Mounir Zili
Darya V. Filatova *Editors*

Stochastic Differential Equations and Processes

Springer

Springer Proceedings in Mathematics

Volume 7

For further volumes:
http://www.springer.com/series/8806

Springer Proceedings in Mathematics

The book series will feature volumes of selected contributions from workshops and conferences in all areas of current research activity in mathematics. Besides an overall evaluation, at the hands of the publisher, of the interest, scientific quality, and timeliness of each proposal, every individual contribution will be refereed to standards comparable to those of leading mathematics journals. It is hoped that this series will thus propose to the research community well-edited and authoritative reports on newest developments in the most interesting and promising areas of mathematical research today.

Mounir Zili • Darya V. Filatova
Editors

Stochastic Differential Equations and Processes

SAAP, Tunisia, October 7-9, 2010

Springer

Editors
Mounir Zili
Preparatory Institute
to the Military Academies
Department of Mathematics
Avenue Marechal Tito
4029 Sousse
Tunisia
zilimounir@yahoo.fr

Darya V. Filatova
Jan Kochanowski University in Kielce
ul. Krakowska 11
25-029 Kielce
Poland
daria_filatova@interia.pl

ISSN 2190-5614
ISBN 978-3-642-22367-9 e-ISBN 978-3-642-22368-6
DOI 10.1007/978-3-642-22368-6
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2011938117

Mathematical Subject Classification (2010): 60HXX, 60GXX, 60FXX, 91BXX, 93BXX, 92DX

Cover design: deblik, Berlin

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Stochastic analysis is currently undergoing a period of intensive research and various new developments, motivated in part by the need to model, understand, forecast, and control the behavior of many natural phenomena that evolve in time in a random way. Such phenomena appear in the fields of finance, telecommunications, economics, biology, geology, demography, physics, chemistry, signal processing, and modern control theory, to mention just a few.

Often, it is very convenient to use stochastic differential equations and stochastic processes to study stochastic dynamics. In such cases, research needs the guarantee of some theoretical properties, such as the existence and uniqueness of the stochastic equation solution. Without a deep understanding of the nature of the stochastic process this is seldom possible. The theoretical background of both stochastic processes and stochastic differential equations are therefore very important.

Nowadays, quite a few stochastic differential equations can be solved by means of exact methods. Even if this solution exists, it cannot necessarily be used for computer simulations, in which the continuous model is replaced by a discrete one. The problems of "ill-posed" tasks, the "stiffness" or "stability" of the system limit numerical approximations of the stochastic differential equation. As a result, new approaches for the numerical solution and, consequently, new numerical algorithms are also very important.

This volume contains 8 refereed papers dealing with these topics, chosen from among the contributions presented at the international conference on *Stochastic Analysis and Applied Probability* (SAAP 2010), which was held at Yasmine-Hammamet, Tunisia, from 7 to 9 October 2010. This conference was organized by the "Applied Mathematics & Mathematical Physics" research unit of the preparatory institute to the military academies of Sousse, Tunisia. It brought together some 60 researchers and PhD students, from 14 countries and 5 continents. Through lectures, communications, and posters, these researchers reported on theoretical, numerical, or application work as well as on significant results obtained for several topics within the field of stochastic analysis and probability, particularly for "Stochastic processes and stochastic differential equations." The conference program was planned by an international committee chaired by Mounir Zili (Preparatory Institute to the Military Academies of Sousse, Tunisia) and consisted of Darya Filatova (Jan Kochanowski University in Kielce, Poland), Ibtissem Hdhiri (Faculty of Sciences of Gabès, Tunisia), Ciprian A. Tudor (University of Lille, France), and Mouna Ayachi (Faculty of Sciences of Monastir, Tunisia).

As this book emphasizes the importance of numerical and theoretical studies of the stochastic differential equations and stochastic processes, it will be useful for a wide spectrum of researchers in applied probability, stochastic numerical and theoretical analysis and statistics, as well as for graduate students.

To make it more complete and accessible for graduate students, practitioners, and researchers, we have included a survey dedicated to the basic concepts of numerical analysis of the stochastic differential equations, written by Henri Schurz. This survey is valuable not only due to its excellent theoretical conception with respect to modern tendencies, but also with regard to its comprehensive concept of the dynamic consistency of numerical methods for the stochastic differential equations. In a second paper, motivated by its applications in econometrics, Ciprian Tudor develops an asymptotic theory for some regression models involving standard Brownian motion and the standard Brownian sheet. The result proved in this paper is an impressive example of convergence in distribution to a non-Gaussian limit. The paper "General shot noise processes and functional convergence to stable processes" by Wissem Jedidi, Jalal Almhana, Vartan Choulakian, and Robert McGorman also addresses the topic of stochastic processes, and the authors consider a model appropriate for the network traffic consisting of an infinite number of sources linked to a unique server. This model is based on a general Poisson shot noise representation, which is a generalization of a compound Poisson process. In the fourth paper of this volume, Charles El-Nouty deals with the lower classes of the sub-fractional Brownian motion, which has been introduced to model some self-similar Gaussian processes, with non-stationary increments. Then, in a paper by Mohamed Erraoui and Youssef Ouknine, the bounded variation of the flow of a stochastic differential equation driven by a fractional Brownian motion and with non-Lipschitz coefficients is studied. In the sixth paper, Antoine Ayache and Qidi Peng develop an extension of several probabilistic and statistical results for stochastic volatility models satisfying some stochastic differential equations for cases in which the fractional Brownian motion is replaced by the multifractional Brownian motion. The advantage of the multifractional stochastic volatility models is that they allow account variations with respect to time of volatility local roughness. The seventh paper was written by Archil Gulisashvili and Josep Vives and addresses two-sided estimates for the distribution density of standard models, perturbed by a double exponential law. The results obtained in this paper can especially be used in the study of distribution densities arising in some stochastic stock price models. And in the last paper in the volume, Mario Lefebvre explicitly solves the problem of maximizing a function of the time spent by a stochastic process by arriving at solutions of some particular stochastic differential equations.

All the papers presented in this book were carefully reviewed by the members of the SAAP 2010 Scientific Committee, a list of which is presented in the appendix.

We would like to thank the anonymous reviewers for their reports and many others who contributed enormously to the high standards for publication in these proceedings by carefully reviewing the manuscripts that were submitted.

Finally, we want to express our gratitude to Marina Reizakis for her invaluable help in the process of preparing this volume edition.

Tunisia
September 2011

Mounir Zili
Daria Filatova

Contents

Contributors

Jalal Almhana GRETI Group, University of Moncton, Moncton, NB E1A3E9, Canada, almhanaj@umoncton.ca

Antoine Ayache U.M.R. CNRS 8524, Laboratory Paul Painlevé, University Lille 1, 59655 Villeneuve d'Ascq Cedex, France, smith@smith.edu

Vartan Choulakian GRETI Group, University of Moncton, Moncton, NB E1A3E9, Canada, choulav@umoncton.ca

Charles El-Nouty UMR 557 Inserm/ U1125 Inra/ Cnam/ Université Paris XIII, SMBH-Université Paris XIII, 74 rue Marcel Cachin, 93017 Bobigny Cedex, France, c.el-nouty@uren.smbh.univ-paris13.fr

Mohamed Erraoui Faculté des Sciences Semlalia Département de Mathématiques, Université Cadi Ayyad BP 2390, Marrakech, Maroc, erraoui@ucam.ac.ma

Archil Gulisashvili Department of Mathematics, Ohio University, Athens, OH 45701, USA, gulisash@ohio.edu

Wissem Jedidi Department of Mathematics, Faculty of Sciences of Tunis, Campus Universitaire, 1060 Tunis, Tunisia, wissem.jedidi@fst.rnu.tn

Mario Lefebvre Département de Mathématiques et de Génie Industriel, École Polytechnique, C.P. 6079, Succursale Centre-ville, Montréal, PQ H3C 3A7, Canada, mlefebvre@polymtl.ca

Robert McGorman NORTEL Networks, 4001 E. Chapel Hill-Nelson Hwy, Research Triangle Park, NC 27709, USA, mcgorman@nortelnetworks.com

Youssef Ouknine Faculté des Sciences Semlalia Département de Mathématiques, Université Cadi Ayyad BP 2390, Marrakech, Maroc, ouknine@ucam.ac.ma

Qidi Peng U.M.R. CNRS 8524, Laboratory Paul Painlevé, University Lille 1, 59655 Villeneuve d'Ascq Cedex, France, smith@smith.edu

Henry Schurz Department of Mathematics, Southern Illinois University, 1245 Lincoln Drive, Carbondale, IL 62901, USA, hschurz@math.siu.edu

Ciprian A. Tudor Laboratoire Paul Painlevé, Université de Lille 1, F-59655 Villeneuve d'Ascq, France, tudor@math.univ-lille1.fr

Josep Vives Departament de Probabilitat, Lògica i Estadística, Universitat de Barcelona, Gran Via 585, 08007-Barcelona (Catalunya), Spain, josep.vives@ub.edu

Chapter 1
Basic Concepts of Numerical Analysis of Stochastic Differential Equations Explained by Balanced Implicit Theta Methods

Henri Schurz

Abstract We present the comprehensive concept of dynamic consistency of numerical methods for (ordinary) stochastic differential equations. The concept is illustrated by the well-known class of balanced drift-implicit stochastic Theta methods and relies on several well-known concepts of numerical analysis to replicate the qualitative behaviour of underlying continuous time systems under adequate discretization. This involves the concepts of consistency, stability, convergence, positivity, boundedness, oscillations, contractivity and energy behaviour. Numerous results from literature are reviewed in this context.

1.1 Introduction

Numerous monographs and research papers on numerical methods of stochastic differential equations are available. Most of them concentrate on the construction and properties of consistency. A few deal with stability and longterm properties. However, as commonly known, the replication of qualitative properties of numerical methods in its whole is the most important issue for modeling and real-world applications. To evaluate numerical methods in a more comprehensive manner, we shall discuss the concept of **dynamic consistency** of numerical methods for stochastic differential equations. For the sake of precise illustration, we will treat the example class of balanced implicit **outer Theta methods**. This class is defined by

H. Schurz (✉)
Southern Illinois University, Department of Mathematics, 1245 Lincoln Drive, Carbondale, IL 62901, USA
e-mail: hschurz@math.siu.edu

M. Zili and D.V. Filatova (eds.), *Stochastic Differential Equations and Processes*, Springer Proceedings in Mathematics 7, DOI 10.1007/978-3-642-22368-6_1,

$$X_{n+1} = \begin{cases} X_n + [\Theta_n a(t_{n+1}, X_{n+1}) + (I - \Theta_n) a(t_n, X_n)]\, h_n + \sum_{j=1}^{m} b^j(t_n, X_n) \Delta W_n^j \\ \qquad + \sum_{j=0}^{m} c^j(t_n, X_n)(X_n - X_{n+1}) |\Delta W_n^j| \end{cases} \tag{1.1}$$

with appropriate (bounded) matrices c^j with continuous entries, where I is the unit matrix in $\mathbb{R}^{d\times d}$ and

$$\Delta W_n^0 = h_n, \quad \Delta W_n^j = W^j(t_{n+1}) - W^j(t_n)$$

along partitions

$$0 = t_0 < t_1 < \ldots < t_n < t_{n+1} < \ldots < t_{n_T} = T < +\infty$$

of finite time-intervals $[0, T]$. These methods are discretizations of d-dimensional ordinary stochastic differential equations (SDEs), [3, 14, 32, 33, 81, 86, 102, 108]

$$dX(t) = a(t, X(t))dt + \sum_{j=1}^{m} b^j(t, X(t)) dW^j(t) \left(= \sum_{j=0}^{m} b^j(t, X(t)) dW^j(t) \right) \tag{1.2}$$

(with $b^0 = a$, $W^0(t) = t$), driven by i.i.d. Wiener processes W^j and started at adapted initial values $X(0) = x_0 \in \mathbb{R}^d$. The vector fields a and b^j are supposed to be sufficiently smooth throughout this survey. All stochastic processes are constructed on the complete probability basis $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$.

The aforementioned Theta methods (1.1) represent a first natural generalization of explicit and implicit Euler methods. Indeed, they are formed by a convex linear combinations of explicit and implicit Euler increment functions of the drift part, whereas the diffusion part is explicitly treated due to the problem of adequate integration within one and the same stochastic calculus. The balanced terms c^j are appropriate matrices and useful to control the pathwise (i.e. almost sure) behaviour and uniform boundedness of those approximations. The parameter matrices $(\Theta_n)_{n\in\mathbb{N}} \in \mathbb{R}^{d\times d}$ determine the degree of implicitness and simplective behaviour (energy- and area-preserving character) of related approximations. Most popular representatives are those with simple scalar choices $\Theta_n = \theta_n I$ where I denotes the unit matrix in $\mathbb{R}^{d\times d}$ and $\theta_n \in \mathbb{R}^1$. Originally, without balanced terms c^j, they were invented by Talay [138] in stochastics, who proposed $\Theta_n = \theta I$ with autonomous scalar choices $\theta \in [0, 1]$. This family with matrix-valued parameters $\Theta \in \mathbb{R}^{d\times d}$ has been introduced by Ryashko and Schurz [116] who also proved their mean square convergence with an estimate of worst case convergence rate 0.5. If $\theta = 0$ then its scheme reduces to the classical **(forward) Euler method** (see Maruyama [90], Golec et al. [35–38], Guo [39, 40], Gyöngy [41, 42], Protter and

Talay [109], Römisch & Wakolbinger [115], Tudor & Tudor [143] among others), if $\theta = 1$ to the **backward Euler method** which is also called *(drift-)implicit Euler method* (Hu [55]), and if $\theta = 0.5$ to the **(drift-implicit) trapezoidal method**, reducing to the scheme

$$X_{n+1} = X_n + \frac{1}{2}\left[a(t_{n+1}, X_{n+1}) + a(t_n, X_n)\right] h_n + \sum_{j=1}^{m} b^j(t_n, X_n)\Delta W_n^j \tag{1.3}$$

without balanced terms c^j. A detailed study of the qualitative dynamic behaviour of these methods can be found in Stuart and Peplow [136] in deterministic numerical analysis (in the sense of spurious solutions), and in Schurz [120] in stochastic numerical analysis.

A slightly different class of numerical methods is given by the balanced implicit **inner Theta methods**

$$X_{n+1} = \begin{cases} X_n + a\left(t_n + \theta_n h_n, \Theta_n X_{n+1} + (I - \Theta_n) X_n\right) h_n + \sum_{j=1}^{m} b^j(t_n, X_n)\,\Delta W_n^j \\ \qquad + \sum_{j=0}^{m} c^j(t_n, X_n)(X_n - X_{n+1})|\Delta W_n^j| \end{cases} \tag{1.4}$$

where $\theta_n \in \mathbb{R}$, $\Theta_n \in \mathbb{R}^{d\times d}$ such that local algebraic resolution can be guaranteed always. The most known representative of this class (1.4) with $\Theta_n = 0.5I$ and without balanced terms c^j is known as the **drift-implicit midpoint method** governed by

$$X_{n+1} = X_n + a\left(\frac{t_{n+1} + t_n}{2}, \frac{X_{n+1} + X_n}{2}\right) \Delta_n + \sum_{j=1}^{m} b^j(t_n, X_n)\,\Delta W_n^j. \tag{1.5}$$

This method is superior for the integration of conservation laws and Hamiltonian systems. Their usage seems to be very promising for the control of numerical stability, area-preservation and boundary laws in stochastics as well. The drawback for their practical implementation can be seen in the local resolution of nonlinear algebraic equations which is needed in addition to explicit methods. However, this fact can be circumvented by its practical implementation through predictor-corrector methods (PCMs), their linear- (LIMs) or partial-implicit (PIMs) derivates (versions). In passing, note that the **partitioned Euler methods** (cf. Strommen–Melbo and Higham [135]) are also a member of stochastic Theta methods (1.1) with the special choice of constant implicitness-matrix

$$\Theta_n = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}.$$

In passing, note that stochastic Theta methods (1.1) represent the simplest class of stochastic Runge-Kutta methods. Despite their simplicity, they are rich enough to cover many aspects of numerical approximations in an adequate manner.

The purpose of this survey is to compile some of the most important facts on representatives of classes (1.1) and (1.4). Furthermore, we shall reveal the goodness of these approximation techniques in view of their dynamic consistency. In the following sections we present and discuss several important key concepts of stochastic numerical analysis explained by Theta methods. At the end we finalize our presentation with a summary leading to the governing concept of dynamic consistency unifying the concepts presented before in a complex fashion. The paper is organized in 12 sections. The remaining part of our introduction reports on auxiliary tools to construct, derive, improve and justify consistency of related numerical methods for SDEs. Topics as consistency in Sect. 1.2, asymptotic stability in Sect. 1.3, convergence in Sect. 1.4, positivity in Sect. 1.5, boundedness in Sect. 1.6, oscillations in Sect. 1.7, energy in Sect. 1.8, order bounds in Sect. 1.9, contractivity in Sect. 1.10 and dynamic consistency in Sect. 1.11 are treated. Finally, the related references are listed alphabetically, without claiming to refer to all relevant citations in the overwhelming literature on those subjects. We recommend also to read the surveys of Artemiev and Averina [5], Kanagawa and Ogawa [66], Pardoux and Talay [106], S. [125] and Talay [140] in addition to our paper. A good introduction to related basic elements is found in Allen [1] and [73] too.

1.1.1 Auxiliary tool: Itô Formula (Itô Lemma) with Operators $\mathscr{L}^j$

Define linear partial differential operators

$$\mathscr{L}^0 = \frac{\partial}{\partial t} + < a(t,x), \nabla_x >_d + \frac{1}{2}\sum_{j=1}^{m}\sum_{i,k=1}^{d} b_i^j(t,x) b_k^j(t,x) \frac{\partial^2}{\partial x_k \partial x_i} \tag{1.6}$$

and $\mathscr{L}^j = < b^j(t,x), \nabla_x >_d$ where $j = 1, 2, \ldots, m$. Then, thanks to the fundamental contribution of Itô [56] and [57], we have the following lemma.

Lemma 1.1.1 (Stopped Itô Formula in Integral Operator Form). *Assume that the given deterministic mapping $V \in C^{1,2}([0,T] \times \mathbb{R}^d, \mathbb{R}^k)$. Let τ be a finite $\mathscr{F}_t$-adapted stopping time with $0 \le t \le \tau \le T$.*
Then, we have

$$V(\tau, X(\tau)) = V(t, X(t)) + \sum_{j=0}^{m} \int_t^\tau \mathscr{L}^j V(s, X(s))\, dW^j(s)\,. \tag{1.7}$$

1.1.2 Auxiliary Tool: Derivation of Stochastic Itô-Taylor Expansions

By iterative application of Itô formula we gain the family of stochastic Taylor expansions. This idea is due to Wagner and Platen [144]. Suppose we have enough smoothness of V and of coefficients a, b^j of the Itô SDE. Remember, thanks to Itô's formula, for $t \geq t_0$

$$V(t, X(t)) = V(t_0, X(t_0)) + \int_{t_0}^{t} \mathscr{L}^0 V(s, X(s))ds + \sum_{j=1}^{m} \int_{t_0}^{t} \mathscr{L}^j V(s, X(s))\, dW^j(s).$$

Now, take $V(t, x) = x$ at the first step, and set $b^0(t, x) \equiv a(t, x)$, $W^0(t) \equiv t$. Then one derives

$$\begin{aligned}
X(t) &= X(t_0) + \int_{t_0}^{t} a(s, X(s))ds + \sum_{j=1}^{m} \int_{t_0}^{t} b^j(s, X(s))\, dW^j(s) \\
V &\equiv b^j \\
&= X(t_0) + \int_{t_0}^{t} \left[a(t_0, X(t_0)) + \sum_{k=0}^{m} \int_{t_0}^{s} \mathscr{L}^k a(u, X(u)) dW^k(u) \right] ds \\
&\quad + \sum_{j=1}^{m} \int_{t_0}^{t} \left[b^j(t_0, X(t_0)) + \sum_{k=0}^{m} \int_{t_0}^{s} \mathscr{L}^k a(u, X(u)) dW^k(u) \right] dW^j(s) \\
b^0 &\equiv a \\
&= X(t_0) + \underbrace{\sum_{j=0}^{m} b^j(t_0, X(t_0)) \int_{t_0}^{t} dW^j(s)}_{\textbf{Euler Increment}} \\
&\quad + \underbrace{\sum_{j,k=0}^{m} \int_{t_0}^{t} \int_{t_0}^{s} \mathscr{L}^k b^j(u, X(u))\, dW^k(u) dW^j(s)}_{\textbf{Remainder Term } \mathbf{R}_E} \\
V &\equiv \mathscr{L}^k b^j \\
&= X(t_0) \\
&\quad + \underbrace{\sum_{j=0}^{m} b^j(t_0, X(t_0)) \int_{t_0}^{t} dW^j(s) + \sum_{j,k=1}^{m} \mathscr{L}^k b^j(t_0, X(t_0)) \int_{t_0}^{t} \int_{t_0}^{s} dW^k(u) dW^j(s)}_{\textbf{Increment of Milstein Method}}
\end{aligned}$$

$$+\sum_{j=1}^{m}\int_{t_0}^{t}\int_{t_0}^{s}\mathscr{L}^0 b^j(u,X(u))\,du\,dW^j(s)$$

$$+\sum_{k=1}^{m}\int_{t_0}^{t}\int_{t_0}^{s}\mathscr{L}^k a(u,X(u))\,dW^k(u)ds$$

$$+\underbrace{\sum_{j,k=1,l=0}^{m}\int_{t_0}^{t}\int_{t_0}^{s}\int_{t_0}^{u}\mathscr{L}^l\mathscr{L}^k b^j(z,X(z))\,dW^l(z)dW^k(u)dW^j(s)}_{\textbf{Remainder Term }\mathbf{R}_M}$$

$$\begin{aligned}V &\equiv \mathscr{L}^k b^j\\ &= X(t_0)\\ &\quad+\underbrace{\sum_{j=0}^{m}b^j(t_0,X(t_0))\int_{t_0}^{t}dW^j(s)+\sum_{j,k=0}^{m}\mathscr{L}^k b^j(t_0,X(t_0))\int_{t_0}^{t}\int_{t_0}^{s}dW^k(u)dW^j(s)}_{\textbf{Increment of 2nd order Taylor Method}}\\ &\quad+\underbrace{\sum_{j,k,l=0}^{m}\int_{t_0}^{t}\int_{t_0}^{s}\int_{t_0}^{u}\mathscr{L}^l\mathscr{L}^k b^j(z,X(z))\,dW^l(z)dW^k(u)dW^j(s)}_{\textbf{Remainder Term }\mathbf{R}_{TM2}}\end{aligned}$$

$$\begin{aligned}V &\equiv \mathscr{L}^r\mathscr{L}^k b^j\\ &= X(t_0)\\ &\quad+\underbrace{\sum_{j=0}^{m}b^j(t_0,X(t_0))\int_{t_0}^{t}dW^j(s)+\sum_{j,k=0}^{m}\mathscr{L}^k b^j(t_0,X(t_0))\int_{t_0}^{t}\int_{t_0}^{s}dW^k(u)dW^j(s)}_{\textbf{Increment of 3rd order Taylor Method}}\\ &\quad+\underbrace{\sum_{j,k,r=0}^{m}\mathscr{L}^r\mathscr{L}^k b^j(t_0,X(t_0))\int_{t_0}^{t}\int_{t_0}^{s}\int_{t_0}^{u}dW^r(v)dW^k(u)dW^j(s)}_{\textbf{Increment of 3rd order Taylor Method}}\\ &\quad+\underbrace{\sum_{j,k,r,l=0}^{m}\int_{t_0}^{t}\int_{t_0}^{s}\int_{t_0}^{u}\int_{t_0}^{v}\mathscr{L}^l\mathscr{L}^r\mathscr{L}^k b^j(z,X(z))dW^l(z)dW^r(v)dW^k(u)dW^j(s)}_{\textbf{Remainder Term }\mathbf{R}_{TM3}}\\ &\quad\ldots\ldots\ldots\ldots\ldots\end{aligned}$$

This process can be continued under appropriate assumptions of smoothness and boundedness of the involved expressions. Thus, this is the place from which most numerical methods systematically originate, and where the main tool for consistency analysis is coming from. One has to expand the functionals in a hierarchical way, otherwise one would loose important order terms, and the implementation would be inefficient. Of course, for qualitative, smoothness and efficiency reasons we do not have to expand all terms in the Taylor expansions at the same time (e.g. cf. Milstein increment versus 2nd order Taylor increments). The *Taylor method* can be read down straight forward by truncation of stochastic Taylor expansion. *Explicit* and *implicit methods*, *Runge-Kutta methods*, *inner and outer Theta methods*, *linear-implicit* or *partially implicit* methods are considered as modifications of Taylor methods by substitution of derivatives by corresponding difference quotients, explicit expressions by implicit ones, respectively. However, it necessitates finding a more efficient form for representing stochastic Taylor expansions and hence Taylor-type methods. For this aim, we shall report on hierarchical sets, coefficient functions and multiple integrals in the subsection below.

In general, Taylor-type expansions are good to understand the systematic construction of numerical methods with certain orders. Moreover, they are useful to prove certain rates of local consistency of numerical methods. However, the rates of convergence (global consistency) of them are also determined by other complex dynamical features of numerical approximations, and "order bounds" and "practical modeling / simulation issues" may decisively limit their usage in practice. To fully understand this statement, we refer to the concept of "dynamic consistency" as developed in the following sections in this paper.

1.1.3 Auxiliaries: Hierarchical Sets, Coefficient Functions, Multiple Integrals

Kloeden and Platen [72] based on the original work of Wagner and Platen [144] have introduced a more compact, efficient formulation of stochastic Taylor expansions. For its statement, we have to formulate what is meant by multiple indices, hierarchical sets, remainder sets, coefficient functions and multiple integrals in the Itô sense.

Definition 1.1.1. A multiple index has the form $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_{l(\alpha)})$ where $l(\alpha) \in \mathbb{N}$ is called the *length* of the multiple index α, and $n(\alpha)$ is the *total number* of zero entries of α. The symbol ν denotes the *empty multiple index* with $l(\nu) = 0$. The operations $\alpha- = (\alpha_1, \ldots, \alpha_{l(\alpha)-1})$ and $-\alpha = (\alpha_2, \ldots, \alpha_{l(\alpha)})$ are called *right-* and *left-subtraction*, respectively (in particular, $(\alpha_1)- = -(\alpha_1) = \nu$). The *set of all multiple indices* is defined to be

$$\mathcal{M}_{k,m} = \left\{\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_{l(\alpha)}) : \alpha_i \in \{k, k+1, \ldots, m\}, i = 1, 2, \ldots, l(\alpha), l(\alpha) \in \mathbb{N}\right\}.$$

A *hierarchical set* $Q \subset \mathscr{M}_{0,m}$ is any set of multiple indices $\alpha \in \mathscr{M}_{0,m}$ such that $\nu \in Q$ and $\alpha \in Q$ implies $-\alpha \in Q$. The hierarchical set Q_k denotes the set of all multiple indices $\alpha \in \mathscr{M}_{0,m}$ with length smaller than $k \in \mathbb{N}$, i.e.

$$Q_k = \{\alpha \in \mathscr{M}_{0,m} : l(\alpha) \leq k\}.$$

The set

$$R(Q) = \{\alpha \in \mathscr{M}_{0,m} \setminus Q : \alpha- \in Q\}$$

is called the *remainder set* $R(Q)$ of the hierarchical set Q. A *multiple (Itô) integral* $I_{\alpha,s,t}[V(.,.)]$ is defined to be

$$I_{\alpha,s,t}[V(.,.)] = \begin{cases} \int_s^t I_{-\alpha,s,u}[V(.,.)]\, dW^{\alpha_1}(u) & \text{if} \quad l(\alpha) > 1 \\ \int_s^t V(u, X_u)\, dW^{\alpha_{l(\alpha)}}(u) & \text{otherwise} \end{cases}$$

for a given process $V(t, X(t))$ where $V \in C^{0,0}$ and fixed $\alpha \in \mathscr{M}_{0,m} \setminus \{\nu\}$. A *multiple (Itô) coefficient function* $V_\alpha \in C^{0,0}$ for a given mapping $V = V(t,x) \in C^{l(\alpha),2l(\alpha)}$ is defined to be

$$V_\alpha(t,x) = \begin{cases} \mathscr{L}^{l(\alpha)} V_{\alpha-}(t,x) & \text{if} \quad l(\alpha) > 0 \\ V(t,x) & \text{otherwise} \end{cases}.$$

Similar notions can be introduced with respect to Stratonovich calculus (in fact, in general with respect to any stochastic calculus), see [72] for Itô and Stratonovich calculus.

1.1.4 Auxiliary Tool: Compact Formulation of Wagner-Platen Expansions

Now we are able to state a general form of Itô-Taylor expansions. Stochastic Taylor-type expansions for Itô diffusion processes have been introduced and studied by Wagner and Platen [144] (cf. also expansions in Sussmann [137], Arous [4], and Hu [54]). Stratonovich Taylor-type expansions can be found in Kloeden and Platen [72]. We will follow the original main idea of Wagner and Platen [144].

An Itô-Taylor expansion for an Itô SDE (1.2) is of the form

$$V(t, X(t)) = \sum_{\alpha \in Q} V_\alpha(s, X(s)) I_{\alpha,s,t} + \sum_{\alpha \in R(Q)} I_{\alpha,s,t}[V_\alpha(.,.)] \tag{1.8}$$

for a given mapping $V = V(t,x) : [0,T] \times \mathbb{R}^d \longrightarrow \mathbb{R}^k$ which is smooth enough, where $I_{\alpha,s,t}$ without the argument $[\cdot]$ is understood to be $I_{\alpha,s,t} = I_{\alpha,s,t}[1]$. Sometimes this formula is also referred to as *Wagner-Platen expansion*. Now, for completeness, let us restate the Theorem 5.1 of Kloeden and Platen [72].

Theorem 1.1.1 (Wagner-Platen Expansion). *Let ρ and τ be two $\mathscr{F}_t$-adapted stopping times with $t_0 \leq \rho \leq \tau \leq T < +\infty$ (a.s.). Assume $V : [0,T] \times I\!R^d \longrightarrow I\!R^k$. Take any hierarchical set $Q \in \mathscr{M}_{0,m}$.*

Then, each Itô SDE with coefficients a, b^j possesses a Itô-Taylor expansion (1.8) with respect to the hierarchical set Q, provided that all derivatives of V, a, b^j (related to Q) exist.

A proof is carried out in using the Itô formula and induction on the maximum length $\sup_{\alpha \in Q} l(\alpha) \in \mathbb{N}$. A similar expansion holds for Stratonovich SDEs.

1.1.5 Auxiliary Tool: Relations Between Multiple Integrals

The following lemma connects different multiple integrals. In particular, its formula can be used to express multiple integrals by other ones and to reduce the computational effort of their generation. The following lemma is a slightly generalized version of an auxiliary lemma taken from Kloeden and Platen [72], see proposition 5.2.3, p. 170.

Lemma 1.1.2 (Fundamental Lemma of Multiple Integrals). *Let $\alpha = (j_1, j_2, \ldots, j_{l(\alpha)}) \in \mathscr{M}_{0,m} \setminus \{\nu\}$ with $l(\alpha) \in I\!N$.*
Then, $\forall k \in \{0, 1, \ldots, m\}$ $\forall t, s : 0 \leq s \leq t \leq T$ we have

$$
\begin{aligned}
(W^k(t) - W^k(s)) I_{\alpha,s,t} &= \sum_{i=0}^{l(\alpha)} I_{(j_1,j_2,\ldots,j_i,k,j_{i+1},\ldots,j_{l(\alpha)}),s,t} \\
&\quad + \sum_{i=0}^{l(\alpha)} \chi_{\{j_i=k\neq 0\}} I_{(j_1,j_2,\ldots,j_{i-1},0,j_{i+1},\ldots,j_{l(\alpha)}),s,t}
\end{aligned}
\tag{1.9}
$$

$$
\begin{aligned}
&= I_{(k,j_1,j_2,\ldots,j_{l(\alpha)}),s,t} + I_{(j_1,k,j_2,\ldots,j_{l(\alpha)}),s,t} + I_{(j_1,j_2,k,j_3,\ldots,j_{l(\alpha)}),s,t} + \ldots + \\
&\quad + I_{(j_1,j_2,j_3,\ldots,j_{l(\alpha)},k),s,t} + \sum_{i=0}^{l(\alpha)} \chi_{\{j_i=k\neq 0\}} I_{(j_1,j_2,\ldots,j_{i-1},0,j_{i+1},\ldots,j_{l(\alpha)}),s,t}
\end{aligned}
$$

where $\chi_{\{.\}}$ denotes the characteristic function of the subscribed set.

Hence, it obviously suffices to generate "minimal basis sets" of multiple integrals. In order to have a more complete picture on the structure of multiple integrals, we note the following assertion.

Lemma 1.1.3 (Hermite Polynomial Recursion of Multiple Integrals). *Assume that the multiple index α is of the form*

$$\alpha = (j_1, j_2, \ldots, j_{l(\alpha)}) \in \mathscr{M}_{0,m} \text{ with } j_1 = j_2 = \ldots = j_{l(\alpha)} = j \in 0, 1, \ldots, m$$

and its length $l(\alpha) \geq 2$.
Then, for all t with $t \geq s \geq 0$ we have

$$I_{\alpha,s,t} = \begin{cases} \dfrac{(t-s)^{l(\alpha)}}{l(\alpha)!}, & j = 0 \\ \dfrac{(W^j(t) - W^j(s))I_{\alpha-,s,t} - (t-s)I_{(\alpha-)-,s,t}}{l(\alpha)!}, & j \geq 1 \end{cases} \tag{1.10}$$

This lemma corresponds to a slightly generalized version of Corollary 5.2.4 (p. 171) in [72]. It is also interesting to note that this recursion formula for multiple Itô integrals coincides with the recursion formula for hermite polynomials. Let us conclude with a list of relations between multiple integrals which exhibit some consequences of Lemmas 1.1.2 and 1.1.3. For more details, see [72]. Take $j, k \in \{0, 1, \ldots, m\}$ and $0 \leq s \leq t \leq T$.

$$\begin{aligned}
I_{(j),s,t} &= W^j(t) - W^j(s) \\
I_{(j,j),s,t} &= \frac{1}{2!}\left(I^2_{(j),s,t} - (t-s)\right) \\
I_{(j,j,j),s,t} &= \frac{1}{3!}\left(I^3_{(j),s,t} - 3(t-s)I_{(j),s,t}\right) \\
I_{(j,j,j,j),s,t} &= \frac{1}{4!}\left(I^4_{(j),s,t} - 6(t-s)I^2_{(j),s,t} + 3(t-s)^2\right) \\
I_{(j,j,j,j,j),s,t} &= \frac{1}{5!}\left(I^5_{(j),s,t} - 10(t-s)I^3_{(j),s,t} + 15(t-s)^2 I_{(j),s,t}\right) \\
&\ldots\ldots\ldots. \\
(t-s)I_{(j),s,t} &= I_{(j,0),s,t} + I_{(0,j),s,t} \\
(t-s)I_{(j,k),s,t} &= I_{(j,k,0),s,t} + I_{(j,0,k),s,t} + I_{(0,j,k),s,t} \\
I_{(j),s,t}I_{(0,j),s,t} &= 2I_{(0,j,j),s,t} + I_{(j,0,j),s,t} + I_{(j,j,0),s,t} \\
I_{(j),s,t}I_{(j,0),s,t} &= I_{(0,j,j),s,t} + I_{(j,0,j),s,t} + 2I_{(j,j,0),s,t} \\
&\ldots\ldots\ldots..
\end{aligned}$$

Some attempts has been made to approximate multiple stochastic integrals. For example, [72] use the technique of Karhunen-Loeve expansion (i.e. the Fourier series expansion of the Wiener process) or [30] exploit Box counting methods and related levy areas. A minimal basis set for multiple integrals is known, see [28, 29]. However, computationally more efficient approximation procedures of multiple stochastic integrals are still a challenge to be constructed and verified (especially in higher dimensions).

1.2 Local Consistency

Throughout this section, fix the time interval $[0, T]$ with finite and nonrandom terminal time T. Let $\|.\|_d$ be the Euclidean vector norm on $\mathbb{R}^d$ and $\mathscr{M}_p([s,t])$ the Banach space of $(\mathscr{F}_u)_{s \leq u \leq t}$-adapted, continuous, $\mathbb{R}^d$-valued stochastic processes X with finite norm $\|X\|_{\mathscr{M}_p} = (\sup_{s \leq u \leq t} \mathbb{E}||X(s)||_d^p)^{1/p} < +\infty$ where $p \geq 1$, $\mathscr{M}([0,s])$ the space of $(\mathscr{F}_s, \mathscr{B}(\mathbb{R}^d))$-measurable stochastic processes and $\mathscr{B}(S)$ the σ-algebra of Borel sets of inscribed set S.

Recall that every (one-step) numerical method Y (difference scheme) defined by

$$Y_{n+1} = Y_n + \Phi_n(Y)$$

with increment functional Φ_n has an associated continuous one-step representation

$$Y_{s,x}(t) = x + \Phi(t|s,x)$$

along partitions

$$0 = t_0 < t_1 < \ldots < t_n < t_{n+1} < \ldots < t_{n_T} = T < +\infty.$$

The continuity modulus of this one-step representation is the main subject of related consistency analysis. For this analysis, the auxiliary tools we presented in the introduction such as Itô formula and relations between multiple integrals are essential in deriving estimates of the one-step representation. For example, the **continuous time one-step representation** of stochastic Theta methods (1.1) is given by

$$\begin{aligned} Y_{s,x}(t) := x &+ [\Theta a(t, Y_{s,x}(t)) + (I - \Theta)a(s,x)]\,(t-s) \\ &+ \sum_{j=1}^{m} b^j(s,x)(W^j(t) \\ &- W^j(s)) + \sum_{j=0}^{m} c^j(s,x)(x - Y_{s,x}(t))|W^j(t) - W^j(s)| \end{aligned} \tag{1.11}$$

driven by stochastic processes W^j, for all $t \geq s \geq 0$ and started at $x \in \mathbb{D}$ at time s.

Definition 1.2.1. A numerical method Y with one-step representation $Y_{s,y}(t)$ is said to be **mean consistent with rate** r_0 on $[0,T]$ iff $\exists$ Borel-measurable function $V : \mathbb{D} \to \mathbb{R}^1_+$ and $\exists$ real constants $K_0^C \geq 0, \delta_0 > 0$ such that $\forall (\mathscr{F}_s, \mathscr{B}(\mathbb{D}))$-measurable random variables $Z(s)$ with $Z \in \mathscr{M}([0,s])$ and $\forall s,t : 0 \leq t - s \leq \delta_0$

$$||\mathbb{E}[X_{s,Z(s)}(t) - Y_{s,Z(s)}(t)|\mathscr{F}_s]||_d \;\leq\; K_0^C \, V(Z(s))\,(t-s)^{r_0}. \tag{1.12}$$

Remark 1.2.1. In the subsections below, we shall show that the balanced Theta methods (1.1) with uniformly bounded weights c^j and uniformly bounded parameters Θ_n are mean consistent with worst case rate $r_0 \geq 1.5$ and moment control function $V(x) = (1 + ||x||_d^2)^{1/2}$ for SDEs (1.2) with global Hölder-continuous and linear growth-bounded coefficients $b^j \in F \subset C^{1,2}([0,T] \times \mathbb{D})$ $(j = 0, 1, 2, \ldots, m)$.

Definition 1.2.2. A numerical method Y with one-step representation $Y_{s,y}(t)$ is said to be ***p*-th mean consistent with rate** r_2 on $[0, T]$ iff $\exists$ Borel-measurable function $V : \mathbb{D} \to \mathbb{R}^1_+$ and $\exists$ real constants $K_p^C \geq 0, \delta_0 > 0$ such that $\forall(\mathscr{F}_s, \mathscr{B}(\mathbb{D}))$-measurable random variables $Z(s)$ with $Z \in \mathscr{M}_p([0, s])$ and $\forall s, t$: $0 \leq t - s \leq \delta_0$

$$\Big(\mathbb{E}[||X_{s,Z(s)}(t) - Y_{s,Z(s)}(t)||_d^p|\mathscr{F}_s]\Big)^{1/p} \;\leq\; K_p^C\, V(Z(s))\,(t-s)^{r_p}. \tag{1.13}$$

If $p = 2$ then we also speak of **mean square consistency** with local mean square rate r_2.

Remark 1.2.2. Below, we shall prove that the balanced Theta methods (1.1) are mean square consistent with worst case rate $r_2 \geq 1.0$ and moment control function $V(x) = (1 + ||x||^2)_d^{1/2}$ for SDEs (1.2) with global Lipschitz-continuous and linear growth-bounded coefficients $b^j \in F \subset C^{1,2}([0,T] \times \mathbb{D})$ $(j = 0, 1, 2, \ldots, m)$.

In the proofs of consistency of balanced Theta methods (1.1) below, it is crucial that one exploits the explicit identity

$$\begin{aligned}
&Y_{s,x}(t) - x\\
&= M_{s,x}^{-1}(t)\,[\Theta a(t, Y_{s,x}(t)) - (I - \Theta)a(s,x)]\,(t-s)\\
&\quad + M_{s,x}^{-1}(t)\sum_{j=1}^{m} b^j(s,x)(W^j(t) - W^j(s))\\
&= M_{s,x}^{-1}(t)\Theta\,[a(t, Y_{s,x}(t)) - a(s,x)]\int_s^t du + M_{s,x}^{-1}(t)\sum_{j=0}^{m} b^j(s,x)\int_s^t dW^j(u)
\end{aligned}$$

where I is the $d \times d$ unit matrix in $\mathbb{R}^{d\times d}$, $b^0(s,x) = a(s,x)$, $W^0(t) = t$, $W^0(s) = s$ and

$$M_{s,x}(t) = I + \sum_{j=0}^{m} c^j(s,x)|W^j(t) - W^j(s)|.$$

1.2.1 *Main Assumptions for Consistency Proofs*

Let $||.||_{d\times d}$ denote a matrix norm on $\mathbb{R}^{d\times d}$ which is compatible to the Euclidean vector norm $||.||_d$ on $\mathbb{R}^d$, and $\langle .,. \rangle_d$ the Euclidean scalar product on $\mathbb{R}^d$.

Furthermore we have to assume that the coefficients a and b^j are Caratheodory functions such that a strong, unique solution $X = (X(t))_{0\le t\le T}$ exists. Recall that $\mathbb{D} \subseteq \mathbb{R}^d$ is supposed to be a nonrandom set. Let $\mathbb{D}$ be simply connected. To guarantee certain rates of consistency of the BTMs (and also its rates of convergence) the following conditions have to be satisfied:

(A0) $\forall s,t \in [0,T] : s < t \implies \mathbb{P}\,(\{X(t) \in \mathbb{D} | X(s) \in \mathbb{D}\}) = \mathbb{P}\,(\{Y_{s,y}(t) \in \mathbb{D} | y \in \mathbb{D}\}) = 1$

(A1) $\exists$ constants $K_B = K_B(T), K_V = K_V(T) \ge 0$ such that

$$\forall t \in [0,T]\ \forall x \in \mathbb{D} : \sum_{j=0}^{m} ||b^j(t,x)||_d^2 \le (K_B)^2\,[V(x)]^2 \quad (1.14)$$

$$\sup_{0\le t\le T} \mathbb{E}[V(X(t))]^2 \le (K_V)^2 \mathbb{E}[V(X(0))]^2 < +\infty \quad (1.15)$$

with appropriate Borel-measurable function $V : \mathbb{D} \to \mathbb{R}^1_+$ satisfying

$$\forall x \in \mathbb{D} : ||x||_d \le V(x)$$

(A2) Hölder continuity of (a, b^j), i.e. $\exists$ real constants L_a and L_b such that

$$\forall s,t : 0 \le t-s \le \delta_0, \forall x,y \in \mathbb{D} : ||a(t,y) - a(s,x)||_d \le L_a(|t-s|^{1/2}$$

$$+||y-x||_d) \quad (1.16)$$

$$\sum_{j=1}^{m} ||b^j(t,y) - b^j(s,x)||_d^2 \le (L_b)^2(|t-s| + ||y-x||_d^2) \quad (1.17)$$

(A3) $\exists$ real constants $K_M = K_M(T) \ge 0$ such that, for the chosen weight matrices $c^j \in \mathbb{R}^{d\times d}$ of balanced Theta methods (1.1), we have

$$\forall s,t : 0 \le t-s \le \delta_0, \forall x \in \mathbb{D} : \exists M_{s,x}^{-1}(t) \text{ with } ||M_{s,x}^{-1}(t)||_{d\times d} \le K_M \quad (1.18)$$

(A4) $\exists$ real constants $K_{ca} = K_{ca}(T) \ge 0$ and $K_{cb} = K_{cb}(T) \ge 0$ such that, for the chosen weight matrices $c^j \in \mathbb{R}^{d\times d}$ of BTMs (1.1), we have

$$\forall s \in [0,T]\ \forall x \in \mathbb{D} : \sum_{j=0}^{m} ||c^j(s,x)a(s,x)||_d \le K_{ca}V(x) \quad (1.19)$$

$$\sum_{k=0}^{m}\sum_{j=0}^{m} ||c^k(s,x)b^j(s,x)||_d^2 \le K_{cb}^2[V(x)]^2 \quad (1.20)$$

(A5) $||\Theta||_{d\times d} \leq K_\Theta$, $||I-\Theta||_{d\times d} \leq K_{I-\Theta}$, and all step sizes $h_n \leq \delta_0$ are uniformly bounded by nonrandom quantity δ_0 such that

$$K_M K_B K_\Theta \delta_0 < 1.$$

Remark 1.2.3. Condition (A3) with uniform estimate (1.18) is guaranteed with the choice of positive semi-definite weight matrices c^j ($j = 0, 1, \ldots, m$) in BTMs (1.1). In this case, we have $K_M \leq 1$. To control boundedness of moments and an appropriate constant K_M for invertible matrices M, it also suffices to take uniformly bounded weights c^0 and vanishing c^j for $j = 1, 2, \ldots, m$ together with sufficiently small step sizes h. Assumption (A5) ensures that the implicit expressions of Y are well-defined, together with the finiteness of some moments and Hölder-continuity (i.p. a guarantee of local resolution).

1.2.2 Rate of Mean Consistency

For simplicity, consider BTMs (1.1) with autonomous implicitness matrices $\Theta \in \mathbb{R}^{d\times d}$ (i.e. Θ is independent of time-variable n).

Theorem 1.2.1 (Mean Consistency of BTMs with Rate $r_0 \geq 1.5$). *Assume that the assumptions (A0)–(A5) are satisfied.*
Then, the BTMs (1.1) with autonomous implicitness matrices $\Theta \in \mathbb{R}^{d\times d}$ and nonrandom step sizes $h_n \leq \delta_0 < 1$ are mean consistent with worst case rate $r_0 \geq 1.5$.

Remark 1.2.4. The proof is based on auxiliary Lemmas 1.2.1 and 1.2.2 as stated and proved below.

Proof. First, rewrite the one-step representations of X and Y in integral form to

$$X_{s,x}(t) = x + \sum_{j=0}^{m} \int_s^t b^j(u, X(u))dW^j(u)$$

$$Y_{s,x}(t) = x + M_{s,x}^{-1}(t)[\Theta a(t, Y_{s,x}(t)) + (I-\Theta)a(s,x)] \int_s^t du$$
$$+ M_{s,x}^{-1}(t) b^j(s,x) \int_s^t dW^j(u).$$

Notice that

$$Y_{s,x}(t)$$
$$= x + \int_s^t a(s,x)du + (M_{s,x}^{-1}(t) - I)\int_s^t a(s,x)du + \sum_{j=1}^{m}\int_s^t b^j(s,x)dW^j(u)$$
$$+ M_{s,x}^{-1}(t)\Theta\int_s^t [a(t, Y_{s,x}(t)) - a(s,x)]du + \sum_{j=1}^{m}(M_{s,x}^{-1}(t) - I)\int_s^t b^j(s,x)dW^j(u).$$

Second, subtracting both representations gives

$$
\begin{aligned}
&X_{s,x}(t) - Y_{s,x}(t) \\
&= \int_s^t [a(u, X(u)) - a(s,x)]du + \sum_{j=1}^{m} \int_s^t [b^j(u, X(u)) - b^j(s,x)]dW^j(u) \\
&\quad + (M_{s,x}^{-1}(t) - I) \int_s^t a(s,x)du + M_{s,x}^{-1}(t)\Theta \int_s^t [a(t, Y_{s,x}(t)) - a(s,x)]du \\
&\quad + \sum_{j=1}^{m} (M_{s,x}^{-1}(t) - I) \int_s^t b^j(s,x)dW^j(u).
\end{aligned}
$$

Recall that the above involved stochastic integrals driven by W^j form martingales with vanishing first moment.
Third, pulling the expectation $\mathbb{E}$ over the latter identity and applying triangle inequality imply that

$$
\begin{aligned}
&||\mathbb{E}[X_{s,x}(t) - Y_{s,x}(t)]||_d \\
&\leq \int_s^t \mathbb{E}||a(u, X(u)) - a(s,x)||_d du + \mathbb{E}||M_{s,x}^{-1}(t) - I||_d \int_s^t ||a(s,x)||_d du \\
&\quad + \mathbb{E}[||M_{s,x}^{-1}(t)\Theta||_d \int_s^t ||a(t, Y_{s,x}(t)) - a(s,x)||_d du] \\
&\quad + \sum_{j=1}^{m} \mathbb{E}[||M_{s,x}^{-1}(t) - I|| \cdot || \int_s^t b^j(s,x)dW^j(u)||_d] \\
&\leq L_a \int_s^t [|u-s|^{1/2} + (\mathbb{E}||X_{s,x}(u) - x||_d^2)^{1/2}]du \\
&\quad + K_M(t-s) \sum_{j=0}^{m} ||c^j(s,x)a(s,x)||_d \mathbb{E}[| \int_s^t dW^j(u)|] \\
&\quad + K_M(t-s)||\Theta||L_a[|t-s|^{1/2} + (\mathbb{E}||Y_{s,x}(t) - x||_d^2)^{1/2}].
\end{aligned}
$$

Note that we used the facts that

$$
M_{s,x}^{-1}(t) - I = -M_{s,x}^{-1}(t) \sum_{k=0}^{m} c^k(s,x) \left| \int_s^t dW^k(u) \right|
$$

and

$$
\mathbb{E}\left[\sum_{k=0}^{m} \sum_{j=1}^{m} M_{s,x}^{-1}(t) c^k(s,x) b^j(s,x) \left| \int_s^t dW^k(u) \right| \int_s^t dW^j(v) \right] = 0
$$