcal{A}_{t_0}\Bigg). \qquad (4.6.15)
\end{aligned}
$$

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

Then by applying Lemma 3.6.1 we obtain

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

$$+4^{l(\alpha)-n(\alpha)} \ \Delta^{l(\alpha)+n(\alpha)-1} \int_{t_{n_t-1}}^{t_{n_t}} V_{t_{n_t-1},u,s(\alpha)}\,du\Big]\Big|\mathcal{A}_{t_0}\Big)$$

$$\leq \ 4\,E\Bigg(\Big[\big|\sum_{n=0}^{n_t-3} \widetilde{I}_\alpha[g(\cdot)]_{t_n,t_{n+1}}\big|^2$$

$$+4^{l(\alpha)-n(\alpha)} \ \Delta^{l(\alpha)+n(\alpha)-1} \int_{t_{n_t-2}}^{t_{n_t-1}} V_{t_{n_t-2},u,s(\alpha)}\,du$$

$$+4^{l(\alpha)-n(\alpha)} \ \Delta^{l(\alpha)+n(\alpha)-1} \int_{t_{n_t-1}}^{t_{n_t}} V_{t_{n_t-2},u,s(\alpha)}\,du\Big]\Big|\mathcal{A}_{t_0}\Bigg), \qquad (4.6.16)$$

where the last passage holds since $V_{t_{n_t-1},u,s(\alpha)} \leq V_{t_{n_t-2},u,s(\alpha)}$. Applying this procedure repetitively and using (4.6.8) we finally obtain

$$E\left(\sup_{t_0\leq z\leq t}\left|\sum_{n=0}^{n_z-1} \widetilde{I}_\alpha[g(\cdot)]_{t_n,t_{n+1}}\right|^2\Bigg|\mathcal{A}_{t_0}\right)$$

$$\leq \ 4^{l(\alpha)-n(\alpha)+1} \ \Delta^{l(\alpha)+n(\alpha)-1}\, E\left(\int_{t_0}^{t} V_{t_0,u,s(\alpha)}\,du\Big|\mathcal{A}_{t_0}\right)$$

$$= \ 4^{l(\alpha)-n(\alpha)+1} \ \Delta^{l(\alpha)+n(\alpha)-1} \int_{t_0}^{t} V_{t_0,u,s(\alpha)}\,du. \qquad (4.6.17)$$

For the second term of equation (4.6.14), by applying the Cauchy-Schwarz inequality, similar steps as the ones used previously and Lemma 3.6.1, we obtain

$$E\left(\sup_{t_0\leq z\leq t}\left|\widetilde{I}_\alpha[g(\cdot)]_{t_{n_z},z}\right|^2\Big|\mathcal{A}_{t_0}\right)$$

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

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

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

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

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

$$= 4^{l(\alpha)-n(\alpha)} \; \Delta^{l(\alpha)+n(\alpha)-1} \int_{t_0}^{t} V_{t_0,u,s(\alpha)} \, du, \tag{4.6.18}$$

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

Therefore, combining equations (4.6.17) and (4.6.18) we finally obtain

$$\begin{aligned} \tilde{F}_t^{\alpha} \;&\leq\; 2 \left(4^{l(\alpha)-n(\alpha)+1} \; \Delta^{l(\alpha)+n(\alpha)-1} \int_{t_0}^{t} V_{t_0,u,s(\alpha)} \, du \right. \\ &\qquad \left. + \, 4^{l(\alpha)-n(\alpha)} \; \Delta^{l(\alpha)+n(\alpha)-1} \int_{t_0}^{t} V_{t_0,u,s(\alpha)} \, du \right) \\ &\leq\; 4^{l(\alpha)-n(\alpha)+2} \; \Delta^{l(\alpha)+n(\alpha)-1} \int_{t_0}^{t} V_{t_0,u,s(\alpha)} du, \end{aligned} \tag{4.6.19}$$

which completes the proof of the lemma for $\tilde{F}_{\tau}^{\alpha}$.

Let us now prove the assertion for F_{τ}^{α}. The case of $l(\alpha) = n(\alpha)$ has been already proved since $F_{\tau}^{\alpha} = \tilde{F}_{\tau}^{\alpha}$.

Consider now the case of $l(\alpha) > n(\alpha) + s(\alpha)$ with $\alpha = (j_1, \ldots, j_l)$. Note that in this case at least one of the elements of the multi-index α belongs to $\{1, \ldots, m\}$ and, thus, the process

$$\{\sum_{n=0}^{k} I_{\alpha}[g(\cdot)]_{t_n,t_{n+1}} \, , k \in \{0, 1 \ldots, n_T - 1\}\} \tag{4.6.20}$$

is a discrete time martingale.

If $j_l = 0$, then using similar steps as in (4.6.14)–(4.6.19) we obtain

$$\begin{aligned} F_t^{\alpha} \;&\leq\; 4^{l(\alpha)-n(\alpha)+2} \; \Delta^{l(\alpha)+n(\alpha)-1} \, \hat{K}^{s(\alpha)} \int_{t_0}^{t} V_{t_0,u,s(\alpha)} du \\ &\leq\; 4^{l(\alpha)-n(\alpha)+2} \; \Delta^{l(\alpha)+n(\alpha)-1} \, \hat{C}^{s(\alpha)} \int_{t_0}^{t} V_{t_0,u,s(\alpha)} du \end{aligned} \tag{4.6.21}$$

where $\hat{K} = \frac{1}{2}(4 + \lambda(T - t_0))$, see Lemma 3.6.1, and $\hat{C} = 4 + \lambda(T - t_0)$.

If $j_l \in \{1, \ldots, m\}$, then by similar steps as in (4.6.10) we obtain

$$\begin{aligned} F_t^\alpha &\leq 4^{l(\alpha)-n(\alpha)+2}\ \Delta^{l(\alpha)+n(\alpha)-1}\ \hat{K}^{s(\alpha)} \int_{t_0}^t V_{t_0,u,s(\alpha)} du \\ &\leq 4^{l(\alpha)-n(\alpha)+2}\ \Delta^{l(\alpha)+n(\alpha)-1}\ \hat{C}^{s(\alpha)} \int_{t_0}^t V_{t_0,u,s(\alpha)} du. \end{aligned} \tag{4.6.22}$$

If $j_l = -1$, then by the decomposition (3.6.1) and the Cauchy-Schwarz inequality we obtain

$$\begin{aligned} F_t^\alpha \leq\ & 2E\left(\sup_{t_0 \leq z \leq t} \left| \int_{t_0}^z \int_{\mathcal{E}} I_{\alpha-}[g(\cdot, v^{s(\alpha)})]_{t_{n_u},u-}\ \widetilde{p}_\phi(dv^{s(\alpha)}, du) \right|^2 \Bigg| \mathcal{A}_{t_0} \right) \\ & + 2E\left(\sup_{t_0 \leq z \leq t} \left| \int_{t_0}^z \int_{\mathcal{E}} I_{\alpha-}[g(\cdot, v^{s(\alpha)})]_{t_{n_u},u}\ \phi(dv^{s(\alpha)}) du \right|^2 \Bigg| \mathcal{A}_{t_0} \right). \end{aligned} \tag{4.6.23}$$

For the first term on the right-hand side of (4.6.23) by similar steps as those used in (4.6.11) and (4.6.12), we obtain

$$\begin{aligned} & E\left(\sup_{t_0 \leq z \leq t} \left| \int_{t_0}^z \int_{\mathcal{E}} I_{\alpha-}[g(\cdot, v^{s(\alpha)})]_{t_{n_u},u-}\ \widetilde{p}_\phi(dv^{s(\alpha)}, du) \right|^2 \Bigg| \mathcal{A}_{t_0} \right) \\ & \quad \leq 4^{l(\alpha)-n(\alpha)}\ \Delta^{l(\alpha)+n(\alpha)-1}\ \hat{K}^{s(\alpha)-1} \int_{t_0}^t V_{t_0,u,s(\alpha)}\, du. \end{aligned} \tag{4.6.24}$$

For the second term on the right-hand side of (4.6.23) by similar steps as those used in (4.6.5)-(4.6.9), we obtain

$$\begin{aligned} & E\left(\sup_{t_0 \leq z \leq t} \left| \int_{t_0}^z \int_{\mathcal{E}} I_{\alpha-}[g(\cdot, v^{s(\alpha)})]_{t_{n_u},u}\ \phi(dv^{s(\alpha)}) du \right|^2 \Bigg| \mathcal{A}_{t_0} \right) \\ & \quad \leq \lambda(t - t_0) 4^{l(\alpha)-n(\alpha)-1}\ \Delta^{l(\alpha)+n(\alpha)-1}\ \hat{K}^{s(\alpha)-1} \int_{t_0}^t V_{t_0,u,s(\alpha)}\, du. \end{aligned} \tag{4.6.25}$$

By combining (4.6.23), (4.6.24) and (4.6.25), we obtain

$$F_t^\alpha \leq \frac{2}{4^3}\big(4 + \lambda(T - t_0)\big) 4^{l(\alpha)-n(\alpha)+2}\ \Delta^{l(\alpha)+n(\alpha)-1}\ \hat{K}^{s(\alpha)-1} \int_{t_0}^t V_{t_0,u,s(\alpha)}\, du$$

$$\leq \; 4^{l(\alpha)-n(\alpha)+2} \; \Delta^{l(\alpha)+n(\alpha)-1} \, \hat{K}^{s(\alpha)} \int_{t_0}^{t} V_{t_0,u,s(\alpha)} \, du$$

$$\leq \; 4^{l(\alpha)-n(\alpha)+2} \; \Delta^{l(\alpha)+n(\alpha)-1} \, \hat{C}^{s(\alpha)} \int_{t_0}^{t} V_{t_0,u,s(\alpha)} \, du. \qquad (4.6.26)$$

Let us finally consider the case of $l(\alpha) = n(\alpha) + s(\alpha)$ with $s(\alpha) \geq 1$. By using the relationship (3.2.8) one can rewrite the multiple stochastic integrals $I_\alpha[g(\cdot)]_{t_n,t_{n+1}}$ appearing in F_t^α as sum of $2^{s(\alpha)}$ multiple stochastic integrals involving integrations with respect to the compensated Poisson measure p_ϕ and to the product of time and the intensity measure $\phi(\cdot)$. Therefore, by applying the Cauchy-Schwarz inequality and similar steps as used before, we obtain

$$\begin{aligned} F_t^\alpha \; &\leq \; 2^{s(\alpha)} 4^{l(\alpha)-n(\alpha)+2} \; \Delta^{l(\alpha)+n(\alpha)-1} \, \hat{K}^{s(\alpha)} \int_{t_0}^{t} V_{t_0,u,s(\alpha)} \, du \\ &= \; 4^{l(\alpha)-n(\alpha)+2} \; \Delta^{l(\alpha)+n(\alpha)-1} \, \hat{C}^{s(\alpha)} \int_{t_0}^{t} V_{t_0,u,s(\alpha)} \, du. \qquad (4.6.27) \end{aligned}$$

This completes the proof of Lemma 4.6.1. □

4.7 Proof of Theorem 4.5.1

Before proceeding to the proof of Theorem 4.5.1, we present a lemma on the second moment estimate of the order γ compensated strong Taylor approximation (4.5.5)

Lemma 4.7.1 *Under the conditions of Theorem 4.5.1, we obtain*

$$E\Big(\sup_{0 \leq z \leq T} |Y_z^\Delta|^2 \Big| \mathcal{A}_0 \Big) \leq C\,(1 + |Y_0^\Delta|^2), \qquad (4.7.1)$$

where Y^Δ is the order γ compensated strong Taylor approximation (4.5.5).

Proof:

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

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

Therefore, we have

$$
\begin{aligned}
E(\sup_{0\le z\le T}|Y_z^\Delta|^2|\mathcal{A}_0) &\le E\left(\sup_{0\le z\le T}(1+|Y_z^\Delta|^2)\middle|\mathcal{A}_0\right)\\
&\le E\Bigg(\sup_{0\le z\le T}\Bigg(1+\Bigg|Y_0^\Delta+\sum_{\alpha\in\mathcal{A}_\gamma\setminus\{v\}}\Bigg\{\sum_{n=0}^{n_z-1}\widetilde{I}_\alpha[\tilde f_\alpha(t_n,Y_{t_n}^\Delta)]_{t_n,t_{n+1}}\\
&\qquad +\widetilde{I}_\alpha[\tilde f_\alpha(t_{n_z},Y_{t_{n_z}}^\Delta)]_{t_{n_z},z}\Bigg\}\Bigg|^2\Bigg)\Bigg|\mathcal{A}_0\Bigg)\\
&\le E\Bigg(\sup_{0\le z\le T}\Bigg(1+2\,|Y_0^\Delta|^2+2\Bigg|\sum_{\alpha\in\mathcal{A}_\gamma\setminus\{v\}}\\
&\qquad \Bigg\{\sum_{n=0}^{n_z-1}\widetilde{I}_\alpha[\tilde f_\alpha(t_n,Y_{t_n}^\Delta)]_{t_n,t_{n+1}}+\widetilde{I}_\alpha[\tilde f_\alpha(t_{n_z},Y_{t_{n_z}}^\Delta)]_{t_{n_z},z}\Bigg\}\Bigg|^2\Bigg)\Bigg|\mathcal{A}_0\Bigg)\\
&\le C_1\left(1+\left|Y_0^\Delta\right|^2\right)+2\,K\sum_{\alpha\in\mathcal{A}_\gamma\setminus\{v\}}E\Bigg(\sup_{0\le z\le T}\\
&\qquad \Bigg|\sum_{n=0}^{n_z-1}\widetilde{I}_\alpha[\tilde f_\alpha(t_n,Y_{t_n}^\Delta)]_{t_n,t_{n+1}}+\widetilde{I}_\alpha[\tilde f_\alpha(t_{n_z},Y_{t_{n_z}}^\Delta)]_{t_{n_z},z}\Bigg|^2\Bigg|\mathcal{A}_0\Bigg),
\end{aligned}
\tag{4.7.2}
$$

where K is a positive constant depending only on the strong order γ of the approximation. By Lemma 4.6.1 and the linear growth condition (4.5.9) we obtain

$$
\begin{aligned}
&E\left(\sup_{0\le z\le T}|Y_z^\Delta|^2\middle|\mathcal{A}_0\right)\\
&\le C_1(1+|Y_0^\Delta|^2)+2\,K_1\sum_{\alpha\in\mathcal{A}_\gamma\setminus\{v\}}\Bigg\{\int_0^T\int_{\mathcal{E}}\cdots\int_{\mathcal{E}}\\
&\quad\times E\left(\sup_{0\le z\le u}\left|\tilde f_\alpha(z,Y_z^\Delta)\right|^2\middle|\mathcal{A}_0\right)\,\phi(dv^1)\ldots\phi(dv^{s(\alpha)})\,du\Bigg\}
\end{aligned}
$$

$$\leq C_1\left(1+|Y_0^\Delta|^2\right)+2\,K_1$$
$$\times \sum_{\alpha\in\mathcal{A}_\gamma\setminus\{v\}}\left\{\int_{\mathcal{E}}\cdots\int_{\mathcal{E}}K_2(v^1,\ldots,v^{s(\alpha)})\phi(dv^1)\ldots\phi(dv^{s(\alpha)})\right.$$
$$\left.\times\int_0^T E\left(\sup_{0\leq z\leq u}\left(1+|Y_z^\Delta|^2\right)\bigg|\mathcal{A}_0\right)du\right\}$$
$$\leq C_1\left(1+|Y^\Delta(0)|^2\right)+C_2\int_0^T E\left(\sup_{0\leq z\leq u}\left(1+|Y_z^\Delta|^2\right)\bigg|\mathcal{A}_0\right)du. \tag{4.7.3}$$

Then by applying the Gronwall inequality, we obtain

$$E\Big(\sup_{0\leq z\leq T}|Y_z^\Delta|^2\Big|\mathcal{A}_0\Big)\leq C\,(1+|Y_0^\Delta|^2), \tag{4.7.4}$$

where C is a positive finite constant independent of Δ. □

Now by using Lemma 4.6.1 and Lemma 4.7.1, we can finally prove Theorem 4.5.1.

Proof:

1. With the Wagner-Platen expansion (3.5.4) we can represent the solution of the SDE (2.1.2) as

$$X_\tau=\sum_{\alpha\in\mathcal{A}_\gamma}\widetilde{I}_\alpha[\tilde{f}_\alpha(\rho,X_\rho)]_{\rho,\tau}+\sum_{\alpha\in\mathcal{B}(\mathcal{A}_\gamma)}\widetilde{I}_\alpha[\tilde{f}_\alpha(\cdot,X_\cdot)]_{\rho,\tau}, \tag{4.7.5}$$

for any two stopping times ρ and τ with $0\leq\rho\leq\tau\leq T$ a.s. Therefore, we can express the solution of the SDE (2.1.2) at time $t\in[0,T]$ as

$$\begin{aligned}X_t &= X_0+\sum_{\alpha\in\mathcal{A}_\gamma\setminus\{v\}}\left\{\sum_{n=0}^{n_t-1}\widetilde{I}_\alpha[\tilde{f}_\alpha(t_n,X_{t_n})]_{t_n,t_{n+1}}+\widetilde{I}_\alpha[\tilde{f}_\alpha(t_{n_t},X_{t_{n_t}})]_{t_{n_t},t}\right\}\\ &\quad+\sum_{\alpha\in\mathcal{B}(\mathcal{A}_\gamma)}\left\{\sum_{n=0}^{n_t-1}\widetilde{I}_\alpha[\tilde{f}_\alpha(\cdot,X_\cdot)]_{t_n,t_{n+1}}+\widetilde{I}_\alpha[\tilde{f}_\alpha(\cdot,X_\cdot)]_{t_{n_t},t}\right\},\end{aligned} \tag{4.7.6}$$

where n_t is defined as in equation (4.1.8).

We recall from (4.5.3) that the order γ compensated strong Taylor approximation Y^Δ at time $t \in [0,T]$ is given by

$$Y_t^\Delta = Y_0^\Delta + \sum_{\alpha\in\mathcal{A}_\gamma\setminus\{v\}} \left\{ \sum_{n=0}^{n_t-1} \tilde{I}_\alpha[\tilde{f}_\alpha(t_n, Y_{t_n}^\Delta)]_{t_n,t_{n+1}} + \tilde{I}_\alpha[\tilde{f}_\alpha(t_{n_t}, Y_{t_{n_t}}^\Delta)]_{t_{n_t},t} \right\}. \tag{4.7.7}$$

From the moment estimate (2.2.13) provided by Theorem 2.2.1 we have

$$E\left(\sup_{0\le z\le T} |X_z|^2 \Big| \mathcal{A}_0\right) \le C\Big(1 + E\left(|X_0|^2\right)\Big). \tag{4.7.8}$$

By Lemma 4.7.1, we obtain a similar uniform estimate for the second moment of the approximation Y^Δ

$$E\Big(\sup_{0\le z\le T} |Y_z^\Delta|^2 \Big| \mathcal{A}_0\Big) \le C\,(1 + |Y_0^\Delta|^2). \tag{4.7.9}$$

2. Let us now analyze the mean square error of the order γ compensated strong Taylor approximation Y^Δ. By (4.7.6), (4.7.7) and Cauchy-Schwarz inequality we obtain

$$\begin{aligned} Z(t) &:= E\left(\sup_{0\le z\le t} |X_z - Y_z^\Delta|^2 \Big| \mathcal{A}_0\right) \\ &= E\Bigg(\sup_{0\le z\le t} \Big| X_0 - Y_0^\Delta \\ &\quad + \sum_{\alpha\in\mathcal{A}_\gamma\setminus\{v\}} \Big\{ \sum_{n=0}^{n_z-1} \tilde{I}_\alpha[\tilde{f}_\alpha(t_n, X_{t_n}) - \tilde{f}_\alpha(t_n, Y_{t_n}^\Delta)]_{t_n,t_{n+1}} \\ &\quad + \tilde{I}_\alpha[\tilde{f}_\alpha(t_{n_z}, X_{t_{n_z}}) - \tilde{f}_\alpha(t_{n_z}, Y_{t_{n_z}}^\Delta)]_{t_{n_z},z} \Big\} \\ &\quad + \sum_{\alpha\in\mathcal{B}(\mathcal{A}_\gamma)} \Big\{ \sum_{n=0}^{n_z-1} \tilde{I}_\alpha[\tilde{f}_\alpha(\cdot, X_\cdot)]_{t_n,t_{n+1}} + \tilde{I}_\alpha[\tilde{f}_\alpha(\cdot, X_\cdot)]_{t_{n_z},z} \Big\} \Big|^2 \Bigg| \mathcal{A}_0 \Bigg) \\ &\le C_3 \left\{ |X_0 - Y_0^\Delta|^2 + \sum_{\alpha\in\mathcal{A}_\gamma\setminus\{v\}} S_t^\alpha + \sum_{\alpha\in\mathcal{B}(\mathcal{A}_\gamma)} U_t^\alpha \right\} \end{aligned} \tag{4.7.10}$$

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

$$\begin{aligned} S_t^\alpha &:= E\Bigg(\sup_{0\le z\le t} \Bigg| \sum_{n=0}^{n_z-1} \widetilde{I}_\alpha \left[\tilde{f}_\alpha(t_n, X_{t_n}) - \tilde{f}_\alpha(t_n, Y_{t_n}^\Delta)\right]_{t_n,t_{n+1}} \\ &\quad + \widetilde{I}_\alpha \left[\tilde{f}_\alpha(t_{n_z}, X_{t_{n_z}}) - \tilde{f}_\alpha(t_{n_z}, Y_{t_{n_z}}^\Delta)\right]_{t_{n_z},z} \Bigg|^2 \Bigg| \mathcal{A}_0 \Bigg), \end{aligned} \tag{4.7.11}$$

$$U_t^\alpha := E\left(\sup_{0\le z\le t} \left| \sum_{n=0}^{n_z-1} \widetilde{I}_\alpha[\tilde{f}_\alpha(\cdot, X_\cdot)]_{t_n,t_{n+1}} + \widetilde{I}_\alpha[\tilde{f}_\alpha(\cdot, X_\cdot)]_{t_{n_z},z} \right|^2 \Bigg| \mathcal{A}_0 \right). \tag{4.7.12}$$

3. By using again Lemma 4.6.1 and the Lipschitz condition (4.5.7) we obtain

$$\begin{aligned} S_t^\alpha &= E\Bigg(\sup_{0\le z\le t} \Bigg| \sum_{n=0}^{n_z-1} \widetilde{I}_\alpha \left[\tilde{f}_\alpha(t_n, X_{t_n}) - \tilde{f}_\alpha(t_n, Y_{t_n}^\Delta)\right]_{t_n,t_{n+1}} \\ &\quad + \widetilde{I}_\alpha \left[\tilde{f}_\alpha(t_{n_z}, X_{t_{n_z}}) - \tilde{f}_\alpha(t_{n_z}, Y_{t_{n_z}}^\Delta)\right]_{t_{n_z},z} \Bigg|^2 \Bigg| \mathcal{A}_0 \Bigg) \\ &\le C_4 \int_0^t \int_{\mathcal{E}} \cdots \int_{\mathcal{E}} E\left(\sup_{0\le z\le u} |\tilde{f}_\alpha(t_{n_z}, X_{t_{n_z}}) - \tilde{f}_\alpha(t_{n_z}, Y_{t_{n_z}}^\Delta)|^2 \Bigg| \mathcal{A}_0 \right) \\ &\quad \times \phi(dv^1)\ldots\phi(dv^{s(\alpha)})du \\ &\le C_4 \int_{\mathcal{E}} \cdots \int_{\mathcal{E}} \left(K_1(v^1,\ldots,v^{s(\alpha)})\right)^2 \phi(dv^1)\ldots\phi(dv^{s(\alpha)}) \\ &\quad \times \int_0^t E\left(\sup_{0\le z\le u} |X_{t_{n_z}} - Y_{t_{n_z}}^\Delta|^2 \Bigg| \mathcal{A}_0 \right) du \\ &\le C_5 \int_0^t Z(u)\, du. \end{aligned} \tag{4.7.13}$$

Applying again Lemma 4.6.1 and the linear growth condition (4.5.9) we obtain

$$\begin{aligned}
U_t^\alpha &= E\left(\sup_{0\le z\le t}\left|\sum_{n=0}^{n_z-1}\widetilde{I}_\alpha[\tilde{f}_\alpha(\cdot,X_\cdot)]_{t_n,t_{n+1}}+\widetilde{I}_\alpha[\tilde{f}_\alpha(\cdot,X_\cdot)]_{t_{n_z},z}\right|^2\middle|\mathcal{A}_0\right)\\
&\le C_5\,\Delta^{\psi(\alpha)}\int_0^t\int_{\mathcal{E}}\cdots\int_{\mathcal{E}}E\left(\sup_{0\le z\le u}|\tilde{f}_\alpha(z,X_z)|^2\middle|\mathcal{A}_0\right)\phi(dv^1)\ldots\phi(dv^{s(\alpha)})du\\
&\le C_5\,\Delta^{\psi(\alpha)}\int_{\mathcal{E}}\cdots\int_{\mathcal{E}}K_2(v^1,\ldots,v^{s(\alpha)})\phi(dv^1)\ldots\phi(dv^{s(\alpha)})\\
&\quad\times\int_0^t E\left(\sup_{0\le z\le u}(1+|X_z|^2)\middle|\mathcal{A}_0\right)du\\
&\le C_6\,\Delta^{\psi(\alpha)}\left(t+\int_0^t E\Big(\sup_{0\le z\le u}|X_z|^2\Big|\mathcal{A}_0\Big)du\right), \qquad (4.7.14)
\end{aligned}$$

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, by applying estimate (2.2.13) of Theorem 2.2.1 we obtain

$$\begin{aligned}
U_t^\alpha &\le C_6\,\Delta^{2\gamma}\left(t+\int_0^t C_1(1+|X_0|^2)du\right)\\
&\le C_7\,\Delta^{2\gamma}\,(1+|X_0|^2). \qquad (4.7.15)
\end{aligned}$$

4. Combining equations (4.7.10), (4.7.13) and (4.7.15) we obtain

$$Z(t)\le C_8\left\{\;|X_0-Y_0^\Delta|^2+C_9\,\Delta^{2\gamma}\left(1+|X_0|^2\right)+C_{10}\int_0^t Z(u)du\right\}. \qquad (4.7.16)$$

By equations (4.7.8) and (4.7.9) $Z(t)$ is bounded. Therefore, by the Gronwall inequality we have

$$Z(T)\le K_4\,\left(1+|X_0|^2\right)\,\Delta^{2\gamma}+K_5\left(|X_0-Y_0^\Delta|^2\right). \qquad (4.7.17)$$

Finally, by assumption (4.5.6), we obtain

$$\sqrt{E(\sup_{0\leq z\leq T}|X_z - Y_z^\Delta|^2\,|\mathcal{A}_0)} = \sqrt{Z(T)} \leq K_3\ \Delta^\gamma, \tag{4.7.18}$$

which completes the proof of Theorem 4.5.1. □

Remark 4.7.2 *Note that the same result as that in Theorem 4.5.1 holds also for the order γ strong Taylor approximation* (4.5.4) *as mentioned in Corollary 4.5.3. The proof uses analogous steps as those used in the proof described above.*

Chapter 5

Regular Strong Itô Approximations

In this chapter we describe regular strong approximations that are more general than the regular strong Taylor approximations presented in the previous chapter. These approximations belong to the class of regular strong Itô schemes, which includes derivative-free, implicit and predictor-corrector schemes. Some of the results to be presented in this chapter have been published in Bruti-Liberati, Nikitopoulos-Sklibosios & Platen (2006). A working paper, Bruti-Liberati & Platen (2007d), on strong predictor-corrector methods is in preparation.

5.1 Introduction

The first types of schemes that we describe in this chapter are the so-called *derivative-free schemes.* Higher order strong Taylor schemes, as the order 1.0 strong Taylor scheme presented in Section 4.3, are rather complex as they involve the evaluation of derivatives of the drift, diffusion and jump coefficients at each time step. For the implementation of general numerical routines for the approximation of jump-diffusion SDEs, without assuming a particular form for the coefficients, this constitutes a serious limitation. In principle, one is required to include a symbolic differentiation into a numerical algorithm. For these reasons, we present in this chapter derivative-free strong schemes that avoid the computation of derivatives.

In the second part of this chapter we present *implicit schemes.* As shown in Hofmann & Platen (1996) and in Higham & Kloeden (2005, 2006), when one has multiplicative noise explicit methods show narrow regions of numerical stability. We emphasize that SDEs with multiplicative noise are typically used when modelling asset prices in finance. They also arise in other important applications such as hidden Markov chain filtering, see Elliott, Aggoun & Moore (1995). In order to construct approximate filters, one needs a strong discrete time approximation of an SDE with multiplicative noise, the Zakai equation. Moreover, in filtering problems for large systems it is often not possible to use 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 numerical stability are crucial. To overcome some of these problems, we describe implicit schemes that have satisfactory numerical stability properties.

As discussed above, explicit schemes have narrower regions of numerical stability than corresponding implicit schemes. For this reason implicit schemes for diffusion and jump-diffusion SDEs have been proposed. Because of their improved numerical stability properties, implicit schemes can be used with much larger time step sizes than those required by explicit schemes. However, implicit schemes carry, in general, an additional computational burden since they usually require the solution of an algebraic equation at each time step. Therefore, in choosing between an explicit and an implicit scheme one faces a trade-off between computational efficiency and numerical stability. Additionally, as will be explained later, when designing an implicit scheme, it is not easy to introduce implicitness in the diffusion coefficient. This is due to problems that arise with the reciprocal of Gaussian random variables. For these reasons, we will present new *predictor-corrector schemes* that aim to combine good numerical stability properties and efficiency. A detailed investigation of the numerical stability properties of predictor-corrector schemes is left for future research.

5.2 Derivative-Free Order 1.0 Scheme

By replacing the derivatives in the order 1.0 strong Taylor scheme, presented in Section 4.3, by the corresponding difference ratios, it is possible to obtain a scheme that does not require the evaluation of derivatives and achieves the same strong order of convergence. However, to construct the difference ratios we need supporting values of the coefficients at additional points.

In the one-dimensional case, $d = m = 1$, the *derivative-free order* 1.0 *strong scheme*, is given by

$$
\begin{aligned}
Y_{n+1} &= Y_n + a\Delta_n + b\Delta W_n + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} c(v)\, p_\phi(dv, dz) \\
&\quad + \frac{\left(b(t_n, \overline{Y_n}) - b\right)}{\sqrt{\Delta_n}} \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} dW(z_1)\, dW(z_2) \\
&\quad + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} \frac{\left(c(t_n, \overline{Y_n}, v) - c(v)\right)}{\sqrt{\Delta_n}}\, dW(z_1)\, p_\phi(dv, dz_2)
\end{aligned}
$$

$$
\begin{aligned}
&+ \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} \int_{\mathcal{E}} \left\{ b(t_n, Y_n + c(v)) - b \right\} p_\phi(dv, dz_1)\, dW(z_2) \qquad (5.2.1) \\
&+ \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} \int_{\mathcal{E}} \left\{ c(t_n, Y_n + c(v_2), v_1) - c(v_1) \right\} p_\phi(dv_1, dz_1) p_\phi(dv_2, dz_2),
\end{aligned}
$$

with the supporting value

$$
\overline{Y_n} = Y_n + b\sqrt{\Delta_n}. \qquad (5.2.2)
$$

The scheme (5.2.1)-(5.2.2) generally achieves a strong order $\gamma = 1.0$ and is a generalization of a corresponding scheme proposed in Platen (1984) for pure diffusions.

We can simplify the double stochastic integrals appearing in (5.2.1), as in Section 4.3, and rewrite the derivative-free order 1.0 strong Taylor scheme as

$$
\begin{aligned}
Y_{n+1} &= Y_n + a\Delta_n + b\Delta W_n + \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} c(\xi_i) + \frac{\left(b(t_n, \overline{Y_n}) - b\right)}{2\sqrt{\Delta_n}} \left((\Delta W_n)^2 - \Delta_n \right) \\
&\quad + \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} \frac{\left(c(t_n, \overline{Y_n}, \xi_i) - c(\xi_i)\right)}{\sqrt{\Delta_n}} \left(W(\tau_i) - W(t_n) \right) \\
&\quad + \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} \left\{ b\left(Y_n + c(\xi_i)\right) - b \right\} \left(W(t_{n+1}) - W(\tau_i) \right) \\
&\quad + \sum_{j=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} \sum_{i=p_\phi(t_n)+1}^{p_\phi(\tau_j)} \left\{ c\left(Y_n + c(\xi_i), \xi_j\right) - c\left(\xi_j\right) \right\}, \qquad (5.2.3)
\end{aligned}
$$

with supporting value (5.2.2), which can be directly used in a scenario simulation.

As discussed in Section 4.4, in the case of mark-independent jump size it is recommended to check the particular structure of the SDE under consideration. Indeed, we derived the jump commutativity condition (4.4.2) under which the order 1.0 strong Taylor scheme (4.3.4) exhibits a computational complexity that is independent of the intensity level of the Poisson random measure.

For the derivative-free order 1.0 strong scheme (5.2.1)-(5.2.2) with mark-independent jump size, the derivative-free coefficient of the multiple stochastic integral $I_{(1,-1)}$, which is of the form

$$\frac{c(t_n, \overline{Y_n}) - c(t_n, Y_n)}{\sqrt{\Delta_n}},$$

depends on the time step size Δ_n. On the other hand, the coefficient of the multiple stochastic integral $I_{(-1,1)}$,

$$b\Big(t_n, Y_n + c(t_n, Y_n)\Big) - b\Big(t_n, Y_n\Big),$$

is independent of Δ_n. Therefore, it is not possible to directly derive a commutativity condition similar to (4.4.2) that permits to identify special classes of SDEs for which the computational efficiency is independent of the jump intensity level.

For instance, for the SDE (2.1.5) with mark-independent jump size $c(t, x, v) = x\,\beta$, with $\beta \geq -1$, the derivative-free order 1.0 strong scheme is given by

$$\begin{aligned} Y_{n+1} &= Y_n + \mu Y_n \Delta_n + \sigma Y_n \Delta W_n + \beta\, Y_n\, \Delta p_n + \frac{\sigma}{\sqrt{\Delta}} \{\overline{Y_n} - Y_n\} I_{(1,1)} \\ &\quad + \frac{\beta}{\sqrt{\Delta}} \{\overline{Y_n} - Y_n\} I_{(1,-1)} + \sigma\,\beta\, Y_n\, I_{(-1,1)} + \beta^2\, Y_n I_{(-1,-1)}, \end{aligned} \tag{5.2.4}$$

with the supporting value

$$\overline{Y_n} = Y_n + \sigma Y_n \sqrt{\Delta_n}. \tag{5.2.5}$$

Since the evaluation of the multiple stochastic integrals $I_{(1,-1)}$ and $I_{(-1,1)}$, as given in (4.3.5), depends on the number of jumps, the computational efficiency of the scheme (5.2.4)-(5.2.5) depends on the total intensity λ of the jump measure.

Let us consider the special class of one-dimensional SDEs satisfying the jump commutativity condition (4.4.2) , which we recall here in the form

$$b(t, x)\frac{\partial c(t, x)}{\partial x} = b\Big(t, x + c(t, x)\Big) - b\Big(t, x\Big), \tag{5.2.6}$$

for all $t \in [0,T]$ and $x \in \mathbb{R}$. Under this condition, we should first derive the order 1.0 strong Taylor scheme, using the relationship

$$I_{(1,-1)} + I_{(-1,1)} = \Delta p_n \, \Delta W_n, \tag{5.2.7}$$

obtaining

$$\begin{aligned} Y_{n+1} &= Y_n + a\Delta_n + b\Delta W_n + c\Delta p_n + \frac{b\,b'}{2}\{(\Delta W_n)^2 - \Delta_n\} \\ &\quad + \{b(t_n, Y_n + c) - b\}\Delta p_n \, \Delta W_n + \frac{\{c(t_n, Y_n + c) - c\}}{2}\{(\Delta p_n)^2 - \Delta p_n\}, \end{aligned} \tag{5.2.8}$$

where we have used again the abbreviation (4.1.9). Then, we replace the derivative b' by the corresponding difference ratio and obtain a *derivative-free order 1.0 strong scheme*

$$\begin{aligned} Y_{n+1} &= Y_n + a\Delta_n + b\Delta W_n + c\Delta p_n \\ &\quad + \frac{\{b(t_n, \overline{Y}_n) - b\}}{2\sqrt{\Delta_n}}\{(\Delta W_n)^2 - \Delta_n\} + \{b(t_n, Y_n + c) - b\}\Delta p_n \, \Delta W_n \\ &\quad + \frac{\{c(t_n, Y_n + c) - c\}}{2}\{(\Delta p_n)^2 - \Delta p_n\}, \end{aligned} \tag{5.2.9}$$

with supporting value given in (5.2.2). The computational efficiency is here independent of the intensity level.

Let us discuss an even more specific example. For the SDE (2.1.5) with $c(t,x,v) = x\,\beta$, for $\beta \geq -1$, we can derive the derivative-free order 1.0 strong scheme, which, due to the multiplicative form of the diffusion coefficient, is the same as the order 1.0 strong Taylor scheme (4.4.6).

In the multi-dimensional case with scalar Wiener process, which means $m = 1$, and mark-dependent jump size, the kth component of the *derivative-free order 1.0 strong scheme* is given by

$$\begin{aligned} Y^k_{n+1} &= Y^k_n + a^k\Delta_n + b^k\Delta W_n + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} c^k(v)\, p_\phi(dv, dz) \\ &\quad + \frac{\left(b^k(t_n, \overline{Y}_n) - b^k\right)}{\sqrt{\Delta_n}} \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} dW(z_1)\, dW(z_2) \end{aligned}$$

$$+\int_{t_n}^{t_{n+1}}\int_{\mathcal{E}}\int_{t_n}^{z_2}\frac{\Big(c^k(t_n,\overline{Y}_n,v)-c^k(v)\Big)}{\sqrt{\Delta_n}}\,dW(z_1)\,p_\phi(dv,dz_2)$$
$$+\int_{t_n}^{t_{n+1}}\int_{t_n}^{z_2}\int_{\mathcal{E}}\Big\{b^k(t_n,Y_n+c(v))-b^k\Big\}\,p_\phi(dv,dz_1)\,dW(z_2)$$
$$+\int_{t_n}^{t_{n+1}}\int_{\mathcal{E}}\int_{t_n}^{z_2}\int_{\mathcal{E}}\Big\{c^k(t_n,Y_n+c(v_2),v_1)-c^k(v_1)\Big\}$$
$$\times\,p_\phi(dv_1,dz_1)\,p_\phi(dv_2,dz_2), \tag{5.2.10}$$

with the supporting value (5.2.2).

As noticed for the one-dimensional situation, even in the case of mark-independent jump size, it is not possible to derive a jump commutativity condition similar to (4.4.10), since the coefficient of the multiple stochastic integral $I_{(1,-1)}$ depends on the time step size Δ_n. However, as shown in Section 4.4, it makes sense to consider the special class of multi-dimensional SDEs with a scalar Wiener process and mark-independent jump size characterized by the jump commutativity condition

$$\sum_{l=1}^{d} b^l(t,x)\,\frac{\partial c^k(,t,x)}{\partial x^l}=b^k\Big(t,x+c(t,x)\Big)-b^k\Big(t,x\Big) \tag{5.2.11}$$

for $k\in\{1,2,\ldots,d\}$, $t\in[0,T]$ and $x\in\mathbb{R}^d$. Using the relationship (5.2.7) one can derive, in this case, an order 1.0 strong Taylor scheme

$$\begin{aligned}Y^k_{n+1} &= Y^k_n+a^k\Delta_n+b^k\Delta W_n+c^k\Delta p_n+\frac{1}{2}\sum_{l=1}^{d}b^l\,\frac{\partial b^k}{\partial x^l}\{(\Delta W_n)^2-\Delta_n\} \qquad (5.2.12)\\ &\quad+\{b^k(t_n,Y_n+c)-b^k\}\Delta p_n\,\Delta W_n+\frac{\{c^k(t_n,Y_n+c)-c^k\}}{2}\{(\Delta p_n)^2-\Delta p_n\},\end{aligned}$$

whose computational complexity is independent of the intensity λ. Replacing the coefficient

$$\sum_{l=1}^{d} b^l\,\frac{\partial b^k}{\partial x^l}$$

by the corresponding difference ratio, we obtain a derivative-free order 1.0 strong

Taylor scheme of the form

$$\begin{aligned} Y_{n+1}^k &= Y_n^k + a^k\Delta_n + b^k\Delta W_n + c^k\,\Delta p_n \\ &+\frac{\{b^k(t_n,\overline{Y}_n) - b^k\}}{2\sqrt{\Delta_n}}\{(\Delta W_n)^2 - \Delta_n\} + \{b^k(t_n, Y_n + c) - b^k\}\Delta p_n\,\Delta W_n \\ &+\frac{\{c^k(t_n, Y_n + c) - c^k\}}{2}\{(\Delta p_n)^2 - \Delta p_n\}, \end{aligned} \tag{5.2.13}$$

with supporting value (5.2.2), for $k \in \{1, 2, \ldots, d\}$. The computational complexity of the scheme (5.2.13) is independent of the jump intensity level.

In the general multi-dimensional case the kth component of the *derivative-free order 1.0 strong scheme* is given by

$$\begin{aligned} Y_{n+1}^k &= Y_n^k + a^k\Delta_n + \sum_{j=1}^m b^{k,j}\Delta W_n^j + \int_{t_n}^{t_{n+1}}\int_{\mathcal{E}} c^k(v)\,p_\phi(dv, dz) \\ &+\frac{1}{\sqrt{\Delta_n}}\sum_{j_1,j_2=1}^m \int_{t_n}^{t_{n+1}}\int_{t_n}^{z_2}\Big\{b^{k,j_1}(t_n,\overline{Y}_n^{j_2}) - b^{k,j_1}\Big\}dW^{j_1}(z_1)dW^{j_2}(z_2) \\ &+\sum_{j_1=1}^m \int_{t_n}^{t_{n+1}}\int_{\mathcal{E}}\int_{t_n}^{z_2}\Big\{c^{k,j_1}(t_n,\overline{Y}_n^{j_2}) - c^{k,j_1}\Big\}\,dW^{j_1}(z_1)p_\phi(dv, dz_2) \\ &+\sum_{j_1=1}^m \int_{t_n}^{t_{n+1}}\int_{\mathcal{E}}\int_{t_n}^{z_2}\Big\{b^{k,j_1}\big(t_n, Y_n + c(v)\big) - b^{k,j_1}\Big\}p_\phi(dv, dz_2)\,dW^{j_1}(z_2) \\ &+\int_{t_n}^{t_{n+1}}\int_{\mathcal{E}}\int_{t_n}^{z_2}\int_{\mathcal{E}}\Big\{c^k\big(t_n, Y_n + c(v_1), v_2\big) - c^k\big(v_2\big)\Big\} \\ &\times p_\phi(dv_1, dz_1)\,p_\phi(dv_2, dz_2), \end{aligned} \tag{5.2.14}$$

for $k \in \{1, 2, \ldots, d\}$, with the vector supporting values

$$\overline{Y}_n^j = Y_n + b^j\sqrt{\Delta_n}, \tag{5.2.15}$$

for $j \in \{1, 2, \ldots, m\}$.

As shown in Section (4.4), for the special class of general multi-dimensional SDEs with mark-independent jump size, satisfying the jump commutativity condition

$$\sum_{l=1}^d b^{l,j_1}(t,x)\,\frac{\partial c^k(t,x)}{\partial x^l} = b^{k,j_1}\Big(t, x + c(t,x)\Big) - b^{k,j_1}\Big(t,x\Big) \tag{5.2.16}$$

for $j_1 \in \{1, 2, \ldots, m\}$, $k \in \{1, 2, \ldots, d\}$, $t \in [0, T]$ and $x \in \mathbb{R}^d$, it is possible to derive an order 1.0 strong Taylor scheme whose computational complexity is independent on the intensity level. Here one needs to use the relationship

$$I_{(j_1,-1)} + I_{(-1,j_1)} = \Delta p_n \, \Delta W_n^{j_1},$$

for $j_1 \in \{1, 2, \ldots, m\}$. Then, by replacing the coefficients involving the derivatives by the corresponding difference ratios, we obtain a derivative-free order 1.0 strong scheme that shows a computational complexity independent of the jump intensity level. This scheme is given by

$$\begin{aligned} Y_{n+1}^k &= Y_n^k + a^k \Delta_n + \sum_{j=1}^m b^{k,j} \Delta W_n^j + c^k \, \Delta p_n \\ &+ \frac{1}{2\sqrt{\Delta_n}} \sum_{j_1, j_2 = 1}^m \left\{ b^{k,j_1}(t_n, \overline{Y}_n^{j_2}) - b^{k,j_1}(t_n, Y_n) \right\} I_{(j_1,j_2)} \\ &+ \sum_{j=1}^m \left\{ b^{k,j_1}(t_n, Y_n + c) - b^{k,j_1} \right\} \Delta p_n \, \Delta W_n^{j_1} \\ &+ \frac{\{c^k(t_n, Y_n + c) - c^k\}}{2} \{(\Delta p_n)^2 - \Delta p_n\}, \end{aligned} \tag{5.2.17}$$

for $k \in \{1, 2, \ldots, d\}$, with the vector supporting values (5.2.15). For the generation of the multiple stochastic integral $I_{(j_1,j_2)}$, for $j_1, j_2 \in \{1, 2, \ldots, d\}$, we refer to Section 4.4.

For the special case of a multi-dimensional SDE satisfying the jump commutativity condition (5.2.16), the diffusion commutativity condition

$$L^{j_1} b^{k,j_2}(t, x) = L^{j_2} b^{k,j_1}(t, x) \tag{5.2.18}$$

for $j_1, j_2 \in \{1, 2, \ldots, m\}$, $k \in \{1, 2, \ldots, d\}$, $t \in [0, T]$ and $x \in \mathbb{R}^d$, and with mark-independent jump size, we obtain an efficiently implementable derivative-free order

1.0 strong scheme. Its kth component is given by

$$\begin{aligned}
Y_{n+1}^k &= Y_n^k + a^k \Delta_n + \sum_{j=1}^m b^{k,j} \Delta W_n^j + c^k \Delta p_n \\
&+ \frac{1}{2} \sum_{j_1,j_2=1}^m \left\{ b^{k,j_1}(t_n, \overline{Y}_n^{j_2}) - b^{k,j_1}(t_n, Y_n) \right\} \left\{ \Delta W_n^{j_1} \Delta W_n^{j_2} - \Delta_n \right\} \\
&+ \sum_{j_1=1}^m \{ b^{k,j_1}(t_n, Y_n + c) - b^{k,j_1} \} \left(\Delta p_n \Delta W_n^{j_1} \right) \\
&+ \frac{1}{2} \{ c^k(t_n, Y_n + c) - c^k \} \left((\Delta p_n)^2 - \Delta p_n \right), \qquad (5.2.19)
\end{aligned}$$

for $k \in \{1, 2, \ldots, d\}$, with the vector supporting values (5.2.15).

5.3 Drift-Implicit Schemes

In general, given an explicit scheme of strong order γ it is usually possible to obtain a similar drift-implicit scheme of the same order. However, since the reciprocal of a Gaussian random variable does not have finite absolute moments, it is not easy to introduce implicitness in the diffusion coefficient. Regions of numerical stability of drift-implicit schemes are typically wider than those of corresponding explicit schemes. Therefore, the former are often more suitable for a range of problems than corresponding explicit schemes. In this section we present drift-implicit strong schemes. In Higham & Kloeden (2005, 2006, 2007), a class of implicit methods of strong order $\gamma = 0.5$ for jump-diffusion SDEs has been proposed. A detailed stability analysis shows that these schemes have good numerical stability properties. In the following we focus on drift-implicit schemes of strong order $\gamma = 0.5$ and $\gamma = 1.0$ for the jump-diffusion SDE (4.1.4).

5.3.1 Drift-Implicit Euler Scheme

In the one-dimensional case, $d = m = 1$, by introducing implicitness in the drift of the Euler scheme (4.2.1), we obtain the *drift-implicit Euler scheme*,

$$Y_{n+1} = Y_n + \{\theta\, a(t_{n+1}, Y_{n+1}) + (1-\theta)\, a\}\, \Delta_n + b\Delta W_n + \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} c^k(\xi_i), \quad (5.3.1)$$

where the parameter $\theta \in [0,1]$ characterizes the degree of implicitness and we have used the abbreviation defined in (4.1.9). For $\theta = 0$ we recover the Euler scheme (4.2.1), while for $\theta = 1$, we obtain a fully drift-implicit Euler scheme. The scheme (5.3.1) achieves a strong order of convergence $\gamma = 0.5$. This scheme was proposed and analyzed in Higham & Kloeden (2006) for SDEs driven by Wiener processes and homogeneous Poisson processes. It generalizes the drift-implicit Euler scheme for pure diffusions presented in Talay (1982b) and Milstein (1988).

By comparing the drift-implicit Euler scheme (5.3.1) with the Euler scheme (4.2.1), one notices that the implementation of the former requires an additional computational effort; an algebraic equation has to be solved at each time step. This can be performed, for instance, by a Newton-Raphson method. In special cases, however, the algebraic equation may admit an explicit solution. Note that the existence and uniqueness of the solution of this algebraic equation is guaranteed by Banach's fixed point theorem, see for instance Evans (1999), for every

$$\Delta \leq \frac{1}{\sqrt{K\theta}},$$

where K is the Lipschitz constant appearing in the Lipschitz condition (2.2.10) for the drift coefficient a.

When we have a mark-independent jump size we obtain the *drift-implicit Euler scheme*

$$Y_{n+1} = Y_n + \{\theta\, a(t_{n+1}, Y_{n+1}) + (1-\theta)\, a\}\, \Delta_n + b\Delta W_n + c\Delta p_n. \quad (5.3.2)$$

In the multi-dimensional case with scalar Wiener process, $m = 1$, and mark-dependent jump size, the kth component of the *drift-implicit Euler scheme* is given by

$$Y_{n+1}^k = Y_n^k + \left\{\theta\, a^k(t_{n+1}, Y_{n+1}) + (1-\theta)\, a^k\right\} \Delta_n + b^k \Delta W_n + \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} c^k(\xi_i),$$

for $k \in \{1, 2, \ldots, d\}$.

In the case of a mark-independent jump size the k-th component of the *drift-implicit Euler scheme* (5.3.3) reduces to

$$Y_{n+1}^k = Y_n^k + \left\{\theta\, a^k(t_{n+1}, Y_{n+1}) + (1-\theta)\, a^k\right\} \Delta_n + b^k \Delta W_n + c^k \Delta p_n, \tag{5.3.3}$$

for $k \in \{1, 2, \ldots, d\}$.

For the general multi-dimensional case with mark-dependent jump size the kth component of the *drift-implicit Euler scheme* is of the form

$$Y_{n+1}^k = Y_n^k + \left\{\theta\, a^k(t_{n+1}, Y_{n+1}) + (1-\theta)\, a^k\right\} \Delta_n + \sum_{j=1}^{m} b^{k,j} \Delta W_n^j + \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} c^k(\xi_i),$$

for $k \in \{1, 2, \ldots, d\}$ and $j \in \{1, 2, \ldots, m\}$.

Finally, in the multi-dimensional case with mark-independent jump size the kth component of the *drift-implicit Euler scheme* is given by

$$Y_{n+1}^k = Y_n^k + \left\{\theta\, a^k(t_{n+1}, Y_{n+1}) + (1-\theta)\, a^k\right\} \Delta_n + \sum_{j=1}^{m} b^{k,j} \Delta W_n^j + c^k \Delta p_n,$$

for $k \in \{1, 2, \ldots, d\}$.

5.3.2 Drift-Implicit Order 1.0 Scheme

In a similar way as for the Euler scheme, by introducing implicitness in the drift of the order 1.0 strong Taylor scheme presented in Section 4.3, we obtain the *drift-implicit order* 1.0 *strong scheme*.

In the one-dimensional case, $d = m = 1$, the *drift-implicit order* 1.0 *strong scheme* is given by

$$
\begin{aligned}
Y_{n+1} &= Y_n + \{\theta\, a(t_{n+1}, Y_{n+1}) + (1-\theta)\, a\}\, \Delta_n + b\Delta W_n \\
&\quad + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} c(v)\, p_\phi(dv, dz) + bb' \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} dW(z_1) dW(z_2) \\
&\quad + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} b\, c'(v) dW(z_1) p_\phi(dv, dz_2) \\
&\quad + \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} \int_{\mathcal{E}} \Big\{ b\big(t_n, Y_n + c(v)\big) - b \Big\} p_\phi(dv, dz_1) dW(z_2) \qquad (5.3.4) \\
&\quad + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} \int_{\mathcal{E}} \Big\{ c\big(t_n, Y_n + c(v_1), v_2\big) - c(v_2) \Big\} p_\phi(dv_1, dz_1)\, p_\phi(dv_2, dz_2),
\end{aligned}
$$

where

$$
b' := b'(t,x) = \frac{\partial\, b(t,x)}{\partial\, x} \quad \text{and} \quad c'(v) := c'(t,x,v) := \frac{\partial c(t,x,v)}{\partial x}. \qquad (5.3.5)
$$

For simplicity, we have used the convention (4.1.9). Here the parameter $\theta \in [0,1]$, characterizes again the degree of implicitness. This scheme achieves a strong order of convergence $\gamma = 1.0$. Note that the degree of implicitness, θ, can, in principle, be also chosen greater than one if this helps to stabilize numerically the scheme. This scheme generalizes a first order drift-implicit scheme for pure diffusions presented in Talay (1982a) and Milstein (1988).

One can simplify the double stochastic integrals appearing in the scheme (5.3.4) as shown for the order 1.0 strong Taylor scheme (4.3.3). This makes the resulting scheme more applicable in scenario simulation.

The jump commutativity condition (4.4.2), presented in Section 4.4, also applies to drift-implicit schemes. Therefore, for the class of SDEs identified by the jump commutativity condition (4.4.2) the computational efficiency of drift-implicit schemes of order $\gamma = 1.0$ is independent of the intensity level of the Poisson measure. For instance, for the SDE (2.1.5) with $c(t,x,v) = x\,\beta$ and $\beta \geq -1$, it is possible to derive a drift-implicit order 1.0 strong scheme, given by

$$
\begin{aligned}
Y_{n+1} &= \frac{Y_n}{1 - \mu\theta\Delta_n} \Big\{ 1 + (1-\theta)\, \mu\Delta_n + \sigma\Delta W_n + \beta\,\Delta p_n + \frac{1}{2}\sigma^2 \left\{ (\Delta W_n)^2 - \Delta_n \right\} \\
&\quad + \sigma\,\beta\,\Delta p_n \Delta W_n + \frac{1}{2}\beta^2 \{ (\Delta p_n)^2 - \Delta p_n \} \Big\}, \qquad (5.3.6)
\end{aligned}
$$

which is efficient also in the case of a high intensity jump measure. Note that θ should be chosen different from $(\mu\Delta_n)^{-1}$, which automatically arises for sufficiently small step sizes Δ_n.

In the multi-dimensional case with scalar Wiener process, $m = 1$, and mark-dependent jump size, the kth component of the *drift-implicit order* 1.0 *strong scheme* is given by

$$\begin{aligned} Y_{n+1}^k &= Y_n^k + \left\{\theta\, a^k(t_{n+1}, Y_{n+1}) + (1-\theta)\, a^k\right\} \Delta_n + b^k \Delta W_n \\ &+ \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} c^k(v)\, p_\phi(dv, dz) + \sum_{l=1}^{d} \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} b^l \frac{\partial b^k}{\partial x^l} dW(z_1) dW(z_2) \\ &+ \sum_{l=1}^{d} \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} b^l \frac{\partial c^k(v)}{\partial x^l}\, dW(z_1) p_\phi(dv, dz_2) \\ &+ \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} \int_{\mathcal{E}} \left\{ b^k\big(t_n, Y_n + c(v)\big) - b^k \right\} p_\phi(dv, dz_1) dW(z_2) \\ &+ \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} \int_{\mathcal{E}} \left\{ c^k\big(t_n, Y_n + c(v_1), v_2\big) - c^k\big(v_2\big) \right\} \\ &\quad \times p_\phi(dv_1, dz_1) p_\phi(dv_2, dz_2), \end{aligned} \tag{5.3.7}$$

for $k \in \{1, 2, \ldots, d\}$.

For the special class of multi-dimensional SDEs with scalar Wiener process and mark-independent jump size, satisfying the jump commutativity condition (5.2.11), the kth component of the *drift-implicit order* 1.0 *strong scheme*, given by

$$\begin{aligned} Y_{n+1}^k &= Y_n^k + \left\{\theta\, a^k(t_{n+1}, Y_{n+1}) + (1-\theta)\, a^k\right\} \Delta_n + b^k \Delta W_n \\ &+ c^k \Delta p_n + \frac{1}{2} \sum_{l=1}^{d} b^l \frac{\partial b^k}{\partial x^l} \{(\Delta W_n)^2 - \Delta_n\} \qquad (5.3.8) \\ &+ \{b^k(t_n, Y_n + c) - b^k\} \Delta p_n\, \Delta W_n + \frac{\{c^k(t_n, Y_n + c) - c^k\}}{2} \{(\Delta p_n)^2 - \Delta p_n\}, \end{aligned}$$

shows a computational efficiency independent of the jump intensity level.

In the general multi-dimensional case the kth component of the *drift-implicit order*

1.0 *strong scheme* is given by

$$\begin{aligned}
Y^k_{n+1} &= Y^k_n + \left\{\theta\, a^k(t_{n+1}, Y_{n+1}) + (1-\theta)\, a^k\right\} \Delta_n + \sum_{j=1}^{m} b^{k,j} \Delta W^j_n \\
&\quad + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} c^k(v)\, p_\phi(dv, dz) \\
&\quad + \sum_{j_1,j_2=1}^{m} \sum_{i=1}^{d} \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} \{b^{i,j_1} \frac{\partial b^{k,j_2}}{\partial x^i}\} dW^{j_1}(z_1) dW^{j_2}(z_2) \\
&\quad + \sum_{j_1=1}^{m} \sum_{i=1}^{d} \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} b^{i,j_1} \frac{\partial c^k(v)}{\partial x^i}\, dW^{j_1}(z_1) p_\phi(dv, dz_2) \\
&\quad + \sum_{j_1=1}^{m} \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} \left\{ b^{k,j_1}\big(t_n, Y_n + c(v)\big) - b^{k,j_1} \right\} \\
&\quad \times p_\phi(dv, dz_2)\, dW^{j_1}(z_2) \\
&\quad + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} \int_{\mathcal{E}} \left\{ c^k\big(t_n, Y_n + c(v_1), v_2\big) - c^k(v_2) \right\} \\
&\quad \times p_\phi(dv_1, dz_1)\, p_\phi(dv_2, dz_2), \qquad (5.3.9)
\end{aligned}$$

for $k \in \{1, 2, \ldots, d\}$.

In the implementation of drift-implicit schemes for multi-dimensional SDEs it is important to exploit the specific structure of the SDE under consideration. For instance, for multi-dimensional SDEs satisfying the diffusion commutativity condition (5.2.18) as well as the jump commutativity condition (5.2.16) and with mark-independent jump size, we obtain an efficiently implementable *drift-implicit order* 1.0 *strong scheme*, whose kth component is given by

$$\begin{aligned}
Y^k_{n+1} &= Y^k_n + \left\{\theta\, a^k(t_{n+1}, Y_{n+1}) + (1-\theta)\, a^k\right\} \Delta_n + \sum_{j=1}^{m} b^{k,j} \Delta W^j_n \\
&\quad + c^k \Delta p_n + \frac{1}{2} \sum_{j_1,j_2=1}^{m} \sum_{i=1}^{d} b^{i,j_1} \frac{\partial b^{k,j_2}}{\partial x^i} \left\{ \Delta W^{j_1}_n\, \Delta W^{j_2}_n - \Delta_n \right\} \\
&\quad + \sum_{j_1=1}^{m} \{b^{k,j_1}(t_n, Y_n + c) - b^{k,j_1}\} \Big(\Delta p_n\, \Delta W^{j_1}_n \Big) \\
&\quad + \frac{1}{2} \{c^k(t_n, Y_n + c) - c^k\} \Big((\Delta p_n)^2 - \Delta p_n \Big), \qquad (5.3.10)
\end{aligned}$$

for $k \in \{1, 2, \ldots, d\}$. The special case of additive diffusion and jump coefficient, which means $b(t,x) = b(t)$ and $c(t,x) = c(t)$, satisfies all the required commutativity conditions and, therefore, leads to an efficient drift-implicit order 1.0 strong scheme.

5.4 Predictor-Corrector Schemes

The following new strong predictor-corrector schemes are designed to retain the numerical stability properties of similar implicit schemes, while avoiding the additional computational effort required for solving an algebraic equation in each time step. This is achieved with the following procedure implemented at each time step: At first an explicit scheme is generated, the so-called predictor, and afterwards a de facto implicit scheme is used as corrector. The corrector is made explicit by using a predicted value $\bar{Y}_{n+1}$, instead of Y_{n+1}. Additionally, with this procedure one avoids the problem of the reciprocal of Gaussian random variables and can introduce "implicitness" also in the diffusion coefficient, as will be shown below.

Another advantage of predictor-corrector methods is that the difference $Z_{n+1} := \bar{Y}_{n+1} - Y_{n+1}$ between the predicted and the corrected value provides an indication of the local error. This can be used to implement more advanced schemes with step size control based on Z_{n+1}.

5.4.1 Predictor-Corrector Euler Scheme

In the one-dimensional case, $d = m = 1$, the *family of predictor-corrector Euler schemes*, is given by the corrector

$$\begin{aligned} Y_{n+1} &= Y_n + \left\{\theta \bar{a}_\eta(t_{n+1}, \bar{Y}_{n+1}) + (1-\theta)\bar{a}_\eta\right\} \Delta_n \\ &\quad + \left\{\eta b(t_{n+1}, \bar{Y}_{n+1}) + (1-\eta) b\right\} \Delta W_n + \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} c(\xi_i), \end{aligned} \tag{5.4.1}$$

where $\bar{a}_\eta = a - \eta b\, b'$, and the predictor

$$\bar{Y}_{n+1} = Y_n + a\Delta_n + b\Delta W_n + \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} c(\xi_i). \tag{5.4.2}$$

Here the parameters $\theta, \eta \in [0,1]$ characterize the degree of implicitness in the drift and in the diffusion coefficients, respectively. We remark that with the choice of $\eta > 0$ one obtains a scheme with some degree of implicitness also in the diffusion coefficient. This was not achievable with the drift-implicit schemes presented in Section 5.3. The scheme (5.4.1)-(5.4.2) has a strong order of convergence $\gamma = 0.5$, as will be shown in Section 5.5.3.

For the general multi-dimensional case, the kth component of the *family of predictor-corrector Euler schemes*, is given by the corrector

$$\begin{aligned} Y^k_{n+1} &= Y^k_n + \left\{\theta \bar{a}^k_\eta(t_{n+1}, \bar{Y}_{n+1}) + (1-\theta)\bar{a}^k_\eta\right\}\Delta_n \\ &\quad + \sum_{j=1}^{m}\left\{\eta b^{k,j}(t_{n+1}, \bar{Y}_{n+1}) + (1-\eta) b^{k,j}\right\}\Delta W^j_n + \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} c^k(\xi_i), \end{aligned} \tag{5.4.3}$$

for $\theta, \eta \in [0,1]$, where

$$\bar{a}_\eta = a - \eta \sum_{j_1,j_2=1}^{m}\sum_{i=1}^{d} b^{k,j_1}\frac{\partial b^{k,j_2}}{\partial x^i}, \tag{5.4.4}$$

and the predictor

$$\bar{Y}^k_{n+1} = Y^k_n + a^k\Delta_n + \sum_{j=1}^{m} b^{k,j}\Delta W^j_n + \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} c^k(\xi_i). \tag{5.4.5}$$

5.4.2 Predictor-Corrector Order 1.0 Scheme

As explained above, it is challenging to design an efficient higher order strong scheme with good numerical stability properties. To enhance the numerical stability properties of the order 1.0 strong Taylor scheme (4.3.1) we have presented the drift-implicit order 1.0 strong Taylor scheme (5.3.4). However, this scheme is computationally expensive, since it generally requires the solution of an algebraic equation at each time step. In the following we propose the *predictor-corrector order* 1.0 *strong scheme* which combines good numerical stability properties and efficiency.

In the one-dimensional case, $d = m = 1$, the *predictor-corrector order* 1.0 *strong* *scheme* is given by the corrector

$$
\begin{aligned}
Y_{n+1} &= Y_n + \big\{\theta\, a(t_{n+1}, \bar{Y}_{n+1}) + (1-\theta)\, a\big\}\, \Delta_n + b\Delta W_n \\
&+ \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} c(v)\, p_\phi(dv, dz) + bb' \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} dW(z_1) dW(z_2) \\
&+ \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} b\, c'(v) dW(z_1) p_\phi(dv, dz_2) \\
&+ \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} \int_{\mathcal{E}} \Big\{ b\big(t_n, Y_n + c(v)\big) - b \Big\} p_\phi(dv, dz_1) dW(z_2) \qquad (5.4.6) \\
&+ \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} \int_{\mathcal{E}} \Big\{ c\big(t_n, Y_n + c(v_1), v_2\big) - c(v_2) \Big\} p_\phi(dv_1, dz_1)\, p_\phi(dv_2, dz_2),
\end{aligned}
$$

and the predictor

$$
\begin{aligned}
\bar{Y}_{n+1} &= Y_n + a\Delta_n + b\Delta W_n + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} c(v)\, p_\phi(dv, dz) \\
&+ bb' \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} dW(z_1) dW(z_2) \\
&+ \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} b\, c'(v) dW(z_1) p_\phi(dv, dz_2) \\
&+ \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} \int_{\mathcal{E}} \Big\{ b\big(t_n, Y_n + c(v)\big) - b \Big\} p_\phi(dv, dz_1) dW(z_2) \qquad (5.4.7) \\
&+ \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} \int_{\mathcal{E}} \Big\{ c\big(t_n, Y_n + c(v_1), v_2\big) - c(v_2) \Big\} p_\phi(dv_1, dz_1) p_\phi(dv_2, dz_2).
\end{aligned}
$$

Here the parameter $\theta \in [0, 1]$ characterizes the degree of implicitness in the drift coefficient. This scheme attains strong order $\gamma = 1.0$, as will be shown in Section 5.5.3. For the generation of the multiple stochastic integrals involved, note that these can be simplified as shown for the order 1.0 strong Taylor scheme (4.3.3).

For the general multi-dimensional case, the kth component of the *predictor-corrector* *order* 1.0 *strong scheme* is given by the corrector

$$
\begin{aligned}
Y_{n+1}^k &= Y_n^k + \left\{\theta\, a^k(t_{n+1}, \bar{Y}_{n+1}) + (1-\theta)\, a^k\right\} \Delta_n + \sum_{j=1}^m b^{k,j} \Delta W_n^j \\
&+ \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} c^k(v)\, p_\phi(dv, dz) \\
&+ \sum_{j_1,j_2=1}^m \sum_{i=1}^d \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} \{b^{i,j_1} \frac{\partial b^{k,j_2}}{\partial x^i}\} dW^{j_1}(z_1) dW^{j_2}(z_2) \\
&+ \sum_{j_1=1}^m \sum_{i=1}^d \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} b^{i,j_1} \frac{\partial c^k(v)}{\partial x^i}\, dW^{j_1}(z_1) p_\phi(dv, dz_2) \\
&+ \sum_{j_1=1}^m \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} \left\{b^{k,j_1}\left(t_n, Y_n + c(v)\right) - b^{k,j_1}\right\} \\
&\times p_\phi(dv, dz_2)\, dW^{j_1}(z_2) \\
&+ \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} \int_{\mathcal{E}} \left\{c^k\left(t_n, Y_n + c(v_1), v_2\right) - c^k\left(v_2\right)\right\} \\
&\times p_\phi(dv_1, dz_1)\, p_\phi(dv_2, dz_2),
\end{aligned} \tag{5.4.8}
$$

and the predictor

$$
\begin{aligned}
\bar{Y}_{n+1}^k &= Y_n^k + a^k \Delta_n + \sum_{j=1}^m b^{k,j} \Delta W_n^j + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} c^k(v)\, p_\phi(dv, dz) \\
&+ \sum_{j_1,j_2=1}^m \sum_{i=1}^d b^{i,j_1} \frac{\partial b^{k,j_2}}{\partial x^i} \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} dW^{j_1}(z_1) dW^{j_2}(z_2) \\
&+ \sum_{j_1=1}^m \sum_{i=1}^d \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} b^{i,j_1} \frac{\partial c^k(v)}{\partial x^i}\, dW^{j_1}(z_1) p_\phi(dv, dz_2)
\end{aligned}
$$

$$
\begin{aligned}
&+ \sum_{j_1=1}^m \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} \left\{b^{k,j_1}\left(t_n, Y_n + c(v)\right) - b^{k,j_1}\right\} p_\phi(dv, dz_2)\, dW^{j_1}(z_2) \\
&+ \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} \int_{\mathcal{E}} \left\{c^k\left(t_n, Y_n + c(v_1), v_2\right) - c^k\left(v_2\right)\right\} \\
&\times p_\phi(dv_1, dz_1)\, p_\phi(dv_2, dz_2),
\end{aligned} \tag{5.4.9}
$$

with $\theta \in [0,1]$, for $k \in \{1,2,\ldots,d\}$.

The considerations mentioned in the discussion on the generation of multiple stochastic integrals for the order 1.0 strong Taylor scheme (4.3.8) apply also here. In particular, if the SDE under analysis satisfies the diffusion commutativity condition (5.2.18) together with the jump commutativity condition (5.2.16) and has mark-independent jump size, then we obtain an efficiently implementable *predictor-corrector order* 1.0 *strong scheme.* Its kth component is given by the corrector

$$\begin{aligned} Y^k_{n+1} &= Y^k_n + \left\{\theta\, a^k(t_{n+1}, \bar{Y}_{n+1}) + (1-\theta)\, a^k\right\} \Delta_n + \sum_{j=1}^{m} b^{k,j} \Delta W^j_n \\ &\quad + c^k \Delta p_n + \frac{1}{2} \sum_{j_1,j_2=1}^{m} \sum_{i=1}^{d} b^{i,j_1} \frac{\partial b^{k,j_2}}{\partial x^i} \left\{ \Delta W^{j_1}_n \, \Delta W^{j_2}_n - \Delta_n \right\} \\ &\quad + \sum_{j_1=1}^{m} \{b^{k,j_1}(t_n, Y_n + c) - b^{k,j_1}\} \left(\Delta p_n \, \Delta W^{j_1}_n \right) \\ &\quad + \frac{1}{2} \{c^k(t_n, Y_n + c) - c^k\} \left((\Delta p_n)^2 - \Delta p_n \right), \end{aligned} \tag{5.4.10}$$

and the predictor

$$\begin{aligned} \bar{Y}^k_{n+1} &= Y^k_n + a^k \Delta_n + \sum_{j=1}^{m} b^{k,j} \Delta W^j_n + c^k \Delta p_n \\ &\quad + \frac{1}{2} \sum_{j_1,j_2=1}^{m} \sum_{i=1}^{d} b^{i,j_1} \frac{\partial b^{k,j_2}}{\partial x^i} \left\{ \Delta W^{j_1}_n \, \Delta W^{j_2}_n - \Delta_n \right\} \\ &\quad + \sum_{j_1=1}^{m} \{b^{k,j_1}(t_n, Y_n + c) - b^{k,j_1}\} \left(\Delta p_n \, \Delta W^{j_1}_n \right) \\ &\quad + \frac{1}{2} \{c^k(t_n, Y_n + c) - c^k\} \left((\Delta p_n)^2 - \Delta p_n \right), \end{aligned} \tag{5.4.11}$$

with $\theta \in [0,1]$, for $k \in \{1,2,\ldots,d\}$. In the special case of additive diffusion and jump coefficient, $b(t,x) = b(t)$ and $c(t,x) = c(t)$, which satisfies all the required commutativity conditions, we obtain an efficient predictor-corrector order 1.0 strong scheme.

We remark that as in the predictor-corrector Euler scheme also in the predictor-corrector order 1.0 strong scheme one can introduce quasi-implicitness into the diffusion part.

It shall be emphasized that predictor-corrector schemes with higher strong order can be easily constructed. They provide numerically stable, efficient and conveniently implementable discrete time approximations of jump diffusions that can be recommended in many cases as simulation studies will demonstrate.

5.5 Convergence Results

In this section we consider general strong schemes, the strong Itô schemes, constructed with the same multiple stochastic integrals underlying the strong Taylor schemes (4.5.3), presented in Section 4.5, but with various approximations for the different coefficients. Under particular conditions on these coefficients, the strong Itô schemes converge to the solution X of the SDE (2.1.2) with the same strong order γ achieved by the corresponding strong Taylor schemes. In principle, we can construct more general strong approximations of any given order as those already presented in this chapter. In particular, we will show that derivative-free, drift-implicit, and predictor-corrector schemes are strong Itô schemes. Again we will discuss two different types of schemes; those based on the Wagner-Platen expansion (3.5.4) and those based on the compensated Wagner-Platen expansion (3.5.5).

For a regular time discretization $(t)_\Delta$ with maximum step size $\Delta \in (0,1)$, as the one introduced in (4.1.5), we define the *order γ strong Itô scheme* by the vector equation

$$Y_{n+1}^{\Delta} = Y_n^{\Delta} + \sum_{\alpha \in \mathcal{A}_\gamma \setminus \{v\}} I_\alpha \left[h_{\alpha,n}\right]_{t_n,t_{n+1}} + R_n\,, \tag{5.5.1}$$

and the *order γ compensated strong Itô scheme* by

$$Y_{n+1}^{\Delta} = Y_n^{\Delta} + \sum_{\alpha \in \mathcal{A}_\gamma \setminus \{v\}} \tilde{I}_\alpha \left[\tilde{h}_{\alpha,n}\right]_{t_n,t_{n+1}} + \tilde{R}_n\,, \tag{5.5.2}$$

with $n \in \{0,1,\ldots,n_T - 1\}$. We assume that the coefficients $h_{\alpha,n}$ and $\tilde{h}_{\alpha,n}$ are $\mathcal{A}_{t_n}$-measurable and satisfy the estimates

$$E\left(\max_{0\le n\le n_T-1} |h_{\alpha,n} - f_\alpha(t_n, Y_n)|^2\right) \le C(u)\,\Delta^{2\gamma-\psi(\alpha)}, \tag{5.5.3}$$

and

$$E\left(\max_{0\le n\le n_T-1}|\tilde{h}_{\alpha,n}-\tilde{f}_\alpha(t_n,Y_n)|^2\right)\le C(u)\,\Delta^{2\gamma-\psi(\alpha)}, \tag{5.5.4}$$

respectively, 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

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

Additionally, R_n and $\widetilde{R}_n$ are assumed to satisfy

$$E\left(\max_{1\le n\le n_T}\left|\sum_{0\le k\le n-1}R_k\right|^2\right)\le K\,\Delta^{2\gamma}, \tag{5.5.5}$$

and

$$E\left(\max_{1\le n\le n_T}\left|\sum_{0\le k\le n-1}\widetilde{R}_k\right|^2\right)\le K\,\Delta^{2\gamma} \tag{5.5.6}$$

where K is a finite positive constant independent of Δ.

Now we formulate a convergence theorem that will enable us to construct strong Itô schemes of any given strong order, including derivative-free, drift-implicit and predictor-corrector schemes.

Theorem 5.5.1 *Let $Y^\Delta=\{Y_n^\Delta, n\in\{0,1,\dots,n_T\}\}$ be a discrete time approximation generated via the order γ compensated strong Itô scheme* (5.5.2)*, for a given regular time discretization $(t)_\Delta$ with maximum time step size $\Delta\in(0,1)$, and for $\gamma\in\{0.5,1,1.5,2,\dots\}$. If the conditions of Theorem 4.5.1 are satisfied, then*

$$\sqrt{E\left(\max_{0\le n\le n_T}|X_{t_n}-Y_n^\Delta|^2\right)}\le K\,\Delta^\gamma, \tag{5.5.7}$$

where K is a finite positive constant independent of Δ.

Proof: Since we have already shown in Theorem 4.5.1 that the compensated strong Taylor scheme (4.5.3) converges with strong order γ, here it will be sufficient to show that the compensated Itô scheme (5.5.2) converges with strong order γ to the corresponding compensated Taylor scheme.

By $\overline{Y}^{\Delta}$ we denote here the compensated strong Taylor scheme (4.5.3). Let us also assume, for simplicity, that $\overline{Y}_0 = Y_0$. Then by application of Jensen's inequality together with the Cauchy-Schwarz inequality, we obtain for all $t \in [0, T]$ the estimate

$$
\begin{aligned}
H_t &:= E\left(\max_{1\le n\le n_t}\left|\overline{Y}_n^{\Delta} - Y_n^{\Delta}\right|^2\right) \\
&= E\left(\max_{1\le n\le n_t}\left|\sum_{k=0}^{n-1}\sum_{\alpha\in\mathcal{A}_\gamma\setminus\{v\}} \widetilde{I}_\alpha\left[\tilde{f}_\alpha(t_k,\overline{Y}_k^{\Delta})\right]_{t_k,t_{k+1}}\right.\right. \\
&\qquad \left.\left. - \sum_{k=0}^{n-1}\left(\sum_{\alpha\in\mathcal{A}_\gamma\setminus\{v\}} \widetilde{I}_\alpha\left[\tilde{h}_{\alpha,k}\right]_{t_k,t_{k+1}} + R_n\right)\right|^2\right) \\
&\le K_1 \sum_{\alpha\in\mathcal{A}_\gamma\setminus\{v\}}\left\{E\left(\max_{1\le n\le n_t}\left|\sum_{k=0}^{n-1}\widetilde{I}_\alpha\left[\tilde{f}_\alpha(t_k,\overline{Y}_k^{\Delta}) - \tilde{f}_\alpha(t_k,Y_k^{\Delta})\right]_{t_k,t_{k+1}}\right|^2\right)\right. \\
&\qquad \left. + E\left(\max_{1\le n\le n_t}\left|\sum_{k=0}^{n-1}\widetilde{I}_\alpha\left[\tilde{f}_\alpha(t_k,Y_k^{\Delta}) - \tilde{h}_{\alpha,k}\right]_{t_k,t_{k+1}}\right|^2\right)\right\} \\
&\qquad + K_1\, E\left(\max_{1\le n\le n_t}\left|\sum_{k=0}^{n-1} R_n\right|^2\right). \qquad (5.5.8)
\end{aligned}
$$

By applying Lemma 4.6.1, condition (5.5.6), the Lipschitz condition (4.5.7), as well as condition (5.5.4), we obtain

$$
\begin{aligned}
H_t &\le K_2 \sum_{\alpha\in\mathcal{A}_\gamma\setminus\{v\}}\left\{\int_0^t\int_{\mathcal{E}}\cdots\int_{\mathcal{E}}\Big(E\big(\max_{1\le n\le n_u}|\tilde{f}_\alpha(t_k,\overline{Y}_k^{\Delta}) - \tilde{f}_\alpha(t_k,Y_k^{\Delta})|^2\big)\right. \\
&\qquad \left. + E\big(\max_{1\le n\le n_u}|\tilde{f}_\alpha(t_k,Y_k^{\Delta}) - \tilde{h}_{\alpha,n}|^2\big)\Big)\phi(dv^1)\dots\phi(dv^{s(\alpha)})\,du\right\}\Delta^{\psi(\alpha)} + K_3\Delta^{2\gamma} \\
&\le K_5\int_0^t E\big(\max_{1\le n\le n_u}|\overline{Y}_k^{\Delta} - Y_k^{\Delta}|^2\big)du \times \sum_{\alpha\in\mathcal{A}_\gamma\setminus\{v\}}\Delta^{\psi(\alpha)} + K_6\,\Delta^{2\gamma} \\
&\le K_7\int_0^t H_u\,du + K_6\,\Delta^{2\gamma}, \qquad (5.5.9)
\end{aligned}
$$

From the second moment estimate (4.7.1) on the compensated strong Taylor scheme

$\overline{Y}^{\Delta}$ in Lemma 4.7.1 and a similar estimate on the compensated strong Itô scheme Y^{Δ}, one can show that H_t is bounded. Therefore, by applying the Gronwall inequality to (5.5.9), we obtain

$$H_t \leq K_5\, \Delta^{2\gamma}\, e^{K_7 t}. \tag{5.5.10}$$

Since we have assumed $\overline{Y}_0^{\Delta} = Y_0^{\Delta}$, we get

$$E\big(\max_{0 \leq n \leq n_T} |\overline{Y}_n^{\Delta} - Y_n^{\Delta}|^2\big) \leq K\, \Delta^{2\gamma}. \tag{5.5.11}$$

Finally, by the estimate of Theorem 4.5.1 we obtain

$$\sqrt{E\big(\max_{0 \leq n \leq n_T} |X_{t_n} - Y_n^{\Delta}|^2\big)} = \sqrt{E\big(\max_{0 \leq n \leq n_T} |X_{t_n} - \overline{Y}_n^{\Delta} + \overline{Y}_n^{\Delta} - Y_n^{\Delta}|^2\big)} \leq K\, \Delta^{\gamma}, \tag{5.5.12}$$

which finalizes the proof of Theorem 5.5.1. □

By the same arguments used above, one can show the following result.

Corollary 5.5.2 *Let $Y^{\Delta} = \{Y_n^{\Delta}, n \in \{0, 1, \ldots, n_T\}\}$ be a discrete time approximation generated by the order γ strong Itô scheme* (5.5.1). *If the conditions of Corollary 4.5.3 are satisfied, then*

$$\sqrt{E\left(\max_{0 \leq n \leq n_T} |X_{t_n} - Y_n^{\Delta}|^2\right)} \leq K\, \Delta^{\gamma}, \tag{5.5.13}$$

where K is a finite positive constant independent of Δ.

5.5.1 Derivative-Free Schemes

The strong Itô schemes (5.5.2) and (5.5.1), and the convergence Theorem 5.5.1 and the Corollary 5.5.2 allow us to asses the strong order of convergence of general approximations. In this section we show how to rewrite derivative-free schemes, including those presented in Section 5.2, as strong Itô schemes.

We recall here that in the one-dimensional case the derivative-free order 1.0 strong scheme, presented in Section 5.2, is given as

$$
\begin{aligned}
Y_{n+1} \quad = \quad & Y_n + a\Delta_n + b\Delta W_n + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} c(v)\, p_\phi(dv, dz) \\
& + \frac{\left(b(t_n, \overline{Y}_n) - b\right)}{\sqrt{\Delta_n}} \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} dW(z_1)\, dW(z_2) \\
& + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} \frac{\left(c(t_n, \overline{Y}_n, v) - c(v)\right)}{\sqrt{\Delta_n}} dW(z_1)\, p_\phi(dv, dz_2) \\
& + \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} \int_{\mathcal{E}} \left\{ b(t_n, Y_n + c(v)) - b \right\} p_\phi(dv, dz_1)\, dW(z_2) \qquad (5.5.14) \\
& + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} \int_{\mathcal{E}} \left\{ c(t_n, Y_n + c(v_2), v_1) - c(v_1) \right\} p_\phi(dv_1, dz_1) p_\phi(dv_2, dz_2),
\end{aligned}
$$

with the supporting value

$$\overline{Y}_n = Y_n + b\sqrt{\Delta_n}. \qquad (5.5.15)$$

From the deterministic Taylor expansion, we obtain

$$
\begin{aligned}
b(t_n, \overline{Y}_n) \quad = \quad & b(t_n, Y_n) + b'(t_n, Y_n) \left\{ \overline{Y}_n - Y_n \right\} \\
& + \frac{b''\left(t_n, Y_n + \hat{\theta}(\overline{Y}_n - Y_n)\right)}{2} \left\{ \overline{Y}_n - Y_n \right\}^2, \qquad (5.5.16)
\end{aligned}
$$

with

$$b'(t, x) := \frac{\partial\, b(t, x)}{\partial\, x} \quad \text{and} \quad b''(t, x) := \frac{\partial^2 b(t, x)}{\partial x^2} \qquad (5.5.17)$$

and

$$
\begin{aligned}
c(t_n, \overline{Y}_n, v) \quad = \quad & c(t_n, Y_n, v) + c'(t_n, Y_n, v) \left\{ \overline{Y}_n - Y_n \right\} \\
& + \frac{c''\left(t_n, Y_n + \hat{\theta}(\overline{Y}_n - Y_n), v\right)}{2} \left\{ \overline{Y}_n - Y_n \right\}^2, \qquad (5.5.18)
\end{aligned}
$$

with

$$c'(t, x, v) := \frac{\partial c(t, x, v)}{\partial x} \quad \text{and} \quad c''(t, x, v) := \frac{\partial^2 c(t, x, v)}{\partial x^2} \qquad (5.5.19)$$

for every $v \in \mathcal{E}$ and some corresponding $\hat{\theta} \in (0, 1)$.

Therefore, we can rewrite the scheme (5.5.14) as

$$
\begin{aligned}
Y_{n+1} \quad = \quad & Y_n + I_{(0)}[h_{(0),n}]_{t_n,t_{n+1}} + I_{(1)}[h_{(1),n}]_{t_n,t_{n+1}} + I_{(-1)}[h_{(-1),n}]_{t_n,t_{n+1}} \\
& + I_{(1,1)}[h_{(1,1),n}]_{t_n,t_{n+1}} + I_{(1,-1)}[h_{(1,-1),n}]_{t_n,t_{n+1}} \\
& + I_{(-1,1),n}[h_{(-1,1),n}]_{t_n,t_{n+1}} + I_{(-1,-1)}[h_{(-1,-1),n}]_{t_n,t_{n+1}}, \qquad (5.5.20)
\end{aligned}
$$

with

$$
\begin{aligned}
& h_{(0),n} = a(t_n, Y_n), \qquad h_{(1),n} = b(t_n, Y_n), \qquad h_{(-1),n} = c(t_n, Y_n, v), \\
& h_{(1,-1),n} = \frac{1}{\sqrt{\Delta_n}} \Big\{ c(t_n, \overline{Y}_n, v) - c(t_n, Y_n, v) \Big\}, \\
& h_{(-1,1),n} = b\Big(t_n, Y_n + c(t_n, Y_n, v)\Big) - b\Big(t_n, Y_n\Big), \\
& h_{(-1,-1),n} = c\Big(t_n, Y_n + c(t_n, Y_n, v_2), v_1\Big) - c\Big(t_n, Y_n, v_1\Big), \\
& h_{(1,1),n} = \frac{1}{\sqrt{\Delta_n}} \Big\{ b(t_n, \overline{Y}_n) - b(t_n, Y_n) \Big\}. \qquad (5.5.21)
\end{aligned}
$$

Only for $\alpha = (1,1)$ and $\alpha = (1,-1)$ the coefficients $h_{\alpha,n}$ are different from the coefficients $f_{\alpha,n}$ of the order 1.0 strong Taylor scheme (4.3.1). Therefore, to prove that the scheme (5.5.14)-(5.5.15) is an order 1.0 strong Itô scheme, it remains to check condition (5.5.3) for these two coefficients.

By the linear growth condition (4.5.9) of Theorem 4.5.1, we have

$$
\begin{aligned}
\left| b(t_n, Y_n)^2 b'' \left(t_n, Y_n + \hat{\theta}\, b(t_n, Y_n) \sqrt{\Delta_n} \right) \right|^2 \quad & \leq \quad K_1 (1 + |Y_n|^4)\, K_2 \left(1 + |Y_n|^2\right) \\
& = \quad C_1 \left(1 + |Y_n|^2 + |Y_n|^4 + |Y_n|^6\right).
\end{aligned}
\qquad (5.5.22)
$$

In a similar way we also obtain

$$
\left| b(t_n, Y_n)^2 c'' \left(t_n, Y_n + \hat{\theta}\, b(t_n, Y_n) \sqrt{\Delta_n}, v \right) \right|^2 \leq C_2(v) \left(1 + |Y_n|^2 + |Y_n|^4 + |Y_n|^6\right), \qquad (5.5.23)
$$

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

Following similar steps as the ones used in the first part of the proof of Theorem

4.5.1, one can show that

$$E\left(\max_{0\le n\le n_T-1} |Y_n|^{2q}\right) \le K\Big(1+E(|Y_0|^{2q})\Big), \tag{5.5.24}$$

for $q \in \mathbb{N}$. Therefore, assuming $E(|Y_0|^6) < \infty$, by conditions (5.5.22), (5.5.23) and (5.5.24), we obtain

$$\begin{aligned} &E\left(\max_{0\le n\le n_T-1} \left|h_{(1,1),n} - f_{(1,1)}(t_n, Y_n)\right|^2\right) \\ &\le E\left(\max_{0\le n\le n_T-1} \left|\frac{b(t_n,Y_n)^2 b''(t_n,Y_n)}{2}\sqrt{\Delta_n}\right|^2\right) \\ &\le K\Delta\Big(1+E(|Y_0|^6)\Big) \\ &\le K\,\Delta^{2\gamma-\psi(\alpha)}. \end{aligned} \tag{5.5.25}$$

We also have

$$E\left(\max_{0\le n\le n_T-1} \left|h_{(1,-1),n} - f_{(1,-1)}(t_n, Y_n)\right|^2\right) \le C(v)\,\Delta^{2\gamma-\psi(\alpha)}, \tag{5.5.26}$$

where $C(v) : \mathcal{E} \to \mathbb{R}$ is a $\phi(dv)$- integrable function, which shows that the scheme (5.5.14) is a strong Itô scheme of order $\gamma = 1.0$.

In a similar way one can show that also some other higher order derivative-free strong schemes for general multi-dimensional SDEs can be expressed as strong Itô schemes and, therefore, achieve the corresponding strong order of convergence.

5.5.2 Drift-Implicit Schemes

As explained in Section 5.3, for any strong Taylor scheme of order γ it is usually possible to obtain a corresponding drift-implicit scheme. Due to problems with the reciprocal of Gaussian random variables in a scheme, one usually introduces implicitness only in the drift terms, see Higham & Kloeden (2005, 2006). An exception are balanced implicit methods where the implicitness is brought also into the diffusion terms. We refer to Milstein, Platen & Schurz (1998), Kahl & Schurz (2006) and Alcock & Burrage (2006) for balanced implicit schemes for pure diffusion SDEs. Drift-implicit schemes of order γ can be derived by an application of a Wagner-Platen expansion to the drift terms of a correspondent strong Taylor scheme of order γ. If we apply the Wagner-Platen expansion (3.5.4) to the drift term a, then we can write

$$\begin{aligned} a(t, X_t) &= a(t+\Delta, X_{t+\Delta}) - L^{(0)}a(t, X_t)\Delta - L^{(1)}a(t, X_t)\left(W(t+\Delta) - W(t)\right) \\ &\quad - \int_t^{t+\Delta}\int_{\mathcal{E}} L_v^{(-1)}a(t, X_t)p_\phi(dv, dz) - R_1(t), \end{aligned} \tag{5.5.27}$$

where

$$\begin{aligned} R_1(t) &= \int_t^{t+\Delta}\left\{\int_t^s L^{(0)}L^{(0)}a(u, X_u)du + \int_t^s L^{(1)}L^{(0)}a(u, X_u)dW_u \right. \\ &\quad \left. + \int_t^s\int_{\mathcal{E}} L_{v_1}^{(-1)}L^{(0)}a(u, X_{u-})p_\phi(dv_1, du)\right\} ds \\ &\quad + \int_t^{t+\Delta}\left\{\int_t^s L^{(0)}L^{(1)}a(u, X_u)du + \int_t^s L^{(1)}L^{(1)}a(u, X_u)dW_u \right. \\ &\quad \left. + \int_t^s\int_{\mathcal{E}} L_{v_1}^{(-1)}L^{(1)}a(X_{u-})p_\phi(dv_1, du)\right\} dW_s \\ &\quad + \int_t^{t+\Delta}\int_{\mathcal{E}}\left\{\int_t^{s-} L^{(0)}L_{v_2}^{(-1)}a(u, X_u)du + \int_t^{s-} L^{(1)}L_{v_2}^{(-1)}a(u, X_u)dW_u \right. \\ &\quad \left. + \int_t^{s-}\int_{\mathcal{E}} L_{v_1}^{(-1)}L_{v_2}^{(-1)}a(u, X_{u-})p_\phi(dv_1, du)\right\} p_\phi(dv_2, ds) \end{aligned} \tag{5.5.28}$$

and the operators $L^{(0)}$, $L^{(1)}$ and $L_v^{(-1)}$ are defined in (3.3.4), (3.3.5) and (3.3.6), respectively.

In the one-dimensional case, for any $\theta \in [0,1]$, we can rewrite the Euler scheme (4.2.1) as

$$\begin{aligned} Y_{n+1} &= Y_n + \{\theta\, a(t_n, Y_n) + (1-\theta)\, a(t_n, Y_n)\}\,\Delta_n + b\Delta W_n \\ &\quad + \int_{t_n}^{t_{n+1}}\int_{\mathcal{E}} c(t_n, Y_n, v)\, p_\phi(dv, ds). \end{aligned} \tag{5.5.29}$$

By replacing the first drift coefficient $a(t_n, Y_n)$ with its implicit expansion (5.5.27), we obtain

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

Here we have written $a = a(t_n, Y_n)$ and $b = b(t_n, Y_n)$ according to the abbreviation introduced in (4.1.9).

The terms in the last line of equation (5.5.30) are not necessary for a scheme with strong order $\gamma = 0.5$. Therefore, they can be discarded when deriving the implicit Euler scheme (5.3.1), which yields

$$
\begin{aligned}
Y_{n+1} &= Y_n + \Big\{\theta\, a(t_{n+1}, Y_{n+1}) + (1-\theta)\, a\Big\}\Delta_n + b\Delta W_n \\
&\quad + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} c(t_n, Y_n, v) p_\phi(dv, ds). \qquad (5.5.30)
\end{aligned}
$$

Note that the drift-implicit Euler scheme (5.5.30) is well defined for time step size

$$
\Delta \leq \frac{1}{\sqrt{K\theta}},
$$

where K is the Lipschitz constant appearing in the Lipschitz condition (2.2.10) for the drift coefficient a. As previously mentioned, Banach's fixed point theorem ensures existence and uniqueness of the solution Y_{n+1} of the algebraic equation in (5.5.30).

By applying the same arguments to every time integral appearing in a higher order strong Taylor scheme, even in the general multi-dimensional case, it is possible to derive multi-dimensional higher order implicit schemes as, for instance, the drift-implicit order 1.0 strong scheme (5.3.9).

To prove the strong order of convergence of such drift-implicit schemes it is sufficient to show that one can rewrite these as strong Itô schemes. The drift-implicit Euler scheme (5.5.30), for instance, can be written as an order 0.5 strong Itô scheme given by

$$Y_{n+1} = Y_n + I_{(0)}[h_{(0),n}]_{t_n,t_{n+1}} + I_{(1)}[h_{(1),n}]_{t_n,t_{n+1}} + I_{(-1)}[h_{(-1),n}]_{t_n,t_{n+1}} + R_n \quad (5.5.31)$$

with

$$h_{(0),n} = a(t_n, Y_n), \qquad h_{(1),n} = b(t_n, Y_n), \qquad h_{(-1),n} = c(t_n, Y_n, v), \qquad (5.5.32)$$

and

$$R_n = \theta\, \Delta_n \, (a(t_{n+1}, Y_{n+1}) - a(t_n, Y_n)). \qquad (5.5.33)$$

Since the coefficients $h_{\alpha,n}$ are the same as those of the Euler scheme (4.2.1), it remains to check condition (5.5.6) for the remainder term R_n. Following similar steps as the ones used in the proof of Lemma 4.7.1, one can derive the second moment estimate

$$E\left(\max_{0\le n\le n_T-1} |Y_n|^2\right) \le K\Big(1 + E(|Y_0|^2)\Big). \qquad (5.5.34)$$

We refer to Higham & Kloeden (2006) for details of the proof of this estimate in the case of the drift-implicit Euler scheme for SDEs driven by Wiener processes and homogeneous Poisson processes.

By applying the Cauchy-Schwarz inequality, the linear growth condition (2.2.11) and the estimate (5.5.34), we obtain

$$\begin{aligned}
&E\left(\max_{1\le n\le n_T}\left|\sum_{0\le k\le n-1} R_k\right|^2\right)\\
&= E\left(\max_{1\le n\le n_T}\left|\sum_{0\le k\le n-1} \theta\,\Delta_k\,\big(a(t_{k+1}, Y_{k+1}) - a(t_k, Y_k)\big)\right|^2\right)\\
&\le E\left(\max_{1\le n\le n_T}\left(\sum_{0\le k\le n-1} |\theta\Delta_k|^2\right)\left(\sum_{0\le k\le n-1} |a(t_{k+1}, Y_{k+1}) - a(t_k, Y_k)|^2\right)\right)\\
&\le K\Delta E\left(\left(\sum_{0\le k\le n_T-1} \Delta_k\right)\left(\sum_{0\le k\le n_T-1} |a(t_{k+1}, Y_{k+1}) - a(t_k, Y_k)|^2\right)\right)
\end{aligned}$$

$$\leq K\,\Delta\,E\left(\sum_{0\leq k\leq n_T-1}|a(t_{k+1},Y_{k+1})-a(t_k,Y_k)|^2\right)$$
$$\leq K\,\Delta\,E\left(\sum_{0\leq k\leq n_T-1}|Y_{k+1}-Y_k|^2\right). \tag{5.5.35}$$

Since $n_T<\infty$ almost surely, see (4.1.8), we can write

$$E\left(\sum_{0\leq k\leq n_T-1}|Y_{k+1}-Y_k|^2\right)$$
$$=\sum_{i=1}^{\infty}E\left(\sum_{0\leq k\leq i-1}|Y_{k+1}-Y_k|^2\right)\times P(n_T=i)$$
$$=\sum_{i=1}^{\infty}E\left(\sum_{0\leq k\leq i-1}E\left(|Y_{k+1}-Y_k|^2\,\big|\mathcal{A}_{t_k}\right)\right)\times P(n_T=i), \tag{5.5.36}$$

where we have used the properties of conditional expectations.

By the Cauchy-Schwarz inequality, Itô's isometry and the linear growth conditions, one obtains

$$E\left(|Y_{k+1}-Y_k|^2\,\big|\mathcal{A}_{t_k}\right)$$
$$\leq K\Big\{E\left(\left|\int_{t_k}^{t_{k+1}}\Big(\theta\big(a(t_{k+1},Y_{k+1})-a(t_k,Y_k)\big)+\widetilde{a}(t_k,Y_k)\Big)dz\right|^2\big|\mathcal{A}_{t_k}\right)$$
$$+\sum_{\alpha\in\{(-1),(1)\}}E\left(\left|\widetilde{I}_\alpha[f_\alpha(t_k,Y_k)]_{t_k,t_{k+1}}\right|^2\big|\mathcal{A}_{t_k}\right)\Big\}$$
$$\leq K\Big\{(t_{k+1}-t_k)$$
$$\times\int_{t_k}^{t_{k+1}}E\Big(|a(t_{k+1},Y_{k+1})|^2+|a(t_k,Y_k)|^2+|\widetilde{a}(t_k,Y_k)|^2\big|\mathcal{A}_{t_k}\Big)dz$$
$$+\sum_{\alpha\in\{(-1),(1)\}}\int_{t_k}^{t_{k+1}}\int_{\mathcal{E}}E\left(\left|\tilde{f}_\alpha(t_k,Y_k)\right|^2\big|\mathcal{A}_{t_k}\right)\phi(dv^{s(\alpha)})dz\Big\}$$

$$
\begin{aligned}
&\leq K\Big\{\Delta \int_{t_k}^{t_{k+1}} E\Big(\big(1+|Y_{k+1}|^2\big)+\big(1+|Y_k|^2\big)\big|\mathcal{A}_{t_k}\Big)dz \\
&+ \sum_{\alpha\in\{(-1),(1)\}} \int_{t_k}^{t_{k+1}} \int_{\mathcal{E}} E\left(\left(1+|Y_k|^2\right)\big|\mathcal{A}_{t_k}\right) \phi(dv^{s(\alpha)})dz\Big\} \\
&\leq K\big(1+\max_{0\leq n\leq i}|Y_n|^2\big)\big(t_{k+1}-t_k\big) \qquad (5.5.37)
\end{aligned}
$$

for $k \in \{1,\ldots,i\}$ with $i \in \mathbb{N}$.

Finally, by combining (5.5.35), (5.5.36) and (5.5.37) and by using the second moment estimate (5.5.34), we obtain

$$
\begin{aligned}
&E\left(\max_{1\leq n\leq n_T}\left|\sum_{0\leq k\leq n-1} R_k\right|^2\right) \\
&\leq K\Delta\sum_{i=1}^{\infty} E\left(\left(1+\max_{0\leq n\leq i}|Y_n|^2\right)\sum_{0\leq k\leq i-1}\big(t_{k+1}-t_k\big)\right)\times P(n_T=i) \\
&\leq K\Delta\big(1+E(|Y_0|^2)\big)T\sum_{i=1}^{\infty}P(n_T=i) \\
&\leq K\Delta = K\Delta^{2\gamma}, \qquad (5.5.38)
\end{aligned}
$$

for $\gamma = 0.5$.

Therefore, the convergence of the drift-implicit Euler scheme follows from Corollary 5.5.2 since we have shown that it can be rewritten as a strong Itô scheme of order $\gamma = 0.5$. In a similar way it is possible to show that the drift-implicit order 1.0 strong scheme (5.3.9) can be rewritten as an order $\gamma = 1.0$ strong Itô scheme.

5.5.3 Predictor-Corrector Schemes

In this section we show how to rewrite predictor-corrector schemes as strong Itô schemes and how to derive the corresponding orders of strong convergence.

Let us consider the family of strong predictor-corrector Euler schemes (5.4.1)-(5.4.2). We recall that in the one-dimensional case, $d = m = 1$, this scheme is given by the corrector

$$\begin{aligned}Y_{n+1} &= Y_n + \left\{\theta\bar{a}_\eta(t_{n+1}, \bar{Y}_{n+1}) + (1-\theta)\bar{a}_\eta\right\}\Delta_n \\ &\quad + \left\{\eta b(t_{n+1}, \bar{Y}_{n+1}) + (1-\eta)b\right\}\Delta W_n + \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} c(\xi_i), \end{aligned} \quad (5.5.39)$$

where $\bar{a}_\eta = a - \eta b\, b'$, and the predictor

$$\bar{Y}_{n+1} = Y_n + a\Delta_n + b\Delta W_n + \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} c(\xi_i). \quad (5.5.40)$$

This scheme can be written as an order 0.5 strong Itô scheme given by

$$Y_{n+1} = Y_n + I_{(0)}[h_{(0),n}]_{t_n,t_{n+1}} + I_{(1)}[h_{(1),n}]_{t_n,t_{n+1}} + I_{(-1)}[h_{(-1),n}]_{t_n,t_{n+1}} + R_n \quad (5.5.41)$$

with

$$h_{(0),n} = a(t_n, Y_n), \qquad h_{(1),n} = b(t_n, Y_n), \qquad h_{(-1),n} = c(t_n, Y_n, v), \quad (5.5.42)$$

and

$$R_n = R_{1,n} + R_{2,n} + R_{3,n} + R_{4,n} \quad (5.5.43)$$

with

$$R_{1,n} = \theta\left\{a(t_{n+1}, \bar{Y}_{n+1}) - a(t_n, Y_n)\right\}\Delta_n, \quad (5.5.44)$$

$$R_{2,n} = -\theta\,\eta\left\{b(t_{n+1}, \bar{Y}_{n+1})\, b'(t_{n+1}, \bar{Y}_{n+1}) - b(t_n, Y_n)\, b'(t_n, Y_n)\right\}\Delta_n, \quad (5.5.45)$$

$$R_{3,n} = -\eta\, b(t_n, Y_n)\, b'(t_n, Y_n)\Delta_n, \quad (5.5.46)$$

and

$$R_{4,n} = -\eta\left\{b(t_{n+1}, \bar{Y}_{n+1}) - b(t_n, Y_n)\right\}\Delta W_n. \quad (5.5.47)$$

Since the coefficients $h_{\alpha,n}$ are the same as those of the Euler scheme, which is the strong Taylor scheme of order $\gamma = 0.5$, one needs to show that the remainder term R_n satisfies condition (5.5.5).

We will also need the following second moment estimate on the numerical approximation Y_n, where

$$E\left(\max_{0\le n\le n_T-1}|Y_n|^2\right)\le K\Big(1+E(|Y_0|^2)\Big). \tag{5.5.48}$$

This estimate can be proved by similar steps as those used in the proof of Lemma 4.7.1.

By the Cauchy-Schwarz inequality, the Lipschitz condition on the drift coefficient and equation (5.5.40), we obtain

$$\begin{aligned}
&E\left(\max_{1\le n\le n_T}\left|\sum_{0\le k\le n-1} R_{1,k}\right|^2\right)\\
&= E\left(\max_{1\le n\le n_T}\left|\sum_{0\le k\le n-1}\theta\left\{a(t_{k+1},\bar{Y}_{k+1})-a(t_k,Y_k)\right\}\Delta_k\right|^2\right)\\
&\le E\left(\max_{1\le n\le n_T}\left(\sum_{0\le k\le n-1}|\theta\Delta_k|^2\right)\left(\sum_{0\le k\le n-1}\left|a(t_{k+1},\bar{Y}_{k+1})-a(t_k,Y_k)\right|^2\right)\right)\\
&\le K\Delta E\left(\left(\sum_{0\le k\le n_T-1}\Delta_k\right)\left(\sum_{0\le k\le n_T-1}\left|a(t_{k+1},\bar{Y}_{k+1})-a(t_k,Y_k)\right|^2\right)\right)\\
&\le K\,\Delta\,E\left(\sum_{0\le k\le n_T-1}\left|a(t_{k+1},\bar{Y}_{k+1})-a(t_k,Y_k)\right|^2\right)\\
&\le K\,\Delta\,E\left(\sum_{0\le k\le n_T-1}\left|\bar{Y}_{k+1}-Y_k\right|^2\right)\\
&\le K\,\Delta\,E\left(\sum_{0\le k\le n_T-1}\left|\sum_{\alpha\in\mathcal{A}_{0.5}\setminus\{v\}}\widetilde{I}_\alpha[\tilde{f}_\alpha(t_k,Y_k)]_{t_k,t_{k+1}}\right|^2\right)\\
&\le K\,\Delta\sum_{\alpha\in\mathcal{A}_{0.5}\setminus\{v\}}E\left(\sum_{0\le k\le n_T-1}\left|\widetilde{I}_\alpha[\tilde{f}_\alpha(t_k,Y_k)]_{t_k,t_{k+1}}\right|^2\right).
\end{aligned} \tag{5.5.49}$$

Since $n_T < \infty$ almost surely, see (4.1.8), we can write

$$\begin{aligned}
&E\left(\sum_{0\leq k\leq n_T-1}\left|\tilde{I}_\alpha[\tilde{f}_\alpha(t_k,Y_k)]_{t_k,t_{k+1}}\right|^2\right)\\
&\quad=\sum_{i=1}^{\infty}E\left(\sum_{0\leq k\leq i-1}\left|\tilde{I}_\alpha[\tilde{f}_\alpha(t_k,Y_k)]_{t_k,t_{k+1}}\right|^2\right)\times P(n_T=i) \qquad (5.5.50)\\
&\quad=\sum_{i=1}^{\infty}E\left(\sum_{0\leq k\leq i-1}E\left(\left|\tilde{I}_\alpha[\tilde{f}_\alpha(t_k,Y_k)]_{t_k,t_{k+1}}\right|^2\big|\mathcal{A}_{t_k}\right)\right)\times P(n_T=i)
\end{aligned}$$

for every $\alpha \in \mathcal{A}_{0.5}\backslash\{v\}$. In the last line of (5.5.50) we have applied standard properties of conditional expectations.

By using the Cauchy-Schwarz inequality for $\alpha = (0)$, the Itô isometry for $\alpha = (j)$ and $j \in \{-1, 1\}$, and the linear growth conditions (4.5.9), we obtain

$$\begin{aligned}
&E\left(\left|\tilde{I}_\alpha[\tilde{f}_\alpha(t_k,Y_{t_k})]_{t_k,t_{k+1}}\right|^2\big|\mathcal{A}_{t_k}\right)\\
&\quad\leq \Delta^{n(\alpha)}\int_{t_k}^{t_{k+1}}\int_{\mathcal{E}}E\left(|\tilde{f}_\alpha(t_k,Y_{t_k})|^2\big|\mathcal{A}_{t_k}\right)\phi(dv^{s(\alpha)})dz\\
&\quad\leq \lambda^{s(\alpha)}\,\Delta^{n(\alpha)}\int_{t_k}^{t_{k+1}}E\left(1+|Y_{t_k}|^2\big|\mathcal{A}_{t_k}\right)dz \qquad (5.5.51)
\end{aligned}$$

for $k \in \{0, 1, \ldots, i\}$ and $i \in \mathbb{N}$.

By (5.5.49), (5.5.50), (5.5.51) and (5.5.48), we obtain

$$\begin{aligned}
&E\left(\max_{1\leq n\leq n_T}\left|\sum_{0\leq k\leq n-1}R_{1,k}\right|^2\right)\\
&\leq K\Delta\sum_{\alpha\in\mathcal{A}_{0.5}\backslash\{v\}}\lambda^{s(\alpha)}\Delta^{n(\alpha)}\\
&\quad\times\sum_{i=1}^{\infty}E\left(\int_0^{t_i}E\left((1+|Y_{n_z}|^2)\big|\mathcal{A}_{t_{n_z}}\right)dz\right)\times P(n_T=i)\\
&\leq K\Delta\sum_{\alpha\in\mathcal{A}_{0.5}\backslash\{v\}}\lambda^{s(\alpha)}\Delta^{n(\alpha)}\sum_{i=1}^{\infty}\int_0^{t_i}E\left(1+\max_{0\leq n\leq i}|Y_n|^2\right)dz\times P(n_T=i)
\end{aligned}$$

$$\leq K\Delta \sum_{\alpha\in\mathcal{A}_{0.5}\setminus\{v\}} \lambda^{s(\alpha)}\Delta^{n(\alpha)}\big(1+E\left(|Y_0|^2\right)\big)T\sum_{i=1}^{\infty}P(n_T=i)$$

$$\leq K\Delta = K\Delta^{2\gamma}, \tag{5.5.52}$$

for $\gamma = 0.5$.

Note that in Theorem 4.5.1 and in Theorem 5.5.1 we assumed that the coefficient function f_α satisfies the Lipschitz condition (4.5.7) for every $\alpha \in \mathcal{A}_\gamma$. If instead we require this condition to hold for every $\alpha \in \mathcal{A}_{0.5} \bigcup \mathcal{B}(\mathcal{A}_{0.5})$, then by considering the coefficient function corresponding to the multi-index $(1,1) \in \mathcal{B}(\mathcal{A}_{0.5})$, we see that the coefficient $b\,b'$ satisfies the Lipschitz type condition (4.5.7). Therefore, with similar steps as in (5.5.49), (5.5.50) and (5.5.51), we obtain

$$E\left(\max_{1\leq n\leq n_T}\left|\sum_{0\leq k\leq n-1} R_{2,k}\right|^2\right) \leq K\,\Delta = K\,\Delta^{2\gamma}, \tag{5.5.53}$$

for $\gamma = 0.5$.

By the linear growth condition on the coefficient $b\,b'$, which follows from condition (4.5.9), and the second moment estimate (5.5.48), one can show that

$$E\left(\max_{1\leq n\leq n_T}\left|\sum_{0\leq k\leq n-1} R_{3,k}\right|^2\right) \leq K\,\Delta = K\,\Delta^{2\gamma}, \tag{5.5.54}$$

for $\gamma = 0.5$.

By Doob's inequality, Itô's isometry, the Lipschitz condition on the diffusion coefficient and (5.5.40), we obtain

$$E\left(\max_{1\leq n\leq n_T}\left|\sum_{0\leq k\leq n-1} R_{4,k}\right|^2\right)$$

$$= E\left(\max_{1\leq n\leq n_T}\left|\sum_{0\leq k\leq n-1}\eta\int_{t_k}^{t_{k+1}}\big(b(t_{k+1},\bar{Y}_{k+1}) - b(t_k,Y_k)\big)dW_z\right|^2\right)$$

$$\leq 4E\left(\left|\int_0^T\big(b(t_{n_z+1},\bar{Y}_{n_z+1}) - b(t_{n_z},Y_{n_z})\big)dW_z\right|^2\right)$$

$$
\begin{aligned}
&= 4 \int_0^T E\left(\left|b(t_{n_z+1}, \bar{Y}_{n_z+1}) - b(t_{n_z}, Y_{n_z})\right|^2\right) dz \\
&\leq K \int_0^T E\left(\left|\bar{Y}_{n_z+1} - Y_{n_z}\right|^2\right) dz \\
&\leq K \int_0^T E\left(\left|\sum_{\alpha \in \mathcal{A}_{0.5} \setminus \{v\}} \widetilde{I}_\alpha[\tilde{f}_\alpha(t_{n_z}, Y_{n_z})]_{t_{n_z}, t_{n_z+1}}\right|^2\right) dz. \qquad (5.5.55)
\end{aligned}
$$

By the Lipschitz condition, the estimate (5.5.51) and the second moment estimate (5.5.48), we have

$$
\begin{aligned}
&E\left(\max_{1 \leq n \leq n_T} \left|\sum_{0 \leq k \leq n-1} R_{4,k}\right|^2\right) \\
&\leq K \sum_{\alpha \in \mathcal{A}_{0.5} \setminus \{v\}} \lambda^{s(\alpha)} \Delta^{n(\alpha)} \int_0^T E\left(\int_{t_{n_z}}^{t_{n_z+1}} E\left(1 + |Y_{t_{n_z}}|^2 \big| \mathcal{A}_{t_{n_z}}\right) ds\right) dz \\
&\leq K \sum_{\alpha \in \mathcal{A}_{0.5} \setminus \{v\}} \lambda^{s(\alpha)} \Delta^{n(\alpha)} \int_0^T E\left(\int_{t_{n_z}}^{t_{n_z+1}} E\left(1 + \max_{0 \leq n \leq n_T} |Y_{t_n}|^2 \big| \mathcal{A}_{t_{n_z}}\right) ds\right) dz \\
&\leq K \sum_{\alpha \in \mathcal{A}_{0.5} \setminus \{v\}} \lambda^{s(\alpha)} \Delta^{n(\alpha)+1} \left(1 + E(|Y_0|^2)\right) T \\
&\leq K\Delta = K\Delta^{2\gamma}, \qquad (5.5.56)
\end{aligned}
$$

for $\gamma = 0.5$.

Finally, by combining the estimates (5.5.52)-(5.5.53)-(5.5.54) and (5.5.56) we have shown that the family of predictor-corrector Euler schemes (5.4.1)-(5.4.2) can be rewritten as an order $\gamma = 0.5$ strong Itô scheme. Therefore, it achieves strong order $\gamma = 0.5$.

Also in the general multi-dimensional case one can analogously rewrite the family of predictor-corrector Euler schemes (5.4.3)-(5.4.5) and similarly the predictor-corrector order 1.0 strong scheme (5.4.10)-(5.4.11) as Itô scheme of order $\gamma = 0.5$ and $\gamma = 1.0$, respectively. Finally, strong predictor-corrector schemes of higher strong order can be derived by similar methods as the ones used above.

Chapter 6

Jump-Adapted Strong Approximations

This chapter describes jump-adapted strong schemes. The term *jump-adapted* refers to the time discretizations used to construct these scheme. These discretizations are called jump-adapted because they include all jump times generated by the Poisson jump measure. The form of the resulting schemes is much simpler than that of the regular schemes presented in Chapters 4 and 5. However, jump-adapted schemes are not efficient for SDEs driven by a Poisson measure with high total intensity. In this case, regular schemes are usually preferred. Some of the results of this chapter have been presented in Bruti-Liberati, Nikitopoulos-Sklibosios & Platen (2006) and in Bruti-Liberati & Platen (2007c).

6.1 Introduction

In principle, by including enough terms from Wagner-Platen expansions, it is possible to derive regular strong Taylor schemes of any given order of strong convergence, as shown by Theorems 4.5.1 and 5.5.1. However, as noticed in Section 4.3, even for a one-dimensional SDE, higher order schemes can be quite complex in that they involve multiple stochastic integrals jointly with respect to the Wiener process, the Poisson random measure and time. 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 the multiple stochastic integrals by using random number generators. To avoid the generation of multiple stochastic integrals with respect to the Poisson jump measure, Platen (1982a) proposed *jump-adapted approximations* that significantly reduce the complexity of higher order schemes. A jump-adapted time discretization makes these schemes easily implementable for scenario simulation also in the case of a mark-dependent jump size. Indeed, between jump times the evolution of the SDE (4.1.4) is that of a diffusion without jumps and can be approximated by standard schemes for pure diffusions, as presented in Kloeden & Platen (1999). At jump times the prescribed jumps are performed. As we will show in this chapter, it is possible to develop tractable jump-adapted higher order strong schemes also in the case of mark-dependent jump sizes. The required multiple stochastic integrals involve only time and Wiener process integrations.

In this chapter we consider a *jump-adapted time discretization* $0 = t_0 < t_1 < \ldots < t_{n_T} = T$, on which we construct a *jump-adapted discrete time approximation* $Y^\Delta = \{Y_t^\Delta, t \in [0,T]\}$ of the solution of the SDE (4.1.4). Here

$$n_t := \max\{n \in \{0,1,\ldots\} : t_n \leq t\} < \infty \quad \text{a.s.} \tag{6.1.1}$$

denotes the largest integer n such that $t_n \leq t$, for all $t \in [0,T]$. We require the jump-adapted time discretization to include all the jump times $\{\tau_1, \tau_2, \ldots, \tau_{p_\phi(T)}\}$ of the Poisson random measure p_ϕ. Moreover, as defined in Section 4.1, for a given maximum step size $\Delta \in (0,1)$ we require the *jump-adapted time discretization*

$$(t)_\Delta = \{0 = t_0 < t_1 < \ldots < t_{n_T} = T\}, \tag{6.1.2}$$

to satisfy the following conditions:

$$P(t_{n+1} - t_n \leq \Delta) = 1, \tag{6.1.3}$$

and

$$t_{n+1} \quad \text{is} \quad \mathcal{A}_{t_n} - \text{measurable}, \tag{6.1.4}$$

for $n \in \{0,1,\ldots,n_T - 1\}$, if it is not a jump time.

For instance, we could consider a jump-adapted time discretization $(t)_\Delta$ constructed by a superposition of the jump times $\{\tau_1, \tau_2, \ldots, \tau_{p_\phi(T)}\}$ of the Poisson random measure p_ϕ to a deterministic equidistant grid with nth discretization time $n\Delta$, $n \in \{0,1,\ldots,N\}$ and time step size $\Delta = \frac{T}{N}$. This means that we add all random jump times to an equidistant grid, as the one presented in Section 4.1. In this way the maximum time step size of the resulting jump-adapted discretization equals Δ.

Within this time grid we can separate the diffusive part of the dynamics from the jumps, because the jumps can arise only at discretization times. Therefore, between discretization points we can approximate the diffusive part with a strong scheme for pure diffusion processes. We add the effect of a jump to the evolution of the approximate solution when we encounter a jump time as discretization time. We remark that with jump-adapted schemes a mark-dependent jump coefficient does not add any additional complexity to the numerical approximation. Therefore, in this section we consider the general case of a jump-diffusion SDE with mark-dependent jump size given in (4.1.4).

For convenience we set $Y_{t_n} = Y_n$ and define

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

as the almost sure left-hand limit at time t_{n+1}. Moreover, to simplify the notation, we will use again the abbreviation

$$f = f(t_n, Y_{t_n}) \tag{6.1.5}$$

for any coefficient function f when no misunderstanding is possible, see (4.1.9). Furthermore, we will omit to mention the initial value Y_0 and the time step numbers $n \in \{0, 1, \ldots, n_T\}$ if this is convenient. However, we will aim to show the dependence on marks if this is relevant.

We will use again the operators

$$\begin{aligned} L^{(0)} f(t,x) \quad &:= \quad \frac{\partial}{\partial t} f(t,x) + \sum_{i=1}^{d} a^i(t,x) \frac{\partial}{\partial x^i} f(t,x) \\ &\quad + \frac{1}{2} \sum_{i,r=1}^{d} \sum_{j=1}^{m} b^{i,j}(t,x) b^{r,j}(t,x) \frac{\partial^2}{\partial x^i \partial x^j} f(t,x) \end{aligned} \tag{6.1.6}$$

$$L^{(k)} f(t,x) := \sum_{i=1}^{d} b^{i,k}(t,x) \frac{\partial}{\partial x^i} f(t,x), \qquad \text{for } k \in \{1, \ldots, m\} \tag{6.1.7}$$

for all $t \in [0, T]$ and $x \in \mathbb{R}^d$, similar as in the pure diffusion case, see Kloeden & Platen (1999).

Essentially, in the remainder of this chapter we review strong discrete time approximations for the diffusion part of the given SDE, as described in Kloeden & Platen (1999). Additionally, at each jump time of the driving Poisson measure we introduce a time discretization point in the jump-adapted discretization and approximate the jump as required.

6.2 Taylor Schemes

In this section we present jump-adapted strong approximations whose diffusion part is given by the truncated Wagner-Platen expansion for pure diffusions.

6.2.1 Euler Scheme

For the one-dimensional case, which means $d = m = 1$, we present the *jump-adapted Euler scheme* given by

$$Y_{t_{n+1}-} = Y_{t_n} + a\Delta_{t_n} + b\Delta W_{t_n} \tag{6.2.1}$$

and

$$Y_{t_{n+1}} = Y_{t_{n+1}-} + \int_{\mathcal{E}} c(t_{n+1}, Y_{t_{n+1}-}, v)\, p_\phi(dv, \{t_{n+1}\}), \tag{6.2.2}$$

where

$$\Delta_{t_n} = t_{n+1} - t_n \tag{6.2.3}$$

is the length of the time step size $[t_n, t_{n+1}]$ and

$$\Delta W_{t_n} = W_{t_{n+1}} - W_{t_n} \tag{6.2.4}$$

is the nth Gaussian $\mathcal{N}(0, \Delta_{t_n})$ distributed increment of the Wiener process W, $n \in \{0, 1, \ldots, N-1\}$. Note that we wrote $a = a(t_n, Y_{t_n})$ and $b = b(t_n, Y_{t_n})$ according to the abbreviation (6.1.5). The impact of jumps is obtained by (6.2.2). If t_{n+1} is a jump time, then

$$\int_{\mathcal{E}} p_\phi(dv, \{t_{n+1}\}) = 1 \tag{6.2.5}$$

and

$$\int_{\mathcal{E}} c(t_{n+1}, Y_{t_{n+1}-}, v)\, p_\phi(dv, \{t_{n+1}\}) = c(t_{n+1}, Y_{t_{n+1}-}, \xi_{p_\phi(t_{n+1})}). \tag{6.2.6}$$

If t_{n+1} is not a jump time one has

$$Y_{t_{n+1}} = Y_{t_{n+1}-}, \tag{6.2.7}$$

as

$$\int_{\mathcal{E}} p_\phi(dv, \{t_{n+1}\}) = 0. \tag{6.2.8}$$

Therefore, the strong order of convergence of the jump-adapted Euler scheme is $\gamma = 0.5$, resulting from the strong order of the approximation (6.2.1) of the diffusive component.

For instance, for the SDE (2.1.5), the jump-adapted Euler scheme is given by

$$Y_{t_{n+1}-} = Y_{t_n} + \mu\, Y_{t_n} \Delta_{t_n} + \sigma\, Y_{t_n} \Delta W_{t_n} \tag{6.2.9}$$

and

$$Y_{t_{n+1}} = Y_{t_{n+1}-} + Y_{t_{n+1}-} \int_{\mathcal{E}} v\, p_\phi(dv, \{t_{n+1}\}). \tag{6.2.10}$$

In the multi-dimensional case with scalar Wiener process, which means $m = 1$, the kth component of the *jump-adapted Euler scheme* is given by

$$Y^k_{t_{n+1}-} = Y^k_{t_n} + a^k \Delta_{t_n} + b^k \Delta W_{t_n} \tag{6.2.11}$$

and

$$Y^k_{t_{n+1}} = Y^k_{t_{n+1}-} + \int_{\mathcal{E}} c^k(t_{n+1}, Y_{t_{n+1}-}, v)\, p_\phi(dv, \{t_{n+1}\}), \tag{6.2.12}$$

for $k \in \{1, 2, \dots, d\}$, where a^k, b^k, and c^k are the kth components of the drift, the diffusion and the jump coefficients, respectively.

For the general multi-dimensional case the kth component of the *jump-adapted Euler scheme* is of the form

$$Y^k_{t_{n+1}-} = Y^k_{t_n} + a^k \Delta_{t_n} + \sum_{j=1}^{m} b^{k,j} \Delta W^j_{t_n} \tag{6.2.13}$$

and

$$Y^k_{t_{n+1}} = Y^k_{t_{n+1}-} + \int_{\mathcal{E}} c^k(t_{n+1}, Y_{t_{n+1}-}, v)\, p_\phi(dv, \{t_{n+1}\}), \tag{6.2.14}$$

where a^k and c^k are the kth components of the drift and the jump coefficients, respectively. Furthermore, $b^{k,j}$ is the component of the kth row and jth column of the diffusion matrix b, for $k \in \{1, 2, \ldots, d\}$, and $j \in \{1, 2, \ldots, m\}$. Additionally,

$$\Delta W^j_{t_n} = W^j_{t_{n+1}} - W^j_{t_n} \tag{6.2.15}$$

is the $\mathcal{N}(0, \Delta_{t_n})$ distributed nth increment of the jth Wiener process.

6.2.2 Order 1.0 Taylor Scheme

As the order of convergence of jump-adapted schemes is, in general, the one induced by the approximation of the diffusive part, by replacing the diffusive part of the jump-adapted Euler scheme with the order 1.0 strong Taylor scheme for diffusions, see Kloeden & Platen (1999), we obtain the jump-adapted order 1.0 strong Taylor scheme.

For a one-dimensional SDE, $d = m = 1$, we can derive the *jump-adapted order* 1.0 *strong Taylor scheme* given by

$$Y_{t_{n+1}-} = Y_{t_n} + a\Delta_{t_n} + b\Delta W_{t_n} + \frac{b\,b'}{2}\Big((\Delta W_{t_n})^2 - \Delta_{t_n}\Big) \tag{6.2.16}$$

and (6.2.2), which achieves strong order $\gamma = 1.0$. This scheme can be interpreted as a jump-adapted version of the Milstein scheme for pure diffusions, see Milstein (1974).

The comparison of the jump-adapted order 1.0 strong scheme (6.2.16) with the regular order 1.0 strong Taylor scheme (4.3.1), 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.

For instance, for the SDE (2.1.5), describing the Merton model, we have the jump-adapted order 1.0 strong Taylor scheme of the form

$$Y_{t_{n+1}-} = Y_{t_n} + \mu\, Y_{t_n}\Delta_{t_n} + \sigma\, Y_{t_n}\Delta W_{t_n} + \frac{\sigma^2\, Y_{t_n}}{2}\left\{ \Big(\Delta W_{t_n}\Big)^2 - \Delta_{t_n}\right\} \tag{6.2.17}$$

and (6.2.10).

For the multi-dimensional case with scalar Wiener process, which means $m = 1$, the kth component of the *jump-adapted order* 1.0 *strong Taylor scheme* is given by

$$Y^k_{t_{n+1}-} = Y^k_{t_n} + a^k \Delta_{t_n} + b^k \Delta W_{t_n} + \frac{1}{2}\sum_{l=1}^{d} b^l \frac{\partial b^k}{\partial x^l}\Big(\big(\Delta W_{t_n}\big)^2 - \Delta_{t_n}\Big) \tag{6.2.18}$$

and (6.2.12), for $k \in \{1, 2, \ldots, d\}$.

In the general multi-dimensional case the kth component of the *jump-adapted order* 1.0 *strong Taylor scheme* is given by

$$Y^k_{t_{n+1}-} = Y^k_{t_n} + a^k \Delta_{t_n} + \sum_{j=1}^{m} b^{k,j} \Delta W^j_{t_n} + \sum_{j_1,j_2=1}^{m} \sum_{i=1}^{d} \Big(b^{i,j_1} \frac{\partial b^{k,j_2}}{\partial x^i}\Big) I_{(j_1,j_2)} \tag{6.2.19}$$

and (6.2.14), for $k \in \{1, 2, \ldots, d\}$. For the generation of the multiple stochastic integrals $I_{(j_1,j_2)}$ we refer to Section 4.4.

If we have a multi-dimensional SDE satisfying the diffusion commutativity condition, where

$$L^{j_1} b^{k,j_2}(t, x) = L^{j_2} b^{k,j_1}(t, x) \tag{6.2.20}$$

for $j_1, j_2 \in \{1, 2, \ldots, m\}$, $k \in \{1, 2, \ldots, d\}$, $t \in [0, T]$ and $x \in \mathbb{R}^d$, then it is possible to express all the multiple stochastic integrals in terms of the increments $\Delta W^{j_1}_{t_n}$ and $\Delta W^{j_2}_{t_n}$ of the Wiener process. Therefore, we obtain an efficiently implementable *jump-adapted order* 1.0 *strong Taylor scheme*, whose kth component is given by

$$Y^k_{t_{n+1}-} = Y^k_{t_n} + a^k \Delta_{t_n} + \sum_{j=1}^{m} b^{k,j} \Delta W^j_{t_n} + \frac{1}{2} \sum_{j_1,j_2=1}^{m} \sum_{i=1}^{d} b^{i,j_1} \frac{\partial b^{k,j_2}}{\partial x^i} \Big\{ \Delta W^{j_1}_{t_n} \, \Delta W^{j_2}_{t_n} - \Delta_{t_n} \Big\}$$

and (6.2.14), for $k \in \{1, 2, \ldots, d\}$. The special case of additive diffusion noise, which means $b(t, x) = b(t)$, satisfies the required commutativity condition and, therefore, leads to an efficient jump-adapted order 1.0 strong Taylor scheme. Note that this holds for general jump coefficients, unlike the case of regular strong schemes in Chapter 4.

6.2.3 Order 1.5 Taylor Scheme

If we approximate the diffusive part of the SDE (4.1.4) with the order 1.5 strong Taylor scheme, see Kloeden & Platen (1999), then we obtain the *jump-adapted* *order* 1.5 *strong Taylor scheme.*

In the autonomous one-dimensional case, $d = m = 1$, the *jump-adapted order* 1.5 *strong Taylor scheme* is given by

$$\begin{aligned} Y_{t_{n+1}-} &= Y_{t_n} + a\Delta_{t_n} + b\Delta W_{t_n} + \frac{bb'}{2}\left((\Delta W_{t_n})^2 - \Delta_{t_n}\right) + a'\,b\Delta Z_{t_n} \\ &+\frac{1}{2}\left(a\,a' + \frac{1}{2}b^2 a''\right)(\Delta_{t_n})^2 + \left(a\,b' + \frac{1}{2}b^2\,b''\right)\left(\Delta W_{t_n}\Delta_{t_n} - \Delta Z_{t_n}\right) \\ &+\frac{1}{2}b\left(b\,b'' + b'^2\right)\left\{\frac{1}{3}(\Delta W_{t_n})^2 - \Delta_{t_n}\right\}\Delta W_{t_n}, \end{aligned} \tag{6.2.21}$$

and (6.2.2), where we have used the abbreviation defined in (6.1.5). Here we need the multiple stochastic integral

$$\Delta Z_{t_n} = \int_{t_n}^{t_{n+1}} \int_{t_n}^{s_2} dW_{s_1}\,ds_2. \tag{6.2.22}$$

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:

$$\Delta W_{t_n} = U_1\sqrt{\Delta_{t_n}} \quad \text{and} \quad \Delta Z_{t_n} = \frac{1}{2}(\Delta_{t_n})^{\frac{3}{2}}\left(U_1 + \frac{1}{\sqrt{3}}U_2\right). \tag{6.2.23}$$

For example, for the SDE (2.1.5) of the Merton model the terms involving the random variable ΔZ_{t_n} cancel out, thus yielding the rather simple jump-adapted order 1.5 strong Taylor scheme

$$\begin{aligned} Y_{t_{n+1}-} &= Y_{t_n} + \mu Y_{t_n}\Delta_{t_n} + \sigma Y_{t_n}\Delta W_{t_n} + \frac{\sigma^2 Y_{t_n}}{2}\left((\Delta W_{t_n})^2 - \Delta_{t_n}\right) \\ &+\frac{\mu^2 Y_{t_n}}{2}(\Delta_{t_n})^2 + \mu\,\sigma\,Y_{t_n}\left(\Delta W_{t_n}\Delta_{t_n}\right) \\ &+\frac{1}{2}\sigma^3\,Y_{t_n}\left\{\frac{1}{3}(\Delta W_{t_n})^2 - \Delta_{t_n}\right\}\Delta W_{t_n} \end{aligned} \tag{6.2.24}$$

and (6.2.10).

For the multi-dimensional case with scalar Wiener process, $m = 1$, the kth component of the *jump-adapted order* 1.5 *strong Taylor scheme* is given by

$$
\begin{aligned}
Y^k_{t_{n+1}-} &= Y^k_{t_n} + a^k \Delta_{t_n} + b^k \Delta W_{t_n} + \frac{1}{2} L^{(1)} b^k \left\{ \left(\Delta W_{t_n}\right)^2 - \Delta_{t_n} \right\} \\
&\quad + L^{(1)} a^k \Delta Z_{t_n} + L^{(0)} b^k \left\{ \Delta W_{t_n} \Delta_{t_n} - \Delta Z_{t_n} \right\} + \frac{1}{2} L^{(0)} a^k \left(\Delta_{t_n}\right)^2 \\
&\quad + \frac{1}{2} L^{(1)} L^{(1)} b^k \left\{ \frac{1}{3} \left(\Delta W_{t_n}\right)^2 - \Delta_{t_n} \right\} \Delta W_{t_n} \qquad (6.2.25)
\end{aligned}
$$

and (6.2.12), for $k \in \{1, 2, \ldots, d\}$, where the differential operators L^0 and L^1 are defined in (6.1.6) and (6.1.7), respectively.

In the general multi-dimensional case the kth component of the *jump-adapted order* 1.5 *strong Taylor scheme* is given by

$$
\begin{aligned}
Y^k_{t_{n+1}-} &= Y^k_{t_n} + a^k \Delta_{t_n} + \frac{1}{2} L^{(0)} a^k \left(\Delta_{t_n}\right)^2 \\
&\quad + \sum_{j=1}^{m} \left(b^{k,j} \Delta W^j_{t_n} + L^{(0)} b^{k,j} I_{(0,j)} + L^{(j)} a^k I_{(j,0)} \right) \qquad (6.2.26) \\
&\quad + \sum_{j_1, j_2 = 1}^{m} L^{(j_1)} b^{k,j_2} I_{(j_1,j_2)} + \sum_{j_1, j_2, j_3 = 1}^{m} L^{(j_1)} L^{(j_2)} b^{k,j_3} I_{(j_1,j_2,j_3)},
\end{aligned}
$$

and (6.2.14), for $k \in \{1, 2, \ldots, d\}$.

The double stochastic integrals appearing in the jump-adapted order 1.5 strong Taylor scheme can be generated as discussed in Section 4.3. We refer to Kloeden & Platen (1999) for the generation of the required triple stochastic integrals and for diffusion commutativity conditions that reduce the complexity of this 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 6.7 that provides the methodology for the construction of jump-adapted schemes of any given strong order.

6.3 Derivative-Free Schemes

As noticed in Section 5.2, it is convenient to develop higher order numerical approximations that do not require the evaluation of derivatives of the coefficient functions. Within 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.

6.3.1 Derivative-Free Order 1.0 Scheme

For a one-dimensional SDE, $d = m = 1$, the *jump-adapted derivative-free order* 1.0 *strong scheme*, which achieves a strong order $\gamma = 1.0$, is given by

$$\begin{aligned} Y_{t_{n+1}-} &= Y_{t_n} + a\Delta_{t_n} + b\Delta W_{t_n} \\ &\quad + \frac{1}{2\sqrt{\Delta_{t_n}}}\{b(t_n, \bar{Y}_{t_n}) - b\} \left((\Delta W_{t_n})^2 - \Delta_{t_n}\right), \end{aligned} \tag{6.3.1}$$

and (6.2.2) with the supporting value

$$\bar{Y}_{t_n} = Y_{t_n} + b\sqrt{\Delta_{t_n}}. \tag{6.3.2}$$

In the multi-dimensional case with scalar Wiener process, which means $m = 1$, the kth component of the *jump-adapted derivative-free order* 1.0 *strong scheme* is given by

$$\begin{aligned} Y^k_{t_{n+1}-} &= Y^k_{t_n} + a^k\Delta_{t_n} + b^k\Delta W_{t_n} \\ &\quad + \frac{1}{2\sqrt{\Delta_{t_n}}}\{b^k(t_n, \bar{Y}_{t_n}) - b^k\} \left((\Delta W_{t_n})^2 - \Delta_{t_n}\right), \end{aligned} \tag{6.3.3}$$

and (6.2.12), with the vector supporting value

$$\bar{Y}_{t_n} = Y_{t_n} + b\sqrt{\Delta_{t_n}}, \tag{6.3.4}$$

for $k \in \{1, 2, \ldots, d\}$.

In the general multi-dimensional case the kth component of the *jump-adapted* *derivative-free order* 1.0 *strong scheme* is given by

$$
\begin{aligned}
Y^k_{t_{n+1}} &= Y^k_{t_n} + a^k \Delta_{t_n} + \sum_{j=1}^{m} b^{k,j} \Delta W^j_{t_n} \\
&+ \frac{1}{\sqrt{\Delta_{t_n}}} \sum_{j_1,j_2=1}^{m} \left\{ b^{k,j_2}(t_n, \overline{Y}^{j_1}_{t_n}) - b^{k,j_2}(t_n, Y_{t_n}) \right\} I_{(j_1,j_2)}, \qquad (6.3.5)
\end{aligned}
$$

and (6.2.14), with the vector supporting value

$$
\overline{Y}^j_{t_n} = Y_{t_n} + b^j \sqrt{\Delta_{t_n}}, \qquad (6.3.6)
$$

for $k \in \{1, 2, \ldots, d\}$ and for $j \in \{1, 2, \ldots, m\}$. The multiple stochastic integrals can be generated, as in Section 4.3, by Karhunen-Loève expansions. The diffusion commutativity condition presented in Section 6.2 may apply also here and, therefore, can lead to efficiently implementable jump-adapted derivative-free schemes.

6.3.2 Derivative-Free Order 1.5 Scheme

In the autonomous one-dimensional case, $d = m = 1$, the *jump-adapted derivative-free order* 1.5 *strong scheme* is given by

$$
\begin{aligned}
Y_{t_{n+1}-} &= Y_{t_n} + b\Delta W_{t_n} + \frac{1}{2\sqrt{\Delta_{t_n}}} \left\{ a(\overline{Y}^+_{t_n}) - a(\overline{Y}^-_{t_n}) \right\} \Delta Z_{t_n} \\
&+ \frac{1}{4} \left\{ a(\overline{Y}^+_{t_n}) + 2a + a(\overline{Y}^-_{t_n}) \right\} \Delta_{t_n} \\
&+ \frac{1}{4\sqrt{\Delta_{t_n}}} \left\{ b(\overline{Y}^+_{t_n}) - b(\overline{Y}^-_{t_n}) \right\} \left((\Delta W_{t_n})^2 - \Delta_{t_n} \right) \\
&+ \frac{1}{2\sqrt{\Delta_{t_n}}} \left\{ b(\overline{Y}^+_{t_n}) + 2b + b(\overline{Y}^-_{t_n}) \right\} \left(\Delta W_{t_n} \Delta_{t_n} - \Delta Z_{t_n} \right) \\
&+ \frac{1}{4\sqrt{\Delta_{t_n}}} \left[b(\overline{\Phi}^+_{t_n}) - b(\overline{\Phi}^-_{t_n}) - b(\overline{Y}^+_{t_n}) + b(\overline{Y}^-_{t_n}) \right] \\
&\quad \times \left\{ \frac{1}{3} (\Delta W_{t_n})^2 - \Delta_{t_n} \right\} \Delta W_{t_n}, \qquad (6.3.7)
\end{aligned}
$$

and (6.2.2), with

$$
\overline{Y}^{\pm}_{t_n} = Y_{t_n} + a\Delta_{t_n} \pm b\Delta W_{t_n}, \qquad (6.3.8)
$$

and

$$\overline{\Phi}^{\pm}_{t_n} = \overline{Y}^{\pm}_{t_n} \pm b(\overline{Y}^{+}_{t_n})\sqrt{\Delta_{t_n}}. \tag{6.3.9}$$

The multiple stochastic integral $\Delta Z_{t_n} = I_{(1,0)}$ can be generated as in (6.2.23).

6.4 Drift-Implicit Schemes

As previously discussed, 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 only.

6.4.1 Drift-Implicit Euler Scheme

For a one-dimensional SDE, $d = m = 1$, the *jump-adapted drift-implicit Euler scheme* is given by

$$Y_{t_{n+1}-} = Y_{t_n} + \left\{\theta\, a(t_{n+1}, Y_{t_{n+1}-}) + (1-\theta)\, a\right\} \Delta_{t_n} + b\Delta W_{t_n}, \tag{6.4.1}$$

and (6.2.2), where the parameter $\theta \in [0,1]$ characterizes the degree of implicitness.

In the multi-dimensional case with scalar Wiener noise, $m = 1$, the kth component of the *jump-adapted drift-implicit Euler scheme* is given by:

$$Y^k_{t_{n+1}-} = Y^k_{t_n} + \left\{\theta\, a^k(t_{n+1}, Y_{t_{n+1}-}) + (1-\theta)\, a^k\right\} \Delta_{t_n} + b^k\Delta W_{t_n}, \tag{6.4.2}$$

and (6.2.12), for $k \in \{1, 2, \ldots, d\}$.

In the multi-dimensional case the kth component of the *jump-adapted drift-implicit Euler scheme* is given by

$$Y^k_{t_{n+1}-} = Y^k_{t_n} + \left\{\theta\, a^k(t_{n+1}, Y_{t_{n+1}-}) + (1-\theta)\, a^k\right\} \Delta_{t_n} + \sum_{j=1}^{m} b^{k,j}\Delta W^j_{t_n},$$

and (6.2.14), for $k \in \{1, 2, \ldots, d\}$.

6.4.2 Drift-Implicit Order 1.0 Scheme

By using a drift-implicit order 1.0 strong scheme for the diffusive part, we obtain a jump-adapted drift-implicit order 1.0 strong scheme.

For a one-dimensional SDE, $d = m = 1$, the *jump-adapted drift-implicit order* 1.0 *strong scheme* is given by

$$\begin{aligned} Y_{t_{n+1}-} &= Y_{t_n} + \left\{\theta\, a(t_{n+1}, Y_{t_{n+1}-}) + (1-\theta)\, a\right\} \Delta_{t_n} + b\, \Delta W_{t_n} \\ &\quad + \frac{b\, b'}{2} \left((\Delta W_{t_n})^2 - \Delta_{t_n} \right) \end{aligned} \tag{6.4.3}$$

and (6.2.2), which achieves strong order $\gamma = 1.0$.

In the multi-dimensional case with scalar Wiener process, $m = 1$, the kth component of the *jump-adapted drift-implicit order* 1.0 *strong scheme* is given by

$$\begin{aligned} Y^k_{t_{n+1}-} &= Y^k_{t_n} + \left\{\theta\, a^k(t_{n+1}, Y_{t_{n+1}-}) + (1-\theta)\, a^k\right\} \Delta_{t_n} + b^k \Delta W_{t_n} \\ &\quad + \frac{1}{2} \sum_{l=1}^{d} b^l \frac{\partial b^k}{\partial x^l} \left((\Delta W_{t_n})^2 - \Delta_{t_n} \right) \end{aligned} \tag{6.4.4}$$

and (6.2.12), for $k \in \{1, 2, \ldots, d\}$.

In the general multi-dimensional case, the kth component of the *jump-adapted drift-implicit order* 1.0 *strong scheme* is given by

$$\begin{aligned} Y^k_{t_{n+1}-} &= Y^k_{t_n} + \left\{\theta\, a^k(t_{n+1}, Y_{t_{n+1}-}) + (1-\theta)\, a^k\right\} \Delta_{t_n} + \sum_{j=1}^{m} b^{k,j} \Delta W^j_{t_n} \\ &\quad + \sum_{j_1, j_2 = 1}^{m} \sum_{i=1}^{d} \left(b^{i,j_1} \frac{\partial b^{k,j_2}}{\partial x^i} \right) I_{(j_1, j_2)} \end{aligned} \tag{6.4.5}$$

and (6.2.14), for $k \in \{1, 2, \ldots, d\}$.

6.4.3 Drift-Implicit Order 1.5 Scheme

We present here a jump-adapted drift-implicit order 1.5 strong Taylor scheme.

In the autonomous one-dimensional case, $d = m = 1$, the *jump-adapted drift-*

implicit order 1.5 *strong scheme* in its simplest form is given by

$$
\begin{aligned}
Y_{t_{n+1}-} &= Y_{t_n} + \frac{1}{2}\left\{a(Y_{t_{n+1}-}) + a\right\}\Delta_{t_n} + b\,\Delta W_{t_n} + \frac{b\,b'}{2}\left((\Delta W_{t_n})^2 - \Delta_{t_n}\right) \\
&\quad + \left(a\,b' + \frac{1}{2}b^2\,b''\right)(\Delta W_{t_n}\Delta_{t_n} - \Delta Z_{t_n}) + a'\,b\left\{\Delta Z_{t_n} - \frac{1}{2}\Delta W_{t_n}\Delta_{t_n}\right\} \\
&\quad + \frac{1}{2}b\left(b\,b'' + (b')^2\right)\left\{\frac{1}{3}(\Delta W_{t_n})^2 - \Delta_{t_n}\right\}\Delta W_{t_n}, \qquad (6.4.6)
\end{aligned}
$$

and (6.2.2), and achieves strong order $\gamma = 1.5$.

Finally, we remark that balanced implicit methods, see Milstein, Platen & Schurz (1998), Kahl & Schurz (2006) and Alcock & Burrage (2006), can be used to obtain a numerically stable approximation of the diffusion part. We do not discuss these methods here, due to the limited space available. In general, when using balanced implicit methods, one has still to solve an algebraic equation at each time step. The following class of predictor-corrector methods avoids this extra computational effort.

6.5 Predictor-Corrector Schemes

As previously discussed, predictor-corrector schemes combine good numerical stability properties with efficiency. In this section we present new jump-adapted predictor-corrector schemes with strong order of convergence $\gamma \in \{0.5, 1\}$.

6.5.1 Predictor-Corrector Euler Scheme

In the one-dimensional case, $d = m = 1$, the *family of jump-adapted predictor-corrector Euler schemes* is given by the corrector

$$
\begin{aligned}
Y_{t_{n+1}-} &= Y_n + \left\{\theta\bar{a}_\eta(t_{n+1}, \bar{Y}_{t_{n+1}-}) + (1-\theta)\bar{a}_\eta\right\}\Delta_{t_n} \\
&\quad + \left\{\eta b(t_{n+1}, \bar{Y}_{t_{n+1}-}) + (1-\eta)b\right\}\Delta W_{t_n}, \qquad (6.5.1)
\end{aligned}
$$

where $\bar{a}_\eta = a - \eta b\,b'$, the predictor

$$
\bar{Y}_{t_{n+1}-} = Y_{t_n} + a\Delta_{t_n} + b\Delta W_{t_n}, \qquad (6.5.2)
$$

and (6.2.2). The parameters $\theta, \eta \in [0,1]$ characterize the degree of implicitness in the drift and in the diffusion coefficients, respectively. This scheme achieves a strong order of convergence $\gamma = 0.5$.

For the general multi-dimensional case, the kth component of the *family of jump-adapted predictor-corrector Euler schemes* is given by the corrector

$$\begin{aligned} Y^k_{t_{n+1}-} &= Y^k_{t_n} + \left\{\theta \bar{a}^k_\eta(t_{n+1}, \bar{Y}_{t_{n+1}-}) + (1-\theta)\bar{a}^k_\eta\right\} \Delta_{t_n} \\ &\quad + \sum_{j=1}^m \left\{\eta b^{k,j}(t_{n+1}, \bar{Y}_{t_{n+1}}) + (1-\eta) b^{k,j}\right\} \Delta W^j_{t_n}, \end{aligned} \tag{6.5.3}$$

for $\theta, \eta \in [0,1]$, where

$$\bar{a}_\eta = a - \eta \sum_{j_1, j_2 = 1}^m \sum_{i=1}^d b^{k,j_1} \frac{\partial b^{k,j_2}}{\partial x^i}, \tag{6.5.4}$$

the predictor

$$\bar{Y}^k_{t_{n+1}-} = Y^k_{t_n} + a^k \Delta_{t_n} + \sum_{j=1}^m b^{k,j} \Delta W^j_{t_n}, \tag{6.5.5}$$

and (6.2.14).

6.5.2 Predictor-Corrector Order 1.0 Scheme

Here we present the *jump-adapted predictor-corrector order* 1.0 *strong scheme.* For a one-dimensional SDE, $d = m = 1$, it is given by the corrector

$$\begin{aligned} Y_{t_{n+1}-} &= Y_{t_n} + \left\{\theta\, a(t_{n+1}, \bar{Y}_{t_{n+1}-}) + (1-\theta)\, a\right\} \Delta_{t_n} \\ &\quad + b\, \Delta W_{t_n} + \frac{b\, b'}{2} \left((\Delta W_{t_n})^2 - \Delta_{t_n}\right), \end{aligned} \tag{6.5.6}$$

the predictor

$$\bar{Y}_{t_{n+1}-} = Y_{t_n} + a\Delta_{t_n} + b\Delta W_{t_n} + \frac{b\, b'}{2}\Big((\Delta W_{t_n})^2 - \Delta_{t_n}\Big) \tag{6.5.7}$$

and (6.2.2). The parameter $\theta \in [0,1]$ characterizes again the degree of implicitness in the drift coefficient. This scheme attains a strong order $\gamma = 1.0$.

In the general multi-dimensional case, the kth component of the *jump-adapted* *predictor-corrector order* 1.0 *strong scheme* is given by the corrector

$$
\begin{aligned}
Y^k_{t_{n+1}-} &= Y^k_{t_n} + \left\{\theta\, a^k(t_{n+1}, \bar{Y}_{t_{n+1}-}) + (1-\theta)\, a^k\right\} \Delta_{t_n} + \sum_{j=1}^{m} b^{k,j} \Delta W^j_{t_n} \\
&\quad + \sum_{j_1,j_2=1}^{m} \sum_{i=1}^{d} \left(b^{i,j_1} \frac{\partial b^{k,j_2}}{\partial x^i} \right) I_{(j_1,j_2)}, \qquad (6.5.8)
\end{aligned}
$$

the predictor

$$
\bar{Y}^k_{t_{n+1}-} = Y^k_{t_n} + a^k \Delta_{t_n} + \sum_{j=1}^{m} b^{k,j} \Delta W^j_{t_n} + \sum_{j_1,j_2=1}^{m} \sum_{i=1}^{d} \left(b^{i,j_1} \frac{\partial b^{k,j_2}}{\partial x^i} \right) I_{(j_1,j_2)} \qquad (6.5.9)
$$

and (6.2.14), for $k \in \{1, 2, \ldots, d\}$. For the generation of the multiple stochastic integrals $I_{(j_1,j_2)}$ we refer again to Section 4.4.

For SDEs satisfying the diffusion commutativity condition (6.2.20), as in the case of an additive diffusion coefficient $b(t,x) = b(t)$, we can express all the multiple stochastic integrals in terms of the increments $\Delta W^{j_1}_{t_n}$ and $\Delta W^{j_2}_{t_n}$ of the Wiener process. This yields an efficiently implementable *jump-adapted predictor-corrector order* 1.0 *strong scheme*, whose kth component is given by the corrector

$$
\begin{aligned}
Y^k_{t_{n+1}-} &= Y^k_{t_n} + \left\{\theta\, a^k(t_{n+1}, \bar{Y}_{t_{n+1}-}) + (1-\theta)\, a^k\right\} \Delta_{t_n} \\
&\quad + \sum_{j=1}^{m} b^{k,j} \Delta W^j_{t_n} + \frac{1}{2} \sum_{j_1,j_2=1}^{m} \sum_{i=1}^{d} b^{i,j_1} \frac{\partial b^{k,j_2}}{\partial x^i} \left\{ \Delta W^{j_1}_{t_n} \Delta W^{j_2}_{t_n} - \Delta_{t_n} \right\},
\end{aligned}
$$

the predictor

$$
\bar{Y}^k_{t_{n+1}-} = Y^k_{t_n} + a^k \Delta_{t_n} + \sum_{j=1}^{m} b^{k,j} \Delta W^j_{t_n} + \frac{1}{2} \sum_{j_1,j_2=1}^{m} \sum_{i=1}^{d} b^{i,j_1} \frac{\partial b^{k,j_2}}{\partial x^i} \left\{ \Delta W^{j_1}_{t_n} \Delta W^{j_2}_{t_n} - \Delta_{t_n} \right\}
$$

and (6.2.14), for $k \in \{1, 2, \ldots, d\}$.

We remark that, as in Section 5.5.3, we can also make the diffusion part quasi-implicit to obtain better numerical stability.

6.6 Exact Schemes

In this section we discuss the strong approximation of a special class of SDEs under which jump-adapted schemes have no discretization error. As will become clear later, in this case it is *not* necessary to request a maximum time step size Δ, see condition (6.1.3). Here the jump-adapted time discretization will be given by a superposition of the jump times generated by the Poisson measure and the times at which we are interested in sampling the simulated values of the solution X. For instance, if one needs the value of the jump-diffusion X only at the final time T, then the jump-adapted time discretization is given by $0 = t_0 < t_1 < \ldots < t_{n_T} = T$, where n_T is defined in (4.1.8) and the sequence $t_1 < \ldots < t_{n_T - 1}$ equals that of the jump times $\tau_1 < \ldots < \tau_{p_\phi(T)}$ of the Poisson measure p_ϕ.

Let us first present an illustrative example. Consider the jump-diffusion SDE with *multiplicative* drift and diffusion coefficients and general jump coefficient $c(t, x, v)$, given by

$$dX_t = \mu X_t dt + \sigma X_t dW_t + \int_{\mathcal{E}} c(t, X_{t-}, v)\, p_\phi(dv, dt). \tag{6.6.1}$$

Because of the general form of the jump coefficient c, the SDE (6.6.1) does rarely admit an explicit solution. However, on the above described jump-adapted time discretization, we can construct the jump-adapted scheme given by

$$Y_{t_{n+1}-} = Y_{t_n} e^{(\mu - \frac{1}{2}\sigma^2)\Delta_{t_n} + \sigma \Delta W_{t_n}}, \tag{6.6.2}$$

and

$$Y_{t_{n+1}} = Y_{t_{n+1}-} + \int_{\mathcal{E}} c(t_{n+1}, Y_{t_{n+1}-}, v)\, p_\phi(dv, \{t_{n+1}\}). \tag{6.6.3}$$

Since we are using the explicit solution of the diffusion part, see (2.1.6), no discretization error is introduced between jump times. Moreover, since by equation (6.6.3) the jump impact is added at the correct jump times, even at the jump times we do not introduce any error. Therefore, we have described a way to express the solution of the jump-diffusion SDE (6.6.1) that does not generate any discretization error.

This approach can be generalized for cases where we have an explicit solution of the diffusion part of the SDE under consideration. In the general case, we recall here the d-dimensional jump-diffusion SDE

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t + \int_{\mathcal{E}} c(t, X_{t-}, v)\, p_\phi(dv, dt), \tag{6.6.4}$$

that we aim to solve. One should then check whether this SDE belongs to the special subclass of jump-diffusion SDEs for which the corresponding diffusion SDE

$$dZ_t = a(t, Zt)dt + b(t, Z_t)dW_t, \tag{6.6.5}$$

admits an explicit solution. If this is the case, then we can construct, as in (6.6.2)-(6.6.3), a jump-adapted scheme without discretization error. For explicit solutions of pure diffusions SDEs we refer to Chapter 4 of Kloeden & Platen (1999).

As previously noticed, if the SDE under consideration is driven by a Poisson measure with high intensity, then a jump-adapted scheme may be computationally too expensive. In such a case, one may prefer to use a regular scheme that entails a discretization error but permits to use a coarser time discretization.

6.7 Convergence Results

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 this chapter.

We consider here a *jump-adapted time discretization* $(t)_\Delta$, as defined in (6.1.2). Let us recall that by "jump-adapted" we mean that the time discretization includes all jump times $\{\tau_1, \tau_2, \dots, \tau_{p_\phi(T)}\}$ of the Poisson measure p_ϕ. As explained in Section 6.1, by construction the jumps arise only at discretization points. Therefore, between discretization 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 compared to the one introduced in Chapter 3, as will be outlined below.

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

$$\widehat{\mathcal{M}}_m = \{(j_1, \dots, j_l) : j_i \in \{0, 1, 2, \dots, m\}, i \in \{1, 2, \dots, l\} \text{ for } l \in \mathbb{N}\} \cup \{v\},$$

where v is the multi-index of length zero.

We also recall the operators

$$\begin{aligned} L^{(0)}f(t,x) \; &:= \; \frac{\partial}{\partial t}f(t,x) + \sum_{i=1}^{d} a^i(t,x)\frac{\partial}{\partial x^i}f(t,x) \\ &+\frac{1}{2}\sum_{i,r=1}^{d}\sum_{j=1}^{m} b^{i,j}(t,x)b^{r,j}(t,x)\frac{\partial^2}{\partial x^i \partial x^j}f(t,x) \end{aligned} \tag{6.7.1}$$

and

$$L^{(k)}f(t,x) := \sum_{i=1}^{d} b^{i,k}(t,x)\frac{\partial}{\partial x^i}f(t,x) \tag{6.7.2}$$

for $k \in \{1,\dots,m\}$ and a function $f(t,x) : [0,T] \times \mathbb{R}^d \longrightarrow \mathbb{R}^d$ from $\mathcal{C}^{1,2}$.

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

$$f_\alpha(t,x) := \begin{cases} f(t,x) & \text{for} \quad l(\alpha) = 0, \\ L^{(j_1)}f_{-\alpha}(t,x) & \text{for} \quad l(\alpha) \geq 1, \end{cases} \tag{6.7.3}$$

assuming that the coefficients of the SDE (2.1.2) have the differentiability needed for the operators in (6.7.3) to be well defined.

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

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

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

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

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

$$Y_{t_{n+1}-} = Y_{t_n} + \sum_{\alpha \in \widehat{\mathcal{A}}_\gamma \backslash \{v\}} f_\alpha(t_n, Y_{t_n}) I_\alpha \tag{6.7.4}$$

and

$$Y_{t_{n+1}} = Y_{t_{n+1}-} + \int_{\mathcal{E}} c(t_n, Y_{t_{n+1}-}, v)\, p_\phi(dv, \{t_{n+1}\}), \tag{6.7.5}$$

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

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

$$Y_t = \sum_{\alpha \in \widehat{\mathcal{A}}_\gamma \setminus \{v\}} I_\alpha[f_\alpha(t_{n_t}, Y_{t_{n_t}})]_{t_{n_t}, t} \tag{6.7.6}$$

since there are no jumps between grid points.

We can now formulate a convergence theorem for jump-adapted schemes similar to a result in Platen (1982a).

Theorem 6.7.1 *For a given $\gamma \in \{0.5, 1, 1.5, 2, \ldots\}$, let $Y^\Delta = \{Y_t^\Delta,\ t \in [0,T]\}$ be the jump-adapted order γ strong Taylor approximation corresponding to a jump-adapted time discretization with maximum step size $\Delta \in (0,1)$. We assume that*

$$E(|X_0|^2) < \infty \quad \text{and} \quad E(|X_0 - Y_0^\Delta|^2) \le C\,\Delta^{2\gamma}. \tag{6.7.7}$$

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

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

$$|f_\alpha(t,x) - f_\alpha(t,y)| \le K_1\,|x-y|. \tag{6.7.8}$$

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

$$f_{-\alpha} \in C^{1,2} \qquad \text{and} \qquad f_\alpha \in \mathcal{H}_\alpha, \tag{6.7.9}$$

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

$$|f_\alpha(t,x)|^2 \le K_2\,(1 + |x|^2). \tag{6.7.10}$$

Then the estimate

$$\sqrt{E(\sup_{0 \le s \le T} |X_s - Y_s^\Delta|^2\,|\mathcal{A}_0)} \le K_3\,\Delta^\gamma \tag{6.7.11}$$

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

Remark 6.7.2 *Instead of conditions* (6.7.8)–(6.7.10) *on the coefficients f_α, one can derive analogous conditions on the coefficients a, b and c of the SDE* (3.5.1), *see also Remark 4.5.2.*

Proof: Since the jump-adapted time discretization contains all jump times of the solution X of the SDE (2.1.2), with the aid of the Wagner-Platen expansion for diffusion processes we can write

$$\begin{aligned} X_t &= X_0 + \sum_{\alpha\in\widehat{\mathcal{A}}_\gamma\setminus\{v\}} \left\{ \sum_{n=0}^{n_t-1} I_\alpha[f_\alpha(t_n, X_{t_n})]_{t_n,t_{n+1}} + I_\alpha[f_\alpha(t_{n_t}, X_{t_{n_t}})]_{t_{n_t},t} \right\} \\ &+ \sum_{\alpha\in\widehat{\mathcal{B}}(\widehat{\mathcal{A}}_\gamma)} \left\{ \sum_{n=0}^{n_t-1} I_\alpha[f_\alpha(\cdot, X_\cdot)]_{t_n,t_{n+1}} + I_\alpha[f_\alpha(\cdot, X_\cdot)]_{t_{n_t},t} \right\} \\ &+ \int_0^t \int_{\mathcal{E}} c(t_{n_z}, X_{t_{n_z}-}, v)\, p_\phi(dv, dz), \end{aligned} \tag{6.7.12}$$

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

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

$$\begin{aligned} Y_t &= Y_0 + \sum_{\alpha\in\widehat{\mathcal{A}}_\gamma\setminus\{v\}} \left\{ \sum_{n=0}^{n_t-1} I_\alpha[f_\alpha(t_n, Y_{t_n})]_{t_n,t_{n+1}} + I_\alpha[f_\alpha(t_{n_t}, Y_{t_{n_t}})]_{t_{n_t},t} \right\} \\ &+ \int_0^t \int_{\mathcal{E}} c(t_{n_z}, Y_{t_{n_z}-}, v)\, p_\phi(dv, dz) \end{aligned} \tag{6.7.13}$$

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

From the estimate of Theorem 2.2.1 we have

$$E\Big(\sup_{0\le z\le T} |X_z|^2 \Big| \mathcal{A}_0\Big) \le C\Big(1 + E\left(|X_0|^2\right)\Big). \tag{6.7.14}$$

Moreover, with similar steps as those used in the proof of Lemma 4.7.1, we can show the estimate

$$E\Big(\sup_{0\le z\le T} |Y_z^\Delta|^2 \Big| \mathcal{A}_0\Big) \le C\Big(1 + E\left(|Y_0^\Delta|^2\right)\Big). \tag{6.7.15}$$

The mean square error is given by

$$
\begin{aligned}
Z(t) &:= E\left(\sup_{0\le z\le t}|X_z - Y_z^{\Delta}|^2 \Big| \mathcal{A}_0\right)\\
&= E\Bigg(\sup_{0\le z\le t}\Bigg|X_0 - Y_0^{\Delta}\\
&\quad + \sum_{\alpha\in\widehat{\mathcal{A}}_\gamma\setminus\{v\}}\Bigg\{\sum_{n=0}^{n_z-1} I_\alpha[f_\alpha(t_n, X_{t_n}) - f_\alpha(t_n, Y_{t_n}^{\Delta})]_{t_n,t_{n+1}}\\
&\quad + I_\alpha[f_\alpha(t_{n_z}, X_{t_{n_z}}) - f_\alpha(t_{n_z}, Y_{t_{n_z}}^{\Delta})]_{t_{n_z},z}\Bigg\}\\
&\quad + \sum_{\alpha\in\widehat{\mathcal{B}}(\widehat{\mathcal{A}}_\gamma)}\Bigg\{\sum_{n=0}^{n_z-1} I_\alpha[f_\alpha(\cdot, X_\cdot)]_{t_n,t_{n+1}} + I_\alpha[f_\alpha(\cdot, X_\cdot)]_{t_{n_z},z}\Bigg\}\\
&\quad + \int_0^z\int_{\mathcal{E}}\Big\{c(t_{n_u}, X_{t_{n_u}-}, v) - c(t_{n_u}, Y_{t_{n_u}-}, v)\Big\}\, p_\phi(dv, du)\Bigg|^2 \Bigg| \mathcal{A}_0\Bigg)\\
&\le C_3\left\{ \left|X_0 - Y_0^{\Delta}\right|^2 + \sum_{\alpha\in\widehat{\mathcal{A}}_\gamma\setminus\{v\}} S_t^\alpha + \sum_{\alpha\in\widehat{\mathcal{B}}(\widehat{\mathcal{A}}_\gamma)} U_t^\alpha + P_t\right\} \qquad (6.7.16)
\end{aligned}
$$

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

$$
\begin{aligned}
S_t^\alpha &:= E\Bigg(\sup_{0\le z\le t}\Bigg|\sum_{n=0}^{n_z-1} I_\alpha\left[f_\alpha(t_n, X_{t_n}) - f_\alpha(t_n, Y_{t_n}^{\Delta})\right]_{t_n,t_{n+1}}\\
&\quad + I_\alpha\left[f_\alpha(t_{n_z}, X_{t_{n_z}}) - f_\alpha(t_{n_z}, Y_{t_{n_z}}^{\Delta})\right]_{t_{n_z},z}\Bigg|^2 \Bigg| \mathcal{A}_0\Bigg), \qquad (6.7.17)
\end{aligned}
$$

$$
U_t^\alpha := E\left(\sup_{0\le z\le t}\left|\sum_{n=0}^{n_z-1} I_\alpha[f_\alpha(\cdot, X_\cdot)]_{t_n,t_{n+1}} + I_\alpha[f_\alpha(\cdot, X_\cdot)]_{t_{n_z},z}\right|^2 \Bigg| \mathcal{A}_0\right), \qquad (6.7.18)
$$

and

$$P_t \;:=\; E\left(\sup_{0\le z\le t}\left|\int_0^z\int_{\mathcal{E}}\left\{c(t_{n_u},X_{t_{n_u}-},v)-c(t_{n_u},Y_{t_{n_u}-},v)\right\}p_\phi(dv,du)\right|^2\middle|\mathcal{A}_0\right). \tag{6.7.19}$$

Therefore, the terms S_t^α and U_t^α can be estimated as in the proof of Theorem 4.5.1, while for P_t, by applying Jensen's and Doob's inequalities, Itô's isometry for jump processes, the Cauchy-Schwarz inequality and the Lipschitz condition (2.2.10), we obtain

$$\begin{aligned}
P_t &= E\Bigg(\sup_{0\le z\le t}\bigg|\int_0^z\int_{\mathcal{E}}\left\{c(t_{n_u},X_{t_{n_u}-},v)-c(t_{n_u},Y_{t_{n_u}-},v)\right\}\widetilde{p}_\phi(dv,du)\\
&\quad+\int_0^z\int_{\mathcal{E}}\left\{c(t_{n_u},X_{t_{n_u}-},v)-c(t_{n_u},Y_{t_{n_u}-},v)\right\}\phi(dv)du\bigg|^2\,\bigg|\mathcal{A}_0\Bigg)\\
&\le 8\,E\left(\left|\int_0^t\int_{\mathcal{E}}\left\{c(t_{n_u},X_{t_{n_u}-},v)-c(t_{n_u},Y_{t_{n_u}-},v)\right\}\widetilde{p}_\phi(dv,du)\right|^2\middle|\mathcal{A}_0\right)\\
&\quad+2\,E\left(\sup_{0\le z\le t}\left|\int_0^z\int_{\mathcal{E}}\left\{c(t_{n_u},X_{t_{n_u}-},v)-c(t_{n_u},Y_{t_{n_u}-},v)\right\}\phi(dv)du\right|^2\middle|\mathcal{A}_0\right)\\
&\le 8\,E\left(\int_0^t\int_{\mathcal{E}}\left|c(t_{n_u},X_{t_{n_u}-},v)-c(t_{n_u},Y_{t_{n_u}-},v)\right|^2\phi(dv)\,du\middle|\mathcal{A}_0\right)\\
&\quad+2\,\lambda\,t\,E\left(\int_0^t\int_{\mathcal{E}}\left|c(t_{n_u},X_{t_{n_u}-},v)-c(t_{n_u},Y_{t_{n_u}-},v)\right|^2\phi(dv)\,du\middle|\mathcal{A}_0\right)\\
&\le K\,E\left(\int_0^t\left|X_{t_{n_u}-}-Y_{t_{n_u}-}\right|^2du\middle|\mathcal{A}_0\right)\\
&\le C\int_0^t Z(u)du.
\end{aligned}\tag{6.7.20}$$

Therefore, since by (6.7.14) and (6.7.15) $Z(t)$ is bounded, by applying the Gronwall inequality to (6.7.16) we can complete the proof of Theorem 6.7.1. □

Theorem 6.7.1 establishes the order of strong convergence of the jump-adapted strong Taylor schemes presented in Section 6.2. To prove the order of strong convergence of the other jump-adapted schemes in this chapter, one can define the *jump-adapted order γ strong Itô scheme*, with $\gamma \in \{0.5, 1, 1.5, \ldots\}$, constructed by the following procedure: The diffusion part is approximated by an order γ strong Itô scheme for pure diffusions, see Kloeden & Platen (1999), and the jump part is generated as in (6.7.5).

One can prove the strong order of convergence of the jump-adapted order γ strong Itô scheme by first showing that this scheme converges, with strong order γ, to the jump-adapted order γ strong Taylor scheme. This can be done by using similar steps as those described in the proof of Theorem 5.5.1. Thus, since Theorem 6.7.1 establishes the strong order of convergence of jump-adapted strong Taylor schemes, this yields also the strong order γ of the corresponding jump-adapted strong Itô scheme. Finally, the strong order of the jump-adapted derivative-free, implicit and predictor-corrector schemes presented in this chapter can be shown by rewriting these schemes as jump-adapted strong Itô schemes.

Chapter 7

Numerical Results on Strong Schemes

This short chapter provides some numerical results for the application of the strong schemes presented in Chapters 4, 5 and 6. We investigate the accuracy of strong schemes, while a study of the numerical stability properties is left for future research.

7.1 Introduction

We study the strong approximation of the one-dimensional linear SDE

$$dX_t = X_{t-}\left(\mu dt + \sigma dW_t + \int_{\mathcal{E}}(v-1)\, p_\phi(dv,dt)\right) \tag{7.1.1}$$

for $t \in [0,T]$ and $X_0 > 0$, which is that of the Merton model introduced in (2.1.5). We recall that this SDE admits the explicit solution

$$X_t = X_0\, e^{(\mu-\frac{1}{2}\sigma^2)t+\sigma W_t} \prod_{i=1}^{p_\phi(t)} \xi_i, \tag{7.1.2}$$

where the marks ξ_i are distributed according to a given probability measure $F(dv) = \frac{\phi(dv)}{\lambda}$ and $p_\phi = \{p_\phi(t), t \in [0,T]\}$ denotes a Poisson process with intensity $\lambda = \phi(\mathcal{E}) < \infty$.

We consider the following schemes with strong order $\gamma = 0.5$: the regular and jump-adapted versions of the Euler, the drift-implicit Euler, and the predictor-corrector Euler schemes. Moreover, we study the following schemes with strong order $\gamma = 1.0$: the regular and jump-adapted order 1.0 Taylor schemes and the jump-adapted drift-implicit order 1.0 scheme. We also present the jump-adapted order 1.5 Taylor scheme, which attains strong order $\gamma = 1.5$. We report the strong error

$$\varepsilon_s(\Delta) = \sqrt{E(|X_T - Y_T^\Delta|^2)}, \tag{7.1.3}$$

as defined (1.2.1), when comparing the results of these strong schemes with the closed form solution (7.1.2). The strong error $\varepsilon_s(\Delta)$ is estimated by running an extremely large number of simulations. The exact number depends on the scheme implemented. It will always be chosen such that the statistical errors become negligible when compared to the systematic errors caused by the time discretization. In the corresponding plots we show the logarithm $\log_2(\varepsilon_s(\Delta))$ of the strong error versus the logarithm $\log_2(\Delta)$ of the time step size. By using log-log plots, the slopes of the estimated error lines will indicate the orders of strong convergence attained. We will first consider the case of a driving jump process with a small intensity $\lambda = 0.05$. Later, to illustrate the impact of frequent jumps on the strong error, we will use a jump process with a higher intensity.

7.2 The Case of Low Intensities

In this section we select the following default parameters: $\mu = 0.05$, $\sigma = 0.15$, $X_0 = 1$, $T = 1$ and $\lambda = 0.05$. At first we consider the case of the SDE (7.1.1) with degenerate marks, that is $\xi_i = \psi > 0$, with $\psi = 0.85$. This reduces the SDE (7.1.1) to an SDE with mark-independent jump coefficient $c(t, x, v) = x(\psi - 1) = -0.15x$.

In Figure 7.2.1, we report the results obtained from the regular and jump-adapted Euler schemes, the regular and jump-adapted drift-implicit Euler schemes, and the regular predictor-corrector Euler scheme. We do not report in Figure 7.2.1 the results of the jump-adapted predictor-corrector scheme because its accuracy is indistinguishable from that of the jump-adapted Euler scheme. Note that, here and in the rest of this chapter, the implicitness parameters are set to $\theta = \eta = 0.5$. All schemes achieve an order of strong convergence of about $\gamma = 0.5$. This is consistent with the strong orders proved in the previous chapters. Moreover, all schemes except the regular predictor-corrector Euler scheme have very similar accuracy. The regular predictor-corrector Euler scheme is significantly more accurate than the other schemes for all time step sizes considered.

Let us now analyze the strong errors generated by the regular and jump-adapted order 1.0 Taylor schemes, the jump-adapted drift-implicit order 1.0 scheme, and the jump-adapted order 1.5 Taylor scheme. In Figure 7.2.2 we report the results for these schemes, along with those of the regular predictor-corrector Euler scheme, already plotted in Figure 7.2.1. We omit from Figure 7.2.2 the result of the regular order 1.0 Taylor scheme, since its accuracy is almost indistinguishable from that of its jump-adapted version. We notice that the jump-adapted order 1.0 Taylor scheme and the jump-adapted drift-implicit order 1.0 scheme achieve strong order one in accordance with the convergence theorems proved in Chapters 4, 5 and 6. The jump-adapted drift-implicit order 1.0 scheme is more accurate. In this plot we can notice that the accuracy of the regular predictor-corrector Euler scheme, for the selected time step sizes, is similar to that of first order schemes. Of course, since its strong order of convergence equals $\gamma = 0.5$, for smaller time step sizes it becomes less accurate than first order schemes. Finally, the most accurate scheme for all time step sizes considered is the jump-adapted order 1.5 Taylor scheme, which achieves an order of strong convergence of about $\gamma = 1.5$.

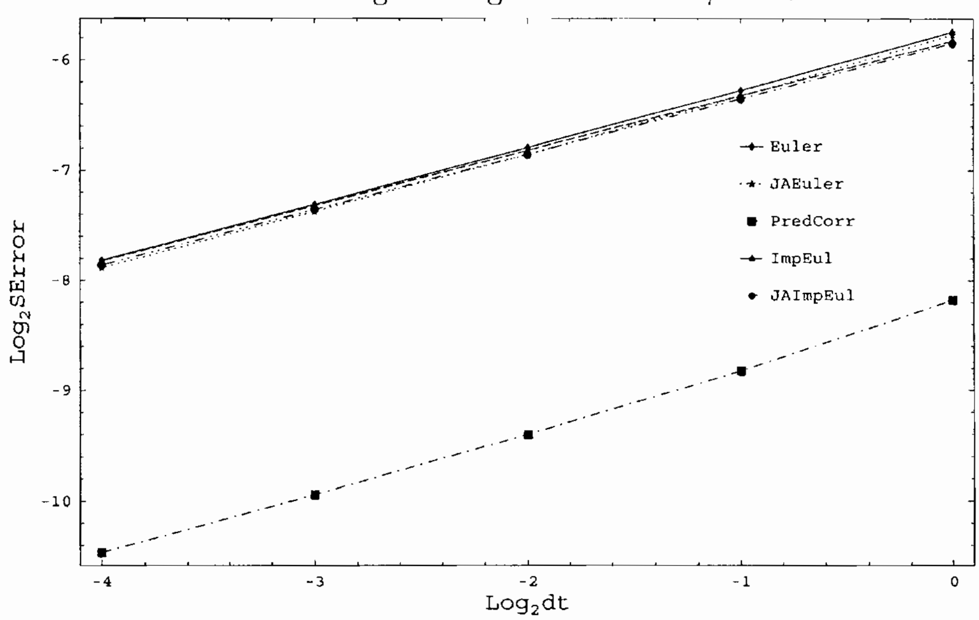

Figure 7.2.1: Log-log plot of strong error versus time step size (constant marks)

Let us now consider the case of lognormally distributed marks. Here the logarithm of mark $\zeta_i = \ln(\xi_i)$ is an independent Gaussian random variable, $\zeta_i \sim \mathcal{N}(\varrho, \varsigma)$, with mean $\varrho = -0.1738$ and standard deviation $\sqrt{\varsigma} = 0.15$. These parameters imply that the expected value of the marks equals $E(\xi) = 0.85$.

In Figure 7.2.3, we plot the results obtained from the regular and jump-adapted Euler schemes, the regular and jump-adapted drift-implicit Euler schemes, and the regular predictor-corrector Euler scheme. Also in this case the results for the jump-adapted predictor-corrector Euler scheme are indistinguishable from those of the jump-adapted Euler scheme and, thus, we omit them. We remark that these results are very similar to those obtained in Figure 7.2.1 for the case of constant marks,

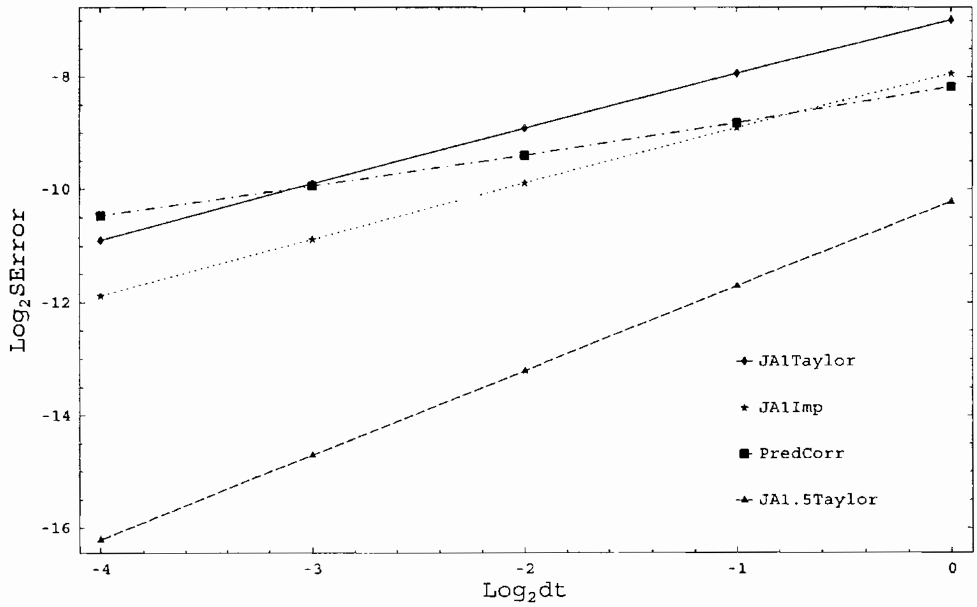

Figure 7.2.2: Log-log plot of strong error versus time step size (constant marks)

thus confirming that the orders of convergence derived in the previous chapters hold also in the case of random marks. Again all schemes considered achieve a strong order of about $\gamma = 0.5$. Moreover, the regular predictor-corrector Euler scheme is the most accurate. The remaining schemes have similar accuracy, with the regular Euler scheme the least accurate and the jump-adapted drift-implicit Euler scheme the most accurate. In Figure 7.2.4 we report the results for the regular predictor-corrector Euler, the jump-adapted order 1.0 Taylor, the jump-adapted drift-implicit order 1.0, and the jump-adapted order 1.5 Taylor schemes. The results are again very similar to those obtained for the case of constant marks, reported in Figure 7.2.2, with all schemes achieving the prescribed orders of strong convergence.

7.3 The Case of High Intensities

Let us now consider the strong errors generated by the strong schemes analyzed in the previous section, when using the relative large intensity $\lambda = 2$. The remaining parameters of the SDE (7.1.1) are set as in the previous section.

In Figure 7.3.5 we show the results for the regular and jump-adapted Euler scheme, the regular and jump-adapted predictor-corrector Euler schemes and the jump-adapted drift-implicit Euler schemes. Note that the error generated by the regular

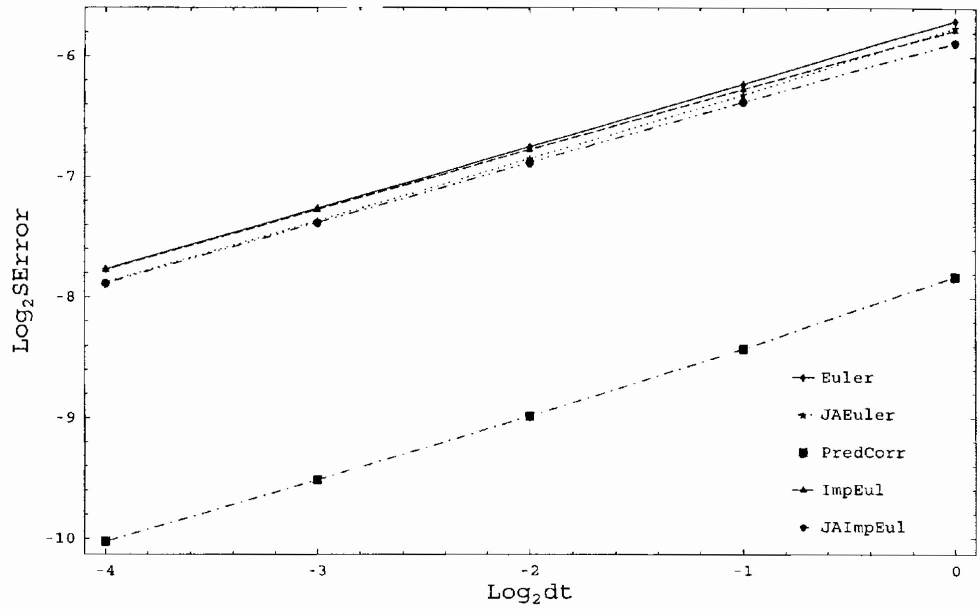

Figure 7.2.3: Log-log plot of strong error versus time step size (lognormal marks)

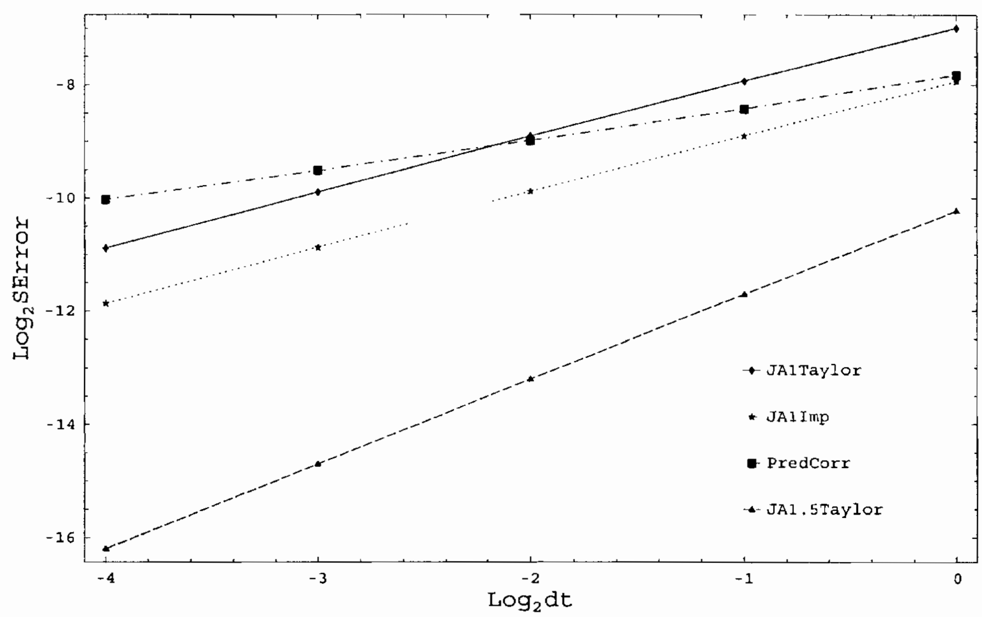

Figure 7.2.4: Log-log plot of strong error versus time step size (lognormal marks)

drift-implicit scheme is very similar to that of the regular Euler and, thus, is here omitted. All schemes achieve orders of strong convergence of about $\gamma = 0.5$. Here we can clearly notice that jump-adapted schemes are more accurate than regular schemes. This is due to the simulation of the jump impact at the correct jump times in jump-adapted schemes. Moreover, we report that drift-implicit schemes are the least accurate, while predictor-corrector schemes are the most accurate schemes. In Figure 7.3.6 we show the results for the jump-adapted predictor-corrector Euler scheme, the jump-adapted order 1.0 Taylor scheme, the jump-adapted drift-implicit order 1.0 scheme, and the jump-adapted order 1.5 Taylor scheme. Also here all schemes achieve the orders of strong convergence expected from the theory. Note that while in the low intensity case the accuracy of the regular and jump-adapted versions of the order 1.0 Taylor scheme were very similar, in the high intensity case the jump-adapted version is more accurate. We point out that the jump-adapted order 1.5 Taylor scheme is the most accurate for all time step sizes considered.

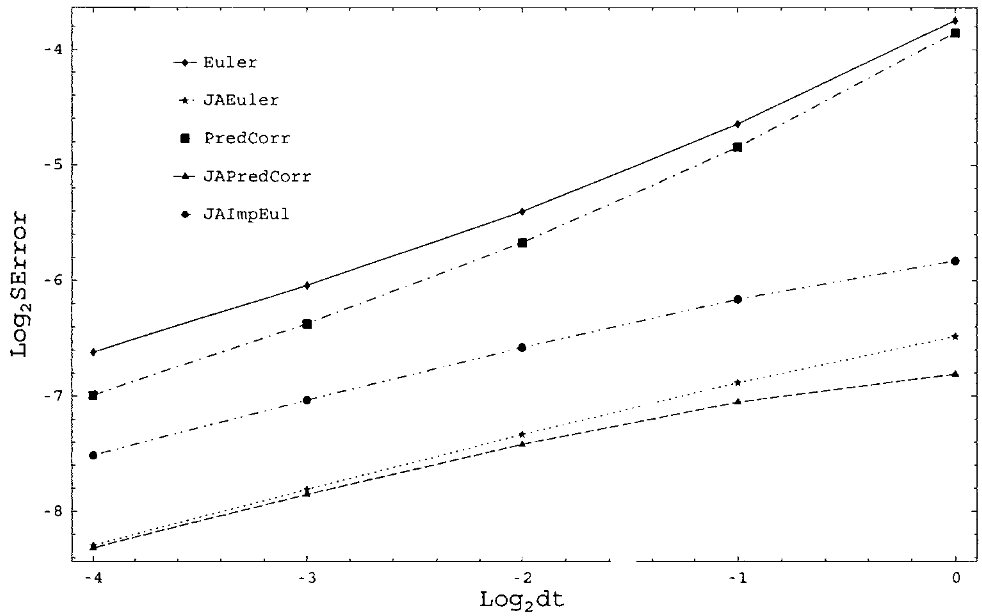

Figure 7.3.5: Log-log plot of strong error versus time step size (constant marks)

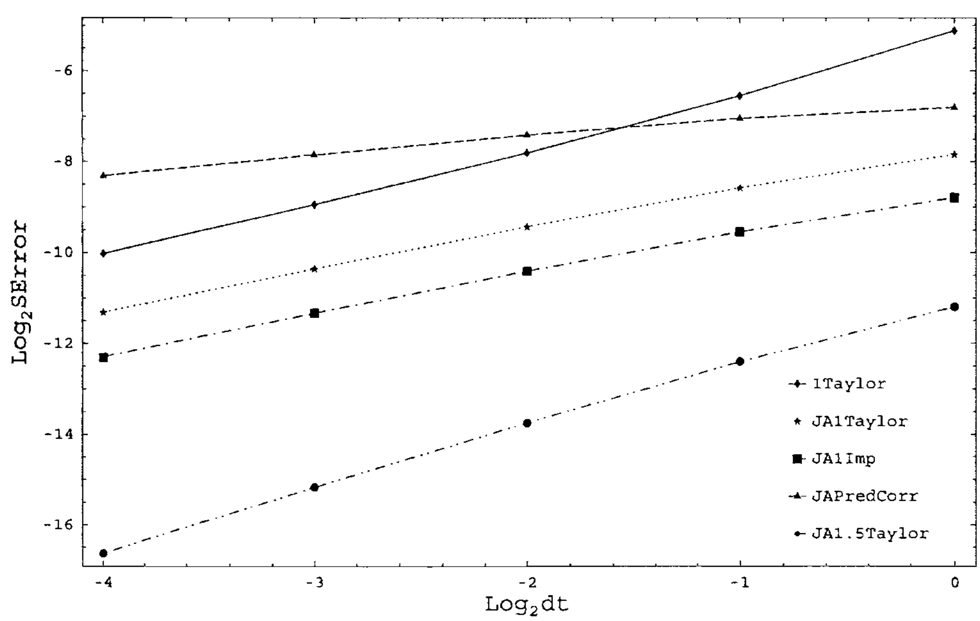

Figure 7.3.6: Log-log plot of strong error versus time step size (constant marks)

Finally, in Figures 7.3.7 and 7.3.8 we report the strong errors for all schemes analyzed in this chapters in the case of lognormal marks. The results are again very similar to those obtained in the case of constant marks. In particular, we report that all schemes achieve the orders of strong convergence proved in Chapters 4, 5 and 6.

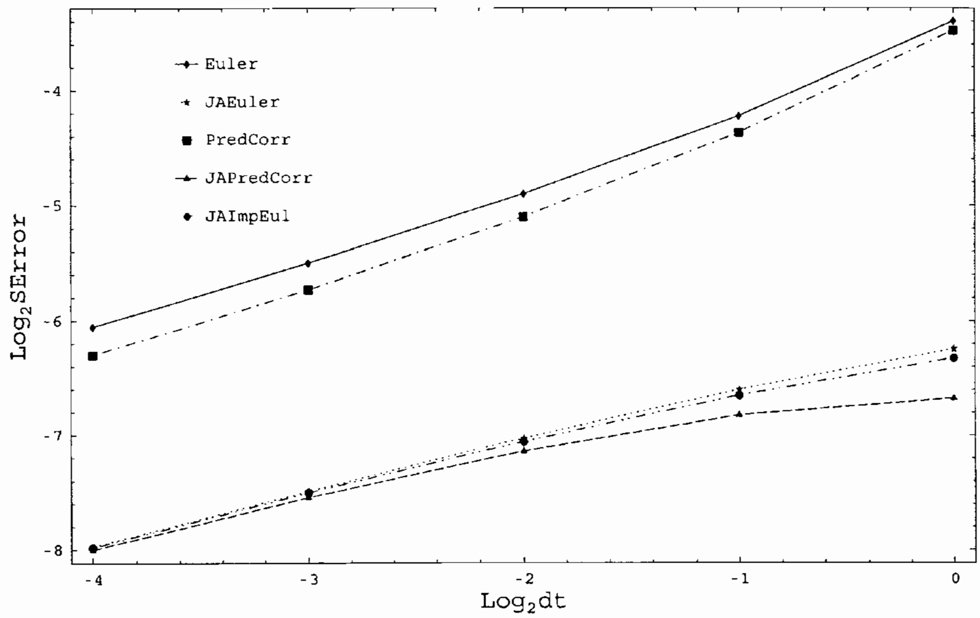

Figure 7.3.7: Log-log plot of strong error versus time step size (lognormal marks)

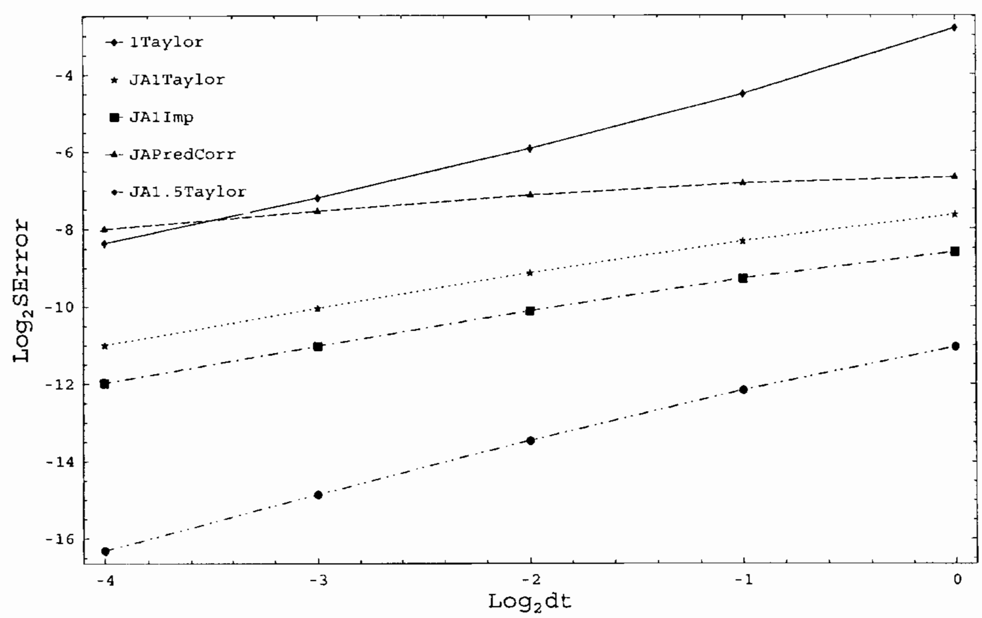

Figure 7.3.8: Log-log plot of strong error versus time step size (lognormal marks)

Chapter 8

Strong Approximation of Pure Jump Processes

In this chapter we consider strong discrete time approximations of pure jump SDEs. The schemes to be presented are special cases of those considered in Chapters 4, 5 and 6 when the drift and the diffusion coefficients equal zero. The particular nature of the pure jump dynamics simplifies the implementation of corresponding higher order strong schemes. Additionally, as we will see at the end of this chapter, strong orders of convergence are derived under weaker assumptions than those needed in the jump-diffusion case. Most results of this chapter have been published in Bruti-Liberati & Platen (2007c).

8.1 Introduction

We now present strong numerical approximations of pure jump SDEs. Such SDEs arise, for instance, when using a birth and death process or, more generally, a continuous time Markov chain. They play an important role in modelling credit rating changes, bio-chemistry reactions and other areas of applications, see Turner, Schnell & Burrage (2004). The piecewise constant nature of pure jump dynamics simplifies the resulting numerical schemes. For instance, jump-adapted approximations, constructed on time discretizations including all jump times, produce no discretization error in this case. Therefore, in the case of low to medium jump intensities one can construct efficient schemes without discretization error. In the case of high intensity jump processes, jump-adapted schemes are often not feasible. However, we will demonstrate that one can derive higher order discrete time approximations whose complexities turn out to be significantly lower than that of numerical approximations of jump diffusions. In the case of SDEs driven by a Poisson process, the generation of the multiple stochastic integrals required for higher order approximations is straightforward, since it involves only one Poisson distributed random variable in each time step. Moreover, the simple structure of pure jump SDEs permits us to illustrate the use of a stochastic expansion in the derivation of higher order approximations. At the end of the current chapter, we show that higher strong orders of convergence of discrete time approximations for SDEs driven purely by a Poisson random measure can be derived under weaker conditions than those typically required for jump diffusions.

8.2 Pure Jump Model

Let us consider a counting process $N = \{N_t,\ t \in [0,T]\}$, which is right-continuous with left-hand limits and counts the arrival of certain events. Most of the following analysis applies for rather general counting processes. However, for simplicity, we take N to be a *Poisson process* with constant intensity $\lambda \in (0,\infty)$ that starts at time $t = 0$ in $N_0 = 0$. It is defined on a filtered probability space $(\Omega, \mathcal{A}_T, \underline{\mathcal{A}}, P)$ with $\underline{\mathcal{A}} = (\mathcal{A}_t)_{t\in[0,T]}$ satisfying the usual conditions. Note that we can rewrite the Poisson process in terms of the Poisson random measure introduced in Chapter 2 by setting

$$N_t = p_\phi(\mathcal{E}, [0,t]), \tag{8.2.1}$$

for $t \in [0,T]$. An alternative representation of the Poisson process is provided by a Poisson random measure with mark space $\mathcal{E} = \{1\}$ and intensity measure $\phi(\{1\}) = \lambda$. The Poisson process $N = \{N_t, t \in [0,T]\}$ generates an increasing sequence $(\tau_i)_{i\in\{1,2,\dots,N_T\}}$ of jump times. For any right-continuous process $Z = \{Z_t,\ t \in [0,T]\}$ we define its *jump size* $\Delta\, Z_t$ at time t as the difference

$$\Delta\, Z_t = Z_t - Z_{t-}, \tag{8.2.2}$$

for $t \in [0,T]$, where Z_{t-} denotes again the left-hand limit of Z at time t. Thus, we can write

$$N_t = \sum_{s\in(0,t]} \Delta\, N_s$$

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

For a pure jump process $X = \{X_t,\ t \in [0,T]\}$ that is driven by the Poisson process N we assume that its value X_t at time t satisfies the SDE

$$dX_t = c(t, X_{t-})\, dN_t \tag{8.2.3}$$

for $t \in [0, T]$ with deterministic initial value $X_0 \in \mathbb{R}$. This is a special case of the SDE (2.1.2), where the drift coefficient a and the diffusion coefficient b both equal zero and the jump coefficient c is mark-independent.

The *jump coefficient* $c : [0, T] \times \mathbb{R} \to \mathbb{R}$ is again assumed to be Borel measurable, Lipschitz continuous, such that

$$|c(t, x) - c(t, y)| \leq K\,|x - y|,$$

and to satisfy the growth condition

$$|c(t, x)|^2 \leq K\,(1 + |x|^2)$$

for $t \in [0, T]$ and $x, y \in \mathbb{R}$ with some constant $K \in (0, \infty)$.

To provide for later illustration a simple, still interesting example, let us consider the linear SDE

$$dX_t = X_{t-}\,\psi\,dN_t \tag{8.2.4}$$

for $t \in [0, T]$ with $X_0 > 0$ and constant $\psi \in \mathbb{R}$. This is a degenerate case of the SDE (2.1.5), with drift coefficient $a(t, x) = 0$, diffusion coefficient $b(t, x) = 0$ and mark-independent jump coefficient $c(t, x) = x\,\psi$. By application of the Itô formula one can demonstrate that the solution $X = \{X_t,\, t \in [0, T]\}$ of the SDE (8.2.4) is a pure jump process with explicit representation

$$X_t = X_0\, \exp\{N_t\, \ln(\psi + 1)\} = X_0\,(\psi + 1)^{N_t} \tag{8.2.5}$$

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

8.3 Jump-Adapted Schemes

We consider a *jump-adapted time discretization* $0 = t_0 < t_1 < \ldots < t_{n_T} = T$, where n_T is defined in (4.1.8) and the sequence $t_1 < \ldots < t_{n_T - 1}$ equals that of the jump times $\tau_1 < \ldots < \tau_{N_T}$ of the Poisson process N. On this jump-adapted time grid we construct the *jump-adapted Euler scheme* by the algorithm

$$Y_{n+1} = Y_n + c\,\Delta N_n, \tag{8.3.6}$$

for $n \in \{0, 1, \ldots, n_T - 1\}$, with initial value $Y_0 = X_0$, where $\Delta N_n = N_{t_{n+1}} - N_{t_n}$ is the nth increment of the Poisson process N. Between discretization times the right-continuous process Y is set to be piecewise constant. Note that here and in the sequel, when no misunderstanding is possible, we use the previously introduced abbreviation $c = c(t_n, Y_n)$.

Since the discretization points are constructed exactly at the jump times of N, and the simulation of the increments $N_{t_{i+1}} - N_{t_i} = 1$ of N can be made exact, the jump-adapted Euler scheme (8.3.6) produces no discretization error. Let us emphasize that this is a particular feature of jump-adapted schemes when applied to pure jump SDEs. In the case of jump-diffusion SDEs, the jump-adapted schemes in Chapter 6 typically produce a discretization error.

For the implementation of the scheme (8.3.6) one needs to compute the jump times τ_i, $i \in \{1, 2 \ldots, N_T\}$, and has then to apply equation (8.3.6) recursively for every $i \in \{0, 1, 2 \ldots, n_T - 1\}$. One can obtain the jump times via the corresponding waiting times between two consecutive jumps by sampling from an exponential distribution with parameter λ.

The computational effort when running the algorithm (8.3.6) is heavily dependent on the intensity λ of the jump process. Indeed, the average number of steps and, thus, of operations is proportional to the intensity λ. Below we will introduce alternative methods suitable for large intensities, based on regular time discretizations.

8.4 Euler Scheme

In this section we develop discrete time strong approximations whose computational complexity is independent of the jump intensity level.

We consider an equidistant time discretization with time step size $\Delta \in (0, 1)$ as in Chapter 4. The simplest strong Taylor approximation $Y = \{Y_t, t \in [0, T]\}$ is the *Euler scheme*, which is given by

$$Y_{n+1} = Y_n + c\,\Delta N_n \tag{8.4.7}$$

for $n \in \{0, 1, \ldots, n_T - 1\}$ with initial value $Y_0 = X_0$ and $\Delta N_n = N_{t_{n+1}} - N_{t_n}$. Between discretization times the right-continuous process Y is assumed to be piecewise constant.

By comparing the scheme (8.4.7) with the algorithm (8.3.6), we notice that the difference in the schemes consists in the time discretization. We emphasize that the average number of operations and, thus, the computational complexity of the Euler scheme (8.4.7) is independent of the jump intensity. Therefore, a simulation based on the Euler scheme (8.4.7) is feasible also in the case of jump processes with high intensity. However, while the jump-adapted Euler scheme (8.3.6) produces no discretization error, the accuracy of the Euler scheme (8.4.7) depends on the size of the time step Δ and the nature of the jump coefficient.

For example, for the linear SDE (8.2.4) the Euler scheme (8.4.7) has the form

$$Y_{n+1} = Y_n + Y_n \, \psi \, \Delta N_n = Y_n \, (1 + \psi \Delta N_n) \tag{8.4.8}$$

for $n \in \{0, 1, \ldots, n_T - 1\}$ with $Y_0 = X_0$. Since the equidistant time discretization $t_n = \Delta n$ of this Euler scheme does not include the jump times of the underlying Poisson process, we have an approximation error.

Theorem 4.5.1 shows that the Euler approximation (8.4.7) achieves strong order of convergence $\gamma = 0.5$. This raises the question of constructing higher order discrete time approximations for the case of pure jump SDEs. The problem can be approached by a stochastic expansion for pure jump SDEs that we will describe below. This expansion is a particular case of the Wagner-Platen expansion (3.5.4) for jump diffusions presented in Chapter 3. Therefore, the resulting strong approximations are particular cases of the strong schemes presented in Chapter 4.

8.5 Wagner-Platen Expansion

Since the use of the Wagner-Platen expansion for pure jump processes is not common in the literature let us first illustrate the structure of this formula for a simple example. For any measurable function $f : \mathbb{R} \to \mathbb{R}$ and a given adapted counting process $N = \{N_t, \, t \in [0, T]\}$ we have the representation

$$f(N_t) = f(N_0) + \sum_{s \in (0,t]} \Delta \, f(N_s) \tag{8.5.9}$$

for all $t \in [0,T]$, where $\Delta f(N_s) = f(N_s) - f(N_{s-})$. We can formally write the equation (8.5.9) in the form of an SDE

$$df(N_t) = \Delta f(N_t) = (f(N_{t-}+1) - f(N_{t-}))\,\Delta N_t$$

for $t \in [0,T]$. This equation can also be obtained from the Itô formula for semimartingales, see Protter (2004), for the case with jumps.

Obviously, the following difference expression $\widetilde{\Delta}_N f(N_{s-})$ defines a measurable function, as long as

$$\widetilde{\Delta}_N\, f(N) = f(N+1) - f(N) \tag{8.5.10}$$

is a measurable function of N. By using this function we can rewrite (8.5.9) in the form

$$f(N_t) = f(N_0) + \int_{(0,t]} \widetilde{\Delta}_N\, f(N_{s-})\, dN_s \tag{8.5.11}$$

for $t \in [0,T]$. Since $\widetilde{\Delta}_N f(N_{s-})$ is a measurable function we can apply the formula (8.5.11) to $\widetilde{\Delta}_N f(N_{s-})$ in (8.5.11), which yields

$$\begin{aligned} f(N_t) &= f(N_0) + \int_{(0,t]} \widetilde{\Delta}_N f(N_0) dN_s + \int_{(0,t]}\int_{(0,s_2)} \left(\widetilde{\Delta}_N\right)^2 f(N_{s_1-}) dN_{s_1} dN_{s_2} \\ &= f(N_0) + \widetilde{\Delta}_N f(N_0) \int_{(0,t]} dN_s + \int_{(0,t]}\int_{(0,s_2)} \left(\widetilde{\Delta}_N\right)^2 f(N_{s_1-}) dN_{s_1} dN_{s_2} \end{aligned} \tag{8.5.12}$$

for $t \in [0,T]$. Here $\left(\widetilde{\Delta}_N\right)^q$ denotes for an integer $q \in \{1,2,\ldots\}$ the q times consecutive application of the function $\widetilde{\Delta}_N$ given in (8.5.10). Note that a double stochastic integral with respect to the counting process N naturally arises in (8.5.12). One can now continue in (8.5.12) to apply the formula (8.5.11) to the measurable function $(\widetilde{\Delta}_N)^2 f(N_{s_1-})$, which yields

$$f(N_t) = f(N_0) + \widetilde{\Delta}_N f(N_0) \int_{(0,t]} dN_s + \left(\widetilde{\Delta}_N\right)^2 f(N_0) \int_{(0,t]}\int_{(0,s_2)} dN_{s_1} dN_{s_2} + \bar{R}_3(t) \tag{8.5.13}$$

with remainder term

$$\bar{R}_3(t) = \int_{(0,t]}\int_{(0,s_3)}\int_{(0,s_2)} \left(\widetilde{\Delta}_N\right)^3 f(N_{s_1-})\, dN_{s_1}\, dN_{s_2}\, dN_{s_3}$$

for $t \in [0,T]$. In (8.5.13) we have obtained a double integral in the expansion part.

Furthermore, we have a triple integral in the remainder term. We call (8.5.13) a Wagner-Platen expansion of the function $f(\cdot)$ with respect to the counting process N. Its expansion part only depends on multiple stochastic integrals with respect to the counting process N. These are weighted by some constant coefficient functions with values taken at the expansion point N_0. It is clear how to proceed to obtain higher order Wagner-Platen expansions by iterative application of formula (8.5.11).

Fortunately, the multiple stochastic integrals that arise can be easily computed. It is straightforward to prove by induction, see Engel (1982), that

$$\int_{(0,t]} dN_s = N_t$$

$$\int_{(0,t]} \int_{(0,s_2)} dN_{s_1}\, dN_{s_2} = \frac{1}{2\,!}\, N_t\,(N_t - 1),$$

$$\int_{(0,t]} \int_{(0,s_3)} \int_{(0,s_2)} dN_{s_1}\, dN_{s_2}\, dN_{s_3} = \frac{1}{3\,!}\, N_t\,(N_t - 1)\,(N_t - 2),$$

$$\int_{(0,t]} \int_{(0,s_l)} \cdots \int_{(0,s_2)} dN_{s_1} \ldots dN_{s_{l-1}}\, dN_{s_l} = \begin{cases} \binom{N_t}{l} & \text{for} \quad N_t \geq l \\ 0 & \text{otherwise} \end{cases} \tag{8.5.14}$$

for $t \in [0,T]$ and $l \in \{1,2\ldots\}$. Here we have used the common combinatorial abbreviation

$$\binom{i}{l} = \frac{i!}{l!(i-l)!} \tag{8.5.15}$$

for $i \geq l$ with $0\,! = 1$.

With (8.5.14) we can rewrite the Wagner-Platen expansion (8.5.13) in the form

$$f(N_t) \;=\; f(N_0) + \widetilde{\Delta}_N\, f(N_0) \binom{N_t}{1} + \left(\widetilde{\Delta}_N\right)^2 f(N_0) \binom{N_t}{2} + \bar{R}_3(t),$$

where

$$\widetilde{\Delta}_N\, f(N_0) = \widetilde{\Delta}_N\, f(0) = f(1) - f(0),$$

$$\left(\widetilde{\Delta}_N\right)^2 f(N_0) = f(2) - 2\,f(1) + f(0).$$

In the given case this leads to the expansion

$$f(N_t) = f(0) + (f(1) - f(0))\, N_t + (f(2) - 2\,f(1) + f(0))\, \frac{1}{2}\, N_t\,(N_t - 1) + \bar{R}_3(t)$$

for $t \in [0,T]$. More generally, by induction it follows the Wagner-Platen expansion

$$f(N_t) = \sum_{k=0}^{l} (\widetilde{\Delta}_N)^k f(N_0) \binom{N_t}{k} + \bar{R}_{l+1}(t) \tag{8.5.16}$$

with

$$\bar{R}_{l+1}(t) = \int_{(0,t]} \cdots \int_{(0,s_2)} (\widetilde{\Delta}_N)^{l+1} f(N_{s_1-})\, dN_{s_1} \ldots dN_{s_{l+1}}$$

for $t \in [0,T]$ and $l \in \{0,1,\ldots\}$, where $(\widetilde{\Delta}_N)^0 f(N_0) = f(N_0)$. By neglecting the remainder term in (8.5.16) one does not consider the occurrence of a higher number of jumps and obtains a useful truncated Taylor approximation of a measurable function f with respect to a counting process N. Note that in (8.5.16) the truncated expansion is exact if no more than l jumps occur until time t in the realization of N. Consequently, if there is a small probability that more than l jumps occur over the given time period, then the truncated Wagner-Platen expansion can be expected to be quite accurate under any reasonable criterion.

Similar to (8.5.16) let us now derive a Wagner-Platen expansion for functions of solutions of the general pure jump SDE (8.2.3). We define similarly as above the measurable function $\widetilde{\Delta}_N f(\cdot)$ such that

$$\widetilde{\Delta}_N f(X_{t-}) = \Delta f(X_t) = f(X_t) - f(X_{t-}) \tag{8.5.17}$$

for all $t \in [0,T]$. In the same manner as previously shown, this leads to the expansion

$$\begin{aligned} f(X_t) &= f(X_0) + \int_{(0,t]} \widetilde{\Delta}_N f(X_{s-})\, dN_s \\ &= f(X_0) + \int_{(0,t]} \left(\widetilde{\Delta}_N f(X_0) + \int_{(0,s_2)} \left(\widetilde{\Delta}_N\right)^2 f(X_{s_1-})\, dN_{s_1} \right) dN_{s_2} \\ &= f(X_0) + \sum_{k=1}^{l} \left(\widetilde{\Delta}_N\right)^k f(X_0) \int_{(0,t]} \cdots \int_{(0,s_2)} dN_{s_1} \cdots dN_{s_k} + \tilde{R}_{f,t}^{l+1} \\ &= f(X_0) + \sum_{k=1}^{l} \left(\widetilde{\Delta}_N\right)^k f(X_0) \binom{N_t}{k} + \tilde{R}_{f,t}^{l+1} \end{aligned} \tag{8.5.18}$$

with

$$\tilde{R}_{f,t}^{l+1} = \int_{(0,t]} \cdots \int_{(0,s_2)} \left(\widetilde{\Delta}_N\right)^{l+1} f(X_{s_1-})\, dN_{s_1} \cdots dN_{s_{l+1}}$$

for $t \in [0,T]$ and $l \in \{1,2,\ldots\}$. One notes that (8.5.18) generalizes (8.5.16) in a simple fashion.

Let us give an illustration. For the particular example of the linear SDE (8.2.4) we obtain for any measurable function f the function

$$\tilde{\Delta}_N f(X_{\tau-}) = f(X_{\tau-}(1+\psi)) - f(X_{\tau-})$$

for the jump times $\tau \in [0,T]$ with $\Delta N_\tau = 1$. Therefore, in the case $l = 2$, we get from (8.5.18) and (8.5.14) the expression

$$\begin{aligned} f(X_t) &= f(X_0) + \big(f(X_0(1+\psi)) - f(X_0)\big)(N_t - N_0) \\ &\quad + \Big(f(X_0(1+\psi)^2) - 2 f(X_0(1+\psi)) + f(X_0)\Big) \\ &\quad \times \frac{1}{2}(N_t - N_0)((N_t - N_0) - 1) + \tilde{R}^3_{f,t} \end{aligned}$$

for $t \in [0,T]$. By neglecting the remainder term $\tilde{R}^3_{f,t}$ we obtain, for this simple example, a truncated Wagner-Platen expansion of $f(X_t)$ at X_0. Let us emphasize that in the derivation of the expansion (8.5.18) only measurability of the function f and the coefficients $\left(\tilde{\Delta}_N\right)^k f(\cdot)$, for $k \in \{1,\ldots,l\}$ is required. This contrasts with the case of diffusion and jump-diffusion SDEs where differentiability conditions are needed to obtain a Wagner-Platen expansion.

8.6 Order 1.0 Strong Taylor Scheme

The Euler scheme (8.4.7) can be interpreted as being derived from the expansion (8.5.18) applied to each time step by setting $f(x) = x$, choosing $l = 1$ and neglecting the remainder term. By choosing $l = 2$ in the corresponding truncated Wagner-Platen expansion, when applied to each time discretization interval $[t_n, t_{n+1}]$ with $f(x) = x$, we obtain the *order* 1.0 *strong Taylor approximation*

$$Y_{n+1} = Y_n + c\,\Delta N_n + \big(c(t_n, Y_n + c) - c\big)\frac{1}{2}(\Delta N_n)(\Delta N_n - 1)$$

for $n \in \{0,1,\ldots,n_T - 1\}$ with $Y_0 = X_0$ and $\Delta N_n = N_{t_{n+1}} - N_{t_n}$.

In the special case of our linear example (8.2.4), the order 1.0 strong Taylor approximation turns out to be of the form

$$Y_{n+1} = Y_n \left\{ 1 + \psi \, \Delta N_n + \frac{\psi^2}{2} \, \Delta N_n \, (\Delta N_n - 1) \right\} \tag{8.6.19}$$

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

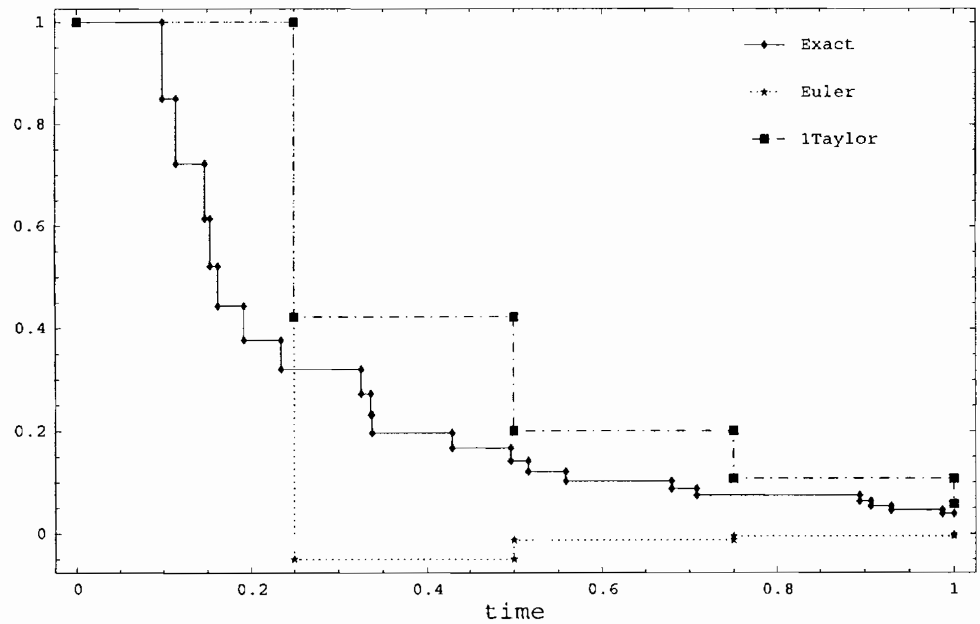

Figure 8.6.1: Exact solution, Euler and order 1.0 Taylor approximations.

For the linear SDE (8.2.4) and a given sample path of the Poisson process, we plot in Figure 8.6.1 the exact solution (8.2.5), the Euler approximation (8.4.8) and the order 1.0 strong Taylor approximation (8.6.19). We selected a time step size $\Delta = 0.25$ and the following parameters: $X_0 = 1$, $T = 1$, $\psi = -0.15$ and $\lambda = 20$. Note in Figure 8.6.1 that the order 1.0 strong Taylor approximation is at the terminal time $t = 1$ rather close to the exact solution. It appears visually better than the Euler approximation, which becomes even negative. Theorem 4.5.1, presented in Chapter 4, and Theorem 8.8.4, to be presented in Section 8.8, provide a firm basis for judging the convergence of such higher order schemes.

8.7 Order 1.5 and 2.0 Strong Taylor Schemes

If we use the truncated Wagner-Platen expansion (8.5.18) with $l = 3$, when applied to each time interval $[t_n, t_{n+1}]$ with $f(x) = x$, we obtain the *order* 1.5 *strong Taylor* *approximation*

$$\begin{aligned} Y_{n+1} &= Y_n + c\,\Delta N_n + \Big\{c\big(t_n, Y_n + c\big) - c\Big\}\binom{\Delta N_n}{2} \\ &\quad + \Big\{c\big(t_n, Y_n + c + c(t_n, Y_n + c)\big) - 2\,c\big(t_n, Y_n + c\big) + c\Big\}\binom{\Delta N_n}{3} \end{aligned}$$

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

In the case of our particular example (8.2.4), the order 1.5 strong Taylor approximation is of the form

$$Y_{n+1} = Y_n\Big\{1 + \psi\,\Delta N_n + \psi^2\binom{\Delta N_n}{2} + \psi^3\binom{\Delta N_n}{3}\Big\}$$

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

To construct an approximation with second order of strong convergence we need to choose $l = 4$ in the truncated expansion (8.5.18) with $f(x) = x$. Then we obtain the *order* 2.0 *strong Taylor approximation*

$$\begin{aligned} Y_{n+1} &= Y_n + c\Delta N_n + \Big\{c\big(Y_n + c(Y_n)\big) - c\big(Y_n\big)\Big\}\binom{\Delta N_n}{2} \\ &\quad + \Big\{c\big(t_n, Y_n + c + c(t_n, Y_n + c)\big) - 2\,c\big(t_n, Y_n + c\big) + c\Big\}\binom{\Delta N_n}{3} \\ &\quad + \Big\{c\Big(t_n, Y_n + c + c\big(t_n, Y_n + c\big) + c\big(t_n, Y_n + c + c(t_n, Y_n + c)\big)\Big) \\ &\quad - 3c\Big(t_n, Y_n + c + c\big(t_n, Y_n + c\big)\Big) + 3\,c\Big(t_n, Y_n + c\Big) - c\Big\}\binom{\Delta N_n}{4} \end{aligned}$$

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

For the linear SDE (8.2.4) the order 2.0 strong Taylor approximation is of the form

$$Y_{n+1} = Y_n\left\{1 + \psi\,\Delta N_n + \psi^2\binom{\Delta N_n}{2} + \psi^3\binom{\Delta N_n}{3} + \psi^4\binom{\Delta N_n}{4}\right\}$$

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

8.8 Convergence Results

It is desirable to be able to construct systematically highly accurate discrete time approximations for solutions of pure jump SDEs. For this purpose we use the Wagner-Platen expansion (8.5.18) to obtain the order γ strong Taylor scheme for pure jump processes, for $\gamma \in \{0.5, 1, 1.5, \ldots\}$.

In this section we consider a pure jump process described by a more general SDE than the SDE (8.2.3) considered so far in this chapter. For the pure jump SDE (8.2.3) driven by one Poisson process it is possible, as shown above, to derive higher order strong schemes that involve only one Poisson random variable in each time step. However, it is important to study also more general multi-dimensional pure jump processes, which allow the modelling of more complex quantities as, for instance, state-dependent intensities. For this reason, we consider here the d-dimensional pure jump SDE

$$dX_t = \int_{\mathcal{E}} c(t, X_{t-}, v)\, p_\phi(dv, dt), \tag{8.8.20}$$

for $t \in [0, T]$, with $X_0 \in \mathbb{R}^d$. Here the jump coefficient c and the Poisson random measure are defined as in (2.1.2). Note that the mark space $\mathcal{E}$ of the Poisson random measure can be made multi-dimensional or split into disjoint subsets and, thus, can generate several sources of jumps. The case of a multi-dimensional SDE driven by several Poisson processes is a specific case of the SDE (8.8.20). The SDE (8.8.20) is equivalent to the jump-diffusion SDE (2.1.2) when the drift coefficient a and the diffusion coefficient b equal zero. Note that birth and death processes and, more generally, continuous time Markov chains can be described by the SDE (8.8.20).

Theorem 4.5.1, presented in Chapter 4, establishes the strong order of convergence of strong Taylor approximations for jump-diffusion SDEs. When specifying the mentioned theorem to the case of SDEs driven by pure jump processes, it will turn out that it is possible to weaken the assumptions on the coefficients of the Wagner-Platen expansion. As we will see below, the Lipschitz and growth conditions on the jump coefficient are already sufficient to establish the convergence of strong Taylor schemes of any given strong order of convergence $\gamma \in \{0.5, 1, 1.5, 2, \ldots\}$. Differentiability of the jump coefficient is not required. This is due to the structure of the increment operator $L_v^{(-1)}$, see (3.3.6), naturally appearing in the coefficient of the Wagner-Platen expansion for pure jump processes.

For a regular time discretization $(t)_\Delta$ with maximum step size $\Delta \in (0,1)$ we define, the *order γ strong Taylor scheme* for pure jump SDEs by

$$\begin{aligned} Y_{n+1}^{\Delta} &= Y_n^{\Delta} + \sum_{k=0}^{2\gamma-1} \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \cdots \int_{t_n}^{s_1} \int_{\mathcal{E}} \left(L^{(-1)}\right)^k c(t_n, Y_n^{\Delta}, v^0) \\ &\qquad p_\phi(dv^0, ds^0) \ldots p_\phi(dv^k, ds^k), \end{aligned} \tag{8.8.21}$$

for $n \in \{0, 1, \ldots, n_T - 1\}$ and $\gamma \in \{0.5, 1, 1.5, \ldots\}$. Here we have used the notation

$$\left(L^{(-1)}\right)^k c(t_n, Y_n^{\Delta}, v^0) := \begin{cases} c(t_n, Y_n^{\Delta}, v^0) & \text{when } k = 0 \\ L_{v^1}^{(-1)} c(t_n, Y_n^{\Delta}, v^0) & \text{when } k = 1 \\ L_{v^k}^{(-1)}\Big(\ldots\big(L_{v^1}^{(-1)} c(t_n, Y_n^{\Delta}, v^0)\big)\Big) & \text{when } k \in \{1, \ldots\}, \end{cases} \tag{8.8.22}$$

where the increment operator $L^{(-1)}$ is defined in (3.3.6).

The following three lemmas show that for SDEs driven by pure jump processes, Lipschitz and growth conditions provide sufficient conditions for Theorem 4.5.1 to guarantee the corresponding strong of strong convergence.

Lemma 8.8.1 *Assume that the jump coefficient satisfies the Lipschitz condition*

$$|c(t,x,u) - c(t,y,u)| \le K|x-y|, \tag{8.8.23}$$

for $t \in [0,T]$, $x, y \in \mathbb{R}^d$ and $u \in \mathcal{E}$, with some constant $K \in (0,\infty)$. Then for any $\gamma \in \{0.5, 1, 1.5, \ldots\}$ and $k \in \{0, 1, 2, \ldots, 2\gamma-1\}$ the kth coefficient $\left(L^{(-1)}\right)^k c(t,x,u)$ of the order γ strong Taylor scheme, satisfies the Lipschitz condition

$$\left| \left(L^{(-1)}\right)^k c(t,x,u) - \left(L^{(-1)}\right)^k c(t,y,u) \right| \le C_k |x-y|, \tag{8.8.24}$$

for $t \in [0,T]$, $x, y \in \mathbb{R}^d$, $u \in \mathcal{E}^k$ and some constant $C_k \in (0,\infty)$ which only depends on k.

Proof: We prove the assertion (8.8.24) by induction with respect to k. For $k = 0$, by the Lipschitz condition (8.8.23) we obtain

$$\left| \left(L^{(-1)}\right)^0 c(t,x,u) - \left(L^{(-1)}\right)^0 c(t,y,u) \right| = \left| c(t,x,u) - c(t,y,u) \right| \le K\,|x-y|.$$

For $k = l+1$, by the induction hypothesis, Jensen's inequality, see Appendix A, and the Lipschitz condition (8.8.23) we obtain

$$
\begin{aligned}
\Big| \left(L^{(-1)}\right)^{l+1} & c(t,x,u) - \left(L^{(-1)}\right)^{l+1} c(t,y,u)\Big| \\
= \quad & \Big| \left(L^{(-1)}\right)^{l} c(t, x + c(t,x,v), u) - \left(L^{(-1)}\right)^{l} c(t,x,u) \\
& - \left(L^{(-1)}\right)^{l} c(t, y + c(t,y,v), u) + \left(L^{(-1)}\right)^{l} c(t,y,u)\Big| \\
\leq \quad & C_l \Big| x - y + \left(c(t,x,v) - c(t,y,v)\right)\Big| + C_l \Big| x - y\Big| \\
\leq \quad & 2\, C_l \Big| x-y\Big| + C_l\, K \Big| x - y\Big| \\
= \quad & C_{l+1} \left| x - y\right|,
\end{aligned}
$$

which completes the proof of Lemma 8.8.1. □

Lemma 8.8.2 *Assume that the jump coefficient satisfies the growth condition*

$$|c(t,x,u)|^2 \leq \widetilde{K}\,(1 + |x|^2) \tag{8.8.25}$$

for $t \in [0,T]$ and $x \in \mathbb{R}^d$ and $u \in \mathcal{E}$, with some constant $\widetilde{K} \in (0,\infty)$. Then for any $\gamma \in \{0.5, 1, 1.5, \ldots\}$ and $k \in \{0, 1, 2, \ldots, 2\gamma{-}1\}$ the kth coefficient $\left(L^{(-1)}\right)^k c(t,x,u)$ of the order γ strong Taylor scheme, satisfies the growth condition

$$\Big| \left(L^{(-1)}\right)^{k} c(t,x,u)\Big|^2 \leq \widetilde{C}_k\,(1 + |x|^2) \tag{8.8.26}$$

for $t \in [0,T]$, $x, y \in \mathbb{R}^d$, $u \in \mathcal{E}^k$ and some constant $\widetilde{C}_k \in (0,\infty)$ which only depends on k.

Proof: We prove the assertion of Lemma 8.8.2 by induction with respect to k. For $k = 0$, by applying the growth condition (8.8.25) we obtain

$$\Big| \left(L^{(-1)}\right)^{0} c(t,x,u)\Big|^2 = \Big| c(t,x,u)\Big|^2 \leq \widetilde{K}\,(1 + |x|^2).$$

For $k = l+1$, by the induction hypotheses, Jensen's inequality and the growth condition (8.8.25) we obtain

$$\begin{aligned}\left|\left(L^{(-1)}\right)^{l+1} c(t,x,u)\right|^2 &= \left|\left(L^{(-1)}\right)^{l} c(t,x+c(t,x,v),u) - \left(L^{(-1)}\right)^{l} c(t,x,u)\right|^2 \\ &\leq 2\left(\widetilde{C}_l\left(1+\left|x+c(t,x,v)\right|^2\right) + \widetilde{C}_l\left(1+|x|^2\right)\right) \\ &\leq 2\left(\widetilde{C}_l\left(1+2\left(|x|^2+\left|c(t,x,v)\right|^2\right)\right) + \widetilde{C}_l\left(1+|x|^2\right)\right) \\ &\leq \widetilde{C}_{l+1}\,(1+|x|^2),\end{aligned}$$

which completes the proof of Lemma 8.8.2. □

Lemma 8.8.3 *Let us assume that*

$$E(|X_0|^2) < \infty \tag{8.8.27}$$

and the jump coefficient satisfies the Lipschitz condition

$$|c(t,x,u) - c(t,y,u)| \leq K_1\,|x-y| \tag{8.8.28}$$

and the growth condition

$$|c(t,x,u)|^2 \leq K_2\,(1+|x|^2) \tag{8.8.29}$$

for $t \in [0,T]$, $x, y \in \mathbb{R}^d$, *and* $u \in \mathcal{E}$, *with constants* $K_1, K_2 \in (0,\infty)$. *Then for any* $\gamma \in \{0.5, 1, 1.5, \ldots\}$ *and* $k \in \{0,1,2,\ldots,2\gamma-1\}$ *the kth coefficient* $\left(L^{(-1)}\right)^k c(t,x,u)$ *of the order* γ *strong Taylor scheme satisfies the integrability condition*

$$\left(L^{(-1)}\right)^k c(\cdot,x,\cdot) \in \mathcal{H}_k,$$

for $x \in \mathbb{R}^d$, *where* $\mathcal{H}_k$ *is the set of adapted stochastic process* $g = \{g(t), t \in [0,T]\}$ *such that*

$$E\left(\int_0^T \int_{\mathcal{E}} \int_0^{s_k} \int_{\mathcal{E}} \cdots \int_0^{s_2} |g(s,v^1,\ldots,v^k,\omega)|^2 \phi(dv^1)ds_1 \ldots \phi(dv^k)\,ds_k\right) < \infty.$$

Proof: By Lemma 8.8.2 for any $\gamma \in \{0.5, 1, 1.5, \ldots\}$ and $k \in \{0,1,2,\ldots,2\gamma-1\}$ the kth coefficient $\left(L^{(-1)}\right)^k c(t,x,u)$ of the order γ strong Taylor scheme satisfies the growth condition

$$\left| \left(L^{(-1)}\right)^k c(t,x,u)\right|^2 \le \widetilde{C}_k \,(1+|x|^2) \tag{8.8.30}$$

for $t \in [0,T]$, $x, y \in \mathbb{R}^d$ and $u \in \mathcal{E}^k$, with the constant $\widetilde{C}_k \in (0,\infty)$. Therefore, for any $\gamma \in \{0.5, 1, 1.5, \ldots\}$ and $k \in \{0, 1, 2, \ldots, 2\gamma - 1\}$, by condition (8.8.30) and Fubini's theorem we obtain

$$E\Big(\int_0^T \int_{\mathcal{E}} \int_0^{s_k} \int_{\mathcal{E}} \cdots \int_0^{s_1} \int_{\mathcal{E}} |\left(L^{(-1)}\right)^k c(t, X_{s_0}, u^0)|^2$$
$$\phi(du^0)\, ds_0 \ldots \phi(du^k)\, ds_k\Big)$$
$$\le E\left(\int_0^T \int_{\mathcal{E}} \int_0^{s_k} \int_{\mathcal{E}} \cdots \int_0^{s_1} \int_{\mathcal{E}} \widetilde{C}_k(1+|X_{s_0}|^2)\phi(du^0)\, ds_0 \ldots \phi(du^k)\, ds_k\right)$$
$$\le \widetilde{C}_k \frac{(T\lambda)^k}{k!} + \widetilde{C}_k \int_0^T \int_0^{s_k} \cdots \int_0^{s_1} E\left(\sup_{0 \le z \le T} |X_z|^2\right) ds_0 \ldots ds_k < \infty.$$

The last passage holds, since conditions (8.8.27), (8.8.28) and (8.8.29) ensure that

$$E\left(\sup_{z\in[0,T]} |X_z|^2\right) < \infty,$$

see Theorem 2.2.1. This completes the proof of Lemma 8.8.3. □

We emphasize that in the case of pure jump SDEs, unlike the more general case of jump diffusions, no extra differentiability conditions on the jump coefficient c are required when deriving higher order approximations.

Theorem 8.8.4 *For given $\gamma \in \{0.5, 1, 1.5, 2, \ldots\}$, let $Y^\Delta = \{Y^\Delta(t), t \in [0,T]\}$ be the order γ strong Taylor scheme* (8.8.21) *for the SDE* (8.8.20) *corresponding to a time discretization $(t)_\Delta$ with maximum step size $\Delta \in (0,1)$. We assume for the jump coefficient $c(t,x,v)$ the Lipschitz condition* (8.8.23) *and the growth condition* (8.8.25). *Moreover, suppose that*

$$E(|X_0|^2) < \infty \qquad \textit{and} \qquad \sqrt{E(|X_0 - Y_0^\Delta|^2)} \le K_1 \Delta^\gamma.$$

Then the estimate

$$\sqrt{E(\max_{0 \le n \le n_T} |X_n - Y_n^\Delta|^2)} \le K\, \Delta^\gamma$$

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

Proof: The proof of Theorem 8.8.4 is a direct consequence of the convergence Theorem 4.5.1 for jump diffusions presented in Chapter 4. This is the case because by the Lemmas 8.8.1, 8.8.2 and 8.8.3, the coefficients of the order γ strong Taylor scheme (8.8.21) satisfy the conditions required by the convergence Theorem 4.5.1. Note that the differentiability condition in Theorem 4.5.1, where $f_{-\alpha} \in \mathcal{C}^{1,2}$ for all $\alpha \in \mathcal{A}_\gamma \bigcup B(\mathcal{A}_\gamma)$, is not required for pure jump SDEs. This condition is used in the proof of the general convergence Theorem 4.5.1 only for the derivation of the Wagner-Platen expansion. In the pure jump case, as shown in Section 8.5, one needs only measurability of the jump coefficient c to derive the corresponding Wagner-Platen expansion. $\square$

Theorem 8.8.4 states that the order γ strong Taylor scheme for pure jump SDEs achieves a strong order of convergence equal to γ. In fact Theorem 8.8.4 states that the strong convergence of order γ is not just at the endpoint T but it is also uniform over all time discretization points. Thus, by including enough terms from the Wagner-Platen expansion (8.5.18) we are able to construct schemes of any given strong order of convergence $\gamma \in \{0.5, 1, 1.5, \ldots\}$. Note that Theorem 8.8.4 applies to solutions of multi-dimensional pure jump SDEs.

For the mark-independent pure jump SDE (8.2.3) driven by one Poisson process, the order γ strong Taylor scheme (8.8.21) reduces to

$$Y_{n+1}^{\Delta} = Y_n^{\Delta} + \sum_{k=1}^{2\gamma} \left(\tilde{\Delta}_N\right)^k f(Y_n^{\Delta}) \binom{\Delta N_n}{k} \tag{8.8.31}$$

for $n \in \{0, 1, \ldots, n_T - 1\}$, with $f(x) = x$, where the operator $\tilde{\Delta}_N$ is defined in (8.5.17). In this case the generation of the multiple stochastic integrals involved is straightforward, since only one Poisson distributed random variable at each time step is required, as we have seen in (8.5.14). This allows the above schemes to be easily implemented. Such an implementation is more complex in the case of genuine jump-diffusion SDEs and it is worth to know the advantages that one has when deriving higher order strong Taylor schemes for pure jump SDEs.

Chapter 9

Regular Weak Taylor Approximations

As pointed out in the introduction, it is a much easier task to approximate the probability measure generated by a jump diffusion than approximating its paths. Only weak approximations are needed to achieve this goal. In this chapter we present regular weak approximations obtained directly from a truncated Wagner-Platen expansion. The desired weak order of convergence determines which terms of the stochastic expansion one has to include in the approximation. These weak Taylor schemes are different from the regular strong Taylor schemes presented in Chapter 4. The construction of weak schemes requires a separate analysis. A convergence theorem, useful to construct weak Taylor approximations of any given weak order of convergence $\beta \in \{1, 2, \ldots\}$, will be presented at the end of this chapter.

9.1 Introduction

As in Chapter 4, we first consider the one-dimensional, $d = m = 1$, jump-diffusion SDE

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t + \int_{\mathcal{E}} c(t, X_{t-}, v)\, p_\phi(dv, dt), \tag{9.1.1}$$

for $t \in [0, T]$, with $X_0 \in \mathbb{R}$, where $W = \{W_t, t \in [0, T]\}$ is a one-dimensional Wiener process and $p_\phi(dv, dt)$ is a Poisson measure. Later, we will consider the autonomous d-dimensional jump-diffusion SDE

$$dX_t = a(X_t)dt + b(X_t)dW_t + \int_{\mathcal{E}} c(X_{t-}, v)\, p_\phi(dv, dt), \tag{9.1.2}$$

for $t \in [0, T]$, with $X_0 \in \mathbb{R}^d$. As in Chapter 4, we will consider both the case of a scalar Wiener process, $m = 1$, and that of an m-dimensional Wiener process. Note that in the case of an autonomous multi-dimensional SDE, we can always recover an SDE with time-dependent coefficients by considering the time t as first component of the process X. Moreover, we remark that we will often treat separately the simpler case of a mark-independent jump coefficient $c(t,x,v) = c(t,x)$.

For the following discrete time approximations, we consider a regular time discretization $(t)_\Delta$ with maximum time step size Δ, as defined in Chapter 4, that does not include the jump times of the Poisson measure. We recall that we use the notation $\Delta_n = t_{n+1} - t_n$.

9.2 Euler Scheme

Due to the nature of the Wagner-Platen expansion, which approximates the diffusion and jump features, the simplest useful weak Taylor approximation coincides with the simplest useful strong Taylor approximation: the Euler scheme (4.2.12), presented in Chapter 4. Nonetheless, we will prove at the end of this chapter that the Euler scheme attains an order of weak convergence $\beta = 1.0$, as opposed to a strong order $\gamma = 0.5$, shown in Chapter 4.

We recall that in the general multi-dimensional case the kth component of the Euler scheme is given by

$$Y_{n+1}^k = Y_n^k + a^k \Delta_n + \sum_{j=1}^{m} b^{k,j} \Delta W_n^j + \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} c^k(\xi_i),$$

for $k \in \{1, 2, \ldots, d\}$, where we have used the abbreviations defined in (4.1.9)–(4.1.11).

9.3 Order 2.0 Taylor Scheme

We have seen that the Euler scheme is the simplest Taylor scheme both for strong and weak approximations. Note that one can simplify the Euler scheme further by using discrete random variables, as will be discussed in Chapter 12. When higher accuracy is required and, thus, a scheme with higher order of convergence is sought, then it is important to distinguish between strong and weak schemes. Indeed, by adding to the Euler scheme the four multiple stochastic integrals appearing in the order 1.0 strong Taylor scheme presented in Section 4.3, we improve the order of strong convergence from $\gamma = 0.5$ to $\gamma = 1.0$. However, it can be shown that the order 1.0 strong Taylor scheme has generally the same order of weak convergence $\beta = 1.0$ as the Euler scheme. This indicates that the construction of efficient higher order weak schemes requires rules that are different from those used for strong schemes. At the end of this chapter we present a convergence theorem for weak Taylor approximations. It provides a rule for selecting from the Wagner-Platen expansion the multiple stochastic integrals needed to achieve a given order of weak convergence $\beta = \{1.0, 2.0, \ldots\}$. In this way, we obtain the *order* 2.0 *weak Taylor scheme*, which, in the one-dimensional case, $d = m = 1$, is given by

$$\begin{aligned}
Y_{n+1} &= Y_n + a\Delta_n + b\Delta W_n + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} c(v)\, p_\phi(dv, dz) \\
&\quad + bb' \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} dW_{z_1} dW_{z_2} \\
&\quad + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} b\, c'(v) dW_{z_1} p_\phi(dv, dz_2) \\
&\quad + \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} \int_{\mathcal{E}} \Big\{ b\big(t_n, Y_n + c(v)\big) - b \Big\} p_\phi(dv, dz_1) dW_{z_2} \\
&\quad + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} \int_{\mathcal{E}} \Big\{ c\big(t_n, Y_n + c(v_1), v_2\big) - c(v_2) \Big\} p_\phi(dv_1, dz_1) p_\phi(dv_2, dz_2) \\
&\quad + \Big(\frac{\partial a}{\partial t} + a\, a' + \frac{a''}{2} b^2 \Big) \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} dz_1 dz_2 + a'\, b \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} dW_{z_1} dz_2 \\
&\quad + \left(\frac{\partial b}{\partial t} + a\, b' + \frac{b''}{2} b^2 \right) \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} dz_1 dW_{z_2} \\
&\quad + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} \left(\frac{\partial c(v)}{\partial t} + a\, c'(v) + \frac{c''(v)}{2} b^2 \right) dz_1 p_\phi(dv, dz_2) \\
&\quad + \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} \int_{\mathcal{E}} \Big\{ a\big(t_n, Y_n + c(v)\big) - a \Big\} p_\phi(dv, dz_1) dz_2, \qquad (9.3.1)
\end{aligned}$$

where we have used the abbreviated notation (4.1.9)–(4.1.11), and we have denoted the partial derivative with respect to time with $\frac{\partial}{\partial t}$. This scheme will be shown to achieve weak order $\beta = 2.0$. Note that to achieve second order of weak convergence, we have added all nine double stochastic integrals to the Euler scheme. The order 2.0 weak Taylor scheme was presented in Liu & Li (2000). It generalizes a scheme for pure diffusion SDEs presented in Milstein (1978) and mentioned in Talay (1984).

The scheme (9.3.1) is rather complex, as it involves all nine possible double stochastic integrals with respect to time, Wiener processes and the Poisson random measure. By Itô's formula, the integration by parts formula and using (9.3.1), we can rewrite the order 2.0 weak Taylor scheme as

$$
\begin{aligned}
Y_{n+1} &= Y_n + a\Delta_n + b\Delta W_n + \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} c(\xi_i) + \frac{b\,b'}{2}\left((\Delta W_n)^2 - \Delta_n\right) \\
&\quad + b \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} c'(\xi_i)\big(W_{\tau_i} - W_{t_n}\big) \\
&\quad + \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} \Big\{b\big(Y_n + c(\xi_i)\big) - b\Big\}\big(W_{t_{n+1}} - W_{\tau_i}\big) \\
&\quad + \sum_{j=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} \sum_{i=p_\phi(t_n)+1}^{p_\phi(\tau_j)} \Big\{c\big(Y_n + c(\xi_i), \xi_j\big) - c(\xi_j)\Big\} \\
&\quad + \frac{1}{2}\Big(\frac{\partial a}{\partial t} + a\,a' + \frac{a''}{2}b^2\Big)(\Delta_n)^2 + a'\,b\,\Delta Z_n \\
&\quad + \left(\frac{\partial b}{\partial t} + a\,b' + \frac{b''}{2}b^2\right)(\Delta W_n \Delta_n - \Delta Z_n) \\
&\quad + \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} \left(\frac{\partial c(\xi_i)}{\partial t} + a\,c'(\xi_i) + \frac{c''(\xi_i)}{2}b^2\right)(\tau_i - t_n) \\
&\quad + \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} \Big\{a\big(Y_n + c(\xi_i)\big) - a\Big\}(t_{n+1} - \tau_i), \qquad (9.3.2)
\end{aligned}
$$

which is readily applicable for weak approximation and , thus, for Monte Carlo simulation. The correlated Gaussian random variables $\Delta W_n = W_{t_{n+1}} - W_{t_n}$ and

$$
\Delta Z_n = \int_{t_n}^{t_{n+1}} \int_{t_n}^{s_2} dW_{s_1}\, ds_2 \qquad (9.3.3)
$$

can be generated as in (6.2.23). The order 2.0 Taylor scheme for the for the SDE

(2.1.5) is given by

$$\begin{aligned}
Y_{n+1} &= Y_n + \mu Y_n \Delta_n + \sigma Y_n \Delta W_n + Y_n \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} (\xi_i - 1) \\
&+ \frac{\sigma^2}{2} Y_n \big((\Delta W_n)^2 - \Delta_n\big) + \sigma Y_n \Delta W_n \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} (\xi_i - 1) \\
&+ Y_n \sum_{j=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} \sum_{i=p_\phi(t_n)+1}^{p_\phi(\tau_j)} (\xi_i - 1)(\xi_j - 1) \\
&+ \frac{\mu^2}{2} Y_n (\Delta_n)^2 + \mu\sigma Y_n \Delta W_n \Delta_n + \mu Y_n \Delta_n \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} (\xi_i - 1). \qquad (9.3.4)
\end{aligned}$$

In the special case of a mark-independent jump coefficient $c(t, x, v) = c(t, x)$, the order 2.0 weak Taylor scheme reduces to

$$\begin{aligned}
Y_{n+1} &= Y_n + a\Delta + b\Delta W_n + c\Delta p_n + b\, b'\, I_{(1,1)} + b\, c'\, I_{(1,-1)} \\
&+ \{b(t_n, Y_n + c) - b\}\, I_{(-1,1)} + \{c\,(t_n, Y_n + c) - c\}\, I_{(-1,-1)} \\
&+ \frac{1}{2}\Big(\frac{\partial a}{\partial t} + a\, a' + \frac{a''}{2} b^2\Big)\, I_{(0,0)} + a'\, b\, I_{(1,0)} + \left(\frac{\partial b}{\partial t} + a\, b' + \frac{b''}{2} b^2\right) I_{(0,1)} \\
&+ \left(\frac{\partial c}{\partial t} + a\, c' + \frac{c''}{2} b^2\right) I_{(0,-1)} + \Big\{a\big(t_n, Y_n + c\big) - a\Big\} I_{(-1,0)}, \qquad (9.3.5)
\end{aligned}$$

where the multiple stochastic integrals $I_{(1,1)}$, $I_{(1,-1)}$, $I_{(-1,1)}$ and $I_{(-1,-1)}$ are defined in (4.3.5) and

$$\begin{aligned}
I_{(0,0)} &:= \int_{t_n}^{t_{n+1}} \int_{t_n}^{s_2} ds_1 ds_2 = \frac{\Delta_n^2}{2} \\
I_{(1,0)} &:= \int_{t_n}^{t_{n+1}} \int_{t_n}^{s_2} dW_{s_1} ds_2 = \Delta Z_n \\
I_{(0,1)} &:= \int_{t_n}^{t_{n+1}} \int_{t_n}^{s_2} ds_1 dW_{s_2} = \Delta_n\, \Delta W_n - I_{(1,0)} \\
I_{(0,-1)} &:= \int_{t_n}^{t_{n+1}} \int_{t_n}^{s_2} ds_1 p_\phi(\mathcal{E}, ds_2) = \sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} \tau_i - \Delta p_n\, t_n
\end{aligned}$$

$$I_{(-1,0)} := \int_{t_n}^{t_{n+1}} \int_{t_n}^{s_2} p_\phi(\mathcal{E}, ds_1) ds_2 = \Delta_n \, \Delta p_n - I_{(0,-1)}. \tag{9.3.6}$$

Even in the case of mark-independent jump sizes, the computational complexity of the order 2.0 weak Taylor scheme depends on the intensity of the Poisson jump measure. Indeed, as already discussed in Section (4.3), for the generation of double stochastic integrals involving the Poisson measure the random variables ΔW_n and Δp_n are not sufficient. In this case one needs also the location of jump times in every time interval $(t_n, t_{n+1}]$, with $n \in \{0, \ldots, n_T - 1\}$.

Let us consider the autonomous multi-dimensional SDE (9.1.2) with scalar Wiener process, $m = 1$. The kth component of the *order* 2.0 *weak Taylor scheme* is given by

$$\begin{aligned}
Y_{n+1}^k &= Y_n^k + a^k \Delta_n + b^k \Delta W_n + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} c^k(v) \, p_\phi(dv, dz) \\
&\quad + \sum_{l=1}^{d} \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} b^l \, \frac{\partial b^k}{\partial x^l} dW_{z_1} dW_{z_2} \\
&\quad + \sum_{l=1}^{d} \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} b^l \, \frac{\partial c^k(v)}{\partial x^l} \, dW_{z_1} p_\phi(dv, dz_2) \\
&\quad + \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} \int_{\mathcal{E}} \Big\{ b^k\big(t_n, Y_n + c(v)\big) - b^k \Big\} p_\phi(dv, dz_1) dW_{z_2} \\
&\quad + \int_{t_n}^{t_{n+1}} \int_{\mathcal{E}} \int_{t_n}^{z_2} \int_{\mathcal{E}} \Big\{ c^k\big(t_n, Y_n + c(v_1), v_2\big) - c^k\big(v_2\big) \Big\} \\
&\quad \times p_\phi(dv_1, dz_1) \, p_\phi(dv_2, dz_2) \\
&\quad + \Big(\sum_{l=1}^{d} a^l \, \frac{\partial a^k}{\partial x^l} + \sum_{i,l=1}^{d} \frac{\partial^2 a^k}{\partial x^l \partial x^i} \frac{b^l b^i}{2} \Big) \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} dz_1 dz_2 \\
&\quad + \sum_{l=1}^{d} b^l \frac{\partial a^k}{\partial x^l} \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} dW_{z_1} dz_2 \\
&\quad + \Big(\sum_{l=1}^{d} a^l \, \frac{\partial b^k}{\partial x^l} + \sum_{i,l=1}^{d} \frac{\partial^2 b^k}{\partial x^l \partial x^i} \frac{b^l b^i}{2} \Big) \int_{t_n}^{t_{n+1}} \int_{t_n}^{z_2} dz_1 dW_{z_2}
\end{aligned}$$

$$+\int_{t_n}^{t_{n+1}}\int_{\mathcal{E}}\int_{t_n}^{z_2}\Big(\sum_{l=1}^{d}a^l\,\frac{\partial c^k(v)}{\partial x^l}+\sum_{i,l=1}^{d}\frac{\partial^2 c^k(v)}{\partial x^l\partial x^i}\frac{b^l b^i}{2}\Big)dz_1 p_\phi(dv,dz_2)$$

$$+\int_{t_n}^{t_{n+1}}\int_{t_n}^{z_2}\int_{\mathcal{E}}\Big\{a^k\big(t_n,Y_n+c(v)\big)-a^k\Big\}p_\phi(dv,dz_1)dz_2, \tag{9.3.7}$$

for $k\in\{1,2,\dots,d\}$, where a^k, b^k, and c^k are the kth components of the drift, diffusion and jump coefficients, respectively. Note that the multiple stochastic integrals can be generated similarly as in the one-dimensional case.

In the general multi-dimensional case, the kth component of the *order* 2.0 *weak Taylor scheme* is given by

$$\begin{aligned}
Y_{n+1}^k \;=\;& Y_n^k + a^k\Delta_n+\sum_{j=1}^{m}b^{k,j}\Delta W_n^j+\int_{t_n}^{t_{n+1}}\int_{\mathcal{E}}c^k(v)\,p_\phi(dv,dz)\\
&+\sum_{j_1,j_2=1}^{m}L^{(j_1)}b^{k,j_2}\int_{t_n}^{t_{n+1}}\int_{t_n}^{z_2}dW_{z_1}^{j_1}dW_{z_2}^{j_2}\\
&+\sum_{j_1=1}^{m}\int_{t_n}^{t_{n+1}}\int_{\mathcal{E}}\int_{t_n}^{z_2}L^{(j_1)}c^k(v)dW_{z_1}^{j_1}p_\phi(dv,dz_2)\\
&+\sum_{j_1=1}^{m}\int_{t_n}^{t_{n+1}}\int_{\mathcal{E}}\int_{t_n}^{z_2}L_v^{(-1)}b^{k,j_1}p_\phi(dv,dz_2)\,dW_{z_2}^{j_1}\\
&+\int_{t_n}^{t_{n+1}}\int_{\mathcal{E}}\int_{t_n}^{z_2}\int_{\mathcal{E}}L_{v_1}^{(-1)}c^k(v_2)p_\phi(dv_1,dz_1)p_\phi(dv_2,dz_2)\\
&+\sum_{j=1}^{m}L^{(j)}a^k\int_{t_n}^{t_{n+1}}\int_{t_n}^{z_2}dW_{z_1}^j dz_2+\sum_{j=1}^{m}L^{(0)}b^{k,j}\int_{t_n}^{t_{n+1}}\int_{t_n}^{z_2}dz_1 dW_{z_2}^j\\
&+L^{(0)}a^k\int_{t_n}^{t_{n+1}}\int_{t_n}^{z_2}dz_1dz_2+\int_{t_n}^{t_{n+1}}\int_{\mathcal{E}}\int_{t_n}^{z_2}L^{(0)}c^k(v)dz_1p_\phi(dv,dz_2)\\
&+\int_{t_n}^{t_{n+1}}\int_{t_n}^{z_2}\int_{\mathcal{E}}L_v^{(-1)}a^k p_\phi(dv,dz_1)dz_2,
\end{aligned}\tag{9.3.8}$$

for $k\in\{1,2,\dots,d\}$, where the operators $L^{(j)}$, with $j\in\{-1,\dots,m\}$, are defined in (3.3.4)–(3.3.6).

In the multi-dimensional case with mark-independent jump size, the kth component of the *order* 2.0 *weak Taylor scheme* simplifies to the algorithm

$$
\begin{aligned}
Y_{n+1}^k \;=\;& Y_n^k + a^k \Delta_n + \sum_{j=1}^{m} b^{k,j} \Delta W_n^j + c^k \Delta p_n \\
&+ \sum_{j_1,j_2=1}^{m} L^{(j_1)} b^{k,j_2}\, I_{(j_1,j_2)} + \sum_{j_1=1}^{m} L^{(j_1)} c^k\, I_{(j_1,-1)} \\
&+ \sum_{j_1=1}^{m} \Big\{ b^{k,j_1}\big(t_n, Y_n + c\big) - b^{k,j_1} \Big\} I_{(-1,j_1)} + \Big\{ c^k\big(t_n, Y_n + c\big) - c^k \Big\} I_{(-1,-1)} \\
&+ \sum_{j=1}^{m} L^{(j)} a^k I_{(j,0)} + \sum_{j=1}^{m} L^{(0)} b^{k,j} I_{(0,j)} + L^{(0)} a^k I_{(0,0)} \\
&+ L^{(0)} c^k I_{(0,-1)} + \Big\{ a^k\big(t_n, Y_n + c\big) - a^k \Big\} I_{(-1,0)},
\end{aligned}
\tag{9.3.9}
$$

for $k \in \{1, 2, \ldots, d\}$. All multiple stochastic integrals that do not involve Wiener processes can be generated as in the one-dimensional case (9.3.5). For those involving one Wiener process, we can use the relations

$$
\begin{aligned}
I_{(j,-1)} &= \textstyle\sum_{i=p_\phi(t_n)+1}^{p_\phi(t_{n+1})} W_{\tau_i}^j - \Delta p_n\, W_{t_n}^j \qquad & I_{(-1,j)} &= \Delta p_n\, \Delta W_n^j - I_{(j,-1)} \\
I_{(j,0)} &= \Delta Z_n^j & I_{(0,j)} &= \Delta W_n^j \Delta_n - \Delta Z_n^j,
\end{aligned}
\tag{9.3.10}
$$

for $j \in \{1, 2, \ldots, m\}$ and $n \in \{0, 1, \ldots, n_T - 1\}$. Recall that, for every $j \in \{1, 2, \ldots, m\}$, the random variable ΔZ_n^j has a Gaussian distribution with mean $E(\Delta Z_n^j) = 0$, variance $E((\Delta Z_n^j)^2) = \frac{1}{3}(\Delta_n)^3$ and covariance $E(\Delta Z_n^j\, \Delta W_n^j) = \frac{1}{2}(\Delta_n)^2$. Therefore, with $2m$ independent $\mathcal{N}(0,1)$ distributed standard Gaussian random variables $U_{1,j}$ and $U_{2,j}$, for $j \in \{1, 2, \ldots, m\}$, we obtain the required random variables by the transformation

$$
\Delta W_n^j = U_{1,j} \sqrt{\Delta_n} \quad \text{and} \quad \Delta Z_n^j = \frac{1}{2} (\Delta_n)^{\frac{3}{2}} \left(U_{1,j} + \frac{1}{\sqrt{3}} U_{2,j} \right). \tag{9.3.11}
$$

Finally, the generation of the multiple stochastic integrals $I_{(j_1,j_2)}$, with $j_1, j_2 \in \{1, 2, \ldots, m\}$, generally requires an approximation such as the Karhunen-Loève expansion proposed in Kloeden & Platen (1999).

9.4 Commutativity Conditions

As discussed in the previous section, the generation of multiple stochastic integrals required in the order 2.0 weak Taylor scheme is computationally demanding. Let us now discuss some commutativity conditions under which the complexity of the order 2.0 weak Taylor scheme is reduced.

The computational complexity of the order 2.0 weak Taylor scheme generally depends on the intensity of the Poisson measure. Indeed, as previously discussed, the generation of the double stochastic integrals $I_{(1,-1)}$, $I_{(-1,1)}$, $I_{(0,-1)}$ and $I_{(-1,0)}$ requires the knowledge of the exact location of the jump times in the time interval $[t_n, t_{n+1}]$. However, the sum of the above first two double integrals is given by

$$I_{(1,-1)} + I_{(-1,1)} = \Delta p_n \Delta W_n, \tag{9.4.1}$$

which yields an expression that is independent of the particular values of the jump times. Additionally, by (9.3.6) the sum of the above last two double integrals is obtained as

$$I_{(0,-1)} + I_{(-1,0)} = \Delta p_n \Delta_n, \tag{9.4.2}$$

which is also independent of the jump times. Therefore, in the one-dimensional case with mark-independent jump coefficient $c(t,x,v) = c(t,x)$, we can formulate the *first jump commutativity condition*

$$b(t,x)\frac{\partial c(t,x)}{\partial x} = b\Big(t, x + c(t,x)\Big) - b\Big(t,x\Big) \tag{9.4.3}$$

and the *second jump commutativity condition*

$$\frac{\partial c(t,x)}{\partial t} + a(t,x)\,\frac{\partial c(t,x)}{\partial x} + \frac{b^2(t,x)}{2}\frac{\partial^2 c(t,x)}{\partial x^2} = a\Big(t, x + c(t,x)\Big) - a\Big(t,x\Big), \tag{9.4.4}$$

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

The first jump commutativity condition (9.4.3) has been already discussed in Chapter 4. We refer to Section 4.4 for a table of selected diffusion coefficients and corresponding jump coefficients satisfying the jump commutativity condition (9.4.3). Note that in this case the implementation of the order 2.0 weak Taylor scheme does not require sampling the Wiener process at all jump times. The second jump commutativity condition (9.4.4) expresses a relationship in the form of a PDE involving the drift, diffusion and jump coefficients. If also this condition is satisfied, then one needs only to sample the Gaussian random variable ΔW_n and the Poisson random variable Δp_n at each time step. In this special case the above algorithm is very efficient also for SDEs driven by a Poisson measure with high intensity.

In the multi-dimensional case with mark-independent jump size, we also need to generate the multiple stochastic integrals $I_{(j,-1)}$ and $I_{(-1,j)}$, for $j \in \{1, \ldots, m\}$. As discussed in Section 4.4, the sum of two multiple stochastic integrals with respect to the jth component of the Wiener process and the Poisson measure is given by

$$I_{(j,-1)} + I_{(-1,j)} = \Delta p_n \, \Delta W_n^j, \tag{9.4.5}$$

which is independent of the particular jump times. Therefore, we obtain the *first jump commutativity condition*

$$\sum_{l=1}^{d} b^{l,j}(t,x) \frac{\partial c^k(t,x)}{\partial x^l} = b^{k,j}\Big(t, x + c(t,x)\Big) - b^{k,j}\Big(t,x\Big) \tag{9.4.6}$$

and the *second jump commutativity condition*

$$L^{(0)} c^k(t,x) = a^k\Big(t, x + c(t,x)\Big) - a^k\Big(t,x\Big), \tag{9.4.7}$$

for $j \in \{1, 2, \ldots, m\}$, $k \in \{1, 2, \ldots, d\}$, $t \in [0,T]$ and $x \in \mathbb{R}^d$. Here the differential operator $L^{(0)}$ is defined in (3.3.4). Note that the above commutativity conditions consist in two systems of $d \times m$ equations each. Therefore, even for simple given drift and diffusion coefficients, there may not exist any jump coefficients satisfying (9.4.6) or (9.4.7).

To simplify also the generation of the double Wiener integrals $I_{(j_1,j_2)}$ for $j_1, j_2 \in \{1, \ldots, m\}$, one should check if the SDE under consideration satisfies the *diffusion commutativity condition*

$$L^{j_1} b^{k,j_2}(t,x) = L^{j_2} b^{k,j_1}(t,x) \tag{9.4.8}$$

for $j_1, j_2 \in \{1, 2, \ldots, m\}$, $k \in \{1, 2, \ldots, d\}$, $t \in [0,T]$ and $x \in \mathbb{R}^d$, see also (4.4.17). In this situation, as discussed in Section 4.4, the double Wiener integrals can be expressed in terms of the increments $\Delta W_n^{j_1}$ and $\Delta W_n^{j_2}$ of the Wiener processes, that is

$$I_{(j_1,j_2)} + I_{(j_2,j_1)} = \Delta W_n^{j_1} \Delta W_n^{j_2}. \tag{9.4.9}$$

For a multi-dimensional SDE satisfying the commutativity conditions (9.4.6)-(9.4.8) and with mark-independent jump size, the order 2.0 weak Taylor scheme reduces to

$$\begin{aligned}
Y_{n+1}^k &= Y_n^k + a^k\Delta_n + \sum_{j=1}^m b^{k,j}\Delta W_n^j + c^k\Delta p_n \\
&+\frac{1}{2}\sum_{j_1,j_2=1}^m\sum_{i=1}^d b^{i,j_1}\frac{\partial b^{k,j_2}}{\partial x^i}\left\{\Delta W_n^{j_1}\,\Delta W_n^{j_2}-\Delta_n\right\} \\
&+\sum_{j_1=1}^m\{b^{k,j_1}(t_n,Y_n+c)-b^{k,j_1}\}\left(\Delta p_n\,\Delta W_n^{j_1}\right) \\
&+\frac{1}{2}\{c^k(t_n,Y_n+c)-c^k\}\left(\left(\Delta p_n\right)^2-\Delta p_n\right)+\sum_{j=1}^m\sum_{l=1}^d b^{l,j}\frac{\partial a^k}{\partial x^l}\Delta Z_n^j \\
&+\sum_{j=1}^m\left(\frac{\partial b^{k,j}}{\partial t}+\sum_{l=1}^d a^l\frac{\partial b^{k,j}}{\partial x^l}+\sum_{i,l=1}^d\sum_{j_1=1}^m\frac{\partial^2 b^{k,j}}{\partial x^i\partial x^l}\frac{b^{i,j_1}b^{l,j_1}}{2}\right)\left\{\Delta W_n^j\Delta_n-\Delta Z_n^j\right\} \\
&+\left(\frac{\partial a^k}{\partial t}+\sum_{l=1}^d a^l\frac{\partial a^k}{\partial x^l}+\sum_{i,l=1}^d\sum_{j_1=1}^m\frac{\partial^2 a^k}{\partial x^i\partial x^l}\frac{b^{i,j_1}b^{l,j_1}}{2}\right)\frac{(\Delta_n)^2}{2} \\
&+\left\{a^k\left(t_n,Y_n+c\right)-a^k\right\}\Delta p_n\Delta_n, \qquad (9.4.10)
\end{aligned}$$

for $k\in\{1,2,\ldots,d\}$. We will prove the weak order of convergence of the above schemes below in Section 9.5.

9.5 Convergence Results

First let us prepare some results that will be used to establish the order of weak convergence of regular weak Taylor approximations.

Consider the Itô process

$$X_t^{z,y} = y + \int_z^t \widetilde{a}(X_u^{z,y})du + \int_z^t b(X_u^{z,y})dW_u + \int_z^t\int_{\mathcal{E}} c(X_{u-}^{z,y},v)\widetilde{p}_\phi(dv,du) \quad (9.5.1)$$

for $z\le t\le T$, which starts at time $z\in[0,T]$ in $y\in\mathbb{R}^d$. This process has the same solution as the Itô process $X = \{X_t, t \in [0,T]\}$, which solves the SDE (2.1.2).

For a given $\beta \in \{1, 2, \ldots\}$ and a function $g \in \mathcal{C}_P^{2(\beta+1)}(\mathbb{R}^d, \mathbb{R})$ we define the functional

$$u(z, y) := E(g(X_T^{z,y})) \tag{9.5.2}$$

for $(z, y) \in [0, T] \times \mathbb{R}^d$. Then we have

$$u(0, X_0) = E(g(X_T^{0,X_0})) = E(g(X_T)). \tag{9.5.3}$$

We will need the following lemma, see Mikulevicius & Platen (1988), Lemma 4.3, and Mikulevicius (1983), involving the operator

$$\widetilde{L}^{(0)} f(z, y) = L^{(0)} f(z, y) + \int_{\mathcal{E}} L_v^{(-1)} f(z, y) \phi(dv) \tag{9.5.4}$$

for a sufficiently smooth function f, see (3.3.7).

Lemma 9.5.1 *Let us assume that the drift, diffusion and jump coefficients of the SDE* (2.1.2) *have components* $a^k, b^{k,j}, c^k \in \mathcal{C}_P^{2(\beta+1)}(\mathbb{R}^d, \mathbb{R})$ *for all* $k \in \{1, 2 \ldots, d\}$ *and* $j \in \{1, 2 \ldots, m\}$ *with uniformly bounded derivatives. Then the functional* u *defined in* (9.5.2) *is the unique solution of the Kolmogorov backward partial integro differential equation (PIDE)*

$$\widetilde{L}^{(0)} u(z, y) = 0 \tag{9.5.5}$$

for all $(z, y) \in (0, T) \times \mathbb{R}^d$ *with terminal condition*

$$u(T, x) = g(x) \tag{9.5.6}$$

for $x \in \mathbb{R}^d$, *with* $\widetilde{L}^{(0)}$ *defined in* (3.3.7). *Moreover, we have*

$$u(z, \cdot) \in \mathcal{C}_P^{2(\beta+1)}(\mathbb{R}^d, \mathbb{R}) \tag{9.5.7}$$

for each $z \in [0, T]$.

Proof: Mikulevicius & Platen (1988) showed that $u(z, \cdot) \in \mathcal{C}_P^{2(\beta+1)}(\mathbb{R}^d, \mathbb{R})$ for each $z \in [0, T]$. Moreover, note that by the Markov property of X we obtain that $u(z) := u(z, X_z^{0,X_0})$ equals $E(g(X_T)|\mathcal{A}_z)$ for $z \in [0, T]$. Therefore, one can show that $u(z)$ forms a martingale. By application of Itô's formula to $u(t, X_t^{s,y})$, for

$0 \leq s < t \leq T$ we obtain

$$\begin{aligned} u(t, X_t^{s,y}) &= u(s, y) + \int_s^t L^{(0)} u(z, X_z^{s,y}) dz \\ &+ \sum_{j=1}^m \int_s^t L^{(j)} u(z, X_z^{s,y}) dW_z^j + \int_s^t \int_{\mathcal{E}} L_v^{(-1)} u(z, X_z^{s,y}) p_\phi(dv, dz), \end{aligned}$$

where the operators $L^{(0)}$, $L^{(j)}$ with $j \in \{1, \ldots, m\}$ and $L_v^{(-1)}$ are defined in (3.3.4), (3.3.5) and (3.3.6), respectively.

By the martingale property of $u(t) = u(t, X_t^{0,X_0})$, the function u satisfies the PIDE $L^{(0)} u(z, y) + \int_{\mathcal{E}} L_v^{(-1)} u(z, y) \phi(dv) = \widetilde{L}^{(0)} u(z, y) = 0$ for all $(z, y) \in (0, T) \times \mathbb{R}^d$. □

Remark 9.5.2 *For simplicity, we have assumed that the Itô process* (9.5.1) *can reach any point in* $\mathbb{R}^d$. *If instead the Itô process* (9.5.1) *can take values only in a subset of* $\mathbb{R}^d$, *then* (9.5.5) *and* (9.5.6) *of Lemma 9.5.1 hold only in this subset of* $\mathbb{R}^d$. *This is sufficient since in the convergence theorems to be presented we will need the relations* (9.5.5) *and* (9.5.6) *only for values of* $(t, x) \in (0, T) \times \mathbb{R}^d$ *that can be reached by the Itô process* (9.5.1).

By an application of Itô's formula, we obtain the following result.

Lemma 9.5.3 *For all* $n \in \{1, \ldots, n_T\}$ *and* $y \in \mathbb{R}^d$, *we have*

$$E\Big(u(t_n, X_{t_n}^{t_{n-1}, y}) - u(t_{n-1}, y) \Big| \mathcal{A}_{t_{n-1}} \Big) = 0. \tag{9.5.8}$$

Proof: By Itô's formula we obtain

$$\begin{aligned} u(t_n, X_{t_n}^{t_{n-1}, y}) &= u(t_{n-1}, y) + \int_{t_{n-1}}^{t_n} L^{(0)} u(z, X_z^{t_{n-1}, y}) dz \\ &+ \sum_{j=1}^m \int_{t_{n-1}}^{t_n} L^{(j)} u(z, X_z^{t_{n-1}, y}) dW_z^j \\ &+ \int_{t_{n-1}}^{t_n} \int_{\mathcal{E}} L_v^{(-1)} u(z, X_z^{t_{n-1}, y}) p_\phi(dv, dz). \end{aligned}$$

Therefore, by applying the expected value, (3.3.4)-(3.3.7), and using (9.5.5) and properties of the Itô integral we obtain (9.5.8). □

By using the notation introduced in Section 3.4, we define for every $\beta \in \{1, 2, \ldots\}$ the hierarchical set

$$\Gamma_\beta = \{\alpha \in \mathcal{M} : l(\alpha) \leq \beta\}, \tag{9.5.9}$$

which will give us the rule for the construction of regular weak Taylor approximations of weak order β.

Given a regular time discretization $(t)_\Delta$, with maximum step size $\Delta \in (0,1)$, we define the *order β weak Taylor scheme* by the vector equation

$$Y_{n+1}^\Delta = Y_n^\Delta + \sum_{\alpha \in \Gamma_\beta \setminus \{v\}} I_\alpha \left[f_\alpha(t_n, Y_n^\Delta)\right]_{t_n, t_{n+1}} = \sum_{\alpha \in \Gamma_\beta} I_\alpha \left[f_\alpha(t_n, Y_n^\Delta)\right]_{t_n, t_{n+1}}, \tag{9.5.10}$$

for $n \in \{0, 1, \ldots, n_T - 1\}$ with $f(t,x) = x$, as defined in (3.3.8). Similarly, we define the *order β compensated weak Taylor scheme* by the vector equation

$$Y_{n+1}^\Delta = Y_n^\Delta + \sum_{\alpha \in \Gamma_\beta \setminus \{v\}} \widetilde{I}_\alpha \left[\tilde{f}_\alpha(t_n, Y_n^\Delta)\right]_{t_n, t_{n+1}} = \sum_{\alpha \in \Gamma_\beta} \widetilde{I}_\alpha \left[\tilde{f}_\alpha(t_n, Y_n^\Delta)\right]_{t_n, t_{n+1}}, \tag{9.5.11}$$

for $n \in \{0, 1, \ldots, n_T - 1\}$ with $f(t,x) = x$ in (3.3.9).

For simplicity, for the next and the following theorems we will assume an autonomous multi-dimensional jump-diffusion SDE. This formulation is not restrictive, since we can always rewrite an SDE with time-dependent coefficients as an autonomous one by modelling the time t as first component of the process X. However, these strong conditions on the time component can be relaxed in a direct proof for the non-autonomous case.

We now present the following convergence theorem which states that, under suitable conditions, for any given $\beta \in \{1, 2, \ldots\}$ the corresponding order β compensated weak Taylor scheme (9.5.11) achieves the weak order of convergence β.

Theorem 9.5.4 *For given $\beta \in \{1, 2, \ldots\}$, let $Y^\Delta = \{Y_n^\Delta, n \in \{0, 1, \ldots, n_T\}\}$ be the order β compensated weak Taylor approximation defined in (9.5.11) corresponding to a regular time discretization with maximum time step size $\Delta \in (0,1)$.*

We assume that $E(|X_0|^i) < \infty$, for $i \in \{1, 2, \ldots\}$, and for any $g \in C_P^{2(\beta+1)}(\mathbb{R}^d, \mathbb{R})$ there exists a positive constant C, independent of Δ, such that

$$|E(g(X_0)) - E(g(Y_0^\Delta))| \leq C\Delta^\beta. \tag{9.5.12}$$

Moreover, suppose that the drift, diffusion and jump coefficients are Lipschitz continuous with components $a^k, b^{k,j}, c^k \in C_P^{2(\beta+1)}(\mathbb{R}^d, \mathbb{R})$ *for all* $k \in \{1, 2 \ldots, d\}$ *and* $j \in \{1, 2 \ldots, m\}$ *and the coefficients* $\tilde{f}_\alpha$, *with* $f(t,x) = x$, *satisfy the linear growth condition*

$$|\tilde{f}_\alpha(t,x)| \leq K(1 + |x|), \tag{9.5.13}$$

with $K < \infty$, *for all* $t \in [0,T]$, $x \in \mathbb{R}^d$ *and* $\alpha \in \Gamma_\beta \bigcup \mathcal{B}(\Gamma_\beta)$.

Then for any function $g \in C_P^{2(\beta+1)}(\mathbb{R}^d, \mathbb{R})$ *there exists a positive constant* C, *independent of* Δ, *such that*

$$|E(g(X_T)) - E(g(Y_{n_T}^\Delta))| \leq C\Delta^\beta. \tag{9.5.14}$$

Remark 9.5.5 *Note that the linear growth condition* (9.5.13), *on the coefficient functions* $\tilde{f}_\alpha$, *is satisfied if, for instance, the drift, diffusion and jump coefficients are uniformly bounded.*

Remark 9.5.6 *By replacing the conditions on the compensated Itô coefficient functions* $\tilde{f}_\alpha$ *with equivalent conditions on the Itô coefficient functions* f_α, *one can show that for a given* $\beta \in \{1, 2, \ldots\}$ *also the order* β *weak Taylor scheme* (9.5.10) *attains the weak order of convergence* β.

Theorem 9.5.4 is an extension of the weak convergence theorem for diffusion SDEs presented in Kloeden & Platen (1999). The following proof of Theorem 9.5.4 has also similarities to the one given in Liu & Li (2000).

Proof: For ease of notation, when no misunderstanding is possible, we write Y for Y^Δ. By (9.5.3) and the terminal condition of the Kolmogorov backward equation (9.5.6) we can write

$$\begin{aligned} H &:= \Big|E\big(g(Y_{n_T})\big) - E\big(g(X_T)\big)\Big| \\ &= \Big|E\big(u(T, Y_{n_T}) - u(0, X_0)\big)\Big|. \end{aligned} \tag{9.5.15}$$

Moreover, by (9.5.12), (9.5.8), (9.5.7) and the deterministic Taylor expansion we obtain

$$\begin{aligned} H &\leq \left| E\left(\sum_{n=1}^{n_T} \left(u(t_n, Y_n) - u(t_{n-1}, Y_{n-1})\right) \right) \right| + K\,\Delta^\beta \\ &= \left| E\left(\sum_{n=1}^{n_T} \left(u(t_n, Y_n) - u(t_n, X_{t_n}^{t_{n-1}, Y_{n-1}})\right) \right) \right| + K\,\Delta^\beta \\ &\leq H_1 + H_2 + K\,\Delta^\beta, \end{aligned} \tag{9.5.16}$$

where

$$H_1 = \left| E\left(\sum_{n=1}^{n_T} \left\{ \sum_{i=1}^{d} \left(\frac{\partial}{\partial y^i} u(t_n, X_{t_n}^{t_{n-1}, Y_{n-1}})\right) (Y_n^i - X_{t_n}^{i; t_{n-1}, Y_{n-1}}) \right\} \right) \right| \tag{9.5.17}$$

and

$$\begin{aligned} H_2 \;=\; &\left| E\left(\sum_{n=1}^{n_T} \left\{ \sum_{i,j=1}^{d} \frac{1}{2} \left(\frac{\partial^2}{\partial y^i \partial y^j} u\big(t_n, X_{t_n}^{t_{n-1}, Y_{n-1}} + \theta_n (Y_n - X_{t_n}^{t_{n-1}, Y_{n-1}})\big)\right) \right.\right.\right. \\ &\left.\left.\left. \times (Y_n^i - X_{t_n}^{i; t_{n-1}, Y_{n-1}})(Y_n^j - X_{t_n}^{j; t_{n-1}, Y_{n-1}}) \right\} \right) \right|. \end{aligned} \tag{9.5.18}$$

Here, according to the notation introduced in Section 2, we have used a superscript to denote vector components of Y and X. Moreover, θ_n is a $d \times d$ diagonal matrix with

$$\theta_n^{k,k} \in (0,1) \tag{9.5.19}$$

for $k \in \{1, \ldots, d\}$.

Note that by Lemma 3.7.1 we have

$$X_{t_n}^{t_{n-1}, Y_{n-1}} - Y_n = \sum_{\alpha \in \mathcal{B}(\Gamma_\beta)} \tilde{I}_\alpha[\tilde{f}_\alpha(X_\cdot^{t_{n-1}, Y_{n-1}})]_{t_{n-1}, t_n}, \tag{9.5.20}$$

where $\mathcal{B}(\Gamma_\beta) = \{\alpha : l(\alpha) = \beta + 1\}$. Moreover, similar to Lemma 3.7.2, by the linear growth condition (9.5.13) on $\tilde{f}_\alpha$ one can show that for every $p \in \{1, 2, \ldots\}$ there exist constants K and r such that for every $q \in \{1, \ldots, p\}$

$$E\big(\max_{0 \leq n \leq n_T} |Y_n|^{2q}\big) \leq K\big(1 + |Y_0|^{2r}\big).$$

Therefore, by (3.7.10) and (9.5.13) for every $\alpha \in \Gamma_\beta \bigcup \mathcal{B}(\Gamma_\beta)$ and $p \in \{1, 2, \ldots\}$ there exist constants K and r such that

$$E\big(|\tilde{f}_\alpha(X_z^{t_{n-1},Y_{n-1}})|^{2q}\big) \leq K\big(1 + |Y_0|^{2r}\big), \tag{9.5.21}$$

for every $n \in \{1, \ldots, n_T\}$ and $z \in [t_{n-1}, t_n]$.

We can now apply (9.5.20) and since (9.5.21) holds, also Lemma 3.6.3, to obtain

$$\begin{aligned} H_1 &\leq E\Bigg(\sum_{n=1}^{n_T}\sum_{i=1}^{d}\sum_{\{\alpha:l(\alpha)=\beta+1\}} \\ &\qquad \Big|E\Big(\big(\frac{\partial}{\partial y^i}u(t_n, X_{t_n}^{t_{n-1},Y_{n-1}})\big)\widetilde{I}_\alpha[\tilde{f}_\alpha^i(X_\cdot^{t_{n-1},Y_{n-1}})]_{t_{n-1},t_n}\big|\mathcal{A}_{t_{n-1}}\Big)\Big|\Bigg) \\ &\leq KE\Bigg(\sum_{n=1}^{n_T}\sum_{i=1}^{d}\sum_{\{\alpha:l(\alpha)=\beta+1\}}(t_n - t_{n-1})^{\beta+1}\Bigg) \\ &\leq K\Delta^\beta. \end{aligned} \tag{9.5.22}$$

To estimate H_2 we can apply (9.5.20) and, since (9.5.21) holds, also Lemma 3.6.4 to obtain

$$\begin{aligned} H_2 &\leq \Bigg|E\Bigg(\sum_{n=1}^{n_T}\Bigg\{\sum_{i,j=1}^{d}\frac{1}{4}\Big(\frac{\partial^2}{\partial y^i \partial y^j}u\big(t_n, X_{t_n}^{t_{n-1},Y_{n-1}} + \theta_n(Y_n - X_{t_n}^{t_{n-1},Y_{n-1}})\big)\Big) \\ &\qquad \times\Big[(Y_n^i - X_{t_n}^{i;t_{n-1},Y_{n-1}})^2 + (Y_n^j - X_{t_n}^{j;t_{n-1},Y_{n-1}})^2\Big]\Bigg\}\Bigg)\Bigg| \\ &= \Bigg|E\Bigg(\sum_{n=1}^{n_T}\Bigg\{\sum_{i,j=1}^{d}\frac{1}{2}\Big(\frac{\partial^2}{\partial y^i \partial y^j}u\big(t_n, X_{t_n}^{t_{n-1},Y_{n-1}} + \theta_n(Y_n - X_{t_n}^{t_{n-1},Y_{n-1}})\big)\Big) \\ &\qquad \times(Y_n^i - X_{t_n}^{i;t_{n-1},Y_{n-1}})^2\Bigg\}\Bigg)\Bigg| \\ &\leq KE\Bigg(\sum_{n=1}^{n_T}\Bigg\{\sum_{i,j=1}^{d}\sum_{\{\alpha:l(\alpha)=\beta+1\}} \\ &\qquad \times E\Big(\Big|\frac{\partial^2}{\partial y^i \partial y^j}u\big(t_n, X_{t_n}^{t_{n-1},Y_{n-1}} + \theta_n(Y_n - X_{t_n}^{t_{n-1},Y_{n-1}}\big)\Big| \end{aligned}$$

$$\times \left| \tilde{I}_\alpha [\tilde{f}^i_\alpha (X^{t_{n-1},Y_{n-1}}_\cdot)]_{t_{n-1},t_n} \right|^2 \Big| \mathcal{A}_{t_{n-1}} \Big) \Bigg\} \Bigg)$$

$$\leq \; K \Delta^\beta. \tag{9.5.23}$$

Finally, by combining (9.5.16), (9.5.22) and (9.5.23) we complete the proof of Theorem 9.5.4. □

Remark 9.5.7 *Note that the proof of Theorem 9.5.4 holds also for the wider class of test functions $g \in C_P^{2k+1}(\mathbb{R}^d, \mathbb{R})$ with $k = \beta + 1 - [\frac{\beta+1}{2}]$, see Lemma 3.6.3.*

Chapter 10

Jump-Adapted Weak Approximations

In this chapter we consider weak approximations constructed on jump-adapted time discretizations similar to those presented in Chapter 6. Since the jump-adapted discretization includes the jump times of the Poisson measure, we can use an approximation for the pure diffusion part between discretization points. As noticed in Chapter 6, higher order jump-adapted schemes avoid multiple stochastic integrals involving the Poisson random measure. Only multiple stochastic integrals with respect to time and Wiener processes, or their simplifications, are required. This leads to easily implementable schemes. However, jump-adapted weak approximations are computationally demanding when the intensity of the Poisson measure is high.

10.1 Introduction

The weak jump-adapted schemes to be presented are constructed on jump-adapted time discretizations. We recall that a jump-adapted time discretization

$$(t)_\Delta = \{0 = t_0 < t_1 < \ldots < t_{n_T} = T\}, \tag{10.1.1}$$

with maximum step size Δ, includes all the jump times $\{\tau_1, \tau_2, \ldots, \tau_{p_\phi(T)}\}$ of the Poisson random measure p_ϕ. Moreover, as previously discussed in Chapter 6, we assume that such a jump-adapted time discretization satisfies the following conditions:

$$P(t_{n+1} - t_n \leq \Delta) = 1, \tag{10.1.2}$$

and

$$t_{n+1} \quad \text{is} \quad \mathcal{A}_{t_n} - \text{measurable}, \tag{10.1.3}$$

for $n \in \{0, 1, \ldots, n_T - 1\}$, if it is not a jump time.

10.2 Taylor Schemes

10.2.1 Euler Scheme

The simplest jump-adapted weak Taylor scheme is the *jump-adapted Euler scheme* presented in Chapter 6. We recall that in the general multi-dimensional case the kth component is given by

$$Y^k_{t_{n+1}-} = Y^k_{t_n} + a^k \Delta_{t_n} + \sum_{j=1}^{m} b^{k,j} \Delta W^j_{t_n} \tag{10.2.1}$$

and

$$Y^k_{t_{n+1}} = Y^k_{t_{n+1}-} + \int_{\mathcal{E}} c^k(t_{n+1}, Y_{t_{n+1}-}, v)\, p_\phi(dv, \{t_{n+1}\}), \tag{10.2.2}$$

where $\Delta_{t_n} = t_{n+1} - t_n$ and $\Delta W^j_{t_n} = W^j_{t_{n+1}} - W^j_{t_n} \sim \mathcal{N}(0, \Delta_{t_n})$ with $j \in \{1, \ldots, m\}$. According to the notation introduced in Chapter 4, we have used the abbreviations $a = a(Y_{t_n})$ and $b = b(Y_{t_n})$. Also in the sequel, when no misunderstanding is possible, for any coefficient function f along with its derivatives we will write $f = f(Y_{t_n})$.

The impact of jumps is generated by equation (10.2.2). If t_{n+1} is a jump time, then $\int_{\mathcal{E}} p_\phi(dv, \{t_{n+1}\}) = 1$ and

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

while if t_{n+1} is not a jump time one has $Y_{t_{n+1}} = Y_{t_{n+1}-}$, as $\int_{\mathcal{E}} p_\phi(dv, \{t_{n+1}\}) = 0$. Therefore, the weak order of convergence of the jump-adapted Euler scheme is $\beta = 1.0$ and, thus, equals the weak order of the Euler scheme used in (10.2.1) for the diffusive component.

As will be shown below, for weak convergence it is possible to replace the Gaussian distributed random variables by simpler multi-point random variables that satisfy certain moment conditions. For instance, in the Euler scheme we can replace the random variables $\Delta W^j_{t_n}$ by simpler random variables $\Delta \widehat{W}^j_{t_n}$ that satisfy the moment condition

$$|E(\Delta \widehat{W}_{t_n})| + |E\big((\Delta \widehat{W}_{t_n})^3\big)| + |E\big((\Delta \widehat{W}_{t_n})^2\big) - \Delta_{t_n}| \leq K \Delta^2 \tag{10.2.3}$$

for some constant K independent of Δ. In this case the order of weak convergence of the Euler scheme is still $\beta = 1.0$. For instance, we can replace the random variables $\Delta W^j_{t_n}$ by the simpler two-point distributed random variables $\Delta \widehat{W}^j_{2,t_n}$, with

$$P(\Delta \widehat{W}^j_{2,t_n} = \pm\sqrt{\Delta_{t_n}}) = \frac{1}{2}, \tag{10.2.4}$$

for $j \in \{1,\ldots,m\}$. This yields the *jump-adapted simplified Euler scheme*, whose kth component is given by

$$Y^k_{t_{n+1}-} = Y^k_{t_n} + a^k \Delta_{t_n} + \sum_{j=1}^{m} b^{k,j} \Delta \widehat{W}^j_{2,t_n} \tag{10.2.5}$$

together with (10.2.2). This scheme achieves weak order of convergence $\beta = 1.0$. In this way, we have obtained a scheme with the same order of weak convergence of the Euler scheme that requires only the generation of simple two-point distributed random variables. As will be discussed in Chapter 12, multi-point distributed random variables as (10.2.4) can be generated very efficiently via random bit generators, see Bruti-Liberati & Platen (2004). Moreover, the use of hardware accelerators, see Bruti-Liberati, Martini, Piccardi & Platen (2007), can further improve the speed of the simulation of multi-point distributed random variables.

10.2.2 Order 2.0 Taylor Scheme

By using an order 2.0 weak Taylor scheme for the diffusive part, we can derive the *jump-adapted order* 2.0 *weak Taylor scheme*. In the one-dimensional case, $d = m = 1$, it is given by

$$\begin{aligned} Y_{t_{n+1}-} &= Y_{t_n} + a\Delta_{t_n} + b\Delta W_{t_n} + \frac{bb'}{2}\Big((\Delta W_{t_n})^2 - \Delta_{t_n}\Big) + a'b\Delta Z_{t_n} \\ &\quad + \frac{1}{2}\left(\frac{\partial a}{\partial t} + aa' + \frac{a''}{2}b^2\right)\Delta^2_{t_n} + \left(\frac{\partial b}{\partial t} + ab' + \frac{b''}{2}b^2\right)\{\Delta W_{t_n}\Delta_{t_n} - \Delta Z_{t_n}\} \end{aligned} \tag{10.2.6}$$

and

$$Y_{t_{n+1}} = Y_{t_{n+1}-} + \int_{\mathcal{E}} c(t_{n+1}, Y_{t_{n+1}-}, v)\, p_\phi(dv, \{t_{n+1}\}), \tag{10.2.7}$$

and achieves weak order $\beta = 2.0$. Here ΔZ_{t_n} represents the double Itô integral defined in (6.2.22). Therefore, the required random variables ΔW_{t_n} and ΔZ_{t_n} can be generated as in (6.2.23). The jump-adapted order 2.0 weak Taylor scheme was first presented in Mikulevicius & Platen (1988). It generalizes a second order weak scheme for pure diffusion presented in Milstein (1978).

If we compare the regular order 2.0 weak Taylor scheme (9.3.1) with the jump-adapted order 2.0 weak Taylor scheme (10.2.6)–(10.2.7), then we notice that the latter is much simpler since it avoids multiple stochastic integrals with respect to the Poisson measure.

Also in this case, since we are constructing weak schemes, we have some freedom in the choice of the underlying random variable. The *jump-adapted simplified order 2.0 weak scheme* is given by

$$\begin{aligned} Y_{t_{n+1}-} &= Y_{t_n} + a\Delta_{t_n} + b\Delta\widehat{W}_{t_n} + \frac{bb'}{2}\left\{\left(\Delta\widehat{W}_{t_n}\right)^2 - \Delta_{t_n}\right\} \\ &+\frac{1}{2}\left(\frac{\partial a}{\partial t} + aa' + \frac{a''}{2}b^2\right)\Delta_{t_n}^2 + \frac{1}{2}\left(\frac{\partial b}{\partial t} + a'b + ab' + \frac{b''}{2}b^2\right)\Delta\widehat{W}_{t_n}\Delta_{t_n} \end{aligned} \tag{10.2.8}$$

and (10.2.7). To obtain second order of weak convergence, the random variable $\Delta\widehat{W}_{t_n}$ should be $\mathcal{A}_{t_{n+1}}$-measurable and should satisfy the moment condition

$$\begin{aligned} &|E(\Delta\widehat{W}_{t_n})| + |E((\Delta\widehat{W}_{t_n})^3)| + |E((\Delta\widehat{W}_{t_n})^5)| \\ &\quad +|E((\Delta\widehat{W}_{t_n})^2) - \Delta_{t_n}| + |E((\Delta\widehat{W}_{t_n})^4) - 3\Delta_{t_n}^2| \leq K\Delta^3 \end{aligned} \tag{10.2.9}$$

for some constant K independent of Δ. Note that if we choose independent random variables $\Delta\widehat{W}_{t_n}$, $n \in \{0, \ldots, n_T - 1\}$, we automatically obtain the required $\mathcal{A}_{t_{n+1}}$-measurability. The moment condition (10.2.9) is satisfied, for instance, by a Gaussian $\mathcal{N}(0, \Delta_{t_n})$ distributed random variable, but also by a three-point distributed random variable $\Delta\widehat{W}_{3,t_n}$, where

$$P(\Delta\widehat{W}_{3,t_n} = \pm\sqrt{3\Delta_{t_n}}) = \frac{1}{6}, \qquad P(\Delta\widehat{W}_{3,t_n} = 0) = \frac{2}{3}. \tag{10.2.10}$$

Note that the jump-adapted simplified order 2.0 weak scheme (10.2.8) requires only one random variable at each time step. Therefore, it is computationally more efficient than the jump-adapted order 2.0 weak Taylor scheme (10.2.6). Moreover, when using three-point distributed random variables $\Delta\widehat{W}_{3,t_n}$, some highly efficient implementations via random bit generators and hardware accelerators can be used, as will be described later in Chapter 12.

In the multi-dimensional case with a scalar Wiener process, $m = 1$, the kth component of the *jump-adapted order* 2.0 *weak Taylor scheme* is given by

$$\begin{aligned} Y^k_{t_{n+1}-} &= Y^k_{t_n} + a^k \Delta_{t_n} + b^k \Delta W_{t_n} + \frac{L^{(1)} b^k}{2}\Big((\Delta W_{t_n})^2 - \Delta_{t_n}\Big) + L^{(1)} a^k \Delta Z_{t_n} \\ &+ \frac{L^{(0)} a^k}{2} \Delta^2_{t_n} + L^{(0)} b^k \left\{ \Delta W_{t_n} \Delta_{t_n} - \Delta Z_{t_n} \right\} \end{aligned} \tag{10.2.11}$$

and (10.2.2), where the operators $L^{(0)}$ and $L^{(1)}$ are defined in (3.3.4) and (3.3.5).

In the general multi-dimensional case, the kth component of the *jump-adapted order* 2.0 *weak Taylor scheme* is of the form

$$\begin{aligned} Y^k_{t_{n+1}-} &= Y^k_{t_n} + a^k \Delta_{t_n} + \frac{L^{(0)} a^k}{2} \Delta^2_{t_n} \qquad (10.2.12) \\ &+ \sum_{j=1}^{m} \left(b^{k,j} \Delta W^j_{t_n} + L^{(0)} b^{k,j} I_{(0,j)} + L^{(j)} a^k I_{(j,0)} \right) + \sum_{j_1,j_2=1}^{m} L^{(j_1)} b^{k,j_2} I_{(j_1,j_2)} \end{aligned}$$

and (10.2.2), where the operators $L^{(0)}$ and $L^{(j)}$, with $j \in \{1,\ldots,m\}$, are as above. The multiple stochastic integrals $I_{(0,j)}$ and $I_{(j,0)}$, for $j \in \{1,2,\ldots,m\}$, can be easily generated as in (9.3.10). However, the generation of the multiple stochastic integrals involving two components of the Wiener process, $I_{(j_1,j_2)}$, with $j_1, j_2 \in \{1,2,\ldots,m\}$, is not straightforward. In the special case of commutative diffusion noise (4.4.17), one can express these multiple stochastic integrals in terms of the Gaussian increments of the Wiener processes $\Delta W^j_{t_n}$. In the general case, as explained in Chapter 4, one could use an approximation such as the Karhunen-Loève expansion. However, since we are interested in a weak approximation, we can instead replace the multiple stochastic integrals by corresponding simple multi-point distributed random variables satisfying certain moment conditions. In this way, we obtain the *jump-adapted simplified order* 2.0 *weak scheme* with kth component of the form

$$\begin{aligned} Y^k_{t_{n+1}-} &= Y^k_{t_n} + a^k \Delta_{t_n} + \frac{L^{(0)} a^k}{2} \Delta^2_{t_n} + \sum_{j=1}^{m} \left\{ b^{k,j} + \frac{\Delta_{t_n}}{2}\left(L^{(0)} b^{k,j} + L^{(j)} a^k\right)\right\} \Delta \widehat{W}^j_{t_n} \\ &+ \frac{1}{2} \sum_{j_1,j_2=1}^{m} L^{(j_1)} b^{k,j_2} \left(\Delta \widehat{W}^{j_1}_{t_n} \Delta \widehat{W}^{j_2}_{t_n} + V^{j_1,j_2}_{t_n} \right) \end{aligned} \tag{10.2.13}$$

and (10.2.2), where $\Delta \widehat{W}^j_{t_n}$, with $j \in \{1,2,\ldots,m\}$, are independent random variables such as (10.2.10), satisfying the moment condition (10.2.9). Additionally,

$V_{t_n}^{j_1,j_2}$ are independent two-point distributed random variables given by

$$P(V_{t_n}^{j_1,j_2} = \pm\Delta_{t_n}) = \frac{1}{2}, \tag{10.2.14}$$

for $j_2 \in \{1, \ldots, j_1 - 1\}$

$$V_{t_n}^{j_1,j_1} = -\Delta_{t_n} \tag{10.2.15}$$

and

$$V_{t_n}^{j_1,j_2} = -V_{t_n}^{j_2,j_1} \tag{10.2.16}$$

for $j_2 \in \{j_1 + 1, \ldots, m\}$ and $j_1 \in \{1, \ldots, m\}$.

10.2.3 Order 3.0 Taylor Scheme

By including in the diffusive component all multiple stochastic integrals of multiplicity three with respect to time and Wiener processes, one obtains the *jump-adapted order* 3.0 *weak Taylor scheme.* In the general multi-dimensional case, its kth component is given by

$$\begin{aligned} Y_{t_{n+1}-}^k &= Y_{t_n}^k + a^k \Delta_{t_n} + \frac{L^{(0)}a^k}{2}\Delta_{t_n}^2 \\ &+ \sum_{j=1}^m \left(b^{k,j}\Delta W_{t_n}^j + L^{(0)}b^{k,j} I_{(0,j)} + L^{(j)}a^k I_{(j,0)} \right) + \sum_{j_1,j_2=1}^m L^{(j_1)}b^{k,j_2} I_{(j_1,j_2)} \\ &+ \sum_{j_1,j_2=0}^m L^{(j_1)}L^{(j_2)}a^k I_{(j_1,j_2,0)} + \sum_{j_1,j_2=0}^m \sum_{j_3=1}^m L^{(j_1)}L^{(j_2)}b^{k,j_3} I_{(j_1,j_2,j_3)} \end{aligned} \tag{10.2.17}$$

and (10.2.2). Because of the difficulties in the generation of multiple stochastic integrals involving different components of the Wiener process, this scheme is usually too complex for a practical implementation. However, in the following we will consider some special cases that lead to implementable schemes with third order of weak convergence. These schemes generalize those in Platen (1984), presented for pure diffusion SDEs.

In the one-dimensional case, $d = m = 1$, by approximating the multiple stochastic integrals with Gaussian random variables, we obtain a *jump-adapted order* 3.0 *weak* *scheme* given by

$$
\begin{aligned}
Y_{n+1} \;=\; & Y_n + a\Delta + b\Delta W_{t_n} + \frac{1}{2}L^{(1)}b\left\{(\Delta W_{t_n})^2 - \Delta_{t_n}\right\} \\
& + L^{(1)}a\Delta Z_{t_n} + \frac{1}{2}L^{(0)}a(\Delta_{t_n})^2 + L^{(0)}b\left\{\Delta W_{t_n}\Delta_{t_n} - \Delta Z_{t_n}\right\} \\
& + \frac{1}{6}\left(L^{(0)}L^{(0)}b + L^{(0)}L^{(1)}a + L^{(1)}L^{(0)}a\right)\left\{\Delta W_{t_n}(\Delta_{t_n})^2\right\} \\
& + \frac{1}{6}\left(L^{(1)}L^{(1)}a + L^{(1)}L^{(0)}b + L^{(0)}L^{(1)}b\right)\left\{(\Delta W_{t_n})^2 - \Delta_{t_n}\right\}\Delta_{t_n} \\
& + \frac{1}{6}L^{(0)}L^{(0)}a(\Delta_{t_n})^3 + \frac{1}{6}L^{(1)}L^{(1)}b\left\{(\Delta W_{t_n})^2 - 3\,\Delta_{t_n}\right\}\Delta W_{t_n}
\end{aligned}
\tag{10.2.18}
$$

and (10.2.2), where ΔW_{t_n} and ΔZ_{t_n} are the correlated Gaussian random variables of the type defined in (6.2.23).

To construct a third order simplified method, the required simplified random variables $\Delta\widehat{W}_n$ need to satisfy the following moment condition

$$
\begin{aligned}
& |E(\Delta\widehat{W}_{t_n})| + |E\big((\Delta\widehat{W}_{t_n})^3\big)| + |E\big((\Delta\widehat{W}_{t_n})^5\big)| + |E\big((\Delta\widehat{W}_{t_n})^7\big)| \\
& + |E\big((\Delta\widehat{W}_{t_n})^2\big) - \Delta_{t_n}| + |E\big((\Delta\widehat{W}_{t_n})^4\big) - 3\Delta_{t_n}^2| \\
& + |E\big((\Delta\widehat{W}_{t_n})^6\big) - 15\Delta_{t_n}^3| \le K\Delta^3,
\end{aligned}
\tag{10.2.19}
$$

see Theorem 10.7.1.

In Hofmann (1994) a four-point distributed random variable $\Delta\widehat{W}_{4,t_n}$ that satisfies condition (10.2.19) was proposed, where

$$
\begin{aligned}
P(\Delta\widehat{W}_{4,t_n} = \pm\sqrt{3+\sqrt{6}}\,\sqrt{\Delta_{t_n}}) &= \frac{1}{12+4\sqrt{6}}, \\
P(\Delta\widehat{W}_{4,t_n} = \pm\sqrt{3-\sqrt{6}}\,\sqrt{\Delta_{t_n}}) &= \frac{1}{12-4\sqrt{6}}.
\end{aligned}
\tag{10.2.20}
$$

However, since the probabilities in (10.2.20) are not rational numbers, the corresponding four-point distributed random variable cannot be efficiently implemented by the method based on random bit generation that will be discussed in Chapter 12. Instead, we present an alternative five-point distributed random variable

$\Delta\widehat{W}_{5,t_n}$, with

$$P(\Delta\widehat{W}_{5,t_n} = \pm\sqrt{6\Delta_{t_n}}) = \frac{1}{30}, \qquad P(\Delta\widehat{W}_{5,t_n} = \pm\sqrt{\Delta_{t_n}}) = \frac{9}{30},$$
$$P(\Delta\widehat{W}_{5,t_n} = 0) = \frac{1}{3}, \tag{10.2.21}$$

that still satisfies condition (10.2.19) and is suitable for a highly efficient implementation based on random bit generators, see Bruti-Liberati, Martini, Piccardi & Platen (2007). This five-point distributed random variable has been independently suggested in Milstein & Tretyakov (2004). The *jump-adapted simplified order* 3.0 *weak scheme* is then given by

$$\begin{aligned} Y_{n+1} &= Y_n + a\Delta + b\Delta\widehat{W}_{5,t_n} + \frac{1}{2}L^{(1)}b\left\{\left(\Delta\widehat{W}_{5,t_n}\right)^2 - \Delta_{t_n}\right\} + \frac{1}{2}L^{(0)}a\Delta_{t_n}^2 \\ &+\frac{1}{2}L^{(1)}a\left\{\Delta\widehat{W}_{5,t_n} + \frac{1}{\sqrt{3}}\Delta\widehat{W}_{2,t_n}\right\}\Delta_{t_n} \\ &+\frac{1}{2}L^{(0)}b\left\{\Delta\widehat{W}_{5,t_n} - \frac{1}{\sqrt{3}}\Delta\widehat{W}_{2,t_n}\right\}\Delta_{t_n} \\ &+\frac{1}{6}\left(L^{(0)}L^{(0)}b + L^{(0)}L^{(1)}a + L^{1}L^{(0)}a\right)\Delta\widehat{W}_{5,t_n}\Delta_{t_n}^2 \\ &+\frac{1}{6}\left(L^{(1)}L^{(1)}a + L^{(1)}L^{(0)}b + L^{(0)}L^{(1)}b\right)\left\{(\Delta\widehat{W}_{5,t_n})^2 - \Delta_{t_n}\right\}\Delta_{t_n} \\ &+\frac{1}{6}L^{(0)}L^{(0)}a\Delta_{t_n}^3 + \frac{1}{6}L^{(1)}L^{(1)}b\left\{(\Delta\widehat{W}_{5,t_n})^2 - 3\,\Delta_{t_n}\right\}\Delta\widehat{W}_{5,t_n}, \end{aligned} \tag{10.2.22}$$

and (10.2.2). It involves the five-point distributed random variables $\Delta\widehat{W}_{5,t_n}$ and the two-point distributed random variables $\Delta\widehat{W}_{2,t_n}$, see (10.2.4), and achieves an order of weak convergence $\beta = 3.0$.

10.3 Derivative-Free Schemes

In the previous section we explained how to construct jump-adapted weak Taylor schemes with higher order of weak convergence β. However, the jump-adapted order 2.0 and order 3.0 weak Taylor schemes require the computation of derivatives of the drift and diffusion coefficients. In this section we present higher order weak schemes that avoid the computation of derivatives.

In the one-dimensional, autonomous case, $d = m = 1$, the *jump-adapted order* 2.0

derivative-free weak scheme is given by

$$
\begin{aligned}
Y_{t_{n+1}-} &= Y_{t_n} + \frac{1}{2}\Big(a(\overline{Y}_{t_n}) + a\Big)\Delta_{t_n} + \frac{1}{4}\Big(b(\overline{Y}^+_{t_n}) + b(\overline{Y}^-_{t_n}) + 2b\Big)\Delta W_{t_n} \\
&\quad + \frac{1}{4}\Big(b(\overline{Y}^+_{t_n}) - b(\overline{Y}^-_{t_n})\Big)\Big((\Delta W_{t_n})^2 - \Delta_{t_n}\Big)\Big(\Delta_{t_n}\Big)^{-\frac{1}{2}}
\end{aligned}
$$

and

$$
Y_{t_{n+1}} = Y_{t_{n+1}-} + \int_{\mathcal{E}} c(t_{n+1}, Y_{t_{n+1}-}, v)\, p_\phi(dv, \{t_{n+1}\}), \tag{10.3.1}
$$

with supporting values

$$
\overline{Y}_{t_n} = Y_{t_n} + a\Delta_{t_n} + b\Delta W_{t_n},
$$

and

$$
\overline{Y}^{\pm}_{t_n} = Y_{t_n} + a\Delta_{t_n} \pm b\sqrt{\Delta_{t_n}}.
$$

This scheme attains weak order $\beta = 2.0$ and generalizes a scheme for pure diffusions presented in Platen (1984).

If we replace the Gaussian random variables ΔW_{t_n} by the three-point distributed random variables $\Delta\widehat{W}_{3,t_n}$ defined in (10.2.10), we obtain the *jump-adapted simplified order* 2.0 *derivative-free weak scheme* that still achieves weak order $\beta = 2.0$.

In the autonomous, multi-dimensional case, the kth component of the *jump-adapted order* 2.0 *derivative-free weak scheme* is given by

$$
\begin{aligned}
Y^k_{t_{n+1}-} &= Y^k_{t_n} + \frac{1}{2}\big(a^k(\overline{Y}_{t_n}) + a^k\big)\Delta_{t_n} \\
&\quad + \frac{1}{4}\sum_{j=1}^{m}\Bigg\{\big(b^{k,j}(\overline{R}^{+,j}_{t_n}) + b^{k,j}(\overline{R}^{-,j}_{t_n}) + 2b^{k,j}\big)\Delta\widehat{W}^j_{t_n} \\
&\qquad + \sum_{\substack{r=1 \\ r\neq j}}^{m} \big(b^{k,j}(\overline{U}^{+,r}_{t_n}) + b^{k,j}(\overline{U}^{-,r}_{t_n}) + 2b^{k,j}\big)\Delta\widehat{W}^j_{t_n}\,\big(\Delta_{t_n}\big)^{-\frac{1}{2}}\Bigg\}
\end{aligned}
$$

$$+\frac{1}{4}\sum_{j=1}^{m}\Bigg\{\left(b^{k,j}(\overline{R}_{t_n}^{+,j}) - b^{k,j}(\overline{R}_{t_n}^{-,j})\right)\left(\left(\Delta\widehat{W}_{t_n}^{j}\right)^2 - \Delta_{t_n}\right) \tag{10.3.2}$$

$$+\sum_{\substack{r=1\\ r\neq j}}^{m}\left(b^{k,j}(\overline{U}_{t_n}^{+,r}) - b^{k,j}(\overline{U}_{t_n}^{-,r})\right)\left(\Delta\widehat{W}_{t_n}^{j}\,\Delta\widehat{W}_{t_n}^{r} + V_{t_n}^{r,j}\right)\Bigg\}$$

and

$$Y_{t_{n+1}}^{k} = Y_{t_{n+1}-}^{k} + \int_{\mathcal{E}} c^k(t_{n+1}, Y_{t_{n+1}-}, v)\, p_\phi(dv, \{t_{n+1}\}), \tag{10.3.3}$$

for $k \in \{1, \ldots, d\}$. The supporting values are

$$\overline{Y}_{t_n} = Y_{t_n} + a\Delta_{t_n} + \sum_{j=1}^{m} b^j \Delta\widehat{W}_{t_n}^{j},$$

$$\overline{R}_{t_n}^{\pm} = Y_{t_n} + a\Delta_{t_n} \pm b^j\sqrt{\Delta_{t_n}},$$

and

$$\overline{U}_{t_n}^{\pm} = Y_{t_n} \pm b^j\sqrt{\Delta_{t_n}}.$$

Moreover, $\Delta\widehat{W}_{t_n}^{j}$ should be an independent $\mathcal{A}_{t_{n+1}}$-measurable random variable satisfying the moment condition (10.2.9). For instance, one can choose independent Gaussian $\mathcal{N}(0, \Delta_{t_n})$ distributed random variables or three-point distributed random variables of the form (10.2.10). Furthermore, $V_{t_n}^{r,j}$ are the previously used two-point distributed random variables defined in (10.2.14)-(10.2.16).

10.4 Predictor-Corrector Schemes

As discussed in Chapter 5, it is important to develop schemes with wide regions of numerical stability. Predictor-corrector schemes have good numerical stability properties as discussed in Chapter 5 for the case of strong approximations. Furthermore, they are computationally efficient. Since the difference between the predicted and the corrected values provides an indication of the local error, one can also design advanced schemes with step size control.

The diffusive component of jump-adapted weak predictor-corrector schemes is derived in the following way. The corrector equation is obtained by using the Wagner-Platen expansion for pure diffusions in implicit form. However, this implicit scheme is made explicit by using the predictor equation to generate the next value. Moreover, the predictor component of the scheme is derived by the Wagner-Platen expansion for pure diffusions.

10.4.1 Order 1.0 Predictor-Corrector Scheme

In the one-dimensional case, $d = m = 1$, the *jump-adapted predictor-corrector Euler scheme* is given by the corrector

$$Y_{t_{n+1}-} = Y_{t_n} + \frac{1}{2}\Big\{a(\bar{Y}_{t_{n+1}-}) + a\Big\}\Delta_{t_n} + b\Delta W_{t_n}, \tag{10.4.1}$$

the predictor

$$\bar{Y}_{t_{n+1}-} = Y_{t_n} + a\Delta_{t_n} + b\Delta W_{t_n}, \tag{10.4.2}$$

and the jump condition

$$Y_{t_{n+1}} = Y_{t_{n+1}-} + \int_{\mathcal{E}} c(t_{n+1}, Y_{t_{n+1}-}, v)\, p_\phi(dv, \{t_{n+1}\}). \tag{10.4.3}$$

By introducing some effects of implicitness in the diffusion part of the corrector, we obtain, for $\theta, \eta \in [0,1]$, a *family of jump-adapted order* 1.0 *weak predictor-corrector schemes* with corrector

$$Y_{t_{n+1}-} = Y_{t_n} + \Big\{\theta\bar{a}(\bar{Y}_{t_{n+1}-}) + (1-\theta)\,\bar{a}\Big\}\Delta_{t_n} + \Big\{\eta b(\bar{Y}_{t_{n+1}-}) + (1-\eta)\,b\Big\}\Delta W_{t_n},$$

where $\bar{a} = a - \eta b b'$, predictor

$$\bar{Y}_{t_{n+1}-} = Y_{t_n} + a\Delta_{t_n} + b\Delta W_{t_n},$$

and (10.4.3). Note that this family of Euler schemes coincides with the family of jump-adapted predictor-corrector Euler schemes (6.5.1)-(6.5.2), which achieve strong order $\gamma = 0.5$. However, the order of weak convergence of this family of schemes equals $\beta = 1.0$. Moreover, if we replace the Gaussian random variables ΔW_{t_n} by the two-point distributed ones $\Delta \widehat{W}_{2,t_n}$ defined in (10.2.4), we obtain a family of *jump-adapted simplified order* 1.0 *weak predictor-corrector schemes* that still attain weak order $\beta = 1.0$. The family of order 1.0 weak predictor-corrector schemes generalizes the family of schemes presented in Platen (1995) for pure diffusion SDEs.

In the general multi-dimensional case, we can construct a *family of jump-adapted order* 1.0 *weak predictor-corrector schemes.* Its kth component is given by the corrector

$$\begin{aligned} Y^k_{t_{n+1}-} &= Y^k_{t_n} + \left\{\theta \bar{a}^k_\eta(t_{n+1}, \bar{Y}_{t_{n+1}-}) + (1-\theta)\bar{a}^k_\eta\right\}\Delta_{t_n} \\ &+ \sum_{j=1}^{m} \left\{\eta b^{k,j}(t_{n+1}, \bar{Y}_{t_{n+1}}) + (1-\eta) b^{k,j}\right\}\Delta \widehat{W}^j_{t_n}, \end{aligned} \tag{10.4.4}$$

for $\theta, \eta \in [0,1]$, where

$$\bar{a}_\eta = a - \eta \sum_{j_1, j_2 = 1}^{m} \sum_{i=1}^{d} b^{k,j_1} \frac{\partial b^{k,j_2}}{\partial x^i}. \tag{10.4.5}$$

Here we use the predictor

$$\bar{Y}^k_{t_{n+1}-} = Y^k_{t_n} + a^k \Delta_{t_n} + \sum_{j=1}^{m} b^{k,j} \Delta \widehat{W}^j_{t_n}, \tag{10.4.6}$$

and the jump condition

$$Y^k_{t_{n+1}} = Y^k_{t_{n+1}-} + \int_{\mathcal{E}} c^k(t_{n+1}, Y_{t_{n+1}-}, v)\, p_\phi(dv, \{t_{n+1}\}). \tag{10.4.7}$$

The random variables $\Delta \widehat{W}^j_{t_n}$, with $j \in \{1, \ldots, m\}$ and $n \in \{0, 1, \ldots n_T - 1\}$, can be chosen, for instance, as independent Gaussian $\mathcal{N}(0, \Delta_{t_n})$ distributed random variables or two-point distributed random variables (10.2.4).

10.4.2 Order 2.0 Predictor-Corrector Scheme

By using higher order weak schemes for the diffusive component of the predictor and the corrector algorithms, we obtain predictor-corrector schemes of higher weak order. In the autonomous one-dimensional case, $d = m = 1$, a *jump-adapted order* 2.0 *weak predictor-corrector scheme* is obtained by setting as corrector

$$Y_{t_{n+1}-} = Y_{t_n} + \frac{1}{2}\Big\{a(\bar{Y}_{t_{n+1}-}) + a\Big\}\Delta_{t_n} + \Psi_{t_n},$$

with

$$\Psi_{t_n} = b\Delta W_{t_n} + \frac{b\,b'}{2}\Big\{\big(\Delta W_{t_n}\big)^2 - \Delta_n\Big\} + \frac{1}{2}\Big\{ab' + \frac{1}{2}b''b^2\Big\}\Delta W_{t_n}\,\Delta_{t_n}$$

and as predictor

$$\bar{Y}_{t_{n+1}-} = Y_{t_n} + a\Delta_{t_n} + \Psi_{t_n} + \frac{1}{2}a'\,b\Delta W_{t_n}\,\Delta_n + \frac{1}{2}\Big\{aa' + \frac{1}{2}a''b^2\Big\}\big(\Delta_{t_n}\big)^2$$

together with (10.4.3). Here the Gaussian random variable ΔW_{t_n} can be replaced, for instance, by the three-point distributed random variable $\Delta\widehat{W}_{3,t_n}$ defined in (10.2.10).

In the general multi-dimensional case, the kth component of the *jump-adapted order* 2.0 *weak predictor-corrector scheme* has corrector

$$Y^k_{t_{n+1}-} = Y^k_{t_n} + \frac{1}{2}\Big\{a^k(t_{n+1}, \bar{Y}_{t_{n+1}-}) + a^k\Big\}\Delta_{t_n} + \Psi^k_{t_n},$$

with

$$\begin{aligned}\Psi^k_{t_n} &= \sum_{j=1}^{m}\Big\{b^{k,j} + \frac{1}{2}L^{(0)}b^{k,j}\Delta_{t_n}\Big\}\Delta\widehat{W}^j_{t_n} \\ &\quad + \frac{1}{2}\sum_{j_1,j_2=1}^{m} L^{(j_1)}b^{k,j_2}\Big(\Delta\widehat{W}^{j_1}_{t_n}\,\Delta\widehat{W}^{j_2}_{t_n} + V^{j_1,j_2}_{t_n}\Big),\end{aligned}$$

predictor

$$\bar{Y}^k_{t_{n+1}-} = Y^k_{t_n} + a^k\Delta_{t_n} + \Psi^k_{t_n} + \frac{1}{2}L^{(0)}a^k\big(\Delta_{t_n}\big)^2 + \frac{1}{2}\sum_{j=1}^{m} L^{(j)}a^k\Delta\widehat{W}^j_{t_n}\,\Delta_{t_n}$$

and uses (10.4.7). The random variables $\Delta\widehat{W}^j_{t_n}$, with $j \in \{1,\dots,m\}$ and $n \in \{0,1,\dots,n_T-1\}$, can be chosen as independent Gaussian $\mathcal{N}(0,\Delta_{t_n})$ distributed random variables or three-point distributed random variables (10.2.10). The two-point distributed random variables $V^{j_1,j_2}_{t_n}$, $j_1, j_2 \in \{1,\dots,m\}$, are defined in (10.2.14)–(10.2.16).

10.5 Exact Schemes

We now discuss a special class of SDEs for which it is possible to develop jump-adapted schemes that do not generate any discretization error. Similar to Section 6.6, we introduce a specific jump-adapted time discretization, obtained by a superposition of jump times generated by the Poisson measure and times at which we are interested in sampling the simulated values of the solution X. Between jump times the SDE is assumed to have an explicit transition density.

Let us consider the d-dimensional jump-diffusion SDE

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t + \int_{\mathcal{E}} c(t, X_{t-}, v)\, p_\phi(dv, dt) \tag{10.5.1}$$

for $t \in [0, T]$ and $X_0 \in \mathbb{R}^d$, that we aim to solve. We have already discussed in Section 6.6 the case when the corresponding diffusion SDE

$$dZ_t = a(t, Zt)dt + b(t, Z_t)dW_t, \tag{10.5.2}$$

admits an explicit solution. Now we assume less because we only require that the transition density of Z_t is explicitly known. Then we can construct an exact jump-adapted weak scheme.

Since we are here interested in weak approximations, we can construct exact jump-adapted schemes for a wider class of jump-diffusion SDEs than in Section 6.6. Let us first present an illustrative example: Consider the jump-diffusion SDE given by

$$dX_t = \alpha(b - X_t)dt + \sigma\sqrt{X_t}dW_t + \int_{\mathcal{E}} c(t, X_{t-}, v)\, p_\phi(dv, dt), \tag{10.5.3}$$

where $\alpha, b, \sigma \in \mathbb{R}$ and $2\alpha b > \sigma^2$. In this case the corresponding diffusion, given by

$$dZ_t = \alpha(b - Z_t)dt + \sigma\sqrt{Z_t}dW_t, \tag{10.5.4}$$

describes a square-root process, see Cox, Ingersoll & Ross (1985). The transition density of Z is known in closed form. The distribution of Z_t given Z_s, with $s < t$, is a non-central chi-square distribution, see Platen & Heath (2006). Therefore, we can construct a jump-adapted weak scheme given by

$$Y_{t_{n+1}-} = \frac{\sigma^2(1 - e^{-\alpha\Delta_{t_n}})}{4\alpha}\chi^2(\delta, l_{t_n}), \tag{10.5.5}$$

and

$$Y_{t_{n+1}} = Y_{t_{n+1}-} + \int_{\mathcal{E}} c(t_{n+1}, Y_{t_{n+1}-}, v)\, p_\phi(dv, \{t_{n+1}\}). \tag{10.5.6}$$

Here $\chi^2(\delta, l_{t_n})$, with $n \in \{0, \ldots, n_T - 1\}$, denote independent non-central chi square distributed random variables with degrees of freedom

$$\delta = \frac{4\alpha b}{\sigma^2}$$

and non-centrality parameter

$$l_{t_n} = \frac{4\alpha e^{-\alpha\Delta_{t_n}}}{\sigma^2(1 - e^{-\alpha\Delta_{t_n}})} Y_{t_n}.$$

In this way, the jump-adapted weak scheme (10.5.5)-(10.5.6) is exact in the sense that it does not generate any weak error. In fact, we obtain the result that at the discretization points the distribution of the numerical approximation Y coincides with that of the solution X of (10.5.3).

This approach can be generalized to the case of d-dimensional jump-diffusion SDEs of the type (10.5.1). In modelling, one should check wether the dynamics under consideration could be modelled by an SDE that belongs to the special subclass of jump-diffusion SDEs whose corresponding diffusion SDE (10.5.2) admits a closed form transition density.

10.6 Convergence of Jump-Adapted Weak Taylor Approximations

In this section we present a convergence theorem for jump-adapted weak Taylor approximations of any weak order of convergence $\beta \in \{1, 2 \ldots\}$. This theorem covers the convergence of the schemes presented in Section 10.3. The results of this section resemble in most parts those from Mikulevicius & Platen (1988) and are here included for completeness. Moreover, they will be needed in the next section, where we will generalize these results to general jump-adapted approximations.

As suggested in Platen (1982a), we define a *jump-adapted time discretization* $0 = t_0 < t_1 < \ldots < t_{n_T} = T$ with maximum step size $\Delta \in (0, 1)$, with n_t defined in (4.1.8). The term *jump-adapted* indicates that all jump times $\{\tau_1, \tau_2, \ldots \tau_{p_\phi(T)}\}$

of the Poisson measure p_ϕ are included in the time discretization. Moreover, we require a maximum step size $\Delta \in (0,1)$, which means that $P(t_{n+1} - t_n \leq \Delta) = 1$ for every $n \in \{0, 1, 2, \ldots, n_T - 1\}$, and, if the discretization time t_{n+1} is not a jump time, then t_{n+1} should be $\mathcal{A}_{t_n}$-measurable. Let us also introduce an additional filtration

$$\widetilde{\mathcal{A}}_{t_n} := \sigma\big(\mathbf{1}_{p_\phi(\mathcal{E},\{t_{n+1}\})\neq 0}\big) \vee \mathcal{A}_{t_n} \tag{10.6.1}$$

for every $n \in \{0, 1, \ldots, n_T\}$. We assume that t_{n+1} is $\widetilde{\mathcal{A}}_{t_n}$-measurable. We also assume a finite number of time discretization points, which means $n_t < \infty$ a.s. for $t \in [0, T]$. The superposition of the jump times with an equidistant time discretization, as discussed in Chapter 4, provides an example of such jump-adapted time discretization.

We recall that for $m \in \mathbb{N}$ we denote the set of all multi-indices α that do not include components equal to -1 by

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

where v is the multi-index of length zero, see Section 6.7.

Given a set $\mathcal{A} \subset \widehat{\mathcal{M}}_m$, the *remainder* set $\widehat{\mathcal{B}}(\mathcal{A})$ of $\mathcal{A}$ is defined by

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

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

$$\widehat{\Gamma}_\beta = \{\alpha \in \widehat{\mathcal{M}}_m : l(\alpha) \leq \beta\}.$$

For a jump-adapted time discretization $(t)_\Delta$, with maximum time step size $\Delta \in (0,1)$, we define the *jump-adapted order β weak Taylor scheme* by

$$Y_{t_{n+1}-} = Y_{t_n} + \sum_{\alpha \in \widehat{\Gamma}_\beta \backslash \{v\}} f_\alpha(t_n, Y_{t_n}) I_\alpha \tag{10.6.2}$$

and

$$Y_{t_{n+1}} = Y_{t_{n+1}-} + \int_{\mathcal{E}} c(t_n, Y_{t_{n+1}-}, v)\, p_\phi(dv, \{t_{n+1}\}), \tag{10.6.3}$$

where I_α is the multiple stochastic integral of the multi-index α over the time period $(t_n, t_{n+1}]$, $n \in \{0, 1, \ldots, n_T - 1\}$, and f_α are the corresponding Itô coefficient functions defined in (3.3.8) with $f(t,x) = x$.

Now we can formulate a convergence theorem for jump-adapted schemes similar to that in Mikulevicius & Platen (1988).

Theorem 10.6.1 *For a given $\beta \in \{1, 2, \dots\}$, let $Y^\Delta = \{Y^\Delta_{t_n}, n \in \{0, 1, \dots, n_T\}\}$ be the order β jump-adapted weak Taylor scheme (10.6.2)–(10.6.3) corresponding to a jump-adapted time discretization with maximum step size $\Delta \in (0,1)$. We assume that $E(|X_0|^i) < \infty$, for $i \in \{1, 2, \dots\}$, and Y^Δ_0 converges weakly to X_0 with order β. Moreover, suppose that the drift, diffusion and jump coefficients have components $a^k, b^{k,j}, c^k \in C_P^{2(\beta+1)}(\mathbb{R}^d, \mathbb{R})$ for all $k \in \{1, 2, \dots, d\}$ and $j \in \{1, 2, \dots, m\}$ and the coefficients f_α, with $f(t,x) = x$, satisfy the linear growth condition $|f_\alpha(t,y)| \le K(1 + |y|)$, with $K < \infty$, for all $t \in [0,T]$, $y \in \mathbb{R}^d$ and $\alpha \in \widehat{\Gamma}_\beta$, see also Remark 9.5.5.*

Then for any function $g \in C_P^{2(\beta+1)}$ there exists a positive constant C, independent of Δ, such that

$$|E(g(X_T)) - E(g(Y^\Delta_{t_{n_T}}))| \le C\Delta^\beta.$$

First we present some results in preparation of the proof of Theorem 10.6.1, see also Mikulevicius & Platen (1988).

Lemma 10.6.2 *For all $n \in \{1, \dots, n_T\}$ and $y \in \mathbb{R}^d$, we have*

$$E\Big(u(t_n, X^{t_{n-1},y}_{t_n-}) - u(t_{n-1}, y) + \int_{t_{n-1}}^{t_n} \int_{\mathcal{E}} L_v^{(-1)} u(z, X_z^{t_{n-1},y}) \phi(dv)dz \Big| \mathcal{A}_{t_{n-1}}\Big) = 0, \tag{10.6.4}$$

where the process $X^{s,y} = \{X^{s,y}_t, t \in [0,T]\}$, for $(s,y) \in [0,t] \times \mathbb{R}^d$, is defined in (9.5.1).

Proof: Note that since all jump times are included in the time discretization, then $X_z^{t_{n-1},y}$ evolves as a diffusion in the time interval (t_{n-1}, t_n), for every $n \in$

$\{1, \ldots, n_T\}$. Therefore, by Itô's formula we obtain

$$\begin{aligned}
u(t_n, X_{t_n-}^{t_{n-1},y}) &= u(t_{n-1}, y) + \int_{t_{n-1}}^{t_n} L^{(0)} u(z, X_z^{t_{n-1},y}) dz \\
&\quad + \sum_{j=1}^{m} \int_{t_{n-1}}^{t_n} L^{(j)} u(z, X_z^{t_{n-1},y}) dW_z^j \\
&= u(t_{n-1}, y) + \int_{t_{n-1}}^{t_n} \widetilde{L}^{(0)} u(z, X_z^{t_{n-1},y}) dz \\
&\quad + \sum_{j=1}^{m} \int_{t_{n-1}}^{t_n} L^{(j)} u(z, X_z^{t_{n-1},y}) dW_z^j \\
&\quad - \int_{t_{n-1}}^{t_n} \int_{\mathcal{E}} L_v^{(-1)} u(z, X_z^{t_{n-1},y}) \phi(dv) dz,
\end{aligned}$$

where the operators $L^{(0)}$, $L^{(j)}$, $L_v^{(-1)}$ and $\widetilde{L}^{(0)}$ are defined in (3.3.4)-(3.3.7).

We complete the proof of the lemma by applying the expected value, using the result (9.5.5) and the properties of Itô's integral. □

Lemma 10.6.3 *For each $p \in \{1, 2, \ldots\}$ there exists a finite constant K such that*

$$E\Big(|X_{t_n-}^{t_{n-1},y} - y|^{2q} \Big| \widetilde{\mathcal{A}}_{t_{n-1}}\Big) \le K(1 + |y|^{2q})(t_n - t_{n-1})^q \tag{10.6.5}$$

for all $q \in \{1, \ldots, p\}$ and $n \in \{1, \ldots, n_T\}$, where $\widetilde{\mathcal{A}}_{t_{n-1}}$ is defined in (10.6.1).

Proof: Since there are no jumps between discretization points, the proof of (10.6.5) follows from that of a similar lemma in the pure diffusion case, see Kloeden & Platen (1999).

Let us also define the process $\eta_{ja} = \{\eta_{ja}(t), t \in [0, T]\}$ by

$$\begin{aligned}
\eta_{ja}(t) &= \eta_{ja}(t_n) + \sum_{\alpha \in \widehat{\Gamma}_\beta \setminus \{v\}} I_\alpha[f_\alpha(t_n, \eta_{ja}(t_n))]_{t_n,t} \\
&\quad + \int_{(t_n,t]} \int_{\mathcal{E}} c(z, \eta_{ja}(z-), v) p_\phi(dv, dz),
\end{aligned} \tag{10.6.6}$$

for $n \in \{0, \ldots, n_T - 1\}$ and $t \in (t_n, t_{n+1}]$, with $\eta_{ja}(0) = Y_0$. Note that

$$\eta_{ja}(t_n) = Y_{t_n} \tag{10.6.7}$$

for every $n \in \{0, \ldots, n_T\}$.

The following result is shown in Mikulevicius & Platen (1988).

Lemma 10.6.4 *For each $p \in \{1, 2, \ldots\}$ there exists a finite constant K such that for every $q \in \{1, \ldots, p\}$*

$$E\Big(\sup_{0 \le t \le T} |\eta_{ja}(t)|^{2q}\Big) \le K\big(1 + |Y_0|^{2q}\big). \tag{10.6.8}$$

We shall also write

$$P_l = \{1, 2, \ldots, d\}^l$$

and

$$F_{\vec{p}}(y) = \prod_{h=1}^{l} y^{p_h} \tag{10.6.9}$$

for all $y = (y^1, \ldots, y^d)^\top \in \Re^d$ and $\vec{p} = (p_1, \ldots, p_l) \in P_l$ where $l \in \{1, 2 \ldots\}$.

Lemma 10.6.5 *For each $p \in \{1, 2 \ldots\}$ there exist finite constants K and $r \in \{1, 2 \ldots\}$ such that*

$$\Big|E\Big(\Big|F_{\vec{p}}\big(\eta_{ja}(z) - Y^\Delta_{t_{n_z}}\big)\Big|^{2q} + \Big|F_{\vec{p}}\big(X_z^{t_{n_z}, Y^\Delta_{t_{n_z}}} - Y^\Delta_{t_{n_z}}\big)\Big|^{2q}\Big|\tilde{\mathcal{A}}_{t_{n_z}}\Big)\Big|$$
$$\le K\big(1 + |Y^\Delta_{t_{n_z}}|^{2r}\big)\,(t_{n_z+1} - t_{n_z})^{ql} \tag{10.6.10}$$

for each $l \in \{1, \ldots, 2(\beta+1)\}$, $q \in \{1, \ldots, p\}$, $\vec{p} \in P_l$ and $z \in [0, T]$, where $F_{\vec{p}}$ is defined in (10.6.9).

Proof: Note that at discretization times the estimate (10.6.10) is trivial since $\eta_{ja}(z) = Y^\Delta_{t_{n_z}}$ and $X_z^{t_{n_z}, Y^\Delta_{t_{n_z}}} = Y^\Delta_{t_{n_z}}$ for $z \in \{t_0, t_1, \ldots, t_{n_T}\}$. Moreover, since jump times arise only at discretization times, we can obtain the estimate (10.6.10) by Itô's formula as in the case of pure diffusion SDEs, see Kloeden & Platen (1999).□

Lemma 10.6.6 *For each $\vec{p} \in P_l$, $l \in \{1, \ldots, 2\beta+1\}$, $n \in \{1, 2, \ldots, n_T\}$ and $z \in [t_{n-1}, t_n)$ there exist two finite constants K and $r \in \{1, 2, \ldots\}$ such that*

$$\Big|E\Big(F_{\vec{p}}\big(\eta_{ja}(z) - Y^\Delta_{t_{n-1}}\big) - F_{\vec{p}}\big(X_z^{t_{n-1}, Y^\Delta_{t_{n-1}}} - Y^\Delta_{t_{n-1}}\big)\Big|\tilde{\mathcal{A}}_{t_{n-1}}\Big)\Big|$$
$$\le K\big(1 + |Y^\Delta_{t_{n-1}}|^r\big)\,\Delta^\beta\,(z - t_{n-1}). \tag{10.6.11}$$

Proof: Since the time discretization includes all jump times, the proof of the lemma follows from that of pure diffusion SDEs, see Kloeden & Platen (1999).□

Now we can prove Theorem 10.6.1.

Proof:[Proof of Theorem 10.6.1] By (9.5.3) and the terminal condition of the Kolmogorov backward equation (9.5.6) we obtain

$$\begin{aligned} H &:= \Big|E\big(g(Y^{\Delta}_{t_{n_T}})\big) - E\big(g(X_T)\big)\Big| \\ &= \Big|E\big(u(T, Y^{\Delta}_{t_{n_T}}) - u(0, X_0)\big)\Big|. \end{aligned} \tag{10.6.12}$$

Note that for the ease of notation we will write Y for Y^{Δ} when no misunderstanding is possible. Since Y_0 converges weakly to X_0 with order β we obtain

$$\begin{aligned} H \leq & \left|E\left(\sum_{n=1}^{n_T} \big(u(t_n, Y_{t_n}) - u(t_n, Y_{t_n-}) + u(t_n, Y_{t_n-}) - u(t_{n-1}, Y^{\Delta}_{t_{n-1}})\big)\right)\right| \\ & + K\,\Delta^{\beta}. \end{aligned} \tag{10.6.13}$$

By (10.6.4) we can write

$$\begin{aligned} H \leq & \Bigg|E\Bigg(\sum_{n=1}^{n_T} \Big[\big\{u(t_n, Y_{t_n}) - u(t_n, Y_{t_n-}) + u(t_n, Y_{t_n-}) - u(t_{n-1}, Y_{t_{n-1}})\big\} \\ & -\big\{u(t_n, X^{t_{n-1}, Y_{t_{n-1}}}_{t_n-}) - u(t_{n-1}, Y_{t_{n-1}}) \\ & + \int_{t_{n-1}}^{t_n} \int_{\mathcal{E}} L^{(-1)}_v u(z, X^{t_{n-1}, Y_{t_{n-1}}}_z)\phi(dv)dz\big\}\Big]\Bigg)\Bigg| + K\,\Delta^{\beta}. \end{aligned}$$

Note that by (10.6.7) and the properties of the stochastic integral with respect to the Poisson measure, we have

$$\begin{aligned} E\Big(\sum_{n=1}^{n_T} \big\{u(t_n, Y_{t_n}) - u(t_n, Y_{t_n-})\big\}\Big) &= E\Big(\int_0^T \int_{\mathcal{E}} L^{(-1)}_v u(z, \eta_{ja}(z-))p_{\phi}(dv, dz)\Big) \\ &= E\Big(\int_0^T \int_{\mathcal{E}} L^{(-1)}_v u(z, \eta_{ja}(z))\phi(dv)dz\Big). \end{aligned}$$

Therefore, we obtain

$$H \leq H_1 + H_2 + K\,\Delta^\beta, \tag{10.6.14}$$

where

$$\begin{aligned} H_1 &= \Bigg| E\Bigg(\sum_{n=1}^{n_T} \Big[\Big(u(t_n, Y_{t_n-}) - u(t_n, Y_{t_{n-1}}) \Big) \\ &\qquad - \Big(u(t_n, X_{t_n-}^{t_{n-1}, Y_{t_{n-1}}}) - u(t_n, Y_{t_{n-1}}) \Big) \Big] \Bigg) \Bigg| \end{aligned} \tag{10.6.15}$$

and

$$\begin{aligned} H_2 &= \Bigg| E\Bigg(\int_0^T \int_{\mathcal{E}} \Big[\Big(L_v^{(-1)} u(z, \eta_{ja}(z)) - L_v^{(-1)} u(z, Y_{t_{n_z}}) \Big) \\ &\qquad - \Big(L_v^{(-1)} u(z, X_z^{t_{n_z}, Y_{t_{n_z}}}) - L_v^{(-1)} u(z, Y_{t_{n_z}}) \Big) \Big] \phi(dv)\, dz \Bigg) \Bigg| \end{aligned} \tag{10.6.16}$$

1. Let us note that by (9.5.7) the function u is smooth enough to apply the deterministic Taylor expansion. Therefore, by expanding the increments of u in H_1 we obtain

$$\begin{aligned} H_1 &= \Bigg| E\Bigg(\sum_{n=1}^{n_T} \Bigg\{ \Big[\sum_{l=1}^{2\beta+1} \frac{1}{l!} \sum_{\vec{p} \in P_l} \big(\partial_y^{\vec{p}} u(t_n, Y_{t_{n-1}}) \big) F_{\vec{p}}(Y_{t_n-} - Y_{t_{n-1}}) + R_n(Y_{t_n-}) \Big] \\ &\qquad - \Big[\sum_{l=1}^{2\beta+1} \frac{1}{l!} \sum_{\vec{p} \in P_l} \big(\partial_y^{\vec{p}} u(t_n, Y_{t_{n-1}}) \big) F_{\vec{p}}(X_{t_n-}^{t_{n-1}, Y_{t_{n-1}}} - Y_{t_{n-1}}) \\ &\qquad + R_n(X_{t_n-}^{t_{n-1}, Y_{t_{n-1}}}) \Big] \Bigg\} \Bigg) \Bigg|, \end{aligned} \tag{10.6.17}$$

where the remainders are

$$\begin{aligned} R_n(Z) &= \frac{1}{2(\beta+1)!} \sum_{\vec{p} \in P_{2(\beta+1)}} \partial_y^{\vec{p}} u\big(t_n, Y_{t_{n-1}} + \theta_{\vec{p},n}(Z)(Z - Y_{t_{n-1}})\big) \\ &\qquad \times F_{\vec{p}}(Z - Y_{t_{n-1}}) \end{aligned} \tag{10.6.18}$$

for $Z = Y_{t_n-}$ and $X_{t_n-}^{t_{n-1}, Y_{t_{n-1}}}$, respectively. Here $\theta_{\vec{p},n}(Z)$ is a $d \times d$ diagonal matrix with

$$\theta_{\vec{p},n}^{k,k}(Z) \in (0,1) \tag{10.6.19}$$

for $k \in \{1, \ldots, d\}$.

Therefore, we have

$$\begin{aligned} H_1 \;\leq\; & E\Bigg(\sum_{n=1}^{n_T} \Bigg\{ \sum_{l=1}^{2\beta+1} \frac{1}{l!} \sum_{\vec{p} \in P_l} \left| \partial_y^{\vec{p}} u(t_n, Y_{t_{n-1}}) \right| & (10.6.20) \\ & \times \left| E\Big(F_{\vec{p}}(Y_{t_n-} - Y_{t_{n-1}}) - F_{\vec{p}}(X_{t_n-}^{t_{n-1}, Y_{t_{n-1}}} - Y_{t_{n-1}}) \Big| \widetilde{\mathcal{A}}_{t_{n-1}} \Big) \right| \\ & + E\Big(\left| R_n(Y_{t_n-}) \right| \widetilde{\mathcal{A}}_{t_{n-1}} \Big) + E\Big(\left| R_n(X_{t_n-}^{t_{n-1}, Y_{t_{n-1}}}) \right| \widetilde{\mathcal{A}}_{t_{n-1}} \Big) \Bigg| \Bigg\} \Bigg). \end{aligned}$$

By (10.6.18), the Cauchy-Schwarz inequality, (9.5.7), (10.6.19) and (10.6.10) we obtain

$$\begin{aligned} & E\Big(\left| R_n(Y_{t_n-}) \right| \widetilde{\mathcal{A}}_{t_{n-1}} \Big) \\ \leq\; & K \sum_{\vec{p} \in P_{2(\beta+1)}} \left[E\Big(\left| \partial_y^{\vec{p}} u\big(t_n, Y_{t_{n-1}} + \theta_{\vec{p},n}(Y_{t_n-})(Y_{t_n-} - Y_{t_{n-1}})\big) \right|^2 \Big| \widetilde{\mathcal{A}}_{t_{n-1}} \Big) \right]^{\frac{1}{2}} \\ & \quad \times \left[E\Big(\left| F_{\vec{p}}(Y_{t_n-} - Y_{t_{n-1}}) \right|^2 \Big| \widetilde{\mathcal{A}}_{t_{n-1}} \Big) \right]^{\frac{1}{2}} \\ \leq\; & K \left[E\Big(1 + |Y_{t_{n-1}}|^{2r} + |Y_{t_n-} - Y_{t_{n-1}}|^{2r} \Big| \widetilde{\mathcal{A}}_{t_{n-1}} \Big) \right]^{\frac{1}{2}} \\ & \quad \times \left[E\Big(\left| F_{\vec{p}}(Y_{t_n-} - Y_{t_{n-1}}) \right|^2 \Big| \widetilde{\mathcal{A}}_{t_{n-1}} \Big) \right]^{\frac{1}{2}} \\ \leq\; & K \big(1 + \left| Y_{t_{n-1}} \right|^{2r} \big) \big(t_n - t_{n-1} \big)^{\beta+1}, \qquad (10.6.21) \end{aligned}$$

using the estimate

$$E(|Y_{t_n-}|^{2r} | \mathcal{A}_{t_{n-1}}) \leq K(1 + |Y_{t_{n-1}}|^{2r}),$$

for every $n \in \{1, \ldots, n_T\}$.

In a similar way, by (10.6.18), the Cauchy-Schwarz inequality, (9.5.7), (10.6.19), (10.6.9) and (10.6.5) we have

$$
\begin{aligned}
& E\big(|R_n(X_{t_n-}^{t_{n-1},Y_{t_{n-1}}})||\tilde{\mathcal{A}}_{t_{n-1}}\big) \\
\leq\ & K \sum_{\vec{p}\in P_{2(\beta+1)}} \Big[E\Big(\big|\partial_y^{\vec{p}} u\big(t_n, Y_{t_{n-1}} \\
& +\theta_{\vec{p},n}(X_{t_n-}^{t_{n-1},Y_{t_{n-1}}})(X_{t_n-}^{t_{n-1},Y_{t_{n-1}}} - Y_{t_{n-1}})\big)\big|^2 \Big|\tilde{\mathcal{A}}_{t_{n-1}}\Big)\Big]^{\frac{1}{2}} \\
& \times\Big[E\Big(\big|F_{\vec{p}}(X_{t_n-}^{t_{n-1},Y_{t_{n-1}}} - Y_{t_{n-1}})\big|^2 \Big|\tilde{\mathcal{A}}_{t_{n-1}}\Big)\Big]^{\frac{1}{2}} \\
\leq\ & K\Big[E\Big(1+|Y_{t_{n-1}}|^{2r} + |X_{t_n-}^{t_{n-1},Y_{t_{n-1}}} - Y_{t_{n-1}}|^{2r}\Big|\tilde{\mathcal{A}}_{t_{n-1}}\Big)\Big]^{\frac{1}{2}} \\
& \times\Big[E\Big(|X_{t_n-}^{t_{n-1},Y_{t_{n-1}}} - Y_{t_{n-1}}|^{4(\beta+1)}\Big|\tilde{\mathcal{A}}_{t_{n-1}}\Big)\Big]^{\frac{1}{2}} \\
\leq\ & K\big(1+|Y_{t_{n-1}}|^{2r}\big)\big(t_n - t_{n-1}\big)^{\beta+1}
\end{aligned}
\tag{10.6.22}
$$

for every $n \in \{1,\ldots,n_T\}$. Here we have used the estimates

$$
E(|X_{t_n-}^{t_{n-1},Y_{t_{n-1}}}|^{2r}|\mathcal{A}_{t_{n-1}}) \leq K(1+|Y_{t_{n-1}}|^{2r}),
$$

and

$$
\Big|F_{\vec{p}}(X_{t_n-}^{t_{n-1},Y_{t_{n-1}}} - Y_{t_{n-1}})\Big|^2 \leq K\Big|X_{t_n-}^{t_{n-1},Y_{t_{n-1}}} - Y_{t_{n-1}}\Big|^{4(\beta+1)},
$$

for every $n \in \{1,\ldots,n_T\}$, with $\vec{p} \in P_{2(\beta+1)}$.

Finally, by applying the Cauchy-Schwarz inequality, (9.5.7), (10.6.7), (10.6.11), (10.6.21), (10.6.22) and (10.6.8) to (10.6.20) we obtain

$$
\begin{aligned}
H_1 &\leq E\Big(K\sum_{n=1}^{n_T}(1+|Y_{t_{n-1}}|^{2r})\Delta^\beta(t_n - t_{n-1})\Big) \\
&\leq K\Delta^\beta\Big(1+E\big(\max_{0\leq n\leq n_T}|Y_{t_n}|^{2r}\big)\Big) \\
&\leq K\Delta^\beta\big(1+|Y_0|^{2r}\big) \leq K\Delta^\beta.
\end{aligned}
\tag{10.6.23}
$$

2. Let us now estimate the term H_2 in (10.6.14). By the smoothness of the function u, see (9.5.7), and of the jump coefficient c, we can apply the deter-

ministic Taylor formula and obtain

$$
\begin{aligned}
H_2 &= \Bigg|E\Bigg(\int_0^T \int_{\mathcal{E}} \Bigg\{\Big[\sum_{l=1}^{2\beta+1} \frac{1}{l!} \sum_{\vec{p}\in P_l} \left(\partial_y^{\vec{p}} L_v^{(-1)} u(z, Y_{t_{n_z}})\right) F_{\vec{p}}(\eta_{ja}(z) - Y_{t_{n_z}}) \\
&\quad + R_{n_z}(\eta_{ja}(z))\Big] \\
&\quad - \Big[\sum_{l=1}^{2\beta+1} \frac{1}{l!} \sum_{\vec{p}\in P_l} \left(\partial_y^{\vec{p}} L_v^{(-1)} u(z, Y_{t_{n_z}})\right) F_{\vec{p}}(X_z^{t_{n_z}, Y_{t_{n_z}}} - Y_{t_{n_z}}) \\
&\quad + R_{n_z}(X_z^{t_{n_z}, Y_{t_{n_z}}})\Big]\Bigg\}\phi(dv)dz\Bigg)\Bigg| \\
&\le \int_0^T \int_{\mathcal{E}} E\Bigg(\sum_{l=1}^{2\beta+1} \frac{1}{l!} \sum_{\vec{p}\in P_l} \left|\partial_y^{\vec{p}} L_v^{(-1)} u(z, Y_{t_{n_z}})\right| \\
&\quad \times \left|E\Big(F_{\vec{p}}(\eta_{ja}(z) - Y_{t_{n_z}}) - F_{\vec{p}}(X_z^{t_{n_z}, Y_{t_{n_z}}} - Y_{t_{n_z}})\Big|\tilde{\mathcal{A}}_{t_{n_z}}\Big)\right| \qquad (10.6.24) \\
&\quad + E\Big(\left|R_{n_z}(\eta_{ja}(z))\right|\Big|\tilde{\mathcal{A}}_{t_{n_z}}\Big) + E\Big(\left|R_{n_z}(X_z^{t_{n_z}, Y_{t_{n_z}}})\right|\Big|\tilde{\mathcal{A}}_{t_{n_z}}\Big)\Bigg)\phi(dv)dz.
\end{aligned}
$$

We can estimate the remainders, as in (10.6.21) and (10.6.22), for every $z \in [0, T]$ by

$$
E\Big(\left|R_{n_z}(\eta_{ja}(z))\right|\Big|\tilde{\mathcal{A}}_{t_{n_z}}\Big) \le K\big(1 + Y_{t_{n_z}}|^{2r}\big)(z - t_{n_z})^{\beta+1} \qquad (10.6.25)
$$

and

$$
E\Big(\left|R_{n_z}(X_z^{t_{n_z}, Y_{t_{n_z}}})\right|\Big|\tilde{\mathcal{A}}_{t_{n_z}}\Big) \le K\big(1 + Y_{t_{n_z}}|^{2r}\big)(z - t_{n_z})^{\beta+1}. \qquad (10.6.26)
$$

Then, by applying the Cauchy-Schwarz inequality, the polynomial growth conditions of $u(t, \cdot)$, see (9.5.7), and of the jump coefficient c, (10.6.19), (10.6.11), (10.6.25), (10.6.26) and (10.6.8) to the estimate (10.6.24), we obtain

$$\begin{aligned} H_2 &\leq K\int_0^T \int_{\mathcal{E}} E\Big(1+|Y_{t_{n_z}}|^{2r}\Big)\Delta^\beta\Big(z-t_{n_z}\Big)\phi(dv)dz \\ &\leq K\Delta^\beta \int_0^T E\Big(1+\max_{0\leq n\leq n_T}|Y_{t_n}|^{2r}\Big)\Big(z-t_{n_z}\Big)dz \\ &\leq K\Delta^\beta. \end{aligned} \tag{10.6.27}$$

3. Finally, by (10.6.14), (10.6.23) and (10.6.27) we obtain

$$\Big|E\big(g(Y_{t_{n_T}})\big)-E\big(g(X_T)\big)\Big| \leq K\Delta^\beta.\square$$

10.7 Convergence of Jump-Adapted Weak Approximations

In this section we present a convergence theorem for general jump-adapted weak approximations of any weak order of convergence $\beta \in \{1, 2\ldots\}$. This theorem requires certain conditions to be satisfied by the increments of a given discrete time approximation in order to obtain a jump-adapted approximation of weak order $\beta \in \{1, 2, \ldots\}$. It covers the convergence of the schemes presented in this chapter.

Let us consider a jump-adapted discrete time approximation $Y^\Delta = \{Y^\Delta_{t_n}, n \in \{0, 1, \ldots, n_T\}\}$ corresponding to a jump-adapted time discretization with maximum step size $\Delta \in (0, 1)$. We simulate the jump impact as before by

$$Y^\Delta_{t_{n+1}} = Y^\Delta_{t_{n+1}-} + \int_{\mathcal{E}} c(t_n, Y^\Delta_{t_{n+1}-}, v)\, p_\phi(dv, \{t_{n+1}\}), \tag{10.7.1}$$

for $n \in \{0, 1, \ldots, n_T - 1\}$. Note that we use here the same notation Y^Δ as that used in (10.6.2)–(10.6.3) for the jump-adapted weak Taylor scheme. However, the jump-adapted approximation considered now is more general since we do not specify its evolution between discretization points. The scheme (10.6.2)–(10.6.3) is a special case of the approximation considered here. The theorem below will state the conditions for obtaining weak order of convergence $\beta \in \{1, 2, \ldots\}$.

First, let us formulate an important condition on the evolution of Y^Δ is the following. Define a stochastic process $\eta_{ja} = \{\eta_{ja}(t), t \in [0,T]\}$ such that for every $n \in \{0,\ldots,n_T\}$

$$\eta_{ja}(t_n) = Y_{t_n} \tag{10.7.2}$$

and

$$\eta_{ja}(t_n-) = Y_{t_n-}. \tag{10.7.3}$$

Assume that the process $\eta_{ja} = \{\eta_{ja}(t), t \in [0,T]\}$ satisfies Lemmas 10.6.4 with respect to the initial value Y_0 of the general jump-adapted approximation considered here. Moreover, we require that the process η_{ja} satisfies also Lemmas 10.6.5 and 10.6.6, where Y is again the general jump-adapted approximation under consideration. Then we can formulate the following convergence theorem.

Theorem 10.7.1 *Let us as assume that $E(|X_0|^i) < \infty$ for $i \in \{1,2,\ldots\}$, and that Y_0^Δ converges weakly to X_0 with order $\beta \in \{1,2,\ldots\}$. Suppose that the drift, diffusion and jump components $a^k, b^{k,j}, c^k$, respectively, belong to the space $C_P^{2(\beta+1)}(\mathbb{R}^d,\mathbb{R})$, for all $k \in \{1,2,\ldots,d\}$ and $j \in \{1,2,\ldots,m\}$, and the coefficients f_α, with $f(t,x) = x$, satisfy the linear growth condition $|f_\alpha(t,y)| \le K(1+|y|)$, with $K < \infty$, for all $t \in [0,T]$, $y \in \mathbb{R}^d$, and $\alpha \in \widehat{\Gamma}_\beta$, see also Remark 9.5.5.*

Moreover, assume that for each $p \in \{1,2\ldots\}$ there exist constants $K < \infty$ and $r \in \{1,2,\ldots\}$, which do not depend on Δ, such that for each $q \in \{1,\ldots,p\}$

$$E\Big(\max_{0\le n\le n_T} |Y_{t_n-}^\Delta|^{2q} | \mathcal{A}_0\Big) \le K\big(1+|Y_0^\Delta|^{2r}\big), \tag{10.7.4}$$

$$E\Big(|Y_{t_{n+1}-}^\Delta - Y_{t_n}^\Delta|^{2q} | \widetilde{\mathcal{A}}_{t_n}\Big) \le K\big(1+\max_{0\le k\le n}|Y_{t_k}^\Delta|^{2r}\big)\,(t_{n+1}-t_n)^q \tag{10.7.5}$$

for $n \in \{0,1,\ldots,n_T-1\}$, and

$$\Big|E\Big(\prod_{h=1}^{l} \big(Y_{t_{n+1}-}^{\Delta,p_h} - Y_{t_n}^{\Delta,p_h}\big) - \prod_{h=1}^{l}\Big(\sum_{\alpha\in\Gamma_\beta\backslash\{v\}} I_\alpha\left[f_\alpha^{p_h}(t_n, Y_{t_n}^\Delta)\right]_{t_n,t_{n+1}}\Big)\Big|\widetilde{\mathcal{A}}_{t_n}\Big)\Big|$$
$$\le K\big(1+\max_{0\le k\le n_T}|Y_{t_k}^\Delta|^{2r}\big)\,\Delta^\beta\,(t_{n+1}-t_n) \tag{10.7.6}$$

for all $n \in \{0,1,\ldots,n_T-1\}$ and $(p_1,\ldots,p_l) \in \{1,\ldots,d\}^l$, where $l \in \{1,\ldots,2\beta+1\}$ and Y^{Δ,p_h} denotes the p_hth component of Y^Δ.

Then for any function $g \in C_P^{2(\beta+1)}(\mathbb{R}^d,\mathbb{R})$ there exists a positive constant C, inde-

pendent of Δ, *such that*

$$|E(g(X_T)) - E(g(Y^{\Delta}_{t_{n_T}}))| \leq C\Delta^{\beta}. \tag{10.7.7}$$

Theorem 10.7.1 states the conditions on a discrete time approximation so that it can be to approximate the diffusion component of a jump-adapted scheme. The most important condition is the estimate (10.7.6). It requires the first $2\beta + 1$ conditional moments of the increments of the diffusion approximation to be close to those of the truncated Wagner-Platen expansion for pure diffusions. Examples of corresponding weak schemes in the pure diffusion case will be given in the next chapter.

To prove Theorem 10.7.1, let us define

$$\eta^{y}_{n,\beta} = y + \sum_{\alpha \in \widehat{\Gamma}_{\beta} \setminus \{v\}} I_{\alpha}[f_{\alpha}(t_n, y)]_{t_{n-1}, t_n}, \tag{10.7.8}$$

for $n \in \{1, \ldots, n_T\}$ and $y \in \mathbb{R}^d$.

Proof:[Proof of Theorem 10.7.1] Note that in the following we will write Y for Y^{Δ}. By (9.5.3) and (9.5.6) we obtain

$$\begin{aligned} H &:= \Big|E\big(g(Y_{t_{n_T}})\big) - E\big(g(X_T)\big)\Big| \\ &= \Big|E\big(u(T, Y_{t_{n_T}}) - u(0, X_0)\big)\Big|. \end{aligned} \tag{10.7.9}$$

Moreover, since Y_0 converges weakly with order β to X_0 we obtain

$$\begin{aligned} H &\leq \left|E\left(\sum_{n=1}^{n_T} \big(u(t_n, Y_{t_n}) - u(t_n, Y_{t_n-}) + u(t_n, Y_{t_n-}) - u(t_{n-1}, Y_{t_{n-1}})\big)\right)\right| \\ &\quad + K\,\Delta^{\beta}. \end{aligned} \tag{10.7.10}$$

By (10.6.4), the definition of η_{ja} and (10.7.2)-(10.7.3), we obtain

$$
\begin{aligned}
H \ \leq \ & \Bigg| E\Bigg(\sum_{n=1}^{n_T} \Big[\big\{ u(t_n, Y_{t_n}) - u(t_n, Y_{t_n-}) + u(t_n, Y_{t_n-}) - u(t_{n-1}, Y_{t_{n-1}}) \big\} \\
& - \big\{ u(t_n, X_{t_n-}^{t_{n-1}, Y_{t_{n-1}}}) - u(t_{n-1}, Y_{t_{n-1}}) \\
& + \int_{t_{n-1}}^{t_n} \int_{\mathcal{E}} L_v^{(-1)} u(z, X_z^{t_{n-1}, Y_{t_{n-1}}}) \phi(dv) dz \big\} \Big] \Bigg) \Bigg| + K\,\Delta^\beta \\
\leq \ & \Bigg| E\Bigg(\int_0^T \int_{\mathcal{E}} \Big[L_v^{(-1)} u(z, \eta_{ja}(z)) - L_v^{(-1)} u(z, X_z^{t_{n_z}, Y_{t_{n_z}}}) \Big] \phi(dv) dz \\
& + \sum_{n=1}^{n_T} \Big[\Big(u(t_n, Y_{t_n-}) - u(t_n, Y_{t_{n-1}}) \Big) - \Big(u(t_n, X_{t_n-}^{t_{n-1}, Y_{t_{n-1}}}) - u(t_n, Y_{t_{n-1}}) \Big) \Big] \Bigg| \\
& + K\,\Delta^\beta \\
\leq \ & H_1 + H_2 + H_3 + K\,\Delta^\beta, \qquad (10.7.11)
\end{aligned}
$$

where

$$
\begin{aligned}
H_1 \ = \ & \Bigg| E\Bigg(\sum_{n=1}^{n_T} \Big[\Big(u(t_n, Y_{t_n-}) - u(t_n, Y_{t_{n-1}}) \Big) \\
& - \Big(u(t_n, \eta_{n,\beta}^{Y_{t_{n-1}}}) - u(t_n, Y_{t_{n-1}}) \Big) \Big] \Bigg) \Bigg|, \qquad (10.7.12)
\end{aligned}
$$

$$
\begin{aligned}
H_2 \ = \ & \Bigg| E\Bigg(\sum_{n=1}^{n_T} \Big[\Big(u(t_n, \eta_{n,\beta}^{Y_{t_{n-1}}}) - u(t_n, Y_{t_{n-1}}) \Big) \\
& - \Big(u(t_n, X_{t_n-}^{t_{n-1}, Y_{t_{n-1}}}) - u(t_n, Y_{t_{n-1}}) \Big) \Big] \Bigg) \Bigg| \qquad (10.7.13)
\end{aligned}
$$

and

$$
\begin{aligned}
H_3 \ = \ & \Bigg| E\Bigg(\int_0^T \int_{\mathcal{E}} \Big[\Big(L_v^{(-1)} u(z, \eta_{ja}(z)) - L_v^{(-1)} u(z, Y_{t_{n_z}}) \Big) \\
& - \Big(L_v^{(-1)} u(z, X_z^{t_{n_z}, Y_{t_{n_z}}}) - L_v^{(-1)} u(z, Y_{t_{n_z}}) \Big) \Big] \phi(dv) dz \Bigg) \Bigg| \qquad (10.7.14)
\end{aligned}
$$

By applying the deterministic Taylor expansion to the increments in H_1 and H_2, we obtain the following estimate, see Kloeden & Platen (1999),

$$H_i \leq K\Delta^\beta \tag{10.7.15}$$

for $i \in \{1,2\}$. The estimate of H_3 follows as in the estimate (10.6.24) of Theorem 10.6.1, since η_{ja} satisfies Lemmas 10.6.4, 10.6.5 and 10.6.6. This completes the proof of Theorem 10.7.1.□

10.7.1 Simplified and Predictor-Corrector Schemes

Theorem 10.7.1 can be used to establish the order of weak convergence of the jump-adapted approximations presented in the current chapter. To do so, we should first construct a stochastic process η_{ja} which satisfies properties (10.7.2)–(10.7.3) and Lemmas 10.6.4, 10.6.5 and 10.6.6. Then, we have to show that the increments of the approximation between discretization points satisfy conditions (10.7.5) and (10.7.6). Finally, we need to check the regularity of the approximation with condition (10.7.4).

Let us consider, for example, the jump-adapted predictor-corrector Euler scheme (10.4.1)–(10.4.2). We can define the following stochastic process $\eta_{ja} = \{\eta_{ja}(t), t \in [0,T]\}$ by

$$\begin{aligned}\eta_{ja}(t) &= \eta_{ja}(t_n) + \frac{1}{2}\Big\{a\big(\xi_{ja}(t)\big) + a\big(\eta_{ja}(t_n)\Big\}(t-t_n) + b\big(\eta_{ja}(t_n)\big)(W_t - W_{t_n}) \\ &\quad + \int_{(t_n,t]}\int_{\mathcal{E}} c(z,\eta_{ja}(z-),v)p_\phi(dv,dz),\end{aligned} \tag{10.7.16}$$

and

$$\xi_{ja}(t) = \xi_{ja}(t_n) + a\big(\eta_{ja}(t_n)\big)\,(t-t_n) + b\big(\eta_{ja}(t_n)\big)\,(W_t - W_{t_n}) \tag{10.7.17}$$

for $n \in \{0,\dots,n_T-1\}$ and $t \in (t_n, t_{n+1}]$, with $\eta_{ja}(0) = Y_0$. By construction, η_{ja} satisfies conditions (10.7.2)–(10.7.3). Conditions (10.7.4) and (10.7.5) can be shown by using the linear growth conditions on the coefficients a and b and the properties of the increments of Wiener processes. Furthermore, condition (10.7.6) holds with $\beta = 1.0$, as in the case of pure diffusion SDEs, see Kloeden & Platen (1999). Finally, one has to show that the process η_{ja} defined in (10.7.16) satisfies

Lemmas 10.6.4, 10.6.5 and 10.6.6 with $\beta = 1.0$. This result can be obtained by using the linear growth conditions (2.2.11) on the coefficients a and b, the Wagner-Platen expansion and the properties of the increments of Wiener processes. Thus, the weak order of convergence $\beta = 1.0$ of the jump-adapted predictor-corrector Euler scheme is established.

The order of weak convergence of other jump-adapted predictor-corrector and derivative-free schemes, as presented in this chapter, can be established in a similar way. Also the general multi-dimensional case can be handled similarly.

To show the order of weak convergence of simplified jump-adapted schemes, based on multi-point distributed random variables, one has to construct a process η_{ja} that satisfies conditions (10.7.2)–(10.7.3). For example, when considering the jump-adapted simplified Euler scheme (10.2.5) in the one-dimensional case, we can define

$$\begin{aligned}\eta_{ja}(t) &= \eta_{ja}(t_n) + a\big(\eta_{ja}(t_n)\big)(t-t_n) + b\big(\eta_{ja}(t_n)\big)\sqrt{t-t_n}\,sign\big(W_t - W_{t_n}\big)\\ &\quad + \int_{(t_n,t]}\int_{\mathcal{E}} c(z,\eta_{ja}(z-),v)p_{\phi}(dv,dz), \qquad (10.7.18)\end{aligned}$$

for $n \in \{0,\ldots,n_T-1\}$ and $t \in (t_n,t_{n+1}]$, with $\eta_{ja}(0) = Y_0$, where we have used the notation

$$sign(x) = \begin{cases} 1 & \text{for} \quad x \geq 0 \\ -1 & \text{for} \quad x < 0. \end{cases} \qquad (10.7.19)$$

Then the process η_{ja} satisfies conditions (10.7.2)–(10.7.3) and Lemmas 10.6.4, 10.6.5 and 10.6.6 with $\beta = 1$.

To establish the order of weak convergence of the jump-adapted simplified order 2.0 weak scheme (10.2.8) in the one-dimensional case, we define

$$\begin{aligned}\eta_{ja}(t) &= \eta_{ja}(t_n) + a\big(\eta_{ja}(t_n)\big)(t-t_n) + b\big(\eta_{ja}(t_n)\big)\omega_n(t)\\ &\quad + \frac{b\big(\eta_{ja}(t_n)\big)b'\big(\eta_{ja}(t_n)\big)}{2}\Big((\omega_n(t))^2 - (t-t_n)\Big)\\ &\quad + \frac{1}{2}\Big(a\big(\eta_{ja}(t_n)\big)a'\big(\eta_{ja}(t_n)\big) + \frac{1}{2}a''\big(\eta_{ja}(t_n)\big)\big(b\big(\eta_{ja}(t_n)\big)\big)^2\Big)(t-t_n)^2\\ &\quad + \frac{1}{2}\Big(a'\big(\eta_{ja}(t_n)\big)b\big(\eta_{ja}(t_n)\big) + a\big(\eta_{ja}(t_n)\big)b'\big(\eta_{ja}(t_n)\big)\\ &\quad + \frac{1}{2}b''\big(\eta_{ja}(t_n)\big)\big(b\big(\eta_{ja}(t_n)\big)\big)^2\Big)\omega_n(t)(t-t_n)\end{aligned}$$

$$+\int_{(t_n,t]}\int_{\mathcal{E}} c(z,\eta_{ja}(z-),v)p_\phi(dv,dz), \tag{10.7.20}$$

where

$$\omega_n(t)=\sqrt{t-t_n}\left(\mathbf{1}_{\{W_t-W_{t_n}<\mathcal{N}^{-1}(\frac{1}{6})\sqrt{t-t_n}\}}-\mathbf{1}_{\{W_t-W_{t_n}>\mathcal{N}^{-1}(\frac{5}{6})\sqrt{t-t_n}\}}\right) \tag{10.7.21}$$

for $n\in\{0,\ldots,n_T-1\}$ and $t\in(t_n,t_{n+1}]$, with $\eta_{ja}(0)=Y_0$. In (10.7.21) we have denoted by $\mathbf{1}$ the indicator function defined in (2.1.8) and by $\mathcal{N}^{-1}(x)$ the inverse of the distribution function of a standard Gaussian random variable. Therefore, the process η_{ja} satisfies conditions (10.7.2)–(10.7.3) and Lemmas 10.6.4, 10.6.5 and 10.6.6 with $\beta=2.0$. In a similar way one can use Theorem 10.7.1 to construct simplified weak schemes with higher order of weak convergence, as the third order weak schemes presented in this chapter, also in the general multi-dimensional case.

Chapter 11

Numerical Results on Weak Schemes

In this short chapter we present some numerical results for the weak schemes presented in Chapters 9 and 10. These complement the numerical results for strong schemes presented in Chapter 7.

11.1 Introduction

We study the weak approximation of the SDE (2.1.5), describing the Merton model, which is

$$dX_t = X_{t-}\left(\mu dt + \sigma dW_t + \int_{\mathcal{E}} (v-1)\, p_\phi(dv, dt)\right), \tag{11.1.1}$$

for $t \in [0, T]$ and $X_0 > 0$. We recall the explicit solution

$$X_t = X_0\, e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t} \prod_{i=1}^{p_\phi(t)} \xi_i, \tag{11.1.2}$$

where the marks ξ_i are distributed according to a given probability measure $F(dv) = \frac{\phi(dv)}{\lambda}$ and $p_\phi = \{p_\phi(t), t \in [0, T]\}$ denotes a Poisson process with intensity $\lambda = \phi(\mathcal{E}) < \infty$.

In this chapter we consider the several schemes with weak order of convergence $\beta \in \{1.0, 2.0\}$. The schemes with weak order $\beta = 1.0$ are the following: the regular Euler scheme, the jump-adapted Euler scheme and the jump-adapted predictor-corrector Euler scheme. Moreover, we consider the following schemes with weak order $\beta = 2.0$: the regular order 2.0 Taylor scheme, the jump-adapted order 2.0 Taylor scheme and the jump-adapted order 2.0 predictor-corrector scheme. We study the weak error

$$\varepsilon_w(\Delta) = |E(g(X_T)) - E(g(Y_T^\Delta))|, \tag{11.1.3}$$

for certain payoff functions g, as defined in (1.2.2). We show several log-log plots with the logarithm of the weak error $\log_2(\varepsilon_w(\Delta))$, versus the logarithm of the maximum time step size $\log_2(\Delta)$. By using log-log plots the achieved orders of weak convergence will appear as the slopes of the estimated error lines obtained. We remark that the aim of this thesis is the study of the error attributable to the time discretization, which we call systematic error, rather than the statistical error due to the Monte Carlo simulation itself. Therefore, we need to generate a large enough number of sample paths to ensure that the statistical error is negligible when compared to the systematic error. We remark that in our study, since we consider the properties of discrete time simulation methods, we have always used a "raw" Monte Carlo simulation. In practice, when solving a specific problem, one should apply variance reduction techniques that can drastically reduce the statistical errors of several orders of magnitude, see Kloeden & Platen (1999) and Glasserman (2004). The problem of variance reduction is not considered in this thesis.

11.2 The Case of a Smooth Payoff

At first, we consider the evaluation of $E(g(X_T))$, where g is a smooth function. The smoothness of the function g is required by the convergence theorems presented in Chapters 9 and 10. Already in the case of pure diffusion SDEs, convergence theorems for weak Taylor approximations typically require a certain degree of smoothness for the function g. For convergence results on the Euler scheme when the function g is non-smooth, we refer to Bally & Talay (1996a, 1996b) and Guyon (2006) for pure diffusion SDEs, and to Hausenblas (2002) for pure jump SDEs. An important application of Monte Carlo simulation with smooth payoff functions is the evaluation of Value at Risk via the simulation of moments, as in Edgeworth expansions and saddle point methods, see Studer (2001). Another application involving smooth payoff functions arises when evaluating expected utilities by simulation.

Let us consider the estimation of moments of the solution X at a final time T. For the SDE (11.1.1), we obtain, via its closed form solution (11.1.2), the expression

$$E\big((X_T)^k\big) = (X_0)^k \exp\left\{kT\big(\mu + \frac{\sigma^2}{2}(k-1)\big) + \lambda T\Big(E\big((\xi)^k\big) - 1\Big)\right\} \qquad (11.2.4)$$

for $k \in \mathbb{N}$, so that the weak error $\varepsilon_w(\Delta)$ can be estimated according to (11.1.3). In the following example we choose the function $g(x) = x^4$, so that we estimate the fourth moment of the solution X at a given maturity date T. We select the following default parameters: $\mu = 0.05$, $\sigma = 0.15$, $\lambda = 0.05$, $X_0 = 1$ and $T = 4$.

We first consider the case of a mark-independent jump coefficient $c(t, x) = x(\psi - 1)$, which is equivalent to the case of constant marks $\xi_i = \psi > 0$, and we set $\psi = 0.85$. Therefore, at each jump time the value of X drops by 15%. In Figure 11.2.1 we report the results for the regular Euler, the jump-adapted Euler, and the jump-adapted predictor-corrector Euler schemes. Here and in the sequel of this chapter, the implicitness parameters have been set to $\theta = 0.5$ and $\eta = 0$. The slopes of the lines in Figure 11.2.1 are about one, which means that these schemes achieve an order of weak convergence $\beta = 1.0$. This is in accordance with the theorems presented in the previous chapters. Furthermore, the accuracies of the regular and the jump-adapted Euler schemes are very similar. The jump-adapted predictor-corrector Euler scheme, instead, is more accurate.

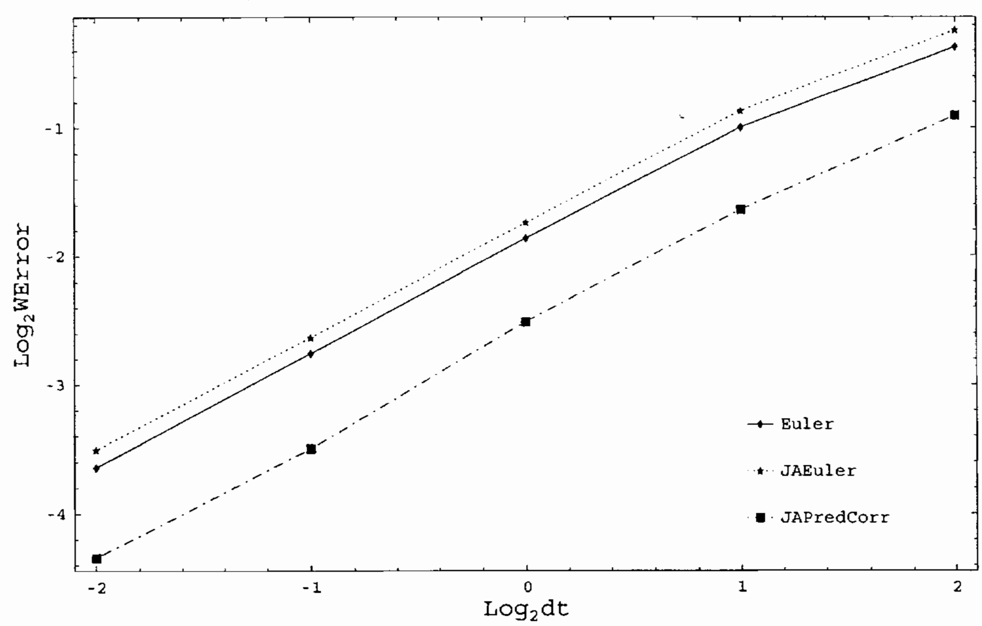

Figure 11.2.1: Log-log plot of weak error versus time step size (fourth moment with constant marks)

In Figure 11.2.2 we show the errors generated by the regular order 2.0 Taylor, the jump-adapted order 2.0 Taylor and the jump-adapted order 2.0 predictor-corrector schemes. Here we can see that the order of convergence achieved is about $\beta = 2.0$, as suggested by the convergence theorems of previous chapters. Also in this case, the predictor-corrector scheme is more accurate than the corresponding explicit

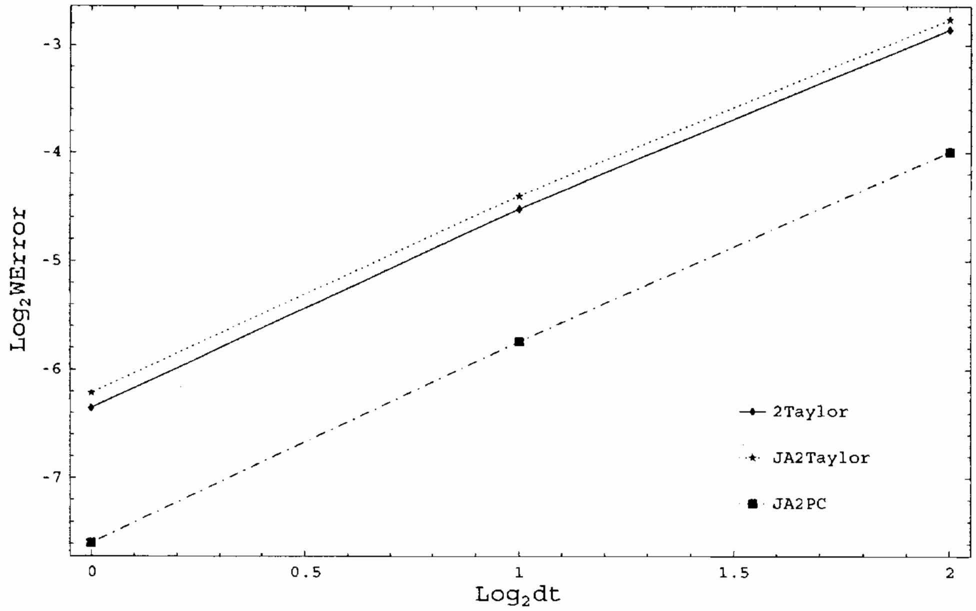

Figure 11.2.2: Log-log plot of weak error versus time step size (fourth moment with constant marks)

schemes. Finally, in Figure 11.2.3 we plot the second order schemes together with the regular Euler and the jump-adapted predictor-corrector Euler schemes. This figure highlights the higher accuracy of schemes with second order of weak convergence. Furthermore, we report that the jump-adapted order 2.0 predictor-corrector scheme is the most accurate among all schemes implemented.

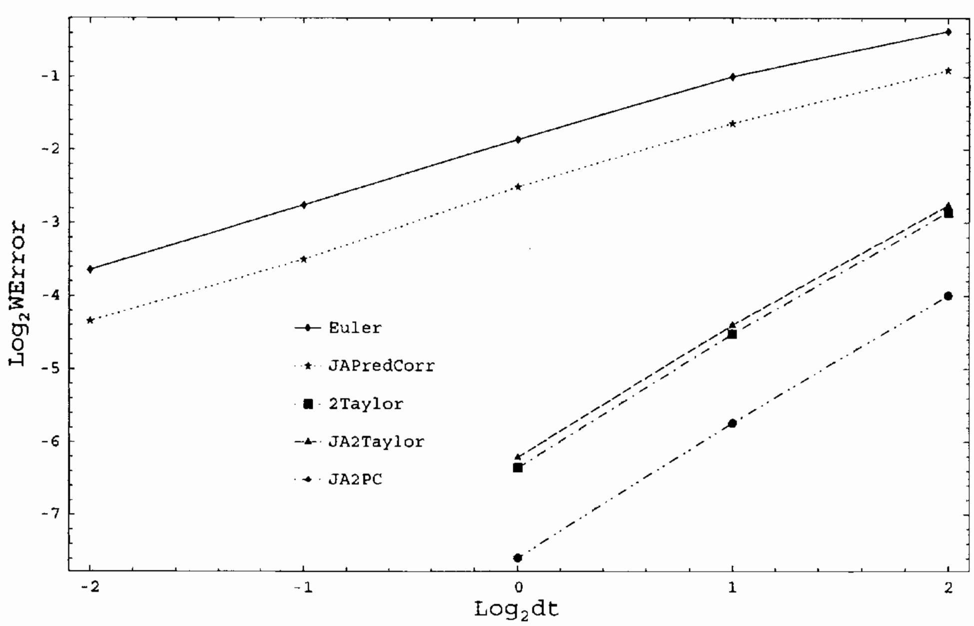

Figure 11.2.3: Log-log plot of weak error versus time step size (fourth moment with constant marks)

Let us now consider the case of lognormally distributed marks, where the logarithm of the mark $\zeta_i = \ln(\xi_i)$ is an independent Gaussian random variable, $\zeta_i \sim \mathcal{N}(\varrho, \varsigma)$, with mean $\varrho = -0.1738$ and standard deviation $\sqrt{\varsigma} = 0.15$. This implies that at a jump date the value of X drops on average by 15%, since $E(\xi) - 1 = -0.15$. In Figure 11.2.4 we report the results for the regular Euler scheme, the jump-adapted predictor-corrector Euler scheme, the regular and jump-adapted order 2.0 Taylor schemes, and the jump-adapted order 2.0 predictor-corrector scheme. The accuracy of these schemes in the case of lognormally distributed jump sizes is almost the same as that achieved in the case of a mark-independent jump coefficient. Here all schemes achieve the orders of weak convergence prescribed by the convergence theorems in Chapters 9 and 10.

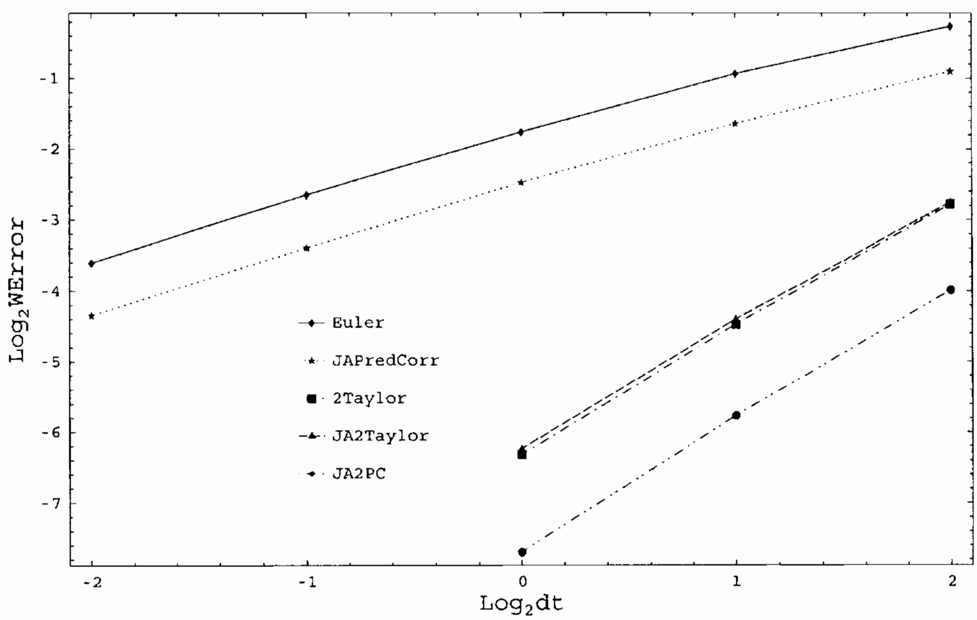

Figure 11.2.4: Log-log plot of weak error versus time step size (fourth moment with lognormal marks)

11.3 The Case of a Non-Smooth Payoff

We now consider the case of a non-smooth payoff function g when computing the price of a European call option. The convergence theorems presented in Chapters 9 and 10 do not cover the case of a non-differentiable payoff as the one studied in this section. It seems that there does not exist in the literature a reasonably general convergence theorem for weak Taylor approximations in the case of non-smooth payoff functions, even when considering SDEs without jumps. The existing results are limited to the Euler scheme for pure diffusion SDEs, see Bally & Talay (1996a, 1996b) and Guyon (2006), and to the Euler scheme for pure jump SDEs, see Hausenblas (2002). It is, however, interesting to obtain some numerical results that could indicate the theoretical performance of the weak schemes presented in this thesis when applied to the non-smooth European call option payoff $g(X_T) = e^{-rT}(X_T - K)^+ = e^{-rT}max(X_T - K, 0)$. Here r is the risk-free rate and K the strike price. Often, the presence of jumps causes market incompleteness and, thus, precludes the possibility of perfect hedging. We refer to the monograph Cont & Tankov (2004) for a discussion on pricing and hedging in incomplete markets with jumps. Here we assume for simplicity, as in Merton (1976), that the jump risk is non-systematic and, thus, diversifiable. In our example the price of a European call option is given by

$$c_{T,X} := E\big(e^{-rT}(X_T - K)^+\big), \tag{11.3.5}$$

where the drift μ of the process X equals $r - q - \lambda\big(E(\xi) - 1\big)$, and q is the continuous dividend yield provided by the security X.

In the case of the SDE (11.1.1), where the logarithm of the mark $\zeta_i = \ln(\xi_i)$ is a Gaussian random variable $\zeta_i \sim \mathcal{N}(\varrho, \varsigma)$ with mean ϱ and variance ς, we obtain a closed form solution, see Merton (1976), given by

$$c_{T,K}(X) = \sum_{j=0}^{\infty} \frac{e^{-\lambda' T}(\lambda' T)^j}{j!} f_j,$$

where $\lambda' = \lambda E(\xi)$. Here

$$f_j = X_0 \mathcal{N}(d_{1,j}) - e^{-\mu_j T} K \mathcal{N}(d_{2,j}),$$

is the Black-Scholes price of a call option with the parameters specified as

$$d_{1,j} = \frac{\ln(\frac{X_0}{K}) + (\mu_j + \frac{\sigma_j^2}{2})T}{\sigma_j\sqrt{T}},$$

$d_{2,j} = d_{1,j} - \sigma_j\sqrt{T}$, $\mu_j = r - q - \lambda(E(\xi) - 1) + \frac{j \ln E(\xi)}{T}$ and $\sigma_j^2 = \sigma^2 + \frac{j\varsigma}{T}$. We recall that we denote by $\mathcal{N}(\cdot)$ the probability distribution of a standard Gaussian random variable.

Let us first consider the case of constant marks $\xi_i = \psi > 0$, with $\psi = 0.85$ as in the previous section. Moreover, we set $r = 0.055$, $q = 0.01125$, $\lambda = 0.05$,

$\sigma = 0.15$, $X_0 = 100$ and $K = 100$. Note that this implies a risk-neutral drift $\mu = r - q - \lambda(\psi - 1) = 0.05$, which equals that used in the previous section.

In Figure 11.3.5 we consider the case of a mark-independent jump coefficient for the regular and jump-adapted Euler schemes, and the jump-adapted predictor-corrector Euler scheme. All these schemes achieve, in our study, an order of weak convergence of about one. The regular and the jump-adapted Euler schemes achieve first order of weak convergence, with the regular scheme being slightly more accurate. The jump-adapted predictor-corrector Euler scheme is far more accurate than the explicit schemes here considered. Its order of convergence is about one, with an oscillatory behavior for large time step sizes.

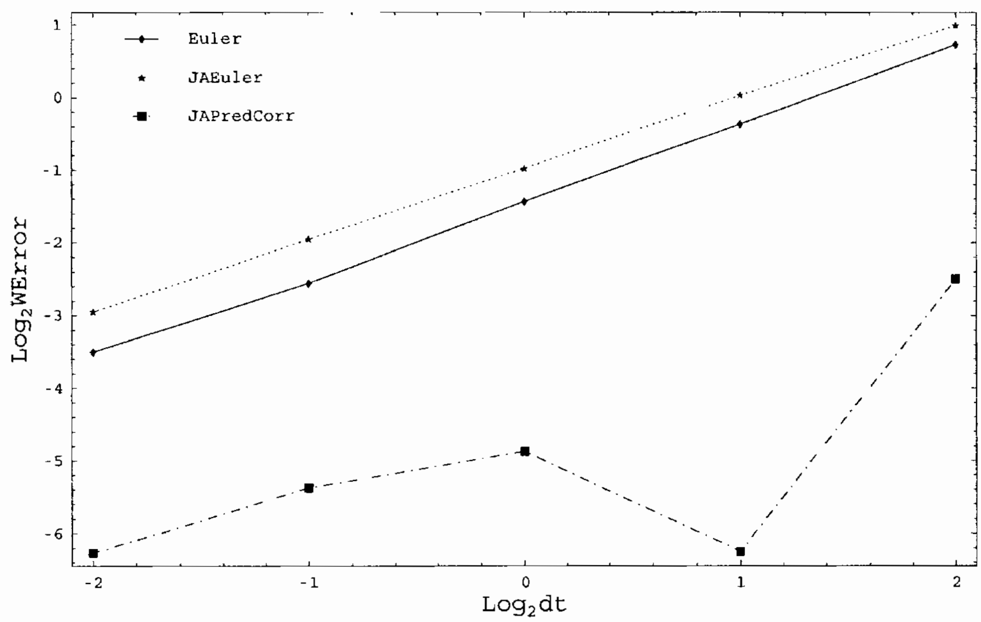

Figure 11.3.5: Log-log plot of weak error versus time step size (call payoff with constant marks)

Figure 11.3.6 shows the accuracy of the regular order 2.0 Taylor, the jump-adapted order 2.0 Taylor and the jump-adapted order 2.0 predictor-corrector schemes. All schemes numerically attain an order of weak convergence equal to $\beta = 2.0$. In this example the jump-adapted order 2.0 Taylor scheme is the most accurate. Finally, in Figure 11.3.7 we report the results of these second order schemes together with those of the regular Euler and jump-adapted predictor-corrector Euler schemes. We clearly notice the difference in accuracy between first order and second order schemes. However, we emphasize the excellent performance of the jump-adapted predictor-corrector Euler scheme already for large step sizes. This scheme has an accuracy similar to that of some second order schemes. The jump-adapted order

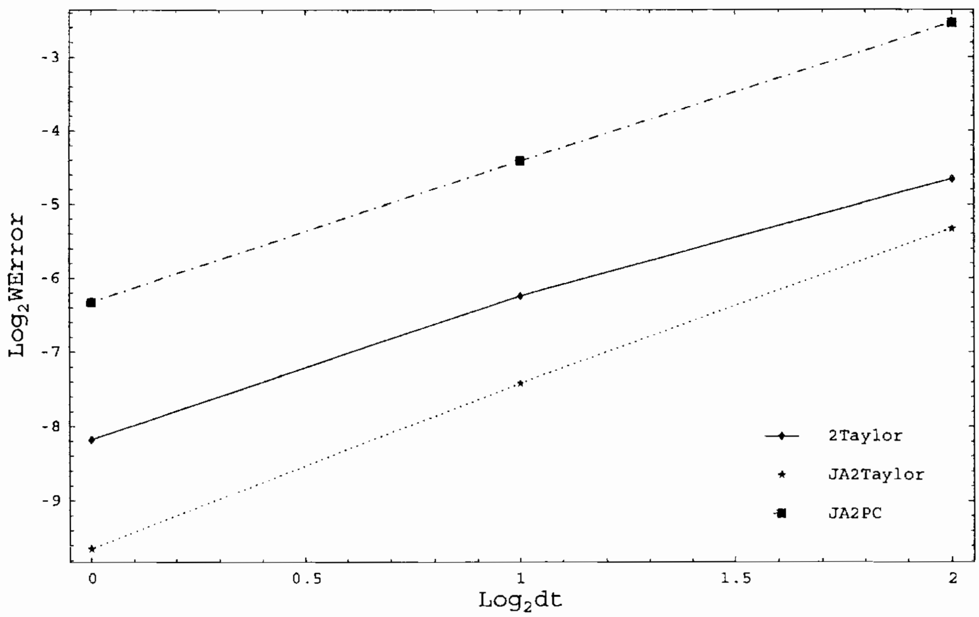

Figure 11.3.6: Log-log plot of weak error versus time step size (call payoff with constant marks)

2.0 Taylor scheme is the most accurate for all time step sizes considered.

We have also tested the above mentioned schemes in the case of lognormal marks, where $\ln(\xi_i) \sim \mathcal{N}(\varrho, \varsigma)$, with mean $\varrho = -0.1738$ and standard deviation $\sqrt{\varsigma} = 0.15$. In Figure 11.3.8 one can notice that the numerical results are practically identical to those reported in Figure 11.3.7 for the case of constant marks.

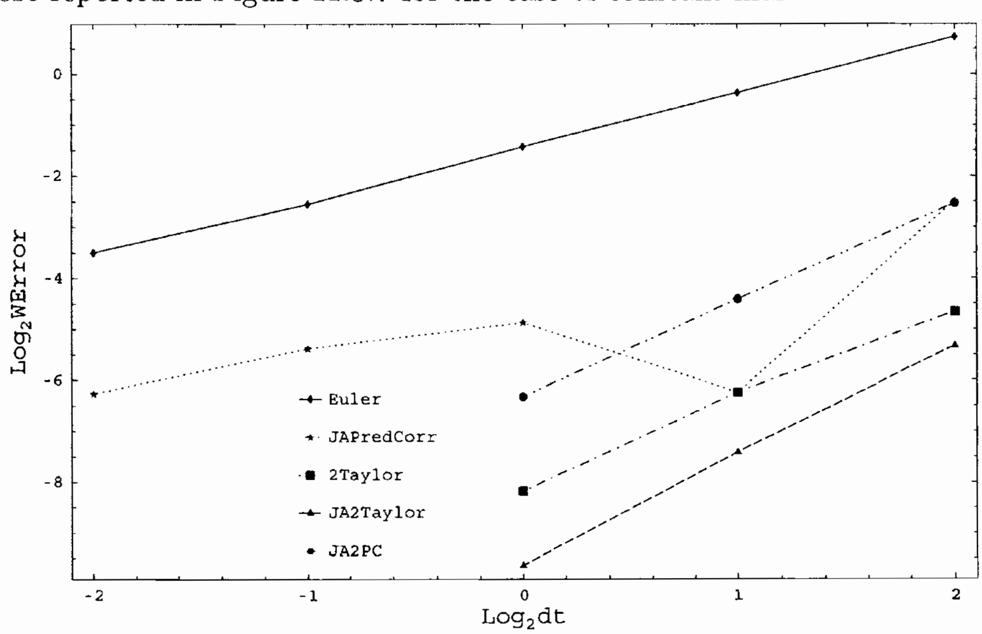

Figure 11.3.7: Log-log plot of weak error versus time step size (call payoff with constant marks)

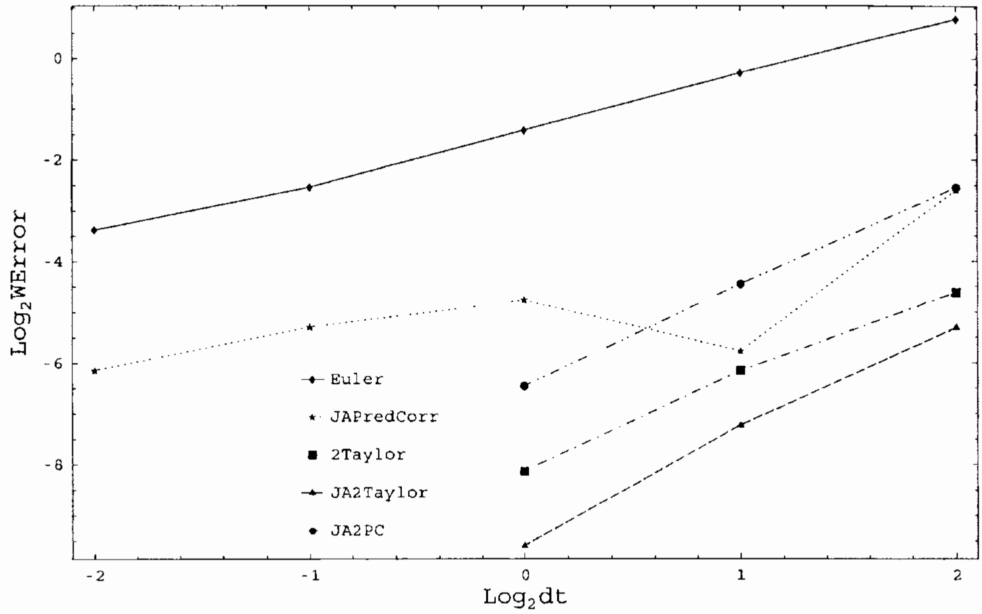

Figure 11.3.8: Log-log plot of weak error versus time step size (call payoff with lognormal marks)

Let us now test the weak schemes proposed when approximating the SDE (11.1.1), for the case of lognormal marks, by using parameter values fitted from real market data. We use here the risk-neutral parameters reported in Andersen & Andreasen (2000) from call options on the S&P 500 in April 1999. The risk-free rate and the dividend yield are given by $r = 0.0559$ and $q = 0.01114$, respectively. With a least-square fit of the Merton model (2.1.5) to the mid implied Black-Scholes volatilities of the S&P 500 in April 1999, Andersen & Andreasen (2000) obtained the following parameters: $\sigma = 0.1765$, $\lambda = 0.089$, $\varrho = -0.8898$ and $\sqrt{\varsigma} = 0.4505$. The last two parameters are the mean and standard deviation of the logarithm of the marks, $\ln(\xi_i)$. These imply that at jump times the S&P 500 drops on average by 54.54%. Note, that these are risk-neutral parameters, and they would have different values under the real-world probability measure. By a general equilibrium analysis Andersen & Andreasen (2000) show that these estimated parameters lead to a reasonable level of risk aversion.

In Figure 11.3.9 we show the results for the regular Euler scheme, the jump-adapted predictor-corrector Euler scheme, the regular and jump-adapted order 2.0 Taylor schemes, and the jump-adapted order 2.0 predictor-corrector schemes, when pricing an at-the-money call option with maturity time $T = 4$. We report that the regular Euler scheme and the jump-adapted Euler scheme, which is not shown in the figure, achieve first order of weak convergence, with the jump-adapted Euler scheme being slightly less accurate than its regular counterpart. The regular and jump-adapted order 2.0 Taylor schemes and the jump-adapted order 2.0 predictor-corrector scheme are significantly more accurate and achieve an order of weak convergence of about $\beta = 2.0$. We emphasize that the jump-adapted predictor-corrector Euler scheme achieves in this example a remarkable accuracy. For large time step sizes it is as good as some second order schemes, with an estimated error slope of about two. By analyzing also other numerical experiments, we conjecture that for smaller time step sizes the jump-adapted predictor-corrector Euler scheme recovers the first order of weak convergence achieved in the case of smooth payoff functions.

In summary, the obtained numerical results indicate that predictor-corrector schemes are very accurate, still when using large time step sizes. This effect is expected to be even more pronounced when approximating non-linear multi-dimensional SDEs. Also, Hunter, Jäckel & Joshi (2001) report that the predictor-corrector Euler scheme is very accurate when pricing interest rate options under the diffusion LIBOR market model. Finally, we remark that in this chapter we have analyzed only the accuracy and not the CPU time needed to run the algorithms. This is of course strongly dependent on the specific implementation and on the problem at hand. However, we remark that predictor-corrector schemes are only slightly computationally more intensive than the corresponding explicit schemes, as they use the same random variables.

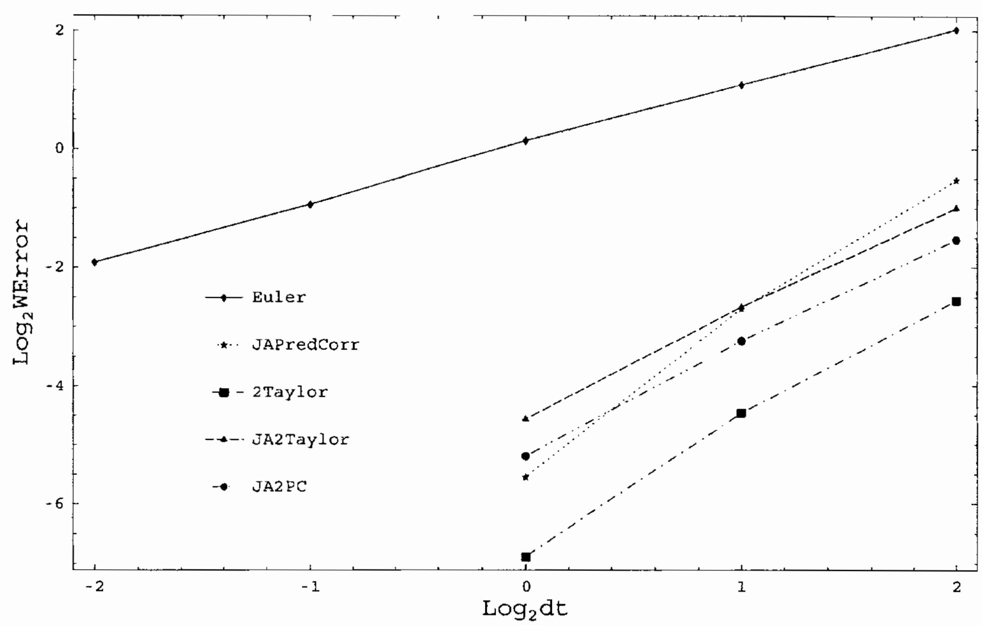

Figure 11.3.9: Log-log plot of weak error versus time step size (call price with lognormal marks)

Chapter 12

Efficiency of Implementation

In this chapter we discuss efficient implementations of discrete time weak approximations for pure diffusion SDEs that exploit the architecture of a digital computer. First we propose an efficient software implementation of simplified weak schemes based on *random bit generators* (RBGs). Then we discuss hardware accelerators based on *field programmable gate arrays* (FPGAs).

As discussed in the previous chapter, a pure diffusion approximation is required between jump times when using a jump-adapted scheme. Therefore, the methods to be presented can be readily applied to the diffusion part in a jump-adapted scheme. Some of the results in this chapter are from Bruti-Liberati & Platen (2004) and Bruti-Liberati, Martini, Piccardi & Platen (2007). The use software-based RBGs for the simulation of simplified weak Taylor schemes has been independently suggested in Milstein & Tretyakov (2004). The analysis of the computational efficiency and the implementation design are the objectives of this chapter.

12.1 Introduction

In order to achieve a required order of weak convergence one can approximate the random variables in a weak Taylor scheme by appropriate discrete random variables, see Chapter 10. For instance, instead of a Gaussian increment one can employ in an Euler scheme a much simpler two-point distributed random variable. The aim of this chapter is to show that an implementation of such simplified schemes based on RBGs significantly increases the computational efficiency of the corresponding Monte Carlo simulation.

A numerical approximation generated by a simplified Euler scheme is equivalent to a random walk. Therefore, its possible states and corresponding probabilities are similar to those of binomial trees. However, while one has to face the curse of dimensionality when generating trees, simplified weak simulation methods employ forward algorithms that generate paths with complexity increasing only polynomially with the dimension. This makes the simplified weak simulation method an efficient tool for high dimensional problems. As we will discuss later in Section 12.4, the numerical properties of simplified methods are similar to those of trees. For instance, we will observe an oscillatory convergence in the case of a Monte Carlo simulation of a European call option, a well-known effect for tree methods, see, for instance, Boyle & Lau (1994).

12.2 Simplified Weak Schemes

Let us consider for simplicity the autonomous one-dimensional pure diffusion SDE

$$dX_t = a(X_t)dt + b(X_t)dW_t \tag{12.2.1}$$

for $t \in [0,T]$, with $X_0 \in \mathbb{R}$, where $W = \{W_t,\, t \in [0,T]\}$ is a standard one-dimensional Wiener process. The SDE (12.2.1) is a special case of the multi-dimensional jump-diffusion SDE (2.1.2) when omitting the jump coefficient, that is when $c(t,x,v) = 0$. Note that the results of this chapter also apply to the diffusive part of simplified jump-adapted approximations of the multi-dimensional jump-diffusion SDE (2.1.2), see Chapter 10.

For simplicity, let us assume an equidistant time discretization with nth discretization time $t_n = n\Delta$ for $n \in \{0,1,\dots,N\}$ where $\Delta = \frac{T}{N}$ and $N \in \{1,2,\dots\}$. In the following we will discuss weak Taylor schemes and their corresponding simplified versions for the one-dimensional pure diffusion SDE (12.2.1). These schemes are special cases of the schemes presented in Chapter 10.

The simplest weak Taylor scheme is the *Euler scheme*, which has weak order of convergence $\beta = 1.0$. It is given by the algorithm

$$Y_{n+1} = Y_n + a\Delta + b\Delta W_n, \tag{12.2.2}$$

where $\Delta W_n = W_{t_{n+1}} - W_{t_n}$ is the Gaussian increment of the Wiener process $W = \{W_t, t \in [0,T]\}$ for $n \in \{0,1,2\dots,N-1\}$ and $Y_0 = X_0$. Here we have used the abbreviations $a = a(Y_n)$ and $b = (Y_n)$ according to the notation introduced in (4.1.9).

If one uses in the above Euler scheme instead of the Gaussian random variables simpler multi-point distributed random variables, then one can still obtain the same weak order of convergence $\beta = 1.0$, see Theorem 14.5.2 in Kloeden & Platen (1999). Furthermore, in Chapter 10 we have shown that this result still holds true when the simplified Euler scheme is used as diffusion approximation in a jump-adapted scheme. For the Euler method these simpler random variables need to satisfy the moment condition (10.2.3). This allows us to replace the Gaussian increment ΔW_n in (12.2.2) by a two-point distributed random variable $\Delta \widehat{W}_{2,n}$, where

$$P(\Delta \widehat{W}_{2,n} = \pm\sqrt{\Delta}) = \frac{1}{2}. \tag{12.2.3}$$

We then obtain the *simplified Euler scheme*

$$Y_{n+1} = Y_n + a\Delta + b\Delta \widehat{W}_{2,n}. \tag{12.2.4}$$

Here the first three moments of the Wiener process increment ΔW_n match those of $\Delta \widehat{W}_{2,n}$, that is

$$\begin{aligned} &E(\Delta W_n) = E(\Delta \widehat{W}_{2,n}) = 0 \qquad E((\Delta W_n)^2) = E((\Delta \widehat{W}_{2,n})^2) = \Delta \\ &E((\Delta W_n)^3) = E((\Delta \widehat{W}_{2,n})^3) = 0. \end{aligned} \tag{12.2.5}$$

The *order* 2.0 *weak Taylor scheme* is given by

$$\begin{aligned} Y_{n+1} &= Y_n + a\Delta + b\Delta W_n + \frac{1}{2}b'b\left\{\left((\Delta W_n)^2\right) - \Delta\right\} + \frac{1}{2}\left(aa' + \frac{1}{2}a''b^2\right)\Delta^2 \\ &\quad + a'b\,\Delta Z_n + \left(ab' + \frac{1}{2}b''b^2\right)\left\{\Delta W_n\,\Delta - \Delta Z_n\right\}, \end{aligned} \tag{12.2.6}$$

where ΔZ_n represents the double Itô integral

$$\Delta Z_n = \int_{t_n}^{t_{n+1}} \int_{t_n}^{s_2} dW_{s_1} ds_2,$$

see also (6.2.22). Note that for SDEs driven by multi-dimensional Wiener processes, the order 2.0 weak Taylor scheme involves certain multiple stochastic integrals that are not Gaussian distributed. However, in the implementation of these schemes one can use Gaussian distributed random variables matching enough moments of these multiple stochastic integrals, see condition (10.7.6). For the purpose of the analysis of this chapter, the term weak Taylor scheme will be used to denote a weak Taylor scheme using Gaussian random variables for the approximation of the underlying multiple stochastic integrals.

In (12.2.6) we can also replace the Gaussian random variables ΔW_n and ΔZ_n by expressions that use a three-point distributed random variable $\Delta \widehat{W}_{3,n}$ with

$$P(\Delta \widehat{W}_{3,n} = \pm\sqrt{3\Delta}) = \frac{1}{6}, \qquad P(\Delta \widehat{W}_{3,n} = 0) = \frac{2}{3}. \tag{12.2.7}$$

Then we obtain the *simplified order* 2.0 *weak scheme*

$$\begin{aligned} Y_{n+1} &= Y_n + a\Delta + b\Delta \widehat{W}_{3,n} + \frac{1}{2} b b' \left\{ \left(\Delta \widehat{W}_{3,n}\right)^2 - \Delta \right\} + \frac{1}{2}\left(a a' + \frac{1}{2} a'' b^2\right)\Delta^2 \\ &\quad + \frac{1}{2}\left(a' b + a b' + \frac{1}{2} b'' b^2\right) \Delta \widehat{W}_{3,n} \Delta. \end{aligned} \tag{12.2.8}$$

Since the three-point distributed random variable $\Delta \widehat{W}_{3,n}$ is such that the first five moments of the increments of the schemes (12.2.6) and (12.2.8) are matched, see condition (10.7.6), then the simplified order 2.0 weak scheme (12.2.8) attains weak order $\beta = 2.0$.

By adding more terms from the Wagner-Platen expansion for pure diffusions and approximating the arising multiple stochastic integrals with Gaussian random variables, we obtain the following *order* 3.0 *weak scheme* given by

$$\begin{aligned} Y_{n+1} &= Y_n + a\Delta + b\Delta W_n + \frac{1}{2} L^{(1)} b \left\{ (\Delta W_n)^2 - \Delta \right\} \\ &\quad + L^{(1)} a \Delta Z_n + \frac{1}{2} L^{(0)} a \Delta^2 + L^{(0)} b \left\{ \Delta W_n \Delta - \Delta Z_n \right\} \\ &\quad + \frac{1}{6}\left(L^{(0)} L^{(0)} b + L^{(0)} L^{(1)} a + L^{(1)} L^{(0)} a \right) \left\{ \Delta W_n \Delta^2 \right\} \\ &\quad + \frac{1}{6}\left(L^{(1)} L^{(1)} a + L^{(1)} L^{(0)} b + L^{(0)} L^{(1)} b \right) \left\{ (\Delta W_n)^2 - \Delta \right\} \Delta \\ &\quad + \frac{1}{6} L^{(0)} L^{(0)} a \, \Delta^3 + \frac{1}{6} L^{(1)} L^{(1)} b \left\{ (\Delta W_n)^2 - 3\Delta \right\} \Delta W_n, \end{aligned} \tag{12.2.9}$$

where $L^{(0)}$ and $L^{(1)}$ are differential operators defined by

$$L^{(0)} = a \frac{\partial}{\partial x} + \frac{1}{2} b^2 \frac{\partial^2}{\partial x^2} \quad \text{and} \quad L^{(1)} = b \frac{\partial}{\partial x}, \tag{12.2.10}$$

see also (3.3.4)-(3.3.5). This scheme achieves an order of weak convergence $\beta = 3.0$.

To construct a third order simplified method, it is sufficient to use multi-point distributed random variables that match the first seven moments of the Gaussian ones, see condition (10.2.19). As discussed in Chapter 10, Hofmann (1994) proposed the four-point distributed random variable (10.2.20) that satisfies this moment condition. However, such a four-point distributed random variable cannot be easily implemented by the method described below because its probability values are not rational numbers. Nonetheless, we can use the five-point distributed random variable $\Delta\widehat{W}_{5,n}$, with

$$P(\Delta\widehat{W}_{5,n} = \pm\sqrt{6\Delta}) = \frac{1}{30}, \qquad P(\Delta\widehat{W}_{5,n} = \pm\sqrt{\Delta}) = \frac{9}{30},$$
$$P(\Delta\widehat{W}_{5,n} = 0) = \frac{1}{3}, \tag{12.2.11}$$

see also Bruti-Liberati, Martini, Piccardi & Platen (2007). This five-point distributed random variable matches the first seven moments of the Gaussian ones and is suitable for a highly efficient implementation based on RBGs. The corresponding *simplified order* 3.0 *weak scheme* is given by

$$\begin{aligned} Y_{n+1} = \; & Y_n + a\Delta + b\Delta\widehat{W}_{5,n} + \frac{1}{2}L^{(1)}b\left\{\left(\Delta\widehat{W}_{5,n}\right)^2 - \Delta\right\} + \frac{1}{2}L^{(0)}a\Delta^2 \\ & + \frac{1}{2}L^{(1)}a\left\{\Delta\widehat{W}_{5,n} + \frac{1}{\sqrt{3}}\Delta\widehat{W}_{2,n}\right\}\Delta + \frac{1}{2}L^{(0)}b\left\{\Delta\widehat{W}_{5,n} - \frac{1}{\sqrt{3}}\Delta\widehat{W}_{2,n}\right\}\Delta \\ & + \frac{1}{6}\left(L^{(0)}L^{(0)}b + L^{(0)}L^{(1)}a + L^{(1)}L^{(0)}a\right)\Delta\widehat{W}_{5,n}\,\Delta^2 \\ & + \frac{1}{6}\left(L^{(1)}L^{(1)}a + L^{(1)}L^{(0)}b + L^{(0)}L^{(1)}b\right)\left\{(\Delta\widehat{W}_{5,n})^2 - \Delta\right\}\Delta \\ & + \frac{1}{6}L^{(0)}L^{(0)}a\Delta^3 + \frac{1}{6}L^{(1)}L^{(1)}b\left\{(\Delta\widehat{W}_{5,n})^2 - 3\,\Delta\right\}\Delta\widehat{W}_{5,n}, \end{aligned} \tag{12.2.12}$$

where $\Delta\widehat{W}_{2,n}$ is the two-point distributed random variable defined in (12.2.3). This order 3.0 simplified weak scheme achieves an order of weak convergence $\beta = 3.0$.

12.3 Multi-Point Random Variables and Random Bit Generators

As discussed above, the random variables appearing in simplified weak schemes are multi-point distributed. Therefore, highly efficient implementations of Monte

Carlo simulations using simplified weak schemes require fast multi-point random number generators. These can be obtained by RBGs. In this case RBGs substitute the Gaussian random number generators needed for weak Taylor schemes.

The multi-point distributed random variables and the corresponding RBGs to be presented can be applied to any weak scheme, including derivative-free, predictor-corrector and implicit schemes. As shown in Chapter 10, the above described schemes can be used as diffusion components in corresponding jump-adapted schemes for the approximation of the multi-dimensional jump-diffusion SDE (2.1.2).

An RBG is an algorithm that generates a bit 0 or 1 with probability 0.5. Random bits can be obtained on a digital computer via the so-called shift register generator. This generator, used in digital communication, see Golomb (1964), relies on the theory of primitive polynomials modulo 2. These are special polynomials of the form

$$y(x) = 1 + c_1 x + \ldots + c_{n-1} x^{n-1} + x^n, \tag{12.3.13}$$

with coefficients $c_i = \{0, 1\}$. A primitive polynomial modulo 2 of order n defines a recurrence relation for obtaining a new bit from the n preceding ones with maximal period, which is $2^n - 1$. The recurrence is given by

$$a_{k+1} = c_1 a_k + c_2 a_{k-1} + \ldots + c_{n-1} a_{k-n+2} + a_{k-n+1} \,(mod\, 2), \tag{12.3.14}$$

where a_{k+1} is the new bit obtained from the preceding ones, a_i, with $i \in \{n, \ldots, 1\}$ and $k > n$. Equation (12.3.14) can be rewritten as

$$a_{k+1} = c_1 a_k \oplus c_2 a_{k-1} \oplus \ldots \oplus c_{n-1} a_{k-n+2} \oplus a_{k-n+1}, \tag{12.3.15}$$

where $\oplus$ is the "exclusive or" operator. Thus, RBGs can be efficiently implemented in C or C++ via bitwise operations, see Press, Teukolsky, Vetterling & Flannery (2002). For a study of random number generators based on primitive polynomials modulo 2 we refer to Tausworthe (1965).

For a first order simplified scheme, as the simplified Euler scheme (12.2.4), each bit obtained from the RBG is used to generate a value for the two-point distributed random variable $\Delta \widehat{W}_{2,n}$ by a simple look-up operation ($0 \to +\sqrt{\Delta}, 1 \to -\sqrt{\Delta}$). For a second order simplified scheme, as (12.2.8), one bit is not sufficient to generate a value for the required three-point distributed random variable $\Delta \widehat{W}_{3,n}$. However, a sequence of three generated random bits can be used to obtain eight equiprobable combinations. Two of these combinations are discarded by an acceptance-rejection method. Then we use four of the remaining six combinations to generate the 0 value for the random variable and one combination each for obtaining the values $+\sqrt{3\Delta}$ and $-\sqrt{3\Delta}$. For the third order simplified scheme (12.2.12), the random variable $\Delta\widehat{W}_{5,n}$ is five-point distributed with the probability distribution (12.2.11). In this case, a sequence of five random bits is used to generate 32 equiprobable combinations. The acceptance-rejection method discards two of them, uses ten to generate the 0 value, nine each for values $+\sqrt{\Delta}$ and $-\sqrt{\Delta}$, and one each for values $+\sqrt{6\Delta}$ and $-\sqrt{6\Delta}$.

12.4 Software Implementation

12.4.1 Random Bit Generators in C++

In the following we report a comparative study on the efficiency of a software implementation in C++ of the above described schemes. The reference personal computer (PC) used in this study is a Pentium 2.4 Ghz (CPU id: x86 Family 15 Model 2 Stepping 7) and the C++ compiler is the Mingw port of GCC (GNU Compiler Collection).

A widely used and efficient method to generate a pair of independent standard Gaussian random variables is the polar Marsaglia-Bray method coupled with a linear congruential random number generator, as described in Press, Teukolsky, Vetterling & Flannery (2002). In our comparative study we use, as Gaussian random number generator, the routine *gasdev*, see p. 293 of Press, Teukolsky, Vetterling & Flannery (2002). Note that the period of this Gaussian random number generator equals 2^{31}. We refer to Devroye (1986) for alternative Gaussian random number generators.

For an unbiased comparison, we have implemented an RBG based on the following primitive polynomial modulo 2 of order 31:

$$y(x) = x^{31} + x^3 + 1. \tag{12.4.16}$$

In this way, the RBG has period $2^{31} - 1$, which is virtually the same as that of the Gaussian random number generator. As explained in the previous section, by simple look-up operations and acceptance-rejection methods, we can obtain the multi-point distributed random numbers required in simplified weak schemes.

The C++ implementation of the two-point random number generator, see (12.2.3), is reported in Figure 12.4.1. It is based on a similar code presented by Press, Teukolsky, Vetterling & Flannery (2002).

```
int rbit1per31(unsigned long & iseed)
{     unsigned long newbit;
      newbit = ((iseed >> 30) & 1)
      ^ ((iseed >> 2) & 1);
      iseed = (iseed << 1) | newbit;
      return int(newbit); }
```

Figure 12.4.1: C++ code of the two-point random number generator.

On the test computer the CPU time needed to generate 100 million random numbers with the polar Marsaglia-Bray method amounts to 41.6 seconds. The two-point random number generator, described above, is almost 47 times faster using only 0.89 seconds.

As discussed in the previous section, for simplified methods of higher order, similar multi-point random number generators can be constructed. For the simplified order 2.0 weak scheme (12.2.8) it is sufficient to use a three-point random number generator. A corresponding code is presented in Figure 12.4.2. It produces three bits coupled with an acceptance-rejection method. On the test computer the CPU time needed to generate 100 million random numbers with this generator amounts to 6.96 seconds, which is still almost 6 times faster than the polar Marsaglia-Bray method.

```
int rbit3per31(unsigned long & iseed)
{     int x1 = 1, x2 = 1, x3 = 0;
      while ( (x1 == 1 && x2 == 1 && x3 == 0)
      || (x1 == 0 && x2 == 1 && x3 == 1))
      { x1 = rbit1per31(iseed);
        x2 = rbit1per31(iseed);
        x3 = rbit1per31(iseed); }
      return x1 - x3; }
```

Figure 12.4.2: C++ code of the three-point random number generator.

Now we present some numerical results for the Euler and the order 2.0 weak Taylor schemes, as well as for their simplified versions. As test dynamics we choose an SDE with multiplicative noise, where

$$dX_t = \mu X_t dt + \sigma X_t dW_t \tag{12.4.17}$$

for $t \in [0, T]$ and $X_0 \in \mathbb{R}$. The SDE admits the closed form solution

$$X_T = X_0 \exp\left\{\left(\mu - \frac{\sigma^2}{2}\right) T + \sigma W_T\right\}. \tag{12.4.18}$$

Let us analyze the CPU time needed to run a Monte Carlo simulation that computes the payoff of a call option with 400000 paths and 128 time steps. We report a CPU time of 22 and 23.5 seconds for the Euler scheme and the order 2.0 weak Taylor scheme, respectively. The corresponding simplified versions only require 0.75 and 4.86 seconds, respectively. Thus, for the Euler method the simplified version is about 29 times faster than the Gaussian one. The simplified order 2.0 weak scheme is nearly five times faster than the order 2.0 weak Taylor scheme.

12.4.2 Experimental Results

We now present some numerical results on the accuracy of the Euler scheme, the order 2.0 weak Taylor scheme and their simplified versions when applied to the SDE (12.4.17). As in the previous chapter, we show several log-log plots with the logarithm, $\log_2(\varepsilon_w(\Delta))$, of the weak error, as defined in (1.2.2), versus the logarithm, $\log_2(\Delta)$, of the maximum time step size Δ. Additionally, we will show some plots of the relative weak error, that is

$$\left|\frac{E\left(g(X_T)\right) - E\left(g(Y_T)\right)}{E\left(g(X_T)\right)}\right|, \tag{12.4.19}$$

with a fixed time step size Δ.

The Case of a Smooth Payoff Function

We first study the computation of $E(g(X_T))$, where g is a smooth function and X_T is the solution of the SDE (12.4.17). Note that the estimation of the expected value of a smooth function of the solution X arises, for instance, when computing expected utilities. Another important application is the calculation of Value at Risk via the simulation of moments, as applied in Edgeworth expansions and saddle point methods, see Studer (2001). Therefore, we consider here the estimation of the kth moment $E((X_T)^k)$ of X_T at time T, for $k \in \mathbb{N}$.

By (12.4.18) we obtain the kth moment of X_T in closed form as

$$E\left((X_T)^k\right) = (X_0)^k \exp\left\{\left(kT(\mu + \frac{\sigma^2}{2}(k-1)\right)\right\} \tag{12.4.20}$$

for $k \in \mathbb{N}$.

The particular structure of the SDE (12.4.17) allows us to obtain a closed form solution also for the estimator of the kth moment provided by the Euler and the order 2.0 weak Taylor schemes, and by their simplified versions.

Let us rewrite the Euler scheme (12.2.2) as

$$Y_T = Y_0 \prod_{n=0}^{N-1} (1 + \frac{\mu T}{N} + \sigma \Delta W_n), \tag{12.4.21}$$

where ΔW_n are i.i.d Gaussian $\mathcal{N}(0, \frac{T}{N})$ random variables and $N = \frac{T}{\Delta}$. Note that

$$E\left((\Delta W_n)^k\right) = \begin{cases} \frac{k!}{(k/2)!2^{k/2}}\left(\frac{T}{N}\right)^{k/2} & \text{for k even} \\ 0 & \text{for k odd.} \end{cases} \tag{12.4.22}$$

By the independence of ΔW_n, for $n \in \{0, 1, \ldots, N-1\}$ and (12.4.22) we obtain

$$\begin{aligned} E\left((Y_T)^k\right) &= (Y_0)^k E\left(\prod_{n=0}^{N-1}(1 + \frac{\mu T}{N} + \sigma\Delta W_n)^k\right) \\ &= (Y_0)^k \prod_{n=0}^{N-1} E\left((1 + \frac{\mu T}{N} + \sigma\Delta W_n)^k\right) \\ &= (Y_0)^k \prod_{n=0}^{N-1}\sum_{i=0}^{k} \binom{k}{i} (1 + \frac{\mu T}{N})^{k-i} \sigma^i E\left((\Delta W_n)^i\right) \qquad (12.4.23) \\ &= (Y_0)^k \left(\sum_{q=0}^{[k/2]} \binom{k}{2q} \left(1 + \frac{\mu T}{N}\right)^{k-2q} \frac{(2q)!}{q!}\left(\frac{\sigma^2 T}{2N}\right)^q\right)^N. \end{aligned}$$

We recall that by $[z]$ we denote the integer part of $z \in \mathbb{R}$ and by $\binom{i}{l}$, for $i \geq l$, the combinatorial coefficient, see (8.5.15).

Similarly, for the simplified Euler scheme (12.2.4) we obtain

$$
\begin{aligned}
E\left((Y_T)^k\right) &= (Y_0)^k\, E\left(\prod_{n=0}^{N-1}(1+\frac{\mu T}{N}+\sigma\Delta\widehat{W}_{2,n})^k\right)\\
&= (Y_0)^k\left(\sum_{q=0}^{[k/2]}\binom{k}{2q}\left(1+\frac{\mu T}{N}\right)^{k-2q}\left(\frac{\sigma^2 T}{N}\right)^q\right)^N, \qquad (12.4.24)
\end{aligned}
$$

where $\Delta\widehat{W}_{2,n}$ is the two-point distributed random variable given in (12.2.3). Here we have used the result

$$
E\left(\left(\Delta\widehat{W}_{2,n}\right)^k\right)=\begin{cases}\left(\dfrac{T}{N}\right)^{k/2} & \text{for k even}\\ 0 & \text{for k odd.}\end{cases} \qquad (12.4.25)
$$

By comparing (12.4.23) to (12.4.24), we notice that the Euler scheme and the simplified Euler scheme give the same estimator for the expectation $E((Y_T)^k)$ with $k \in \{1,2,3\}$.

Let us now rewrite the order 2.0 weak Taylor scheme (12.2.6) for the SDE (12.4.17) as

$$
Y_T = Y_0\prod_{n=0}^{N-1}\left(h_1+h_2\Delta W_n+h_3\left(\Delta W_n\right)^2\right), \qquad (12.4.26)
$$

with

$$
h_1 = 1+\left(\mu-\frac{\sigma^2}{2}\right)\frac{T}{N}+\frac{\mu^2}{2}\frac{T^2}{N^2}, \quad h_2=\sigma+\mu\sigma\frac{T}{N} \quad \text{and} \quad h_3=\frac{\sigma^2}{2}. \qquad (12.4.27)
$$

By the independence of ΔW_n, for $n \in \{0,1,\ldots,N-1\}$ and (12.4.22) we obtain

$$
\begin{aligned}
E\left((Y_T)^k\right) &= (Y_0)^k\prod_{n=0}^{N-1}E\left(\left(h_1+h_2\Delta W_n+h_3\left(\Delta W_n\right)^2\right)^k\right)\\
&= (Y_0)^k\prod_{n=0}^{N-1}\sum_{i=0}^{k}\binom{k}{i}h_1^{k-i}E\left(\left(h_2\Delta W_n+h_3\left(\Delta W_n\right)^2\right)^i\right)\\
&= (Y_0)^k\prod_{n=0}^{N-1}\sum_{i=0}^{k}\binom{k}{i}h_1^{k-i}\sum_{j=0}^{i}\binom{i}{j}h_2^{i-j}h_3^jE\left(\left(\Delta W_n\right)^{i+j}\right)
\end{aligned}
$$

$$
\begin{aligned}
&= (Y_0)^k \prod_{n=0}^{N-1} \left(\sum_{q=0}^{[k/2]} \binom{k}{2q} h_1^{k-2q} \sum_{l=0}^{q} \binom{2q}{2l} h_2^{2(q-l)} h_3^{2l} E\left((\Delta W_n)^{2(q+l)} \right) \right. \\
&\quad + \sum_{q=0}^{[(k-1)/2]} \binom{k}{2q+1} h_1^{k-(2q+1)} \sum_{l=0}^{q} \binom{2q+1}{2l+1} h_2^{2(q-l)} h_3^{2l+1} \\
&\quad \left. \times E\left((\Delta W_n)^{2(q+l+1)} \right) \right) \\
&= (Y_0)^k \left(\sum_{q=0}^{[k/2]} \binom{k}{2q} h_1^{k-2q} \sum_{l=0}^{q} \binom{2q}{2l} h_2^{2(q-l)} h_3^{2l} \frac{(2(q+l))!}{(q+l)!} \left(\frac{T}{2N} \right)^{q+l} \right. \\
&\quad + \sum_{q=0}^{[(k-1)/2]} \binom{k}{2q+1} h_1^{k-(2q+1)} \sum_{l=0}^{q} \binom{2q+1}{2l+1} h_2^{2(q-l)} h_3^{2l+1} \\
&\quad \left. \times \frac{(2(q+l+1))!}{(q+l+1)!} \left(\frac{T}{2N} \right)^{q+l+1} \right)^N . \qquad (12.4.28)
\end{aligned}
$$

For the order 2.0 simplified weak scheme (12.2.8), with the three-point distributed random variable $\Delta\widehat{W}_{3,n}$, see (12.2.7), we obtain

$$
\begin{aligned}
E\left((Y_T)^k \right) &= (Y_0)^k \prod_{n=0}^{N-1} E\left(\left(h_1 + h_2 \Delta\widehat{W}_{3,n} + h_3 \left(\Delta\widehat{W}_{3,n} \right)^2 \right)^k \right) \\
&= (Y_0)^k \left(h_1^k + \sum_{q=1}^{[k/2]} \binom{k}{2q} h_1^{k-2q} \sum_{l=0}^{q} \binom{2q}{2l} h_2^{2(q-l)} h_3^{2l} \frac{1}{3} \left(\frac{3T}{N} \right)^{q+l} \right. \\
&\quad + \sum_{q=0}^{[(k-1)/2]} \binom{k}{2q+1} h_1^{k-(2q+1)} \sum_{l=0}^{q} \binom{2q+1}{2l+1} h_2^{2(q-l)} h_3^{2l+1} \\
&\quad \left. \times \frac{1}{3} \left(\frac{3T}{N} \right)^{q+l+1} \right)^N , \qquad (12.4.29)
\end{aligned}
$$

where we have used the result

$$
E\left(\left(\Delta\widehat{W}_{3,n} \right)^k \right) = \begin{cases} \frac{1}{3} \left(\frac{3T}{N} \right)^{k/2} & \text{for k even} \\ 0 & \text{for k odd,} \end{cases} \qquad (12.4.30)
$$

for $k\{1, 2 \ldots\}$. Therefore, the order 2.0 weak Taylor scheme and the simplified order 2.0 weak scheme provide the same estimate for $E((Y_T)^k)$ with $k \in \{1, 2\}$.

In the following, we consider the estimation of the fifth moment, that is $E((Y_T)^5)$. We consider the following parameters: $X_0 = 1$, $\mu = 0.1$, $\sigma = 0.15$, $T = 1$. Note that by comparing the closed form solution (12.4.20) with the estimators (12.4.23), (12.4.28), (12.4.24) and (12.4.29), we can compute explicitly the weak error $\varepsilon_w(\Delta)$ for the Euler and the order 2.0 weak Taylor schemes, as well as for their simplified versions.

In Figure 12.4.3 we show the logarithm $log_2(\varepsilon_w(\Delta))$ of the weak error for the Euler and the order 2.0 weak Taylor schemes together with their simplified versions versus the logarithm $log_2(\Delta)$ of the time step size. Note that the Euler and the simplified Euler schemes reproduce in the log-log plot the theoretically predicted weak order $\beta = 1.0$. Furthermore, the order 2.0 weak Taylor scheme and its simplified version achieve a weak order of $\beta = 2.0$, as expected. The errors generated by the simplified schemes are very similar to those of the corresponding weak Taylor schemes.

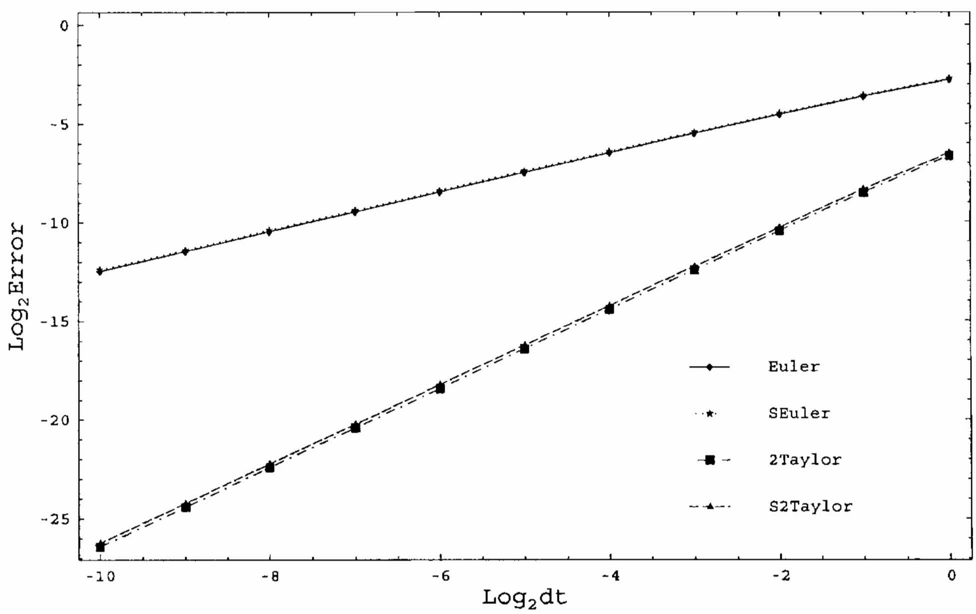

Figure 12.4.3: Log-log plot of the weak error for the Euler, the simplified Euler, the order 2.0 weak Taylor and the simplified order 2.0 weak schemes

We know that simplified schemes achieve the same order of weak convergence as their counterparts based on Gaussian random variables. However, it is useful to check whether there is a significant loss in accuracy when the time step size Δ is large.

In the following we analyze the relative weak error when estimating the fifth moment in our example. This means that we consider the quantity

$$\left| \frac{E\left((X_T)^5\right) - E\left((Y_T)^5\right)}{E\left((X_T)^5\right)} \right|, \tag{12.4.31}$$

where the parameters are set as before with $T \in [1/365, 3]$ and $\sigma \in [0.05, 0.5]$. Moreover, we use only one time step, which means that the time step size Δ used for the discrete time approximation Y_T equals T. In Figure 12.4.4 we report the relative error generated by the Euler scheme. For small values of the time to maturity T and of the volatility σ, the Euler scheme is very precise even by using only one time step, as considered here. For instance, when the maturity time is set to 2/12 and the volatility to 0.1 we obtain a relative error of 0.13%. When the time to maturity and the volatility increase, the accuracy of the Euler scheme is not satisfactory. For instance, for $T = 3$ and $\sigma = 0.5$ the relative weak error amounts to 99.6%. In Figure 12.4.5 we report the relative weak error generated by the simplified Euler scheme for the same parameter values. The results are similar to those obtained in Figure 12.4.4 for the Euler scheme based on Gaussian random variables. Finally, Figure 12.4.6 reports the difference of the relative errors generated by the simplified Euler scheme and by the Euler scheme. Here we can notice that the loss in accuracy due to the use of simplified schemes does not exceed 4%.

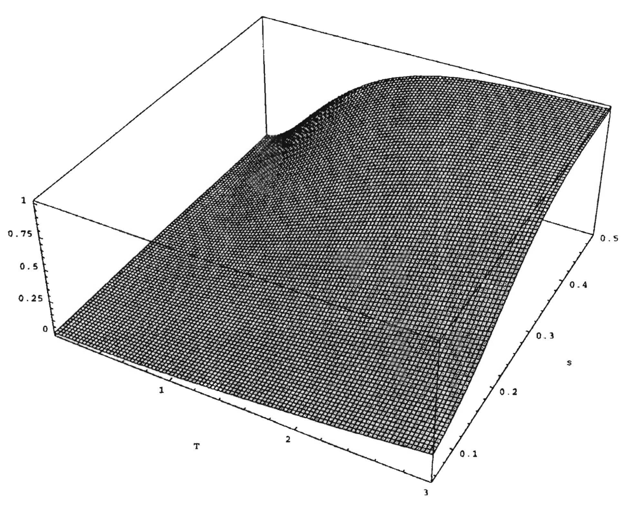

Figure 12.4.4: Relative error for the Euler scheme with $\Delta = T$

The relative errors generated by the order 2.0 weak Taylor scheme and its simplified version are significantly smaller than those reported for the Euler and the simplified Euler schemes. However, the qualitative behavior with respect to the values of the time to maturity T and the volatility σ is similar. In Figure 12.4.7 we report the

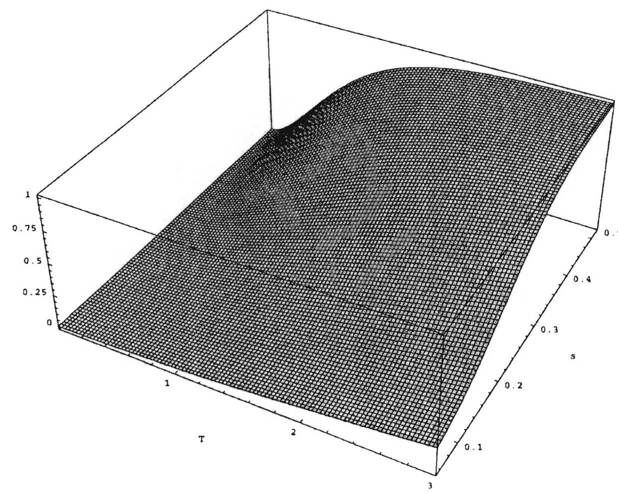

Figure 12.4.5: Relative error for the simplified Euler scheme with $\Delta = T$

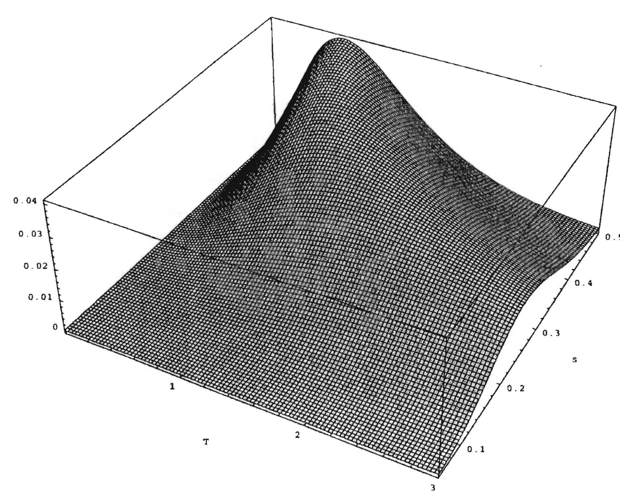

Figure 12.4.6: Relative error of the simplified Euler scheme minus the relative error of the Euler scheme

relative error of the simplified order 2.0 weak scheme minus the relative error of the order 2.0 weak Taylor scheme for the same parameters considered in the previous plots. Also in this case the loss in accuracy generated by the multi-point distributed random variables is limited for all parameter values tested.

We also report that we obtained similar results when estimating higher moments. For instance, in Figure 12.4.8 we plot the difference in the relative errors of the simplified Euler scheme and of the Euler scheme for the tenth moment.

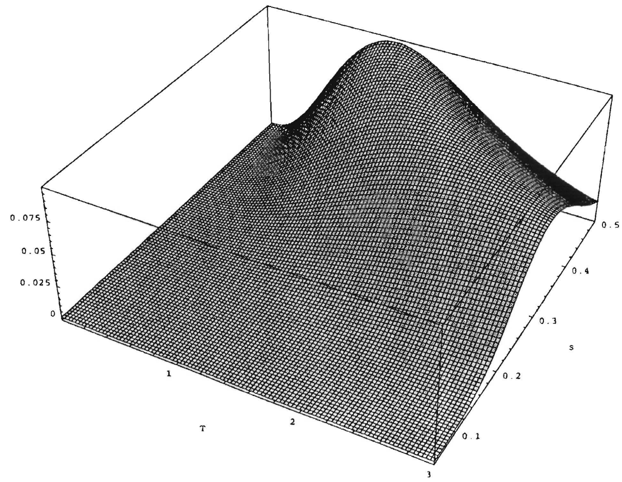

Figure 12.4.7: Relative error of the simplified order 2.0 weak scheme minus the relative error of the order 2.0 weak Taylor scheme

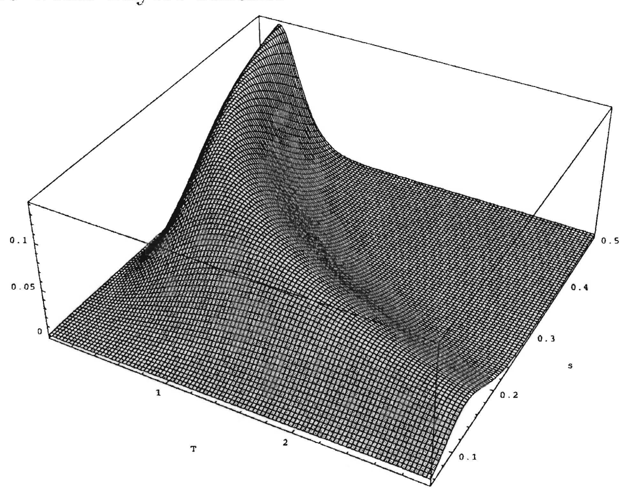

Figure 12.4.8: Relative error of the simplified Euler scheme minus the relative error of the Euler scheme for the tenth moment

The Case of a Non-Smooth Payoff

In option pricing we are confronted with the computation of expectations of non-smooth payoffs. Note that typically weak convergence theorems for discrete time approximations require smooth payoff functions. We refer to Bally & Talay (1996a, 1996b) and Guyon (2006) for weak convergence theorems for the Euler scheme in the case of non-smooth payoff functions.

To give a simple example, let us compute the price of a European call option. In this case we assume that the dynamics of the SDE (12.4.17) is specified under the risk neutral measure, so that the drift coefficient μ equals the risk-free rate r. In this case the price of the European call option is given by the expected value of the continuous but only piecewise differentiable payoff $e^{-\mu T}(X_T - K)^+ = e^{-\mu T}\max(X_T - K, 0)$ with strike price K. By (12.4.18), we obtain

$$c_{T,K}(X) := E(e^{-\mu T}(X_T - K)^+) = X_0\mathcal{N}(d_1) - Ke^{-\mu T}\mathcal{N}(d_2), \tag{12.4.32}$$

where $d_1 = \frac{\ln(\frac{X_0}{K}) + (\mu + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}$ and $d_2 = d_1 - \sigma\sqrt{T}$.

For this particular example the estimator of the simplified Euler scheme and that of the simplified order 2.0 weak scheme can be obtained in closed form. For the simplified Euler scheme we obtain

$$\begin{aligned} c_{T,K}(Y) &= \frac{e^{-\mu T}}{2^N}\sum_{i=0}^{N}\binom{N}{i}\ \Bigg(Y_0\Big(1 + \frac{\mu T}{N} + \sqrt{\frac{T}{N}}\Big)^{N-i} \\ &\qquad \times\Big(1 + \frac{\mu T}{N} - \sqrt{\frac{T}{N}}\Big)^{i} - K\Bigg)^{+}. \end{aligned} \tag{12.4.33}$$

For the simplified order 2.0 weak scheme we have

$$\begin{aligned} c_{T,K}(Y) &= \frac{e^{-\mu T}}{6^N}\sum_{j=0}^{N}\binom{N}{j}4^{N-j}\sum_{i=0}^{j}\binom{j}{i}\Bigg(Y_0 h_1^{N-j}\Big(h_1 + h_2\sqrt{\frac{3T}{N}} + h_3\frac{3T}{N}\Big)^{j-i} \\ &\qquad \times\Big(h_1 - h_2\sqrt{\frac{3T}{N}} + h_3\frac{3T}{N}\Big)^{i} - K\Bigg)^{+}, \end{aligned} \tag{12.4.34}$$

where h_1, h_2 and h_3 are defined in (12.4.27).

For the Euler scheme and the order 2.0 Taylor scheme, we resort to Monte Carlo simulation to obtain an estimate of the expected value $E(e^{-\mu T}(Y_T - K)^+)$. Note that the number of generated sample paths used in the following numerical experiments is large enough to ensure that the statistical error is negligible when compared to the systematic error. In the special case of one time step, that is $\Delta = T$, the estimator of the call price for the Euler scheme and for the order 2.0 weak Taylor scheme can be obtained in closed form. For the Euler scheme with one time step we have

$$c_{T,K}(Y) = e^{-\mu T}\Big(\big(Y_0(1 + \mu T) - K\big)\mathcal{N}(b) + Y_0\sigma\sqrt{T}\mathcal{N}'(b)\Big),$$

with

$$b = \frac{K - Y_0(1+\mu T)}{Y_0 \sigma \sqrt{T}} \qquad \text{and} \qquad \mathcal{N}'(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}},$$

for $x \in \Re$. For the order 2.0 weak Taylor scheme with one time step, if

$$Y_0\big(Y_0(h_2)^2 - 4h_3(Y_0 h_1 - K)\big) > 0$$

we obtain

$$\begin{aligned} c_{T,K}(Y) &= e^{-\mu T}\Big(\big(Y_0(h_1+h_3) - K\big)\big(\mathcal{N}(b_-) + \mathcal{N}(b_+)\big) \\ &\quad + Y_0\Big(\big(\mathcal{N}'(b_+) - \mathcal{N}'(b_-)\big)\big(h_2 + h_3(b_+ - b_-)\big)\Big)\Big), \end{aligned}$$

where

$$b_{\pm} = \frac{-Y_0 h_2 \pm \sqrt{(Y_0 h_2)^2 - 4X_0 h_3(Y_0 h_1 - K)}}{2X_0 h_3},$$

and h_1, h_2 and h_3 are defined in (12.4.27). If instead

$$Y_0\big(Y_0(h_2)^2 - 4h_3(Y_0 h_1 - K)\big) \leq 0,$$

then we have

$$c_{T,K}(Y) = Y_0(h_1 + h_3) - K.$$

These closed form solutions allow us to show the weak error generated by these schemes, when $\Delta = T$, avoiding any statistical error that would arise from a Monte Carlo simulation. Note that for some range of parameters used in our study the weak error is very small and it would be unfeasible to use Monte Carlo simulation to obtain an estimate with a negligible statistical error.

In Figure 12.4.9 we show the log-log weak error plot of an at-the-money call payoff with strike $K = X_0$. The Euler and the simplified Euler schemes have a very similar accuracy and generate a weak order $\beta = 1.0$ with the log-error forming a perfect line in dependence on the log-time step size. The order 2.0 weak Taylor scheme is more accurate and achieves an order of weak convergence of about $\beta = 2.0$. The accuracy of the simplified order 2.0 weak scheme is comparable to that of its Gaussian counterpart, but its convergence is much more erratic.

We report that when testing these schemes with different sets of parameters, we

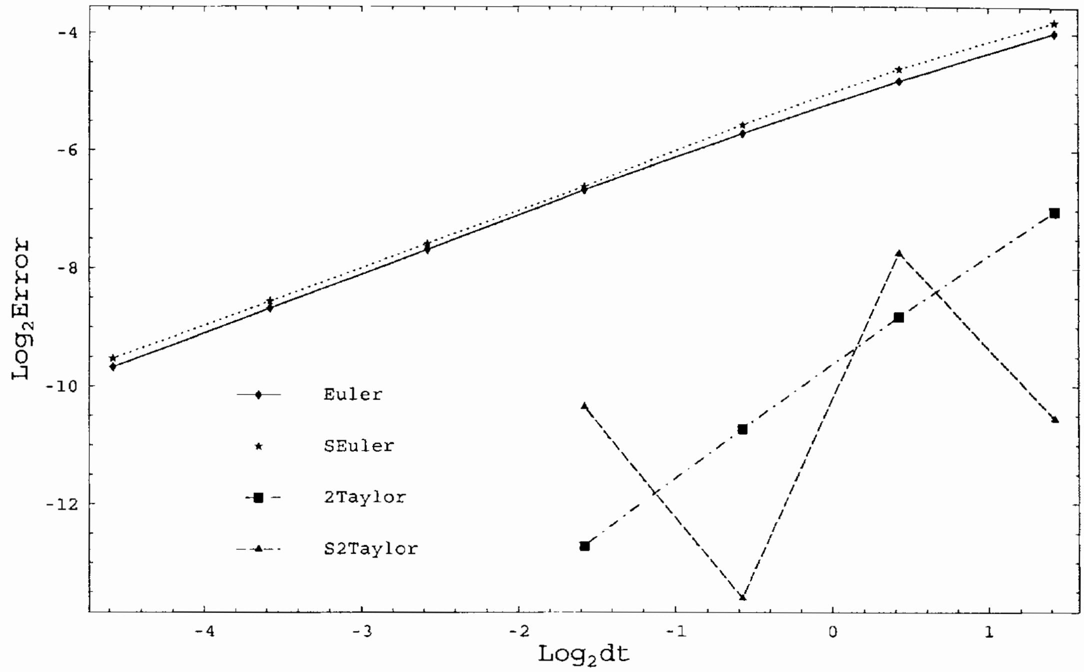

Figure 12.4.9: Log-log plot of the weak error for the Euler, the simplified Euler, the order 2.0 weak Taylor and the simplified order 2.0 weak Taylor schemes

notice that simplified schemes have an accuracy similar to that of corresponding schemes based on Gaussian random variables, but their convergence exhibits oscillations. This effect seems to be more pronounced in the simplified order 2.0 weak Taylor scheme, but it is also present when using the simplified Euler scheme.

As mentioned earlier, the simplified Euler scheme and the simplified order 2.0 weak schemes are approximately equivalent to certain binomial and trinomial trees, respectively. The erratic behavior of the weak error is due to the discrete nature of the multi-point distributed random variables used. This appears to be the same effect that was noticed for tree methods in Boyle & Lau (1994).

As in the previous section, we now compare for a wide range of parameters the accuracy of weak Taylor schemes to that of simplified weak Taylor schemes when using only one time step. We consider the relative weak error,

$$\left| \frac{c_{T,K}(X) - c_{T,K}(Y)}{c_{T,K}(X)} \right|, \tag{12.4.35}$$

with $X_0 = Y_0 = 1$, $\mu = 0.1$, $\sigma \in [0.05, 0.5]$, $K \in [0.88, 1.12]$ and $T = 3/12$. In Figure 12.4.10 we report the relative error generated by the Euler scheme. We notice that for small values of the volatility and large values of the strike the relative error increases by up to 70%. Figure 12.4.11 shows the relative error for the simplified Euler scheme, which is extreme for large strikes and small volatilities. Also in this case the loss in accuracy due to the use of the two-point distributed random variables (12.2.3) is rather small outside a region of high strike prices and low volatilities. Note that since we are using only one time step, for every value of $K \geq Y_0(1 + \mu T + \sigma\sqrt{T})$ all sample paths of the simplified Euler scheme finish out-the-money and, thus, the expected call payoff is evaluated to zero. On the other hand, some of the paths generated by the Euler scheme end up in-the-money, providing a positive value for the expected call payoff which is, thus, closer to the exact solution. Note, however, that in this region the exact value of the expected payoff is very small. For instance, for $K = 1.05$ and $\sigma = 0.05$ the exact value equals 0.00236768. Thus, the absolute error generated by the simplified Euler scheme remains under control.

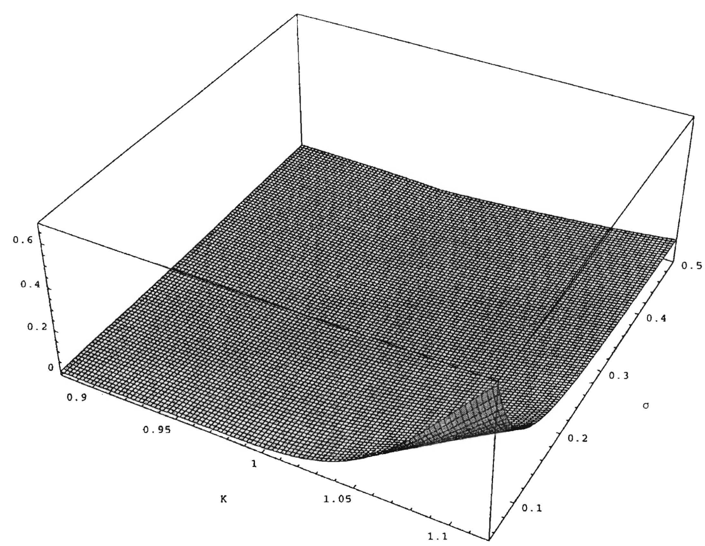

Figure 12.4.10: Relative error for the Euler scheme with $\Delta = T$

In Figures 12.4.12 and 12.4.13 we report the difference between the relative error of the simplified Euler scheme and that of the Euler Scheme with a medium maturity $T = 1$ and a long maturity $T = 4$, respectively. Again we can see that the accuracy of schemes based on simplified random variables is similar to that of schemes based on Gaussian random variables. Furthermore, we notice that for certain sets of parameters the simplified Euler scheme is even more accurate than the Euler scheme.

Finally, we have conducted similar numerical experiments comparing the accuracy of the order 2.0 weak Taylor scheme to that of the simplified order 2.0 weak scheme.

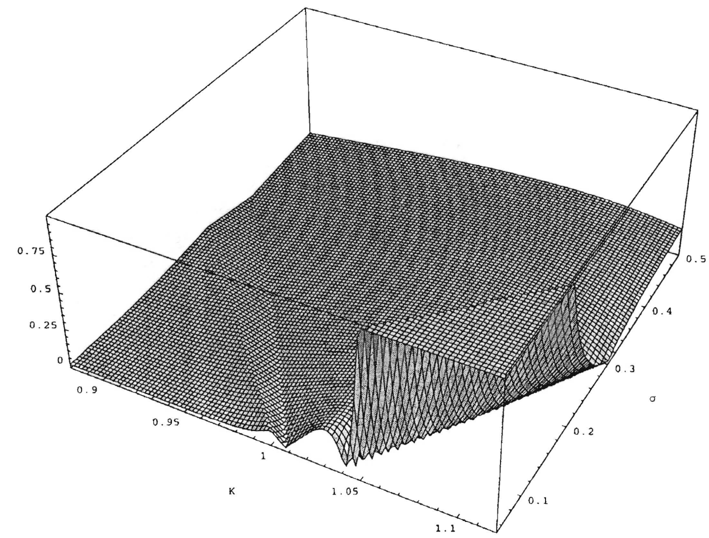

Figure 12.4.11: Relative error for the simplified Euler scheme with $\Delta = T$

We report that also in this case the loss in accuracy due to the use of the three -point distributed random variable (12.2.7) is quite small.

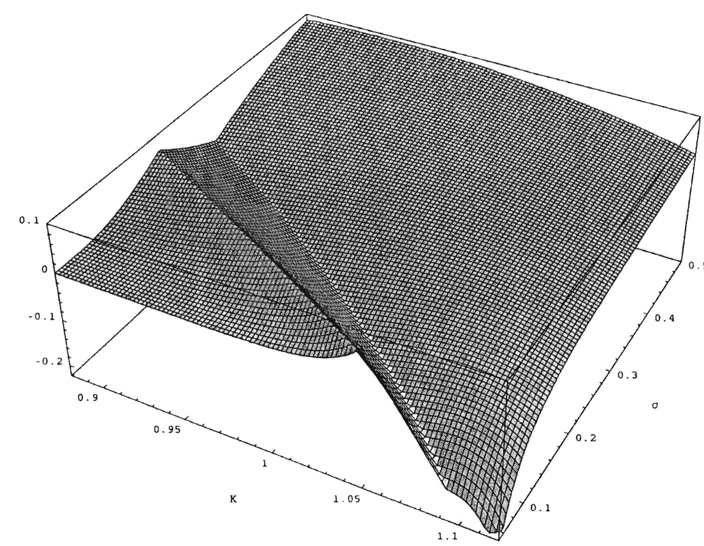

Figure 12.4.12: Relative error of the simplified Euler scheme minus the relative error of the Euler scheme for $T = 1$

The experimental results of this section suggest that the accuracy of weak schemes based on multi-point distributed random variables is similar to that of weak Taylor schemes based on Gaussian random variables. The loss in accuracy, which is due to the use of multi-point random variables, is for typical parameters below 10% when measured in terms of relative errors. We have also reported that in some particular cases the accuracy of simplified weak schemes is superior to that of weak Taylor schemes. Therefore, thanks to high speedups in execution, the proposed RBGs

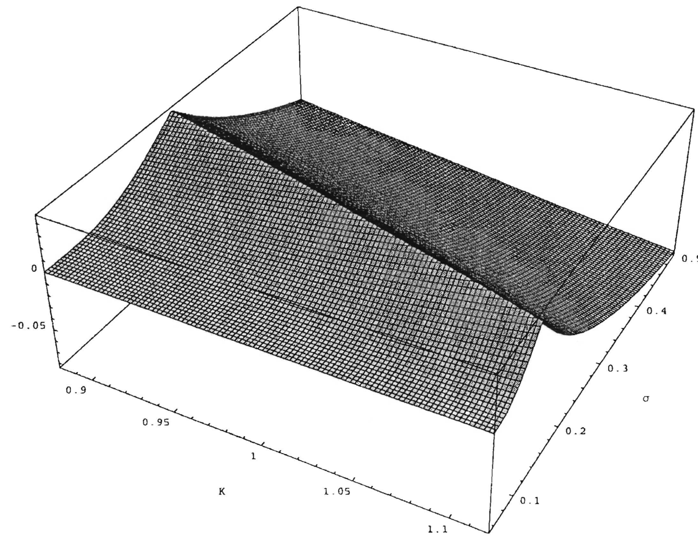

Figure 12.4.13: Relative error of the simplified Euler scheme minus the relative error of the Euler scheme for $T = 4$

when combined with simplified schemes can significantly enhance the efficiency of typical Monte Carlo simulations in finance.

12.5 Hardware Accelerators

In this section we propose a fast generator of multi-point distributed random numbers on a field programmable gate array (FPGA) and describe its system performance in a PC architecture. The proposed approach has been tested over a wide variety of parameters, including different multi-point random variables and corresponding weak Taylor schemes. It proved capable of achieving speedups of up to ten times with respect to an optimized software-only implementation.

12.5.1 System Architecture

Our ultimate goal is to substantially speed up the above described Monte Carlo simulations by moving the random number generation from software to a dedicated hardware platform. More precisely, we aim to move the whole generation of multi-point distributed random numbers from the host processor to a dedicated hardware unit. This approach is advantageous in applications in which a relevant percentage of time is taken by the generation of random numbers. However, there exist critical performance challenges at the system level. As the typical generation time for a software implementation can be as short as a few nanoseconds per number, the dedicated hardware solution must avoid any system-level bottlenecks to prove competitive.

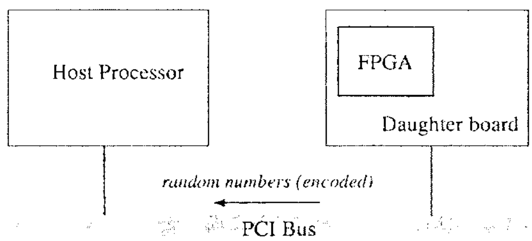

Figure 12.5.14: The system architecture

Our basic idea for a PC environment is that of implementing the hardware "accelerator" as a daughter board on the PCI (peripheral component interconnect) bus, Revision 2.2. For details on the PCI Bus Revision 2.2 we refer to *PCI Local Bus Specification Revision 2.2*, PCI-SIG, 2000. The daughter board hosts the random number generator, not to be confused with the RBG which is just a part of it. The random number generator is implemented on an FPGA and returns the generated numbers to the simulation software through the PCI bus. Figure 12.5.14 shows our proposed system architecture for a PC platform. We can divide the system operations in four phases. In phase 1, the FPGA generates a set of random numbers in an average time T_{FPGA} per number. In phase 2, the FPGA transfers such a set in a compact, combinatorially encoded format to the host memory via burst bus cycles operated under DMA (direct memory access) for maximum communication efficiency in an average time T_{comm} per number. The combinatorial encoding works as follows. Any given multi-point distributed random variable has a small finite set of n possible values, with each value typically represented as a 32-bit floating point datum. We can encode each value by combinatorial encoding with $\lceil \log_2 n \rceil$ bits, where $\lceil a \rceil$ denotes the smallest integer greater than or equal to a. Accordingly, the amount of random numbers that we are able to pack and transfer in a single PCI data phase (32 bits of data over 30 ns) is much larger than that possible with the native floating point representation (only one number per data phase). In phase 3, the host processor is ready to serve the requests for random numbers from the simulation software. At each request, the host processor decodes one encoded number and returns it to the caller in an average time T_{dec} per number. In phase 4, the host processor uses the random numbers in an average time T_{use} per number. In this way, the simulation software sees the system through the same function interface of a conventional software-only implementation and requires no further modifications. More importantly, we coordinate these phases into a pipeline so that the simulation software uses the current set of generated random numbers (phase 4) while the FPGA concurrently produces a new set (phase 1), thus obtaining a significant speedup. Figure 12.5.15 shows how the various phases occur with respect to time. It can be seen that if T_{FPGA} is less than T_{use}, then phase 1 is completely hidden by phase 4 and thus adds no time to the total execution time. With a more aggressive implementation, also phases 2 and 3 could have been considered for pipelining with other phases. In particular, phase 2 could be overlapped with phase 1 by means of a double-buffer implementation on the FPGA. At its turn, phase 3 could be overlapped with phases 1 and 2 by a double-buffer implementation in the host memory. Note that phase 3 cannot overlap with phase 4 as they both require the same resource, the host processor. It can be shown that such changes could result in hiding T_{comm} completely in the overall execution time. On the other hand, T_{dec} will instead increase due to the increased complexity of a multiple-buffer implementation, thus compromising the speedup. For this reason, we decided to limit pipelining to the two main phases, 1 and 4.

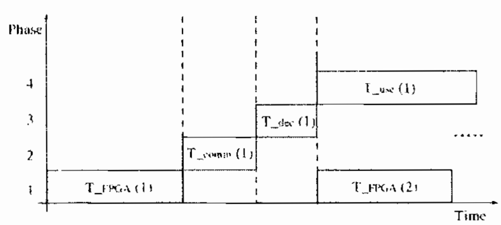

Figure 12.5.15: The various phases with respect to time

The complete time models for the simulation are given in the following. First, we can define T_{gen} as the average time spent for generating a multi-point distributed random number and T_{use} as the average time spent by the rest of the simulation software in using it. If generation and use are sequential, we can write:

$$T_{\text{exe}} = T_{\text{use}} + T_{\text{gen}}, \tag{12.5.36}$$

where T_{exe} is the average total execution time per number.

In the case of a conventional software implementation the above model holds and

$T_{\text{gen}} = T_{\text{gen}_{\text{SW}}}$ is the time taken by the execution of a function that generates and returns one multi-point distributed random number to the caller.

With our system, instead, the simulation software uses the current set of random numbers while the FPGA concurrently generates a new set. In this way, if the generation time on the FPGA, T_{FPGA}, is shorter than the use time, T_{use}, the former does not add up to the total execution time. Such a constraint was largely satisfied in all our experiments. Hence, (12.5.36) still holds with $T_{\text{gen}} = T_{\text{gen}_{\text{HW}}}$ given by:

$$T_{\text{gen}_{\text{HW}}} = T_{\text{comm}} + T_{\text{dec}}, \quad \text{if} \quad T_{\text{FPGA}} < T_{\text{use}}. \tag{12.5.37}$$

12.5.2 FPGA Implementation

A fast and flexible implementation of the random number generator is the main requirement in this application. FPGAs enjoy several features such as quasi-ASIC (application specific integrated circuit) speed and programmer-level flexibility, which makes them the most suitable option for the hardware platform. Accordingly, we have chosen to implement our generator on a high-performance FPGA, the Altera Stratix EP1S1OB672C6. Simulation tools for this device are available in the Altera Quartus II development environment. We have used the Web Edition Software Version 4.2 of such tools. Moreover, all the circuits for the FPGA have been developed in VHDL (Very High Speed Integrated Circuit Hardware Description Language).

Figure 12.5.16 shows a simplified representation of the random number generator. In Figure 12.5.16.(a) the random number generator is shown together with the output FIFO (first in first out) queue (some signals have been omitted for simplicity). The generation of the encoded random numbers is synchronous with the main clock signal (CK), with one number generated per clock cycle over signals RN[0:2]. Each encoded random number generated by the random number generator is input in the FIFO queue which, in turn, allows for asynchronous reading from an external master with 32-bit data parallelism. The writing on the queue is clocked by the BUFF_WR signal, which is synchronous with CK. However, the queue can suspend the random number generation when full by raising FIFO_FULL. The generator needs an initial seed of arbitrary length which can be uploaded asynchronously, in one or more steps, through the SEED[0:31] and WR signals. Figure 12.5.16.(b)

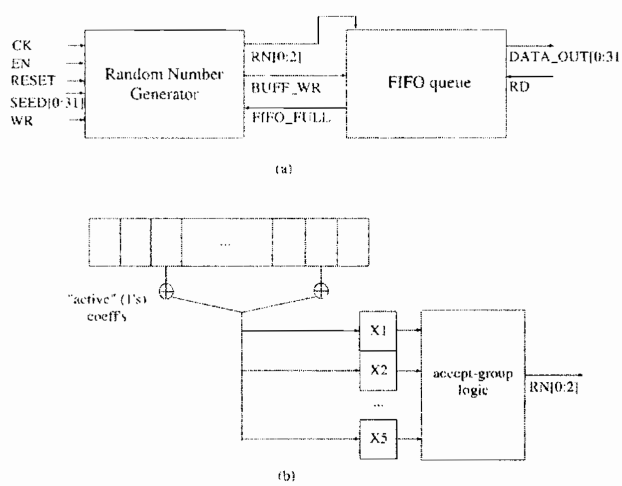

Figure 12.5.16: A simplified representation of the random number generator

shows details of the generator. The generation of the random bits is performed by a shift register generator of programmable length equal to that of the generating polynomial. The "active" (i.e. non-null) coefficients can also be programmed by the user. The orders considered for the weak Taylor scheme range from $\beta = 1.0$ to $\beta = 3.0$, although higher orders can also be straightforwardly implemented. When the selected order is $\beta = 1.0$, the random number generator generates numbers sampled from a two-point distributed random variable with the probabilities described in (12.2.3). Each single bit generated by the shift register generator represents a valid encoded number. When the selected order is $\beta = 2.0$, the random number generator generates numbers sampled from a three-point distributed random variable with the probability distribution (12.2.7). In this case, a sequence of three generated random bits, X1:X3, is used to generate eight equiprobable combinations. As described in Section 12.3, the accept/group logic discards two of them, uses four to generate a 0 value for the random variable and uses one each for values $+\sqrt{3\Delta}$ and $-\sqrt{3\Delta}$. When the selected order is $\beta = 3.0$, the random variable is five-point distributed with the probability distribution (12.2.11). In this case, a sequence of five random bits, X1:X5, is used to generate 32 equiprobable combinations. The accept/group logic discards two of them, uses ten to generate a 0 value, nine each for values $+\sqrt{\Delta}$ and $-\sqrt{\Delta}$, and one each for values $+\sqrt{6\Delta}$ and $-\sqrt{6\Delta}$. All combinatorial functions in the accept/group logic are optimized.

12.5.3 Experimental Results

Table 12.1 shows the main performance results of the proposed implementation for a polynomial order of 31 and the different weak Taylor scheme orders. F_{ck} (in MHz) is the maximal clock frequency obtained for the random number generator. T_{FPGA}, the time (in ns) for generating one multi-point distributed random number, is computed as $\alpha\, 10^3/F_{\text{ck}}$. The α term accounts for the fact that some of the clock cycles generate a random number that should be rejected; such a factor is 8/6 for the three-point distributed random variable and 32/30 for the five-point distributed one. T_{use}, the time spent by the application in using a random number, is measured on the option pricing problem with a lognormal dynamics discussed in Section 12.4.2. From Table 12.1, it is possible to see that the constraint of (12.5.37) is always easily satisfied. Although T_{use} obviously depends on the application, its range will be similar for comparable Monte Carlo simulations. T_{comm} is the average time for transferring one random number from the FIFO queue to the

Table 12.1: Performance results of the proposed implementation, polynomial order 31

Scheme Order	F_{ck}	T_{FPGA}	T_{use}	T_{comm}	T_{dec}
$\beta = 1$	162	6.17	18.06	2	3.85
$\beta = 2$	158	8.42	30.44	4	4.4
$\beta = 3$	116	9.12	68.60	6.4	6.27

host memory over the PCI bus. This time increases proportionally to the size in bits of the encoded random numbers. Moreover, in some cases the data require extra-alignment bits to match the 32-bit PCI data size. For instance, this applies to the case of the 3-bit encoded numbers sampled from a five-point distributed random variable. In Table 12.1, T_{comm} is computed based on a transfer rate of 66 MB/s. However, there exist several implementations over the PCI bus which can almost saturate its peak rate of 132 MB/s; hence, even smaller values for T_{comm} are achievable. Moreover, the upcoming PCI Express$^{\text{TM}}$ bus carries the potential to further decrease T_{comm} by at least a factor of 4. Based on these parameters and thanks to our design choice of combinatorial encoding for the generated random numbers, we have proved herewith that data communication is not a performance bottleneck in our system. Moreover, we have implemented highly-optimized C macros to perform the decoding operation on the host side, thus also limiting T_{dec}, the average time that the host processor takes to decode one encoded random number and return it to the requesting application.

Table 12.2 shows the main performance results for a much higher polynomial order of 521. T_{use}, T_{comm}, and T_{dec} are not influenced by the polynomial order. It can also be seen that the FPGA performance does not suffer from the increased polynomial length and in some cases even slightly exceeds that of the polynomial order 31. As the implementation still uses a very small fraction of the FPGA resources, we cannot see any practical upper bound on the choice of the polynomial length.

To provide a comparative analysis between software and hardware performance, we have implemented both software and hardware versions of the random number generator for a comprehensive variety of parameters. In order for the performance comparison to be unbiased, we have implemented all software functions as highly speed-optimized C macros. The reference PC is a Mobile Pentium 4 2.0 GHz and

Table 12.2: Performance results of the proposed implementation, polynomial order 521

Scheme Order	F_{ck}	T_{FPGA}	T_{use}	T_{comm}	T_{dec}
$\beta = 1$	167	5.99	18.06	2	3.85
$\beta = 2$	157	8.47	30.44	4	4.4
$\beta = 3$	135	7.90	68.60	6.4	6.27

the C compiler used is the Microsoft Visual Studio 6.0 with -O2 optimizations. In the following, the three dimensions of the polynomial order, number of non-null coefficients, and number of points of the random variable are discussed.

Polynomial order

A polynomial order n, for a primitive polynomial modulo 2, guarantees a period of $2^n - 1$ for the generated random sequence. It is known that the accuracy of a simulation, based on a pseudo-random sequence, is compromised when the sequence length is substantial compared with the period of the random number generator. In the light of this, high order polynomials should be preferred. However, in a software implementation one faces an increase in generation time when using high order polynomials, since they cannot be mapped onto a single primitive-type operand. Instead, the hardware implementation does not suffer from any predefined operand size. Figure 12.5.17 shows the generation time, T_{gen}, for the software and hardware implementations as a function of the polynomial order. For the software implementation, T_{gen} remains approximately stable up to 63 ns and then starts to grow with the polynomial order. Yet, the time for the hardware implementation always remains constant. In our tests, even larger polynomial sizes did not introduce any further delay in the FPGA operations.

Number of non-null coefficients

The "randomness" of the random bits, which is crucial for an effective Monte Carlo simulation, is strictly related not only to the order of the generating polynomial but also to the choice of its (non-null) coefficients, see Niederreiter (1992). However, in a software implementation a programmer is tempted to use the polynomial

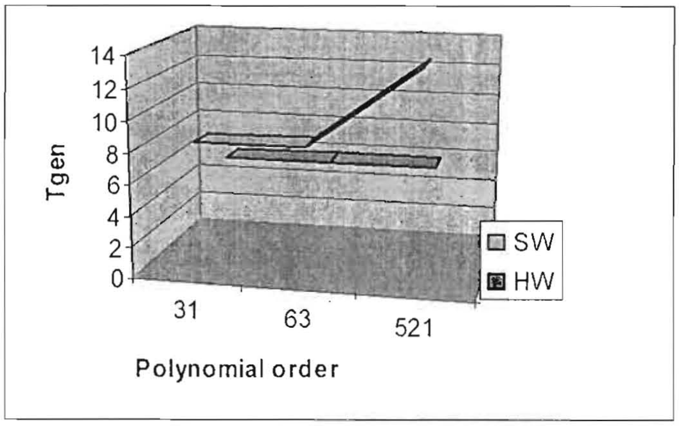

Figure 12.5.17: The generation time as a function of the polynomial order

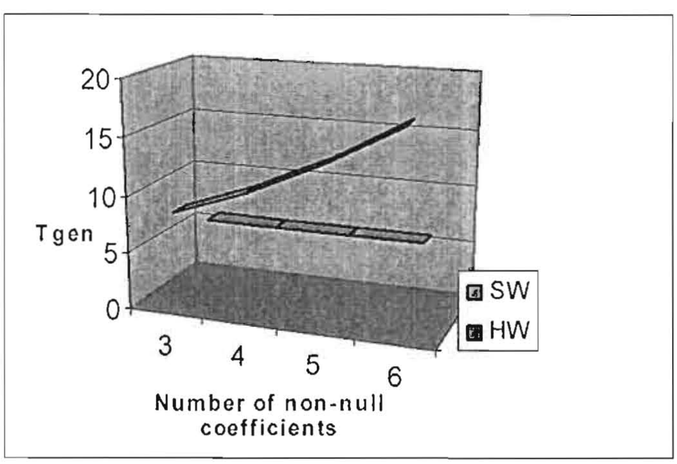

Figure 12.5.18: The generation time as a function of the number of non-null coefficients for a polynomial order of 31

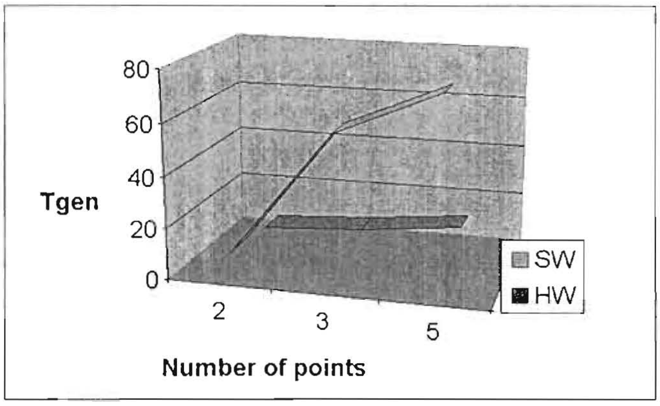

Figure 12.5.19: The generation time as a function of the number of points of the random variable for a polynomial order of 31

with the smallest number of coefficients, as each introduces an additional computational load. Figure 12.5.18 shows that the software implementation suffers from a proportional delay. Again, the time instead remains constant for our hardware implementation as T_{FPGA} remains less than T_{use} in all cases of interest.

Multi-point random variables

When high accuracy is required, higher orders of the weak Taylor schemes will eventually increase the computational efficiency, even though both the scheme and the multi-point distributed random variables are more complex. In any case, speeding up the computation of the random variables has a dramatic impact on the simulation time. Figure 12.5.19 shows the generation time, T_{gen}, for the software and hardware implementations as a function of the number of points in the multi-point distributed random variable, which refers to $\Delta\widehat{W}_{2,n}$, $\Delta\widehat{W}_{3,n}$ and $\Delta\widehat{W}_{5,n}$, defined in (12.2.3),(12.2.7) and (12.2.11), for a polynomial order of 31. Once again, the software time grows steadily, up to 80 ns per value for a five-point distributed random variable. The hardware time, instead, increases negligibly. Actually, the increase in T_{gen} is due only to the larger size of the encoded random numbers. The size of the encoded random numbers grows as $\lceil \log_2 n \rceil$, where n is the number of points representing the possible values of the multi-point distributed variable, and this has an impact on the transfer time, T_{comm}, and the decoding time, T_{dec}, see Tables 12.1 and 12.2. Figure 12.5.20 shows that the trend is similar for a polynomial order of 521.

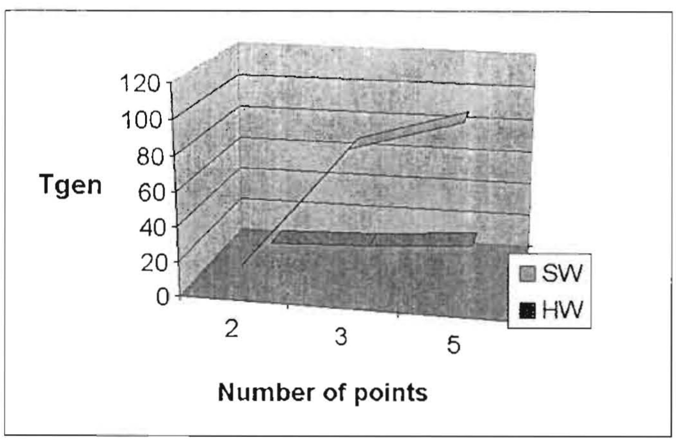

Figure 12.5.20: The generation time as a function of the number of points of the random variable for a polynomial order of 521

Table 12.3: The generation speedup between hardware and software

	$S_{\text{gen}}(31)$	$S_{\text{FPGA}}(31)$	$S_{\text{gen}}(521)$	$S_{\text{FPGA}}(521)$
$\Delta\widehat{W}_{2,n}$	1.44	1.36	2.33	2.27
$\Delta\widehat{W}_{3,n}$	6.84	6.83	9.94	9.86
$\Delta\widehat{W}_{5,n}$	5.88	8.18	8.01	12.84

Speedup

Table 12.3 reports the speedups achieved with the proposed hardware solutions with respect to the optimised software implementation, when considering the multi-point distributed random variables $\Delta\widehat{W}_{2,n}$, $\Delta\widehat{W}_{3,n}$ and $\Delta\widehat{W}_{5,n}$ and polynomial orders of 31 and 521. $S_{\text{gen}} = T_{\text{gen}_{\text{SW}}}/T_{\text{gen}_{\text{HW}}}$ is the speedup between the generation in hardware and that in software. As explained in (12.5.37), $T_{\text{gen}_{\text{HW}}}$ does not account for the generation time on the FPGA device, but it consists, rather, of communication and decoding times. The units responsible for such times are mainly the PCI bus and the host processor. While S_{gen} is the main performance figure in our system, it is important to report also $S_{\text{FPGA}} = T_{\text{gen}_{\text{SW}}}/T_{\text{FPGA}}$, which is the speedup between the generation on the FPGA alone and that in software. This speedup is important to express the relative performance of the FPGA device and the host processor in the generation of multi-point distributed random variables in view of a possible transfer of the whole simulation to FPGAs. Table 12.3 shows that such a speedup is as high as 12.84 and could possibly increase by using FPGA development tools providing further optimizations.

Table 12.4 reports the application speedup of the proposed hardware solutions with respect to the optimized software implementation $S_{\text{exe}} = T_{\text{exe}_{\text{SW}}}/T_{\text{exe}_{\text{HW}}}$ when the option pricing problem with lognormal dynamics discussed in Section 12.4.2 is considered. Table 12.4 shows that the overall application strongly benefits from the hardware acceleration, up to almost three times in some cases. This is due to the large percentage of the total execution time typically spent on the random number generation by the software. Moreover, the speedup increases with the order of the polynomial and also with the number of its non-null coefficients, which is not shown in the table. Therefore, these speedups become more significant in the case of high accuracy simulations.

Table 12.4: The application speedup between hardware and software

Scheme Order	$S_{\mathsf{exe}}(31)$	$S_{\mathsf{exe}}(521)$
$\beta = 1$	1.11	1.32
$\beta = 2$	2.26	2.93
$\beta = 3$	1.76	2.09

From Table 12.4, it appears that the application speedup for a weak Taylor scheme of order $\beta = 2.0$ is greater than that for order $\beta = 3.0$. However, such a result is not general since the measured times and speedups can depend significantly on the compiler used. To verify that, we measured T_{usc} also with another compiler, the Mingw port of GCC. Here we obtained 14, 47 and 57 ns per random number for a weak Taylor scheme of order 1, 2, and 3, respectively. Such times, when compared to those obtained by the Microsoft compiler and reported in Tables 12.1 and 12.2, seem to be in better proportion with the complexity of the operations in (12.2.4), (12.2.8) and (12.2.12). With such times, the application speedup for the weak Taylor scheme of order $\beta = 2.0$ is equivalent to that of order $\beta = 3.0$.

Chapter 13

Conclusions and Further Directions of Research

13.1 Conclusions

Discrete time approximations for SDEs with jumps have been presented in this thesis. New moment estimates of multiple stochastic integrals involving Poisson jump measures and their compensated counterparts have been derived. Using these estimates and the Wagner-Platen expansion, we have proved convergence theorems for strong and weak higher order approximations. Additionally, new, efficient strong approximations, including derivative-free, drift-implicit and predictor-corrector schemes, have been proposed. The new strong predictor-corrector schemes appear to be of particular interest because of their efficiency and numerical stability properties. Various numerical experiments have confirmed the theoretically derived numerical properties of the new strong and weak higher order schemes. Furthermore, efficient implementations of simplified weak schemes based on random bit generators and dedicated hardware accelerators have been developed and tested. The thesis demonstrates that progress has been made in the development of the theory of numerical solution of SDEs with jumps via simulation. This progress raises new interesting and challenging research problems for the development of an advanced theory.

13.2 Further Directions of Research

There are several possible directions of future research which directly derive from the results obtained. First, extensive numerical experiments with the new methods should be conducted to obtain clear measurements of their numerical efficiency and stability when applied to specific SDEs in finance. An important application of discrete time approximations arises in the pricing and hedging of complex interest rate derivatives under the LIBOR market model with jumps, see Glasserman & Kou (2003). As discussed in Chapter 2, the corresponding high-dimensional non-linear dynamics require the use of efficient discrete time approximations. Glasserman & Merener (2003a) analyzed certain schemes for these types of dynamics. A comparative analysis of the performances of the schemes presented in this thesis, when applied to the LIBOR market model with jumps, is therefore a topic of preferred future research. Furthermore, several authors, including Higham & Mao (2005), Kahl & Jäckel (2005), Lord, Koekkoek & van Dijk (2006), Broadie & Kaya (2006), and Andersen (2007), have recently analyzed the problem of the numerical approximation of the SDEs describing the Heston model, see Heston (1993). In Heston (1993) a transform method was proposed, which allows the fast evaluation of European type derivatives. However, for the pricing of exotic derivatives, simulation methods are usually employed. It remains a challenge to derive efficient higher order schemes for these particular dynamics when jumps are included. The derivation of numerically stable, higher order schemes for the Bates models, see Bates (1996), which is a jump augmented Heston model, is a research topic with strong practical background.

Numerical stability is of crucial importance for the effectiveness of discrete time approximations also in the presence of jumps. Therefore, a second important topic of future research is a detailed analysis of the numerical stability properties of the schemes presented in this thesis. Such analysis could follow the lines of Higham & Kloeden (2005, 2006, 2007) for implicit strong schemes. Of particular interest are the numerical stability properties of the new strong predictor-corrector schemes.

The derivation of a convergence theorem for regular weak approximations in the case of pure jump SDEs is another challenging project of future research. As in the case of strong schemes, this research would aim to relax the differentiability conditions on the jump coefficient. Ideas of Glasserman & Merener (2003a), who obtained weak convergence for schemes up to weak order $\beta = 2.0$ for jump-diffusion SDEs under mild conditions on the jump coefficient, could be useful for this problem.

Another topic of future research is the design of regular weak derivative-free, predictor-corrector and simplified schemes. In particular, a moment condition for regular weak schemes, similar to that presented for pure diffusion SDEs in Kloeden & Platen (1999), would be needed to assess the convergence of such general regular weak approximations. A remark in Liu & Li (2000) suggests a moment condition to hold for regular weak schemes, without providing any proof. A topic of future research is therefore the derivation of a full proof under which the above mentioned moment condition is sufficient to obtain the desired order of weak convergence for given classes of SDEs with jumps. Note that because of the peculiar properties of higher order moments of increments of the Poisson processes, which differ from those of diffusions, the jump-diffusion case requires particular care and it is not obvious that the conjecture in Liu & Li (2000) holds.

The design of efficient weak schemes involving automatic step size control is another interesting topic for future research. The work of Mordecki, Szepessy, Tempone & Zouraris (2006) addresses this problem in the case of the Euler scheme. Of particular interest would be the design of weak predictor-corrector schemes with step size control.

Also hardware accelerators for weak schemes should be further explored. One possible extension of our current research is the implementation of entire Monte Carlo simulation algorithms on field programmable gate arrays. Furthermore, the Monte Carlo simulation of SDEs is generally suitable for parallel implementations, see Petersen (1987). A parallel implementation of the weak schemes presented in this thesis can lead to high speedups.

Another field of investigation is that of the numerical approximation of stochastic partial differential equations with jumps, see Hausenblas (2003, 2006). Typically, after an initial space discretization, which can be performed, for instance, by Galerkin or finite differences methods, see Kloeden (2002), one obtains a system of high-dimensional SDEs. The schemes developed in this thesis could be useful to obtain efficient approximations of stochastic partial differential equations with jumps. Also in the numerical solution of forward-backward SDEs with jumps, methods similar to those developed in this thesis could be employed. A method for forward-backward SDEs with jumps based on an Euler scheme has been proposed in Bouchard & Romuald (2005).

Another interesting direction of future research is the design of discrete time approximations for SDEs driven by Wiener processes and infinite activity jump processes. A proper truncation of the small jumps may lead to efficient schemes based on those developed in this thesis. This would also cover the simulation of Lévy processes.

The last important direction of future research that we would like to mention regards the smoothness conditions on the payoff function usually required by weak convergence theorems to guarantee higher order of weak convergence. Weak convergence theorems for non-smooth payoffs for the Euler scheme have been presented in Bally & Talay (1996a, 1996b) and Guyon (2006), in the case of pure diffusion SDEs, and in Hausenblas (2002), in the case of pure jump SDEs. An extension of these results to higher order schemes for jump diffusions would be highly desirable for applications such as option pricing.

Of broader scope are the following two directions of research. First, of great importance is the development of efficient simulation methods for high-dimensional systems of SDEs with jumps. These typically arise in important financial applications. Second, the development of variance reduction techniques is crucial for the efficiency of Monte Carlo simulations. We remark that in this thesis we have addressed the problem of designing efficient methods to reduce the systematic error generated by the time discretization. However, as previously discussed, to obtain accurate results one should develop efficient variance reduction techniques, thus reducing the statistical error generated in practical applications of the Monte Carlo simulation methods proposed.

Appendix A

Appendix: Inequalities

We present here some inequalities, see Ash (1972) and Ikeda & Watanabe (1989), that we have used in the thesis.

A.1 Finite Inequalities

Consider a sequence of pairs of real numbers $\{(a_i, b_i), i \in \{1, \ldots, n\}\}$, with $n \in \mathbb{N}$.

Lemma A.1.1 ***(Cauchy-Schwarz Inequality)***

$$\Big(\sum_{i=1}^{n} a_i b_i\Big)^2 \leq \sum_{i=1}^{n} a_i^2 \sum_{i=1}^{n} b_i^2. \tag{1.1.1}$$

Corollary 1

$$\Big(\sum_{i=1}^{n} a_i\Big)^2 \leq n \sum_{i=1}^{n} a_i^2. \tag{1.1.2}$$

Lemma A.1.2 ***(Hölder Inequality)***

Let $1 < p < \infty$, $1 < q < \infty$ *and* $(1/p) + (1/q) = 1$. *Then*

$$\Big|\sum_{i=1}^{n} a_i b_i\Big| \leq \Big(\sum_{i=1}^{n} |a_i|^p\Big)^{\frac{1}{p}} \Big(\sum_{i=1}^{n} |b_i|^q\Big)^{\frac{1}{q}}. \tag{1.1.3}$$

Corollary 2 *Let* $1 < p < \infty$, *then*

$$\Big(\sum_{i=1}^{n} a_i\Big)^p \leq n^{p-1} \sum_{i=1}^{n} a_i^p. \tag{1.1.4}$$

A.2 Integral Inequalities

Let $(\Omega, \mathcal{A}, \mu)$ be a measurable space and consider two functions $f, g : \Omega \to \mathbb{R}$.

Lemma A.2.1 ***(Cauchy-Schwarz Inequality)***

If f and $g \in L^2(\mu)$, then $fg \in L^1(\mu)$ and

$$\Big|\int_\Omega fg d\mu\Big| \le \Big(\int_\Omega f^2 d\mu\Big)^{\frac{1}{2}} \Big(\int_\Omega g^2 d\mu\Big)^{\frac{1}{2}}. \tag{1.2.1}$$

Corollary 3

$$\Big(\Big|\int_\Omega f d\mu\Big|\Big)^2 \le \mu(\Omega) \int_\Omega f^2 d\mu. \tag{1.2.2}$$

Lemma A.2.2 ***(Hölder Inequality)***

Let $1 < p < \infty$, $1 < q < \infty$ and $(1/p) + (1/q) = 1$. If $f \in L^p(\mu)$ and $g \in L^q(\mu)$, then $fg \in L^1(\mu)$ and

$$\Big|\int_\Omega fg d\mu\Big| \le \Big(\int_\Omega |f|^p d\mu\Big)^{\frac{1}{p}} \Big(\int_\Omega |g|^q d\mu\Big)^{\frac{1}{q}}. \tag{1.2.3}$$

Corollary 4 *Let $1 < p < \infty$ and $f \in L^p(\mu)$. Then*

$$\Big(\Big|\int_\Omega f d\mu\Big|\Big)^p \le \mu(\Omega)^{p-1} \int_\Omega |f|^p d\mu. \tag{1.2.4}$$

Lemma A.2.3 ***(Gronwall Inequality 1)***

Let $\alpha, \beta : [t_0, T] \to \mathbb{R}$ be integrable with

$$0 \le \alpha(t) \le \beta(t) + C \int_{t_0}^t \alpha(s) ds \tag{1.2.5}$$

for $t \in [t_0, T]$ and $C > 0$. Then

$$\alpha(t) \le \beta(t) + C \int_{t_0}^t e^{C(t-s)} \beta(s) ds. \tag{1.2.6}$$

Lemma A.2.4 ***(Gronwall Inequality 2)***

Let $f : [t_0, T] \to \mathbb{R}$ be a nonnegative integrable function and $\alpha, \beta : [t_0, T] \to \mathbb{R}$ be continuous functions on $[t_0, T]$. If

$$\alpha(t) \le \beta(t) + \int_{t_0}^t f(s)\, \alpha(s) ds \tag{1.2.7}$$

for $t \in [t_0, T]$, then

$$\alpha(t) \le \beta(t) + \int_{t_0}^t f(s)\beta(s) ds \exp\Big(\int_s^t f(u) du\Big). \tag{1.2.8}$$

Moreover, if β is non-decreasing, then

$$\alpha(t) \leq \beta(t) \exp\Big(\int_{t_0}^{t} f(u)du \Big). \tag{1.2.9}$$

Lemma A.2.5 ***(Jensen's Inequality)***

Let X be a random variable with finite first moment and $g : \mathbb{R} \to \mathbb{R}$ be a convex function. Then we have

$$g\Big(E(X)\Big) \leq E\Big(g(X)\Big). \tag{1.2.10}$$

A.3 Martingale Inequalities

Let us consider a right-continuous martingale $X = \{X_t, t \in [0,T]\}$ such that $E(|X_t|^p) < \infty$ for $t \in [0,T]$. Then we have the following inequalities.

Lemma A.3.1 ***(Maximal Martingale Inequality)***

$$P\Big(\big\{\omega \in \Omega : \sup_{0 \leq s \leq t} |X_s(\omega)| \geq a\big\}\Big) \leq \frac{1}{a^p} E\big(|X_t|^p\big) \tag{1.3.1}$$

for $a > 0$, $p \geq 1$ and $t \in [0,T]$.

Lemma A.3.2 ***(Doob's Inequality)***

$$E\Big(\sup_{0 \leq s \leq t} |X_s|^p\Big) \leq \big(\frac{p}{p-1}\big)^p E\big(|X_t|^p\big) \tag{1.3.2}$$

for $p > 1$ and $t \in [0,T]$.

Bibliography

Alcock, J. T. & K. Burrage (2006). A note on the balanced method. *BIT Numerical Mathematics* **46**, 689–710.

Andersen, L. B. (2007). Efficient simulation of the Heston stochastic volatility model. Technical report, Banc of America Securities.

Andersen, L. B. & J. Andreasen (2000). Jump-diffusion processes: Volatility smile fitting and numerical methods for option pricing. *Rev. Derivatives Res.* **4**, 231–262.

Ash, R. B. (1972). *Real Analysis and Probability.* Academic Press, New York.

Athreya, K. B., W. H. Kliemann, & G. Koch (1988). On sequential construction of solutions of stochastic differential equations with jump terms. *Syst. Control Lett.* **10**, 141 – 146.

Babbs, S. & N. Webber (1995). When we say jump ... *Risk* **8**(1), 49–53.

Bally, V. & D. Talay (1996a). The law of the Euler scheme for stochastic differential equations I. Convergence rate of the distribution function. *Probab. Theory Related Fields* **104**(1), 43–60.

Bally, V. & D. Talay (1996b). The law of the Euler scheme for stochastic differential equations II. Convergence rate of the density. *Monte Carlo Methods Appl.* **2**(2), 93–128.

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

Bell, E. T. (1934). Exponential polynomials. *Ann. Math.* **35**, 258 – 277.

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

Bouchard, B. & E. Romuald (2005). Discrete time approximation of decoupled forward-backward SDE with jumps. Technical report, LPMA and CEREMADE.

Boyle, P. P. & S. H. Lau (1994). Bumping up against the barrier with the binomial method. *J. Derivatives*, 6–14.

Broadie, M. & O. Kaya (2006). Exact simulation of stochastic volatility and other affine jump diffusion processes. *Oper. Res.* **54**, 217–231.

Bruti-Liberati, N., F. Martini, M. Piccardi, & E. Platen (2007). A hardware generator for multi-point distributed random numbers for Monte Carlo simulation. *Math. Comput. Simulation*. Doi:10.1016/j.matcom.2007.01.031. (To appear).

Bruti-Liberati, N., C. Nikitopoulos-Sklibosios, & E. Platen (2006). First order strong approximations of jump diffusions. *Monte Carlo Methods Appl.* **12**(3), 191–209.

Bruti-Liberati, N. & E. Platen (2004). On the efficiency of simplified weak Taylor schemes for Monte Carlo simulation in finance. In *Computational Science - ICCS 2004*, Volume 3039 of *Lecture Notes in Comput. Sci.*, pp. 771–778. Springer.

Bruti-Liberati, N. & E. Platen (2007a). Approximation of jump diffusions in finance and economics. *Computational Econ.* **29**(3-4), 283–312.

Bruti-Liberati, N. & E. Platen (2007b). On weak predictor-corrector schemes for jump-diffusion processes in finance. In *Numerical Methods in Finance*, Financial Mathematics Series. Chapman & Hall/CRC. (To appear).

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

Bruti-Liberati, N. & E. Platen (2007d). Strong predictor-corrector methods with jumps. Technical report, University of Technology, Sydney.

Bruti-Liberati, N., E. Platen, F. Martini, & M. Piccardi (2005). A multi-point distributed random variable accelerator for Monte Carlo simulation in finance. In *Proceedings of the Fifth International Conference on Intelligent Systems Design and Applications*, pp. 532–537. IEEE Computer Society Press.

Burrage, K., P. M. Burrage, & T. Tian (2004). Numerical methods for strong solutions of SDES. In *Proceeding of the Royal Society London*, Volume 460, pp. 373–402.

Chan, N. H. & H. Y. Wong (2006). *Simulation Techniques in Financial Risk Management.* Wiley.

Cont, R. & P. Tankov (2004). *Financial Modelling with Jump Processes.* Financial Mathematics Series. Chapman & Hall/CRC.

Cox, J. C., J. E. Ingersoll, & S. A. Ross (1985). A theory of the term structure of interest rates. *Econometrica* **53**, 385–407.

Cyganowski, S., L. Grüne, & P. E. Kloeden (2002). MAPLE for jump-diffusion stochastic differential equations in finance. In *Programming Languages and Systems in Computational Economics and Finance*, pp. 441–460. Kluwer Academic Publishers.

Cyganowski, S., P. E. Kloeden, & J. Ombach (2001). *From Elementary Probability to Stochastic Differential Equations with MAPLE.* Universitext. Springer.

Devroye, L. (1986). *Non-Uniform Random Variate Generation.* Springer-Verlag.

Duffie, D. & R. Kan (1994). Multi-factor term structure models. *Philos. Trans. Roy. Soc. London Ser.* A **347**, 577–580.

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.

Engel, D. (1982). The multiple stochastic integral. *Mem. Amer. Math. Soc.* **38**, 265.

Evans, L. C. (1999). *Partial Differential Equations*, Volume 19 of *Graduate Studies in Mathematics.* American Mathematical Society. Second printing.

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

Geman, H. & A. Roncoroni (2006). Understanding the fine structure of electricity prices. *J. Business* **79**(3), 1225–1261.

Glasserman, P. (2004). *Monte Carlo Methods in Financial Engineering*, Volume 53 of *Appl. Math.* Springer.

Glasserman, P. & S. G. Kou (2003). The term structure of simple forward rates with jump risk. *Math. Finance* **13**(3), 383–410.

Glasserman, P. & N. Merener (2003a). Convergence of a discretization scheme for jump-diffusion processes with state-dependent intensities. In *Proceedings of the Royal Society*, Volume 460, pp. 111–127.

Glasserman, P. & N. Merener (2003b). Numerical solution of jump-diffusion LIBOR market models. *Finance Stoch.* **7**(1), 1–27.

Golomb, S. W. (1964). *Digital Communications with Space Applications.* Prentice-Hall, Englewood Cliffs.

Guyon, J. (2006). Euler scheme and tempered distributions. *Stochastic Process. Appl.* **116**(6), 877–904.

Hausenblas, E. (2002). Error analysis for approximation of stochastic differential equations driven by Poisson random measures. *SIAM J. Numer. Anal.* **40**(1), 87–113.

Hausenblas, E. (2003). Weak approximation of stochastic partial differential equations. In *Stochastic analysis and related topics VIII. Silivri workshop, Progress in Probability.* Birkhauser.

Hausenblas, E. (2006). A numerical approximation of parabolic stochastic differential equations driven by a Poisson random measure. *BIT Numerical Mathematics* **46**, 773–811.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Financial Studies* **6**(2), 327–343.

Higham, D. (2001). An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM Rev.* **43**(3), 525–546.

Higham, D. (2004). *An Introduction to Financial Option Valuation: Mathematics, Stochastics and Computation.* Cambridge University Press.

Higham, D. & P. Kloeden (2002). MAPLE and MATLAB for stochastic differential equations in finance. In *Programming Languages and Systems in Computational Economics and Finance*, pp. 233–270. Kluwer Academic Publishers.

Higham, D. & P. Kloeden (2005). Numerical methods for nonlinear stochastic differential equations with jumps. *Numer. Math.* **110**(1), 101–119.

Higham, D. & P. Kloeden (2006). Convergence and stability of implicit methods for jump-diffusion systems. *Internat. J. of Numer. Analysis & Modeling* **3**(2), 125–140.

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

Higham, D. & X. Mao (2005). Convergence of Monte Carlo simulations involving the mean-reverting square root process. *J. Comput. Finance* **8**, 35–61.

Hofmann, N. (1994). *Beiträge zur schwachen Approximation stochastischer Differentialgleichungen.* Ph. D. thesis, Dissertation A, Humboldt Universität Berlin.

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

Hunter, C. J., P. Jäckel, & M. S. Joshi (2001). Getting the drift. *Risk* **14**(7), 81–84.

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

Jäckel, P. (2002). *Monte Carlo Methods in Finance.* Wiley.

Jacod, J. (2004). The Euler scheme for Lévy driven stochastic differential equations: limit theorems. *Ann. Probab.* **32**(3A), 1830–1872.

Jacod, J., T. Kurtz, S. Méléard, & P. Protter (2005). The approximate Euler method for Lévy driven stochastic differential equations. *Ann. Inst. H. Poincaré Probab. Statist.* **41**(3), 523–558.

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

Janicki, A. (1996). *Numerical and Statistical Approximation of Stochastic Differential Equations with Non-Gaussian Measures.* H. Steinhaus Center for Stochastic Methods in Science and Technology, Wroclaw, Poland.

Jarrow, R., D. Lando, & S. Turnbull (1997). A Markov model for the term structure of credit risk spreads. *Rev. Financial Studies* **10**(2), 481–523.

Jimenez, J. C. & F. Carbonell (2006). Local linear approximations for jump diffusion processes. *J. Appl. Probab.* **43**, 185–194.

Johannes, M. (2004). The statistical and economic role of jumps in continuous-time interest rate models. *Journal of Finance* **59**(1), 227 – 260.

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

Joshi, M. S. & A. M. Stacey (2007). New and robust drift approximations for the LIBOR market model. *Quant. Finance.*. (To appear).

Kahl, C. & P. Jäckel (2005). Fast strong approximation Monte-Carlo schemes for stochastic volatility models. Technical report, University of Wuppertal and ABN-AMRO.

Kahl, C. & H. Schurz (2006). Balanced Milstein methods for ordinary SDEs. *Monte Carlo Methods Appl.* **12**(2), 143–170.

Kloeden, P. E. (2002). The systematic derivation of higher order numerical schemes for stochastic differential equations. *Milan Journal of Mathematics* **70**(1), 187–207.

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

Kloeden, P. E., E. Platen, & H. Schurz (2003). *Numerical Solution of SDE's Through Computer Experiments.* Universitext. Springer. Third corrected printing.

Knuth, D. E. (1981). *The Art of Computer Programming. Vol. 2. Seminumerical Algorithms* (2nd ed.). Ser. Computer Science and Inform. Processing. Addison-Wesley, Reading MA.

Kohatsu-Higa, A. & P. Protter (1994). The Euler scheme for SDEs driven by semimartingales. In H. Kunita and H. H. Kuo (Eds.), *Stochastic Analysis on Infinite Dimensional Spaces*, pp. 141–151. Pitman.

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

Krylov, N. V. (1980). *Controlled Diffusion Processes*, Volume 14 of *Appl. Math.* Springer.

Kubilius, K. & E. Platen (2002). Rate of weak convergence of the Euler approximation for diffusion processes with jumps. *Monte Carlo Methods Appl.* **8**(1), 83–96.

Kurtz, T. G. & P. Protter (1991). Wong-Zakai corrections, random evolutions and simulation schemes for SDEs. In E. M. E. Meyer-Wolf and A. Schwartz (Eds.), *Stochastic Analysis*, pp. 331–346. Academic Press.

Lee, D., W. Luk, J. Villasenor, & P. Cheung (2004). A Gaussian noise generator for hardware-based simulations. *IEEE Trans. on Comp.* **53**(12), 1523–1534.

Lee, D., J. Villasenor, W. Luk, & P. Leong (2006). A hardware Gaussian noise generator using the Box-Muller method and its error analysis. *IEEE Trans. on Comp.* **55**(6), 659–671.

Li, C. W. (1995). Almost sure convergence of stochastic differential equations of jump-diffusion type. In *Seminar on Stochastic Analysis, Random Fields and Applications*, Volume 36 of *Progr. Probab.*, pp. 187–197. Birkhäuser Verlag.

Li, C. W. & X. Q. Liu (1997). Algebraic structure of multiple stochastic integrals with respect to Brownian motions and Poisson processes. *Stochastics Stochastics Rep.* **61**, 107–120.

Li, C. W. & X. Q. Liu (2000). Almost sure convergence of the numerical discretisation of stochastic jump diffusions. *Acta Appl. Math.* **62**, 225–244.

Li, C. W., S. C. Wu, & X. Q. Liu (1998). Discretization of jump stochastic differential equations in terms of multiple stochastic integrals. *J. Comput. Appl. Math.* **16**(4), 375–384.

Liu, X. Q. & C. W. Li (1999). Product expansion for stochastic jump diffusions and its application to numerical approximation. *J. Comput. Appl. Math.* **108**(1-2), 1–17.

Liu, X. Q. & C. W. Li (2000). Weak approximation and extrapolations of stochastic differential equations with jumps. *SIAM J. Numer. Anal.* **37**(6), 1747–1767.

Lord, R., R. Koekkoek, & D. van Dijk (2006). A comparison of biased simulation schemes for stochastic volatility models. Technical report, Erasmus University Rotterdam, Rabobank International and Robeco Alternative Investments.

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.

Maghsoodi, Y. & C. J. Harris (1987). In-probability approximation and simulation of nonlinear jump-diffusion stochastic differential equations. *IMA J. Math. Control Inform.* **4**(1), 65–92.

Martin, P. (2002). An analysis of random number generators for a hardware implementation of genetic programming using FPGAs and Handel-C. Technical report, University of Essex, Department of Computer Science. CSM-358.

Martini, F., M. Piccardi, N. Bruti-Liberati, & E. Platen (2005). A hardware generator for multi-point distributed random variables. In *Proceedings of the IEEE International Symposium on Circuits and Systems, 2005 (ISCAS 2005)*, Volume 2, pp. 1702–1705. IEEE Computer Society Press.

McLeish, D. L. (2005). *Monte Carlo Simulation and Finance.* Wiley.

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

Mikulevicius, R. (1983). On some properties of solutions of stochastic differential equations. *Liet. Mat. Rink.* **4**, 18–31.

Mikulevicius, R. & E. Platen (1988). Time discrete Taylor approximations for Ito processes with jump component. *Math. Nachr.* **138**, 93–104.

Milstein, G. N. (1974). Approximate integration of stochastic differential equations. *Theory Probab. Appl.* **19**, 557–562.

Milstein, G. N. (1978). A method of second order accuracy integration of stochastic differential equations. *Theory Probab. Appl.* **23**, 396–401.

Milstein, G. N. (1988). *Numerical Integration of Stochastic Differential Equations.* Urals Univ. Press, Sverdlovsk. (in Russian).

Milstein, G. N. (1995). *Numerical Integration of Stochastic Differential Equations.* Mathematics and Its Applications. Kluwer.

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

Milstein, G. N. & M. V. Tretyakov (2004). *Stochastic Numerics for Mathematical Physics.* Springer.

Mordecki, E., A. Szepessy, R. Tempone, & G. E. Zouraris (2006). Adaptive weak approximation of diffusions with jumps. Technical report, Florida State University, Department of Mathematics. FSU06-23.

Niederreiter, H. (1992). *Random Number Generation and Quasi-Monte-Carlo Methods.* SIAM, Philadelphia, PA.

Petersen, W. P. (1987). Numerical simulation of Ito stochastic differential equations on supercomputers. In *Random Media*, Volume 7 of *IMA Vol. Math. Appl.*, pp. 215–228. Springer.

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. (1984). Beiträge zur zeitdiskreten Approximation von Itôprozessen. Habilitation, Academy of Sciences, Berlin.

Platen, E. (1995). On weak implicit and predictor-corrector methods. *Math. Comput. Simulation* **38**, 69–76.

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

Platen, E. (2001). A minimal financial market model. In *Trends in Mathematics*, pp. 293–301. Birkhäuser.

Platen, E. & D. Heath (2006). *A Benchmark Approach to Quantitative Finance.* Springer Finance. Springer.

Platen, E. & R. Rebolledo (1985). Weak convergence of semimartingales and discretization methods. *Stochastic Process. Appl.* **20**, 41–58.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, & B. P. Flannery (2002). *Numerical Recipes in C++. The Art of Scientific Computing* (2nd ed.). Cambridge University Press.

Protter, P. (1985). Approximations of solutions of stochastic differential equations driven by semimartingales. *Ann. Probab.* **13**, 716–743.

Protter, P. (2004). *Stochastic Integration and Differential Equations* (2nd ed.). Springer.

Protter, P. & D. Talay (1997). The Euler scheme for Lévy driven stochastic differential equations. *Ann. Probab.* **25**(1), 393–423.

Rubenthaler, S. (2003). Numerical simulation of the solution of a stochastic differential equation driven by a Lévy process. *Stochastic Process. Appl.* **103**(2), 311–349.

Rubenthaler, S. & M. Wiktorsson (2003). Improved convergence rate for the simulation of stochastic differential equations driven by subordinated Lévy processes. *Stochastic Process. Appl.* **108**(2), 1–26.

Samuelides, Y. & E. Nahum (2001). A tractable market model with jumps for pricing short-term interest rate derivatives. *Quant. Finance.* **1**, 270–283.

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

Seydel, R. (2006). *Tools for Computational Finance* (3rd ed.). Universitext. Springer.

Situ, R. (2005). *Theory of stochastic differential equations with jumps and applications.* Springer.

Studer, M. (2001). *Stochastic Taylor expansions and saddlepoint approximations for risk management.* Ph. D. thesis, ETH Zurich.

Talay, D. (1982a). *Analyse Numérique des Equations Différentielles Stochastiques.* Ph. D. thesis, Université de Provence, Centre Saint Charles. Thèse 3ème Cycle.

Talay, D. (1982b). Convergence for each trajectory of an approximation scheme of SDE. *Computes Rendus Acad. Sc. Paris, Séries I Math.* **295**(3), 249–252. (in French).

Talay, D. (1984). Efficient numerical schemes for the approximation of expectations of functionals of the solution of an SDE and applications. In *Filtering and Control of Random Processes*, Volume 61 of *Lecture Notes in Control and Inform. Sci.*, pp. 294–313. Springer.

Tausworthe, R. C. (1965). Random numbers generated by linear recurrence modulo two. *Math. Comp.* **19**, 201–209.

Tsoi, K. H., K. H. Leung, & P. H. W. Leong (2003). Compact FPGA-based true and pseudo random number generators. In *Proc. of the 11th Annual IEEE Symposium on Field-Programmable Custom Computing Machines*, pp. 1–13.

Turner, T., S. Schnell, & K. Burrage (2004). Stochastic approaches for modelling in vivo reactions. *Comput. Biol. Chem.* **28**(3), 165–178.

Wright, D. J. (1980). Digital simulation of Poisson stochastic differential equations. *Internat. J. Systems. Sci.* **11**, 781–785.

Zhang, G. L., P. H. W. Leong, C. H. Ho, K. H. Tsoi, C. C. C. Cheung, D.-U. Lee, R. C. C. Cheung, & W. Luk (2005). Reconfigurable acceleration for Monte Carlo based financial simulation. In *Proceedings of the International Conference on Field Programmable Technology (FPT)*, pp. 215–222.