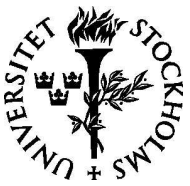
KUNGL
TEKNISKA
HÖGSKOLAN



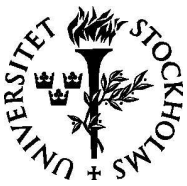
Numerical Complexity Analysis of Weak Approximation of Stochastic Differential Equations

Raúl Tempone Olariaga

TRITA-NA-0220

Doctoral Dissertation
Royal Institute of Technology
Department of Numerical Analysis and Computer Science

KUNGL
TEKNISKA
HÖGSKOLAN



Numerical Complexity Analysis of Weak Approximation of Stochastic Differential Equations

Raúl Tempone Olariaga

Akademisk avhandling som med tillstånd av Kungl Tekniska Högskolan framlägges till offentlig granskning för avläggande av teknisk doktorsexamen Fredagen den 11 oktober 2002 kl 10.00 i D3, Lindstedtsvägen 5, Entreplan, Tekniska Högskolan, Stockholm.

ISBN 91-7283-350-5, TRITA-NA-0220, ISSN 0348-2952, ISRN KTH/NA/R-02/20,

Royal Institute of Technology, Stockholm 2002

Akademisk avhandling som med tillstånd av Kungl Tekniska Högskolan framläges till offentlig granskning för avläggande av teknisk doktorsexamen Fredagen den 11 oktober 2002 kl 10.00 i D3, Lindstedtsvägen 5, Entreplan, Tekniska Högskolan, Stockholm.

ISBN 91-7283-350-5

TRITA-NA-0220
ISSN 0348-2952
ISRN KTH/NA/R-02/20

KTH, Stockholm 2002

Abstract

The thesis consists of four papers on numerical complexity analysis of weak approximation of ordinary and partial stochastic differential equations, including illustrative numerical examples. Here by numerical complexity we mean the computational work needed by a numerical method to solve a problem with a given accuracy. This notion offers a way to understand the efficiency of different numerical methods.

The first paper develops new expansions of the weak computational error for Itô stochastic differential equations using Malliavin calculus. These expansions have a computable leading order term in a posteriori form, and are based on stochastic flows and discrete dual backward problems. Beside this, these expansions lead to efficient and accurate computation of error estimates and give the basis for adaptive algorithms with either deterministic or stochastic time steps. The second paper proves convergence rates of adaptive algorithms for Itô stochastic differential equations. Two algorithms based either on stochastic or deterministic time steps are studied. The analysis of their numerical complexity combines the error expansions from the first paper and an extension of the convergence results for adaptive algorithms approximating deterministic ordinary differential equations. Both adaptive algorithms are proven to stop with an optimal number of time steps up to a problem independent factor defined in the algorithm. The third paper extends the techniques to the framework of Itô stochastic differential equations in infinite dimensional spaces, arising in the Heath Jarrow Morton term structure model for financial applications in bond markets. Error expansions are derived to identify different error contributions arising from time and maturity discretization, as well as the classical statistical error due to finite sampling.

The last paper studies the approximation of linear elliptic stochastic partial differential equations, describing and analyzing two numerical methods. The first method generates iid Monte Carlo approximations of the solution by sampling the coefficients of the equation and using a standard Galerkin finite elements variational formulation. The second method is based on a finite dimensional Karhunen-Loève approximation of the stochastic coefficients, turning the original stochastic problem into a high dimensional deterministic parametric elliptic problem. Then, a deterministic Galerkin finite element method, of either h or p version, approximates the stochastic partial differential equation. The paper concludes by comparing the numerical complexity of the Monte Carlo method with the parametric finite element method, suggesting intuitive conditions for an optimal selection of these methods.

2000 Mathematics Subject Classification. Primary 65C05, 60H10, 60H35, 65C30, 65C20; Secondary 91B28, 91B70.

Keywords and phrases. Adaptive methods, a posteriori error estimates, stochastic differential equations, weak approximation, Monte Carlo methods, Malliavin Calculus, HJM model, option price, bond market, stochastic elliptic equation, Karhunen-Loève expansion, perturbation estimates, numerical complexity

ISBN 91-7283-350-5 • TRITA-NA-0220 • ISSN 0348-2952 • ISRN KTH/NA/R-02/20

Acknowledgments

First I wish to thank my advisor, Anders Szepessy, for all his support, both enthusiastic and continuous, and for his enlightening guidance during my PhD studies. Without his help this dissertation could hardly be written.

I am also grateful to Georgios Zouraris and Kyoung-Sook Moon for many fruitful discussions, scientific and others, and for their nice company at different scientific meetings.

All my good friends at NADA have contributed to make my stay a pleasant one, and I am thankful for that. Among them, I wish to particularly thank Jesper Oppelstrup and Ingrid Melinder. I would like to thank Germund Dalhquist for sharing his exciting views of numerical analysis and finding time for scientific discussions. Special thanks also to my friends Ulf and Christer Andersson, for their infinite patience and invaluable computer consulting.

Ivo Babuška arranged the financial support for my two visits to the Texas Institute for Computational and Applied Mathematics (TICAM) and gave me the opportunity to learn from him there. Many thanks to him and to the people at TICAM for their warm hospitality.

Tomas Björk introduced me to mathematical finance. Here goes my appreciation for sharing his expertise in the area.

Last but not least, I want to specially thank my dear wife Roxana for all her love and patience during these years.

This work was financially supported by the Swedish council for Engineering Science (TFR), grant # 22-148, UdelaR and UdeM in Uruguay and by the National Science Fundation with the grant DMS-9802367. All these grants are gratefully acknowledged.

Contents

List of Papers

Paper I: A. Szepessy, R. Tempone and G. Zouraris. Adaptive weak approximation of stochastic differential equations. *Communications on Pure and Applied Mathematics*, 54,(10): 1169–1214, 2001.

Paper II: K-S. Moon, A. Szepessy, R. Tempone and G. Zouraris. Convergence Rates for Adaptive Weak Approximation of Stochastic Differential Equations.
Presented in the following scientific meetings:
First SIAM-EMS Conference "Applied Mathematics in our Changing World" in Berlin on September 2-6, 2001.
"Third International Conference on Applied Mathematics for Industrial Flow Problems" Lisbon, Portugal, on April 17-20 2002.

Paper III: T. Björk, A. Szepessy R. Tempone and G. Zouraris. Monte Carlo Euler approximation of HJM term structure financial models.
Presented in the following scientific meetings:
Stochastic Numerics 2001 at ETH, Zurich, Switzerland. February 19 - 21, 2001.
First SIAM-EMS Conference "Applied Mathematics in our Changing World" in Berlin on September 2-6, 2001.
The 2nd World Congress of the Bachelier Finance Society at Crete, June 12-15, 2002.

Paper IV: I. Babuška, R. Tempone and G. Zouraris. Galerkin finite element approximations of stochastic elliptic partial differential equations.
Presented in "Current and future trends in numerical pde's: Where is the field, and where is it going?" February 8-9, 2002 Texas Institute for Computational and Applied Mathematics, The University of Texas at Austin, Austin, Texas, USA.

A Roxana.

Chapter 1

Introduction

This work studies numerical methods for weak approximation of stochastic differential equations. In particular, it focuses on the numerical complexity of different numerical methods. Here by numerical complexity we mean the computational work needed by a numerical method to solve a problem with a given accuracy. This work first considers adaptive numerical methods for weak approximation of Itô Stochastic Differential Equations (SDEs) and then analyzes Galerkin finite element approximations for Stochastic Partial Differential Equations (SPDEs) that are linear and elliptic.

SDEs are often part of mathematical models that describe the evolution of dynamical systems under uncertainty. A mathematical model establishes mathematical relations between the relevant variables of a given system. For example, a differential equation modeling the temperature of a hot metal surface subject to water cooling, describes the relation between the given initial condition –the initial temperature–, the flux function –which tells how the heat is convected in the system–, and the final value –the final temperature we want to know–. The purpose of a mathematical model is to predict the outcome of events –which can be past, present or future–, for example the result of a certain physical experiment, and possibly take advantage of that knowledge, since an accurate mathematical model may be used as a basic tool to control the outcomes.

Mathematical models can be deterministic or stochastic. The first case arises when the data and the relations described in the model are deterministic, like in the case of an ordinary differential equation with fixed data, whereas in the second either the data or the relations between the variables are stochastic. As an example of this, consider an ordinary differential equation, whose initial value is not deterministic, but follows a given probability distribution. Another example is to consider a perturbation of an ordinary differential equation, where the evolution itself is affected by some "noise". Setting a formal description of the this intuitive

notion leads to the concept of Stochastic Differential Equations. The field of applications is quite wide, e.g. it comprises ground water flow and financial markets [KP92, Øks98].

Uncertainty comes basically from two sources, namely the lack of complete information about the dynamics of the system to model, or the fact that, for fixed data, the system does not always offer the same outcome. As an example of the second case, when rolling a fair dice a sufficiently large number of times, we tend to observe that all the values appear in similar proportions in the outcomes. It is possible then to use this statistical information within a probability model to answer questions related to dice games. This second step is related with the approach pursued here, that is, we shall assume that the stochastic model has been properly identified by some statistical procedure and *is given*, and then try to compute some related quantities. The need to compute expected values or averages –functionals– depending on the solution of an SDE will guide us towards the notion of *weak convergence*, as opposite to *strong convergence*, where good approximation of realization paths is required.

The discretization of an SDE can be more subtle than for ODEs, for example forward and backward differences do not in general converge to the same limit. Therefore, the model must also include information on the discretization to be used.

Numerical methods offer approximate computable solutions to mathematical problems, and are usually applied when the exact solution is either unknown or its computation is costly or involved. In particular, *adaptive numerical methods* aim for efficient use of computational resources by trying to minimize the degrees of freedom in the numerical discretizations, as well as to provide accurate estimates of the different sources of error present in the computations, like the time discretization error in the solution of an ordinary differential equation. Efficient adaptive numerical methods rely on *a posteriori* information, i.e. information offered by the computable numerical solution, both to estimate the error present in the numerical solution and to apply a refinement criterion when adding degrees of freedom to a given discretization. On the other hand, *a priori* information, i.e. information about the unknown and usually non computable exact solution, is of qualitative kind, e.g. provides smoothness properties of the exact solution, and may be used to prove convergence of numerical approximations, to identify the order of such convergence, as well as to select an appropriate numerical method [EEHJ95, EEHJ96].

Regarding applications, mathematical finance is an area where stochastic modeling with SDEs has obtained a sound success, in particular when dealing with contingent claims pricing theory. A *derivative product* or a *contingent claim* is a financial contract whose value depends on a risk factor, also known as *the underlying*, such as the price of a bond, commodity, currency, share, a yield or rate of interest, an index of prices or yields, etc.

These contracts are also known as "derivatives", for short, and are common in financial markets. The application of derivatives is increasing, consider for example the case of energy derivatives [Boh98, EL00] currently traded in many new regional

markets, arising from deregulation of former national monopolies. A simple example of a derivative, is the so called *European Call Option*, which gives to its owner the right, but not the obligation, to buy the underlying asset at the previously agreed-upon price on the expiration date [Hul93]. The usual valuation method assumes that the financial markets are efficient, that is, that there is no opportunity of making riskless profits, or in other words, that there exist *no arbitrage opportunity in the market*. This assumption leads to a consistency requirement between the price of the underlying, e.g. a stock price, which can be observed directly in the market, and a fair price for a related derivative product, e.g. the call option introduced above. Mathematically this consistency relation can be expressed by the existence of a probability measure Q such that given today's date and today's stock price, the price of the derivative is the expectation *under* Q of its discounted final payoff. The relevant point is that the expectations must be taken under Q, which is known as a *martingale measure*, and *not* under the objective probability measure [BR96, Bjö98].

After the celebrated work of Black and Scholes [BS73], stochastic differential equations have been playing a major role in financial applications. Black and Scholes' model can be used to fit observed data through implied quantities, and the related valuation formula can be interpreted as a nonlinear interpolation procedure to estimate derivative prices. Even though few of the model's assumptions are fully respected in practice, e.g. constant volatility and constant riskless interest rate, the model is quite robust, specially for relatively short maturity options. However, when the life of the option becomes larger, extensions of the Black and Scholes model, e.g. allowing stochastic volatility, are of practical use to explain the so called *volatility smile* effect observed in the market [FSP00, SP99, WO97]. On the other hand, since only relatively few stochastic differential equations have explicit solutions, as the financial models get more and more refined the need for deeper understanding and better numerical methods increases.

Thanks to Kolmogorov's stochastic representation formulae, see [KS88], numerical methods for weak approximation of SDEs can be based either on the numerical solution of a Kolmogorov backward partial differential equation, see [Bjö94a, Bjö94b], using finite differences schemes [WHD95], [Wil98], or the finite element method [BS94], or by the time discretization of the SDE and the computation of sample averages by the Monte Carlo Euler method, see [KP92]. The convergence properties of finite difference schemes and the finite element method make them the best tools whenever the dimension of the given system of SDEs is low, say less or equal than four, since their computational cost increases exponentially with such dimension. On the other hand, the computational cost of Monte Carlo methods is only polynomial in the dimension of the SDE system, making them a feasible alternative to compute with large systems of SDEs. Tree methods [CRR79] are popular and have pedagogical advantages. They may be thought of as a special case of explicit finite difference schemes, although they are non optimal, i.e. with the same amount of computational work there exist other finite difference schemes with better convergence properties.

This work uses the Euler Monte Carlo method for weak approximation of SDEs, developing a posteriori error approximations proposing and analyzing related adaptive numerical methods for weak approximation that are well suited to solve problems with systems of SDEs.

Stochastic Partial Differential Equations (SPDEs) are also used to describe the behavior of systems under uncertainty. Due to the great development in computational resources and scientific computing techniques, more mathematical models can be solved efficiently. Ideally, these techniques could be used to solve many classical partial differential equations (PDEs) to high accuracy. However, in many cases, the information available to solve a given problem is far from complete. This is the case when solving a partial differential equation whose coefficients depend on material properties that are known to some accuracy. The same may occur with its boundary conditions, and even with the geometry of its domain, see for example the work [BCa, BCb]. Naturally, since the current engineering trends are toward more reliance on computational predictions, the need for assessing the level of accuracy in the results grows accordingly. More than ever, the goal then becomes to represent and propagate the uncertainties from the available data to the desired result through our partial differential equation. By uncertainty we mean either intrinsic variability of physical quantities or simply lack of knowledge about some physical behavior, cf. [Roa98]. If variability is interpreted as randomness then naturally we can apply probability theory. To be fruitful, probability theory requires considerable empirical information about the random quantities in question, generally in the form of probability distributions or their statistical moments. Uncertainties may arise at different levels. They could appear in the mathematical model, e.g. if we are not sure about the linear behavior of some material, or in the variables that describe the model, e.g. if the linear coefficient that describes the material is not completely known. Here we shall discuss the second alternative, and use a probabilistic description for the coefficient variability, leading us to the study of stochastic partial differential equations.

Regarding the approximation of SPDEs, this thesis describes and analyzes two numerical methods for a linear elliptic problem with stochastic coefficients and homogeneous Dirichlet boundary conditions. The first method generates iid approximations of the solution by sampling the coefficients of the equation and using a standard Galerkin finite elements variational formulation. The Monte Carlo method then uses these approximations to compute corresponding sample averages. The second method is based on a finite dimensional approximation of the stochastic coefficients, turning the original stochastic problem into a deterministic parametric elliptic problem. A Galerkin finite element method, of either h or p version, then approximates the corresponding deterministic solution yielding approximations of the desired statistics. We include a comparison of the computational work required by each method to achieve a given accuracy, the numerical complexity, to illustrate their nature and possible use.

The thesis is organized as follows: Section 1.1 describes the problem of weak approximation of Itô stochastic differential equations and the contributions from

papers I, II and III. Finally, Section 1.2 describes a problem from linear elliptic SPDEs and describes the contribution from paper IV.

1.1 Itô Stochastic Differential Equations

1.1.1 Weak Approximation of SDEs

Let $(\Omega, \mathcal{F}, P)$ be a probability space, where Ω is a set of outcomes, $\mathcal{F}$ is a set of events in Ω, $P : \mathcal{F} \to [0,1]$ is a probability measure; and then let $W : \mathbb{R} \times \Omega \to \mathbb{R}^{\ell_0}$ be a Wiener process on $(\Omega, \mathcal{F}, P)$. On what follows, $\{\mathcal{F}_t^W\}_{t\in[0,T]}$ denotes the natural filtration, i.e. the filter structure of σ-algebras generated by W, or equivalently the filter generated by the random variables $\{W(s) : 0 \le s \le t\}$.

Let $a(t,x) \in \mathbb{R}^d$ and $b^\ell(t,x) \in \mathbb{R}^d$, $\ell = 1, \ldots, \ell_0$, be given drift and diffusion fluxes and consider the Itô stochastic differential equation in $\mathbb{R}^d$

$$\begin{aligned} dX_k(t) =& a_k(t, X(t))dt + \sum_{\ell=1}^{\ell_0} b_k^\ell(t, X(t))dW^\ell(t), \quad k = 1, \ldots, d, \quad t > 0, \\ X_k(0) =& X_{0,k}, \quad k = 1, \ldots, d. \end{aligned} \tag{1.1}$$

A classical reference for SDEs is the book [KS88]. An existence proof for strong solutions of SDEs, based on Piccard iterations and Lipschitz continuity of the drift and diffusion coefficients can be found in e.g. [Øks98], while a description of an alternative proof based on the Euler method can be found in [MSTZ00b].

The weak approximation of SDEs consists in approximating the expectation $E[g(X(T))]$, where $g : \mathbb{R}^d \to \mathbb{R}$ is a given function, T is a given positive number and the stochastic process X is the solution of (1.1) with initial datum $X(0)$. In finance applications the function g can be a discounted payoff function of a T-contingent claim, and the fluxes a, b describe the dynamics of the underlying process, e.g. a vector of stock values X. Figure 1.1 shows a simple example of such a problem, with $g(x) = \max(X - 0.9, 0)$ corresponding to the payoff diagram of a call option with strike price 0.9, maturity time $T = 1$ and an underlying process that follows a geometric Brownian motion, in this case $dX(t) = \frac{X(t)}{10}dt + \frac{X(t)}{5}\, dW(t)$, with initial condition $X(0) = 1$. The functional to compute is then

$$E[g(X(1))] = \int_{\mathbb{R}} \max(x - 0.9, 0)p(1, x; 0, X_0)dx, \tag{1.2}$$

where $p(1, \cdot; 0, X_0)$ is the probability density of $X(1)$.

A first step towards the development of numerical solutions of the weak approximation problem is the forward Euler method, which is a time discretization of

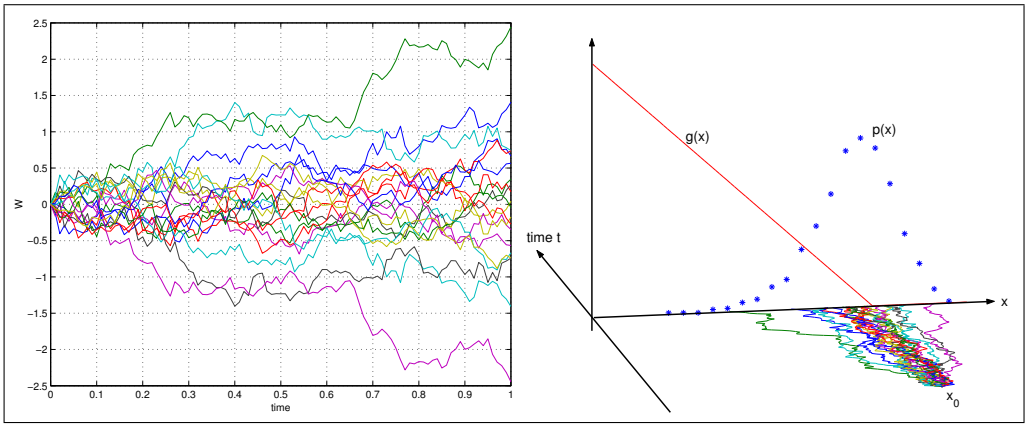


Figure 1.1. Weak approximation example. Left: Realizations for the Wiener process, $\Delta t = 0.01$. Right: Realizations of the process X(t), the function g and a final sample density $p(1, x; 0, X_0)$ corresponding to $M = 1000$ realizations.

(1.1). Consider the time nodes $0 = t_0 < t_1 < \cdots < t_N = T$ and define the discrete time stochastic process $\bar{X}$ by

$$\begin{aligned} \bar{X}(t_{n+1}) &= \bar{X}(t_n) + a(t_n, \bar{X}(t_n))\Delta t_n + \sum_{\ell=1}^{\ell_0} b^\ell(t_n, \bar{X}(t_n))\Delta W_n^\ell, \ 0 \le n \le N-1 \\ \bar{X}(0) &= X_0. \end{aligned} \tag{1.3}$$

Even though a realization of $\bar{X}(t_n)$ is computable, the expectation $E[g(\bar{X}(T)]$ is in general not; however, $E[g(\bar{X}(T)]$ can be approximated by a sample average of M independent realizations, $\frac{1}{M}\sum_{j=1}^{M} g(\bar{X}(T;\omega_j))$, which is the basis of Monte Carlo methods [KP92].

Therefore, the exact computational error, $\mathcal{E}_C$, naturally separates into the two parts

$$\begin{aligned} \mathcal{E}_C &\equiv E[g(X(T))] - \frac{1}{M}\sum_{j=1}^{M} g(\bar{X}(T;\omega_j)) \\ &= E\big[g(X(T)) - g(\bar{X}(T))\big] + \big[E[g(\bar{X}(T))] - \frac{1}{M}\sum_{j=1}^{M} g(\bar{X}(T;\omega_j))\big] \equiv \mathcal{E}_T + \mathcal{E}_S, \end{aligned} \tag{1.4}$$

where the first term, $\mathcal{E}_T \equiv E\big[g(X(T)) - g(\bar{X}(T))\big]$, is the time discretization error, and the second, $\mathcal{E}_s \equiv \big[E[g(\bar{X}(T))] - \frac{1}{M}\sum_{j=1}^{M} g(\bar{X}(T;\omega_j))\big]$, is the statistical error. The time steps for the trajectories $\bar{X}$ are determined from statistical approximations of the time discretization error $\mathcal{E}_T$. The number of realizations, M of $\bar{X}$, are

determined from the statistical error $\mathcal{E}_S$. Therefore, the number of realizations can be asymptotically determined by the Central Limit Theorem

$$\sqrt{M}\mathcal{E}_S \rightharpoonup \chi,$$

where the stochastic variable χ has the normal distribution, with mean zero and variance var$[g(\bar{X}(T))]$. The objective here is to choose the time nodes, which may be different for different realizations of W,

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

and the number of realizations, M, so that the absolute value of the computational error is below a given tolerance, $|\mathcal{E}_C| \leq \text{TOL}$, with probability close to one and with as few time steps and realizations as possible.

Other aspects of the use of the Euler method for the weak approximation of SDEs have been addressed before. Milstein [Mil78] proved that the weak order of the Euler method is 1, i.e. that for uniform deterministic time steps $\Delta t = \frac{T}{N}$, $E[g(X(T)) - g(\bar{X}(T))] = \mathcal{O}\left(\frac{1}{N}\right)$, where N is the number of time steps. Later, Talay and Tubaro [TT90] proved that for uniform deterministic time steps there is an a priori expansion

$$E\big[g(X(T)) - g(\bar{X}(T))\big] = \int_0^T \frac{T}{N} E[\Psi(s, X(s))]ds + \mathcal{O}(\frac{1}{N^2}),$$

where

$$\begin{aligned}\Psi(t,x) \equiv & \tfrac{1}{2}(a_k a_n \partial_{kn} u)(t,x) + (a_i d_{jk} \partial_{ijk} u)(t,x) + \tfrac{1}{2}(d_{ij} d_{kn} \partial_{ijkn} u)(t,x) \\ & + \tfrac{1}{2}\frac{\partial}{\partial t}u(t,x) + (a_i \frac{\partial}{\partial t}\partial_i u)(t,x) + (d_{ij}\frac{\partial}{\partial t}\partial_{ij} u)(t,x),\end{aligned}$$

is based on the definition of the conditional expectation

$$u(t,x) \equiv E[g(X(T))|\ X(t) = x]$$

and the following notation

$$\begin{aligned} d_{ij} &\equiv \frac{1}{2} b_i^\ell b_j^\ell, \\ \partial_k &\equiv \frac{\partial}{\partial x_k}, \\ \partial_{ki} &\equiv \frac{\partial^2}{\partial x_k \partial x_i}, \\ &\vdots \end{aligned}$$

with the summation convention, i.e., if the same subscript appears twice in a term, the term denotes the sum over the range of this subscript, e.g.

$$c_{ik}\partial_k b_j \equiv \sum_{k=1}^{d} c_{ik}\partial_k b_j.$$

For a derivative ∂_α the notation $|\alpha|$ is its order. Kloeden and Platen [KP92] extended the results of Talay and Tubaro on the existence of leading order error expansion in a priori form, for first and second order schemes, to general weak approximations of higher order. Bally and Talay [BT95, BT96] extended the proof to the case where the payoff function g is not smooth, using Malliavin calculus [Nua95]. This expansion motivates the use of Richardson's extrapolation for the development of higher order methods. The case of killed diffusions, e.g. arising in the computation of barrier options, where the distribution of X is not absolutely continuous with respect to the Lebesgue measure, was analyzed by Gobet [Gob00].

An introduction to numerical approximation of SDEs and an extensive review of the literature can be found in the inspiring book by Kloeden and Platen [KP92], including information about the construction and the analysis of the convergence order for higher order methods, either implicit or explicit.

Asymptotical optimal adapted adaptive methods for strong approximation of stochastic differential equations, are analyzed in [HMGR00] and [MG00], which include the hard problem to obtain lower error bounds for any method based on the number of evaluations of W and requires roughly the L^2norm in time of the diffusion $\max_i d_{ii}$ to be positive pathwise. The work [GL97] treats a first study on strong adaptive approximation.

1.1.2 Overview of Paper 1

The main result is new expansions of the computational error, with computable leading order term in a posteriori form, based on stochastic flows and discrete dual backward problems. The expansions lead to efficient and accurate computation of error estimates. In the first simpler expansion, the size of the time steps Δt_n may vary in time but they are deterministic, i.e. the mesh is fixed for all samples. This is useful for solutions with singularities, or approximate singularities, at deterministic times or for problems with small noise. The second error expansion uses time steps which may vary for different realizations of the solution $\bar{X}$. Stochastic time steps are advantageous for problems with singularities at random times. Stochastic time steps use Brownian bridges and require more work for a given number of time steps. The optimal stochastic steps depend on the whole solution $\bar{X}(t),\ 0 < t < T$, and in particular the step $\Delta t(t)$ at time t depends also on $W(\tau),\ \tau > t$. In stochastic analysis the concept adapted to W means that a process at time t only depends on events generated by $\{W(s),\ s < t\}$. In numerical analysis an adaptive method means that the approximate solution is used to control the error, e.g. to determine the time steps. Our stochastic steps are in this sense adaptive non adapted, since $\Delta t(t)$ depends slightly on $W(\tau),\ \tau > t$.

The number of realizations needed to determine the deterministic time steps is asymptotically at most $\mathcal{O}(\mathrm{TOL}^{-1})$, while the number of realizations for the Monte Carlo method to approximate $E[g(\bar{X}(T))]$ is $\mathcal{O}(\mathrm{TOL}^{-2})$. Therefore, the additional work to determine optimal deterministic time steps becomes negligible as the error tolerance tends to zero.

Efficient adaptive time stepping methods, with theoretical basis, use a posteriori error information, since the a priori knowledge usually cannot be as precise as the a posteriori. This work develops adaptive time stepping methods by proving in Theorems 2.2 and 3.3 error estimates of $\mathcal{E}_T$ with leading order terms in computable a posteriori form. Theorem 2.2 uses deterministic time steps, while Theorem 3.3 also holds for stochastic time steps, which are not adapted.

The main new idea here is the efficient use of stochastic flows and dual functions to obtain the error expansion with computable leading order term in Theorems 2.2 and 3.3, including also non adapted adaptive time steps. The use of dual functions is standard in optimal control theory and in particular for adaptive mesh control for ordinary and partial differential equations, see [BMV83], [Joh88], [JS95], [EEHJ96], and [BR96]. The authors are not aware of other error expansions in a posteriori form or adaptive algorithms for weak approximation of stochastic differential equations. In particular error estimates with stochastic non adapted time steps seem to not have been studied before.

Theorem 2.2 describes a computable error expansion, with deterministic time steps, to estimate the computational error,

$$E[g(X(T)) - g(\bar{X}(T))] \simeq \sum_n E[\rho(t_n, \omega)](\Delta t_n)^2 \tag{1.5}$$

where $\rho(t_n, \omega)\Delta t_n$ is the corresponding error density function. Section 3 proves in Theorem 3.3 an analogous error expansion

$$E[g(X(T)) - g(\bar{X}(T))] \simeq E[\sum_n \rho(t_n, \omega)(\Delta t_n)^2] \tag{1.6}$$

which can be used also for stochastic time steps. The leading order terms of the expansion have less variance compared to the expansion in Theorem 2.2, but use upto the third variation, which requires more computational work per realization.

The focus in the paper is on computable error estimates for weak convergence of stochastic differential equations. The technique used here is based on the transition probability density and Kolmogorov's backward equation, which was developed in [SV69] and [SV79] to analyze uniqueness and dependence on initial conditions for weak solutions of stochastic differential equations. The analogous technique for deterministic equations was introduced in [Grö67] and [Ale61].

1.1.3 Overview of Paper 2

Convergence rates of adaptive algorithms for weak approximations of Itô stochastic differential equations are proved for the Monte Carlo Euler method.

Here the focus is on the adaptivity procedures, and we derive convergence rates of two algorithms including dividing and merging of time steps, with either stochastic or deterministic time steps. The difference between the two algorithms is that the stochastic time steps may use different meshes for each realization, while the deterministic time steps use the same mesh for all realizations. The construction and

the analysis of the adaptive algorithms are inspired by the related work [MSTZ00a], on adaptive algorithms for deterministic ordinary differential equations, and use the error estimates from [STZ01]. The main step in the extension is the proof of the almost sure convergence of the error density. Both adaptive algorithms are proven to stop with optimal number of steps up to a problem independent factor defined in the algorithm.

There are two main results on efficiency and accuracy of the adaptive algorithms described in Section 3. In view of accuracy with probability close to one, the approximation errors in (1.4) are asymptotically bounded by the specified error tolerance times a problem independent factor as the tolerance parameter tends to zero. In view of efficiency, both the algorithms with stochastic steps and deterministic steps stop with the optimal expected number of final time steps and optimal number of final time steps respectively, up to a problem independent factor. The number of final time steps is related to the numerical effort needed to compute the approximation. To be more precise, the total work for deterministic steps is roughly $M \cdot N$ where M is the final number of realizations and N is the final number of time steps, since the work to determine the mesh turns out to be negligible. On the other hand, the total work with stochastic steps is on average bounded by $M \cdot E[N_{\text{tot}}]$, where the total number, N_{tot}, of steps including all refinement levels is bounded by $\mathcal{O}(N \log N)$ with N steps in the final refinement; for each realization it is necessary to determine the mesh, which may vary for each realization.

The accuracy and efficiency results are based on the fact that the error density, ρ which measures the approximation error for each interval following (1.5,1.6), converges *almost surely* or *a.s.* as the error tolerance tends to zero. This convergence can be understood by the *a.s.* convergence of the approximate solution, $\bar{X}$, as the maximal step size tends to zero. Once this convergence is established, the techniques to develop the accuracy and efficiency results are similar to those from [MSTZ00a]. Although the time steps are not adapted to the standard filtration generated by W for the stochastic time stepping algorithm, the work [STZ01] proved that the corresponding approximate solution converges to the correct adapted solution X. This result makes it possible to prove the martingale property of the approximate error term with respect to a specific filtration, see Lemma 4.2. Therefore Theorem 4.1 and 4.4 use Doob's inequality to prove the *a.s.* convergence of $\bar{X}$. Similar results of pointwise convergence with constant step sizes, adapted to the standard filtration, are surveyed by Talay in [Tal95].

This work can be easily modified following [MSTZ02] yielding adaptive algorithms with no merging that have several theoretical advantages.

1.1.4 Weak Approximation of an Infinite Dimensional SDE

The extension of the weak approximation problem described in Section 1.1.1 to the infinite dimensional case is here motivated by financial applications, in particular the valuation of contingent claims that have the market interest rate as the underlaying.

Due to the relatively long life span of these products, the interest rate is modeled as a stochastic process, which may be Markovian, like in the case of spot rate models. In the context of interest rates products, the most elementary one is the zero coupon bond with maturity time τ that gives to its owner the right to receive one unit of currency on the date τ. The price at time $t < \tau$ of such a contract is denoted by $P(t, \tau)$.

Since we can choose in principle any possible values of $\tau > t$ we have an unlimited number of bonds. However, due to the assumption of absence of arbitrage in the market, the bond prices corresponding to different maturities are not independent. Mathematical models that take into account the joint evolution of zero coupon bonds with different maturities can use the so called forward rate,

$$f(t, \tau) \equiv -\partial_\tau \log P(t, \tau), \text{ for } t \in [0, \tau] \text{ and } \tau \in [0, \tau_{max}],$$

see [BR96, Bjö94b, Bjö98, Hul93]. Intuitively, we can think of the instantaneous forward rate $f(t, \tau)$ as the risk-free rate at which we can borrow and lend money over the infinitesimal time interval $[\tau, \tau + d\tau]$, provided that the contract is written at time t.

In such models, the absence of arbitrage and friction in the market implies that the drift and the diffusion of the forward rate dynamics must fulfill the Heath-Jarrow-Morton condition [HJM90, HJM92], i.e. that under a risk neutral probability measure, the forward rate $f(t, \tau)$ follows an infinite dimensional Itô stochastic differential equation of the form

$$\begin{aligned} df(t, \tau) &= \sum_{j=1}^{J} \sigma^j(t, \tau) \Big(\int_t^\tau \sigma^j(t, s) ds \Big) dt + \sum_{j=1}^{J} \sigma^j(t, \tau) dW^j(t), \\ f(0, \tau) &= f_0(\tau), \quad \tau \in [0, \tau_{max}]. \end{aligned} \tag{1.7}$$

Here $W(t) = (W^1(t), \ldots, W^J(t))$, is a J-dimensional Wiener process with independent components, and $\sigma^j(t, T)$, $j = 1, \ldots, J$ are stochastic processes, adapted to the filtration generated by W and that may depend on f. Beside this, the initial data $f_0 : [0, \tau_{max}] \to \mathbb{R}$, is a given deterministic C^1 function, obtained from the observable prices $P(0, \tau)$. Figure 1.2 depicts a typical realization of the surface $f(t, \tau)$. Observe that the t-sections, $f(t, \cdot)$, are smooth functions of τ, while the τ-sections, $f(\cdot, \tau)$, are continuous but not smooth.

A basic contract to price is a call option, with exercise time t_{max} and strike price K, on a zero coupon bond with maturity τ_{max}. The price of this option can be written as

$$E\Big[\exp\Big(- \int_0^{t_{max}} f(s, s) ds \Big) \max\Big\{ \exp\Big(- \int_{t_{max}}^{\tau_{max}} f(t_{max}, \tau) d\tau \Big) - K, 0 \Big\} \Big].$$

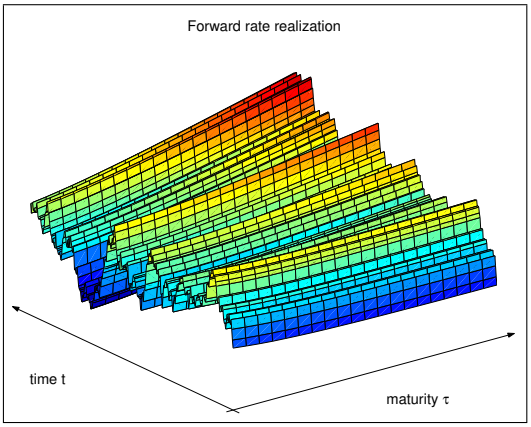


Figure 1.2. Forward rate modeling: a typical realization of $f(t,\tau)$.

With this motivation, we consider computation of the functionals

$$\mathcal{F}(f) \equiv E\Big[F\Big(\int_0^{t_{max}} f(s,s)ds\Big) \quad G\Big(\int_{\tau_a}^{\tau_{max}} Q(f(t_{max},\tau))d\tau\Big) + \int_0^{t_{max}} F\Big(\int_0^s f(v,v)dv\Big)U(f(s,s))ds\Big]$$

The functions $F:\mathbb{R}\to\mathbb{R}$, $G:\mathbb{R}\to\mathbb{R}$, $Q:\mathbb{R}\to\mathbb{R}$, $U:\mathbb{R}\to\mathbb{R}$, and their derivatives up to a sufficiently large order m_0 are assumed to have polynomial growth. Beside this $0 < t_{max} \le \tau_a < \tau_{max}$ are given positive numbers.

Here the aim is to provide a computable approximation of the above functional $\mathcal{F}(f)$. This is accomplished in two steps, namely by a t and τ discretization of (1.7), yielding a numerical solution $\overline{\overline{f}}$, and then the computation of sample averages by the Monte Carlo method.

As in the previous Section, an important issue is to estimate the different sources of computational error. Here the new ingredient is the analysis of the τ discretization error, which appears together with the time discretization error and statistical error introduced in Section 1.1.1.

1.1.5 Overview of Paper 3

This work studies the problem introduced in Section 1.1.4. considering numerical solutions, based on the so called Monte Carlo Euler method, for the price of financial instruments in the bond market, using the Heath Jarrow Morton model for the forward rate [HJM90, HJM92]. The main contribution is to provide rigorous error expansions, with leading error term in computable a posteriori form, offering computational reliability in the use of more complicated HJM multifactor models,

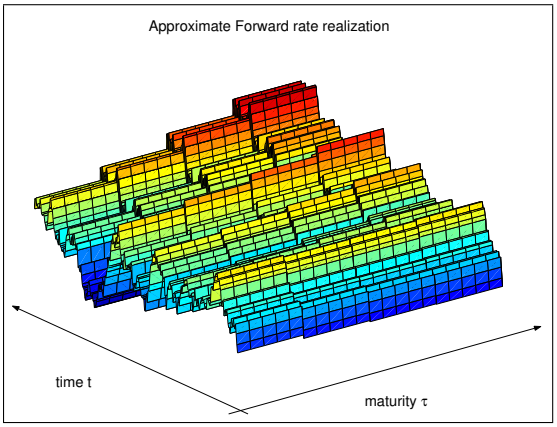


Figure 1.3. Forward rate modeling: a typical realization of $\bar{f}(t,\tau)$, an approximation for $f(t,\tau)$.

where no explicit formula can be found for the pricing of contingent claims. These error estimates can be used to handle *simultaneously* different sources of error, e.g. time discretization, maturity discretization, and finite sampling. To develop error estimates we use a Kolmogorov backward equation in an extended domain and carry out further the analysis in [STZ01], from general weak approximation of Itô stochastic differential equations in $\mathbb{R}^n$, to weak approximation of the HJM Itô stochastic differential equations in infinite dimensional spaces. Therefore the main new ingredient here is to provide error estimates useful for adaptive refinement not only in time t but also in maturity time τ, see Figure 1.3. Another contribution is the removal of the error in the numerical approximations, produced by the representation of the initial term structure in a finite maturity partition. Finally, the formulae to compute sharp error approximations are simplified by exploiting the structure of the HJM model, reducing the work to compute such error estimates.

1.2 Stochastic Partial Differential Equations

Let D be a convex bounded polyhedral domain in $\mathbb{R}^d$ and $(\Omega, \mathcal{F}, P)$ be a complete probability space. Here Ω is the set of outcomes, $\mathcal{F} \subset 2^\Omega$ is the σ-algebra of events and $P : \mathcal{F} \to [0,1]$ is a probability measure. Consider the stochastic linear elliptic boundary value problem: find a stochastic process, $u : \Omega \times \overline{D} \to \mathbb{R}$, such that P-almost everywhere in Ω, or in other words almost surely (a.s.), we have

$$\begin{aligned} -\nabla \cdot (a(\omega,\cdot)\nabla u(\omega,\cdot)) &= f(\omega,\cdot) \quad \text{on} \quad D, \\ u(\omega,\cdot) &= 0 \quad \text{on } \partial D. \end{aligned} \tag{1.8}$$

Here $a, f : \Omega \times D \to \mathbb{R}$ are stochastic processes that are correlated in space. If we denote by $B(D)$ the Borel σ-algebra generated by the open subsets of D, then a, f

are jointly measurable with the σ-algebra $(\mathcal{F} \otimes B(D))$. It is usual to assume that the elliptic operator is bounded and uniformly coercive, i.e.

$$\exists\ a_{\min}, a_{\max} \in (0, +\infty) : a(\omega, x) \in [a_{\min}, a_{\max}], \quad \forall x \in D \quad a.s. \tag{1.9}$$

To ensure regularity of the solution u we assume also that a has a uniformly bounded and continuous first derivative, i.e. there exists a real deterministic constant C such that

$$a(\omega, \cdot) \in C^1(\overline{D}) \quad \text{and} \quad \max_{\overline{D}} |\nabla_x a(\omega, \cdot)| < C \text{ a.s.} \tag{1.10}$$

In addition, the right hand side in (1.8) satisfies

$$\int_\Omega \int_D f^2(\omega, x) dx\, dP(\omega) < +\infty \quad \text{which implies} \quad \int_D f^2(\omega, x) dx < +\infty \ \ a.s. \tag{1.11}$$

Stochastic Sobolev spaces are used to obtain existence and uniqueness results for the solution of (1.8). As in the deterministic space, cf. [BS94], the tools are Hilbert spaces, a suitable weak formulation and then the application of the Lax-Milgram's lemma. All the necessary definitions then have to be extended to the stochastic setting using tensor product of Hilbert spaces, cf. [Lar86] and [ST02], which is a standard procedure.

Definition 1.1. Let H_1, H_2 be Hilbert spaces. The tensor space $H_1 \otimes H_2$ is the completion of formal sums $u(y, x) = \sum_{i=1,\dots,n} v_i(y) w_i(x)$, $\{v_i\} \subset H_1$, $\{w_i\} \subset H_2$, with respect to the inner product $(u, \hat{u})_{H_1 \otimes H_2} = \sum_{i,j} (v_i, \hat{v}_j)_{H_1} (w_i, \hat{w}_j)_{H_2}$.

For example, let us consider two domains, $y \in \Gamma, x \in D$ and the tensor space $L^2(\Gamma) \otimes H^1(D)$, with tensor inner product

$$(u, \hat{u})_{L^2(\Gamma) \otimes H^1(D)} = \int_\Gamma \left(\int_D u(y, x) \hat{u}(y, x) dx \right) dy + \int_\Gamma \left(\int_D \nabla_x u(y, x) \cdot \nabla_x \hat{u}(y, x) dx \right) dy.$$

Thus, if $u \in L^2(\Gamma) \otimes H^k(D)$ then $u(y, \cdot) \in H^k(D)$ a.e. on Γ and $u(\cdot, x) \in L^2(\Gamma)$ a.e. on D. Moreover, we have the isomorphism

$$L^2(\Gamma) \otimes H^k(D) \simeq L^2(\Gamma; H^k(D)) \simeq H^k(D; L^2(\Gamma))$$

with the definitions

$$L^2(\Gamma; H^k(D)) = \left\{ v : \Gamma \times D \to \mathbb{R} \mid v \text{ is jointly measurable and } \int_\Gamma \|v(y, \cdot)\|^2_{H^k(D)} < +\infty \right\},$$

$$H^k(D; L^2(\Gamma)) = \Big\{ v : \Gamma \times D \to \mathbb{R} \mid v \text{ is jointly measurable, } \forall |\alpha| \le k \ \ \exists \ \ \partial_\alpha v \in L^2(\Gamma) \otimes L^2(D) \text{ and}$$
$$\int_D \partial_\alpha v(y, x) \phi(x) dx = (-1)^{|\alpha|} \int_\Gamma \int_D v(y, x) \partial_\alpha \phi(x) dx, \quad \forall \phi \in C_0^\infty(D) \text{ a.e. on } \Gamma \Big\}.$$

Intuitively, a function $v(\omega, x)$ that belongs to the stochastic Sobolev space $\widetilde{W}^{s,q}(D)$ will have its realizations accordingly regular, i.e. $v(\omega, \cdot) \in W^{s,q}(D)$ a.s. We first recall the definition of stochastic weak derivatives. Let $v \in L^2_P(\Omega) \otimes L^2(D)$, then the α stochastic weak derivative of v, $w = \partial_\alpha v \in L^2_P(\Omega) \otimes L^2(D)$, satisfies

$$\int_D v(\omega, x)\partial^\alpha \phi(x)dx = (-1)^{|\alpha|} \int_D w(\omega, x)\phi(x)dx, \quad \forall \phi \in C_0^\infty(D), \text{ a.s.}$$

We shall work with stochastic Sobolev spaces $\widetilde{W}^{s,q}(D) = L^q_P(\Omega, W^{s,q}(D))$ containing stochastic processes, $v : \Omega \times D \to \mathbb{R}$, that are measurable with respect to the product σ-algebra $\mathcal{F} \otimes B(D)$ and equipped with the averaged norms

$$\|v\|_{\widetilde{W}^{s,q}(D)} = E[\|v\|^q_{W^{s,q}(D)}]^{1/q} = E\left[\sum_{|\alpha|\le s} \int_D |\partial^\alpha v|^q dx\right]^{1/q}, \quad 1 \le q < +\infty$$

and

$$\|v\|_{\widetilde{W}^{s,\infty}(D)} = \max_{|\alpha|\le s} \left(\text{ess sup}_{\Omega\times D} |\partial^\alpha v|\right).$$

Observe that if $v \in \widetilde{W}^{s,q}(D)$ then $v(\omega, \cdot) \in W^{s,q}(D)$ a.s. and $\partial^\alpha v(\cdot, x) \in L^q_P(\Omega)$ a.e. on D for $|\alpha| \le s$. Whenever $q = 2$, the above space is a Hilbert space, i.e. $\widetilde{W}^{s,2}(D) = \widetilde{H}^s(D) \simeq L^2_P(\Omega) \otimes H^s(D)$.

Now we recall the definition of weak solutions for (1.8). Consider the bilinear form, $\mathcal{B} : \widetilde{H}^1_0(D) \times \widetilde{H}^1_0(D) \to \mathbb{R}$,

$$\mathcal{B}(v, w) \equiv E\left[\int_D a\nabla v \cdot \nabla w dx\right], \quad \forall v, w \in \widetilde{H}^1_0(D).$$

The standard assumption (1.10) yields both the continuity and the coercivity of $\mathcal{B}$, i.e.

$$|\mathcal{B}(v, w)| \le a_{max} \, \|v\|_{\widetilde{H}^1_0(D)} \, \|w\|_{\widetilde{H}^1_0(D)}, \quad \forall v, w \in \widetilde{H}^1_0(D), \tag{1.12}$$

and

$$a_{min} \quad \|v\|^2_{\widetilde{H}^1_0(D)} \le \mathcal{B}(v, v), \quad \forall v \in \widetilde{H}^1_0(D). \tag{1.13}$$

A direct application of the Lax Milgram's lemma, see [BS94], implies the existence and uniqueness for the solution to the variational formulation: find $u \in \widetilde{H}^1_0(D)$ such that

$$\mathcal{B}(u, v) = \mathcal{L}(v), \quad \forall v \in \widetilde{H}^1_0(D). \tag{1.14}$$

Here $\mathcal{L}(v) \equiv E[\int_D fvdx], \forall v \in \widetilde{H}^1_0(D)$ defines a bounded linear functional since the random field f satisfies (1.11). Moreover, standard arguments from measure theory show that the solution to (1.14) also solves (1.8).

Usually in practical problems the information about the stochastic processes a and f is only limited. For example, we may only have approximations for their

expectations and covariance functions to use in the implementation of a numerical method for (1.8).

The Karhunen-Loève expansion, known also as the Proper Orthogonal Decomposition (POD), is a suitable tool for the approximation of stochastic processes. This expansion is used extensively in the fields of detection, estimation, pattern recognition, and image processing as an efficient tool to approximate random processes.

Now we describe the Karhunen-Loève expansion of a stochastic process. Consider a stochastic process a with continuous covariance function, $Cov[a] : D \times D \to \mathbb{R}$. Besides this, let $\{(\lambda_i, b_i)\}_{i=1}^{+\infty}$ denote the sequence of eigenpairs associated with the compact self adjoint operator that maps

$$f \in L^2(D) \mapsto \int_D Cov[a](x, \cdot) f(x) dx \in L^2(D).$$

The real and non-negative eigenvalues

$$\lambda_1 \geq \lambda_2 \geq \dots$$

satisfy

$$0 \leq \lambda_i \leq \sqrt{\int_D \int_D \left(Cov[a](x_1, x_2)\right)^2 dx_1 \; dx_2}, \quad i = 1, \dots$$

$$\sum_{i=1}^{+\infty} \lambda_i = \int_D Var[a](x) dx,$$

and $\lambda_i \to 0$. The corresponding eigenfunctions are orthonormal, i.e.

$$\int_D b_i(x) b_j(x) dx = \delta_{ij}$$

and by Mercer's theorem, cf. [RSN90] p. 245,

$$\| Cov[a](x, y) - \sum_{n=1}^{N} \lambda_n b_n(x) b_n(y) \|_{L^\infty(D \times D)} \to 0. \tag{1.15}$$

The Karhunen-Loève expansion of the stochastic process a, cf. [Lév92, Yag87a, Yag87b], is

$$a_N(\omega, x) = E[a](x) + \sum_{i=1}^{N} \sqrt{\lambda_i} b_i(x) Y_i(\omega)$$

where $\{Y_i\}_{i=1}^{+\infty}$ is a sequence of uncorrelated real random variables, with mean zero and unit variance. These random variables are uniquely determined by

$$Y_i(\omega) = \frac{1}{\sqrt{\lambda_i}} \int_D (a(\omega, x) - E[a](x)) b_i(x) dx$$

for $\lambda_i > 0$. Then applying (1.15) yields the uniform convergence

$$\sup_{x\in D} E[(a - a_N)^2](x) = \sup_{x\in D} \left(Var[a](x) - \sum_{i=1}^{N} \lambda_i b_i^2(x) \right) \to 0, \quad \text{as} \quad N \to \infty.$$

A standard approach then is to approximate the stochastic coefficients from (1.8) using the Karhunen-Loève expansion and then solve the resulting stochastic partial differential equation.

In this thesis we study the approximation of some statistical moments of the solution from (1.8), e.g. the deterministic function $E[u]$. For example, we are interested in approximating this function using either $L^2(D)$ or $H^1(D)$.

Depending on the structure of the noise that drives an elliptic partial stochastic differential equation, there are different numerical approximations. For example, when the size of the noise is relatively small, a Neumann expansion around the mean value of the equation's operator is a popular alternative. It requires only the solution of standard deterministic partial differential equations, the number of them being equal to the number of terms in the expansion. Equivalently, a Taylor expansion of the solution around its mean value with respect to the noise yields the same result. Similarly, the work [KH92] uses formal Taylor expansions up to second order of the solution but does not study their convergence properties. Recently, the work [BC02] proposed a perturbation method with successive approximations. It also proves that the condition of uniform coercivity of the diffusion is sufficient for the convergence of the perturbation method.

When only the load is stochastic, it is also possible to derive deterministic equations for the moments of the solution. This case was analyzed in [Bab61, Lar86] and more recently in the work [ST02], where a new method to solve these equations with optimal complexity is presented.

On the other hand, the work by Babuška et al. [Deb00, DBO01] and by Ghanem and Spanos [GS91] address the general case where all the coefficients are stochastic. Both approaches transform the original stochastic problem into a deterministic one with higher dimensions, and they differ in the choice of the approximating functional spaces. The work [Deb00] uses finite element to approximate the noise dependence of the solution, while [GRH99, GS91] uses a formal expansion in terms of Hermite polynomials.

Monte Carlo methods are both general and simple to code and they are naturally suited for parallelization. They generate a set of independent identically distributed (iid) approximations of the solution by sampling the coefficients of the equation, using a spatial discretization of the partial differential equation, e.g. by a Galerkin finite element formulation. Then, using these approximations we can compute corresponding sample averages of the desired statistics. The drawback of Monte Carlo methods is their slow rate of convergence. It is worth mentioning that in particular cases their convergence can be accelerated by variance reduction techniques [JCM01] or even the convergence rate improved with Quasi Monte Carlo methods [Caf98, Sob94, Sob98]. Moreover, if the probability density of the

random variable is smooth, the convergence rate of the Monte Carlo method for the approximation of the expected value can be improved, cf. [Nov88, TW98].

Another way to provide a notion of stochastic partial differential equations is based on the Wick product and the Wiener chaos expansion, see [HØUZ96] and [Våg98]. This approach yields solutions in Kondratiev spaces of stochastic distributions which are not the same as those from (1.14). The choice between (1.14) and [HØUZ96] is a modeling decision, based on the physical situation under study. For example, with the Wick product we have $E[a \diamond u] = E[a]E[u]$ regardless of the correlation between a and f, whereas this is in general not true with the usual product. A numerical approximation for Wick stochastic linear elliptic partial differential equations is studied in [The00], yielding a priori convergence rates.

1.2.1 Overview of Paper 4

We describe and analyze two numerical methods for the linear elliptic problem (1.8). Here the aim of the computations is to approximate the expected value of the solution. Since the approximation of the stochastic coefficients a and f by the Karhunen-Loève expansion is in general not exact, we derive related a priori error estimates. The first method generates iid approximations of the solution by sampling the coefficients of the equation and using a standard Galerkin finite elements variational formulation. The Monte Carlo method then uses these approximations to compute corresponding sample averages. More explicitely, we follow:

1. Give a number of realizations, M, and a piecewise linear finite element space on D, X_h^d.

2. For each $j = 1, \ldots, M$ sample iid realizations of the diffusion $a(\omega_j, \cdot)$ and the load $f(\omega_j, \cdot)$ and find a corresponding approximation $u_h(\omega_j, \cdot) \in X_h^d$ such that

$$\int_D a(\omega_j, x)\nabla u_h(\omega_j, x)\nabla\chi(x)dx = \int_D f(\omega_j, x)\chi(x)dx, \quad \forall\chi \in X_h^d.$$

3. Finally use the sample average $\frac{1}{M}\sum_{j=1}^{M} u_h(\omega_j, \cdot)$ to approximate $E[u]$.

Here we only consider the case where $X_h^d \subseteq H_0^1(D)$ is the same for all realizations, i.e. the spatial triangulation is deterministic.

The second method is based on a finite dimensional approximation of the stochastic coefficients, turning the original stochastic problem into a deterministic parametric elliptic problem. In many problems the source of the randomness can be approximated using just a small number of mutually uncorrelated, sometimes mutually independent, random variables. Take for example the case of a truncated Karhunen-Loève expansion described previously.

Assumption 1.2. *Whenever we apply some numerical method to solve* (1.8) *we assume that the coefficients used in the computations satisfy*

$$a(\omega, x) = a(Y_1(\omega), \ldots, Y_N(\omega), x) \quad \text{and} \quad f(\omega, x) = f(Y_1(\omega), \ldots, Y_N(\omega), x) \quad (1.17)$$

where $\{Y_j\}_{j=1}^N$ *are real random variables with mean value zero, unit variance, are mutually independent, and their images,* $\Gamma_{i,N} \equiv Y_i(\Omega)$ *are bounded intervals in* $\mathbb{R}$ *for* $i = 1, \ldots, N$*. Moreover, we assume that each* Y_i *has a density function* $\rho_i : \Gamma_{i,N} \to \mathbb{R}^+$ *for* $i = 1, \ldots, N$*. Use the notations* $\rho(y) = \Pi_{i=1}^N \rho_i(y_i) \quad \forall y \in \Gamma$*, for the joint probability density of* $(Y_1, \ldots, Y_N)$*, and* $\Gamma \equiv \Pi_{i=1}^N \Gamma_{i,N} \subset \mathbb{R}^N$*, for the support of such probability density.*

After making assumption (1.17), we have by Doob-Dynkin's lemma, cf. [Øks98], that u, the solution corresponding to the stochastic partial differential equation (1.8) can be described by just a finite, hopefully small, number of random variables, i.e. $u(\omega, x) = u(Y_1(\omega), \ldots, Y_N(\omega), x)$. Now the goal is to approximate the function $u(y, x)$. The stochastic variational formulation (1.14) now has a deterministic equivalent in the following: find $u \in L^2_\rho(\Gamma) \otimes H^1_0(D)$ such that

$$\int_\Gamma \rho(y) \int_D a(y,x) \nabla u(y,x) \cdot \nabla v(y,x) dx dy = \int_\Gamma \rho(y) \int_D f(y,x) v(y,x) dx dy, \quad \forall\, v \in L^2_\rho(\Gamma) \otimes H^1_0(D). \quad (1.18)$$

A Galerkin finite element method, of either h or p version, then approximates the corresponding deterministic solution yielding approximations of the desired statistics, i.e. we seek $\bar{u}_h \in Z^p_k \otimes X^d_h$, that satisfies

$$\int_\Gamma \rho(y) \int_D a(y,x)\, \nabla \bar{u}_h(y,x) \cdot \nabla \chi(y,x) dx \quad dy = \int_\Gamma \rho(y) \int_D f(y,x) \chi(y,x) dx dy, \quad \forall\, \chi \in Z^p_k \otimes X^d_h. \quad (1.19)$$

Here $Z^p_k \subseteq L^2_\rho(\Gamma)$ is a finite element space that contains tensor products of polynomials with degree less than or equal to p on a mesh with size k. Section 7 explains how to use the tensor product structure to efficiently compute $\bar{u}_h$ from (1.19). Finally, Section 8 uses the a priori convergence rates of the different discretizations to compare the computational work required by (1.16) and (1.19) to achieve a given accuracy in the approximation of $E[u]$, i.e. their numerical complexity, offering a way to understand the efficiency of each numerical method.

Bibliography

[Ale61] V.M. Alekseev. An estimate for the peturbations of the solutions of ordinary differential equations. II. *Vestnik Moskov. Univ. Ser. I Mat. Mech*, 3:3–10, 1961.

[Bab61] I. Babuška. On randomized solutions of Laplace's equation. *Casopis Pest. Mat.*, 1961.

[BCa] I. Babuška and J. Chleboun. Effects of uncertainties in the domain on the solution of Dirichlet boundary value problems in two spatial dimensions. To appear in Num. Math.

[BCb] I. Babuška and J. Chleboun. Effects of uncertainties in the domain on the solution of Neumann boundary value problems in two spatial dimensions. To appear in Math. Comp.

[BC02] I. Babuška and P. Chatzipantelidis. On solving linear elliptic stochastic partial differential equations. *Comput. Methods Appl. Mech. Engrg.*, 191:4093–4122, 2002.

[Bjö94a] T. Björk. Stokastisk kalkyl och kapitalmarknadsteori, del 1 grunderna. Matematiska Institutionen, KTH, Stockholm, 1994.

[Bjö94b] T. Björk. Stokastisk kalkyl och kapitalmarknadsteori, del 2 specialstudier. Matematiska Institutionen, KTH, Stockholm, 1994.

[Bjö98] T. Björk. *Arbitrage Theory in Continuous Time.* Oxford University Press Inc., New York, 1998.

[BMV83] I. Babuška, A. Miller, and M. Vogelius. Adaptive computational methods for partial differential equations. *Adaptive computational methods for partial differential equations*, pages 57–73, 1983. (College Park, Md., 1983), Philadelphia, Pa.

[Boh98] M. Bohlin. Pricing of options with electricity futures as underlying assets. Master's thesis, Royal Institute of Technology, 1998.

[BR96] M. Baxter and A. Rennie. *Financial Calculus: An Introduction to Derivate Pricing*. Cambridge University Press, Cambridge, UK, 1996.

[BS73] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–659, 1973.

[BS94] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer–Verlag, New York, 1994.

[BT95] V. Bally and D. Talay. The Euler scheme for stochastic differential equations: Error analysis with Malliavin calculus. *Math. Comput. Simulation*, 38:35–41, 1995.

[BT96] V. Bally and D. Talay. The law of the Euler scheme for stochastic differential equations. II. convergence rate of the density. *Monte Carlo Methods Appl.*, 2:93–128, 1996.

[Caf98] R. E. Caflisch. Monte Carlo and quasi-Monte Carlo methods. In *Acta numerica, 1998*, pages 1–49. Cambridge Univ. Press, Cambridge, 1998.

[CRR79] J.C. Cox, S.A. Ross, and M. Rubinstein. Option pricing: A simplified approach. *Journal of Financial Economics*, 7:229–263, 1979.

[DBO01] M.-K. Deb, I. M. Babuška, and J. T. Oden. Solution of stochastic partial differential equations using Galerkin finite element techniques. *Comput. Methods Appl. Mech. Engrg.*, 190:6359–6372, 2001.

[Deb00] M.-K. Deb. *Solution of stochastic partial differential equations (SPDEs) using Galerkin method: theory and applications*. PhD thesis, The University of Texas at Austin, 2000.

[EEHJ95] K. Eriksson, D. Estep, P. Hansbo, and K. Johnson. Introduction to adaptive methods for differential equations. *Acta Numerica*, pages 105–158, 1995.

[EEHJ96] K. Eriksson, D. Estep, P. Hansbo, and K. Johnson. *Computational Differential Equations*. Cambridge University Press, 1996.

[EL00] F. Ericsson and M. Lüttgen. Risk managment on the nordic electricity market. Master's thesis, Stockholm School of Economics, January 2000.

[FSP00] J-P Fouque, K. R. Sircar, and G. Papanicolaou. Mean reverting stochastic volatility. *International Journal of Theoretical and Applied Finance*, 3:101–142, 2000.

[GL97] J.G. Gaines and T.J. Lyons. Variable step size control in the numerical solution of stochastic differential equations. *SIAM J. Appl. Math.*, 57:1455–1484, 1997.

[Gob00] E. Gobet. Weak approximation of killed diffusion using Euler schemes. *Stochastic Processes and their applications*, 87:167–197, 2000.

[GRH99] R. Ghanem and J. Red-Horse. Propagation of probabilistic uncertainty in complex physical systems using a stochastic finite element approach. *Phys. D*, 133(1-4):137–144, 1999. Predictability: quantifying uncertainty in models of complex phenomena (Los Alamos, NM, 1998).

[Grö67] W. Gröbner. *Die Lie-Reihen und ihre Anwendungen*. VEB Deutscher Verlag der Wissenschaften, Berlin, 1967.

[GS91] R. G. Ghanem and P. D. Spanos. *Stochastic finite elements: a spectral approach*. Springer-Verlag, New York, 1991.

[HJM90] D. Heath, R. Jarrow, and A. Morton. Bond pricing and the term structure of interest rates: A discrete time approximation. *Journal of Financial and Quantitative Analysis*, 25:419–440, 1990.

[HJM92] D. Heath, R. Jarrow, and A. Morton. Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica*, 60:77–105, 1992.

[HMGR00] N. Hofmann, T. Müller-Gronbach, and K. Ritter. Optimal approximation of stochastic differential equations by adaptive step-size control. *Math. Comp.*, 69(231):1017–1034, 2000.

[HØUZ96] H. Holden, B. Øksendal, J. Ubøe, and T. Zhang. *Stochastic partial differential equations*. Birkhäuser Boston Inc., Boston, MA, 1996. A modeling, white noise functional approach.

[Hul93] J. Hull. *Options, Futures and Other Derivatives*. Prentice Hall, Upper Saddle River, NJ, 1993.

[JCM01] E. Jouini, J. Cvitanić, and M. Musiela, editors. *Option Pricing, Interest Rates and Risk Management*. Cambridge University Press, 2001.

[Joh88] C. Johnson. Error estimates and adaptive time-step control for a class of one-step methods for stiff ordinary differential equations. *SIAM J. Numer. Anal.*, 25:908–926, 1988.

[JS95] C. Johnson and A. Szepessy. Adaptive finite element methods for conservation laws based on a posteriori error estimates. *Comm. Pure Appl. Math.*, 48:199–234, 1995.

[KH92] M. Kleiber and T-D Hien. *The stochastic finite element method*. John Wiley & Sons Ltd., Chichester, 1992. Basic perturbation technique and computer implementation, With a separately available computer disk.

[KP92] P. E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*. Springer-Verlag, Berlin, 1992.

[KS88] I. Karatzas and S.E. Shreve. *Brownian Motion and Stochastic Calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer–Verlag, 1988.

[Lar86] S. Larsen. *Numerical analysis of elliptic partial differential equations with stochastic input data*. PhD thesis, University of Maryland, MD, 1986.

[Lév92] P. Lévy. *Processus stochastiques et mouvement Brownien*. Éditions Jacques Gabay, Paris, 1992.

[MG00] T. Müller-Gronbach. The optimal uniform approximation of systems of stochastic differential equations. Preprint, Freie Universität Berlin, 2000.

[Mil78] G. N. Milstein. A method with second order accuracy for the integration of stochastic differential equations. *Teor. Verojatnost. i Primenen.*, 23(2):414–419, 1978.

[MSTZ00a] K-S. Moon, A. Szepessy, R. Tempone, and G. Zouraris. Adaptive approximation of differential equations based on global and local errors. Technical Report TRITA-NA-0006, NADA, KTH, Stockholm, Sweden, February 2000.

[MSTZ00b] K-S. Moon, A. Szepessy, R. Tempone, and G. Zouraris. Stochastic and partial differential equations with adaptive numerics, 2000. Lecture Notes.

[MSTZ02] K-S. Moon, A. Szepessy, R. Tempone, and G. Zouraris. Convergence rates for adaptive approximation of ordinary differential equations based on global and local errors. 2002.

[Nov88] E. Novak. Stochastic properties of quadrature formulas. *Numer. Math.*, 53(5):609–620, 1988.

[Nua95] D. Nualart. *The Malliavin Calculus and Related Topics*. Probability and its Applications. Springer–Verlag, 1995.

[Øks98] B. Øksendal. *Stochastic differential equations*. Springer-Verlag, Berlin, fifth edition, 1998. An introduction with applications.

[Roa98] P. J. Roache. *Verification and Validation in Computational Science and Engineering*. Hermosa, Albuquerque, New Mexico, 1998.

[RSN90] F. Riesz and B. Sz.-Nagy. *Functional analysis*. Dover Publications Inc., New York, 1990.

[Sob94] I. M. Sobol′. *A primer for the Monte Carlo method*. CRC Press, Boca Raton, FL, 1994.

[Sob98] I. M. Sobol′. On quasi-Monte Carlo integrations. *Math. Comput. Simulation*, 47(2-5):103–112, 1998. IMACS Seminar on Monte Carlo Methods (Brussels, 1997).

[SP99] K. R. Sircar and G. Papanicolaou. Stochastic volatility, smile and asymptotics. *Mathematical Finance*, 6:107–145, 1999.

[ST02] C. Schwab and R.-A. Todor. Finite element method for elliptic problems with stochastic data. Technical report, Seminar for Applied Mathematics, ETHZ, 8092 Zurich, Switzerland, April 2002. Research Report 2002-05.

[STZ01] A. Szepessy, R. Tempone, and G. E. Zouraris. Adaptive weak approximation of stochastic differential equations. *Comm. Pure Appl. Math.*, 54(10):1169–1214, 2001.

[SV69] D.W. Strook and S.R.S. Varadhan. Diffusion processes with continuous coefficients, i and II. *Comm. Pure Appl. Math.*, 22:345–400 and 479–530, 1969.

[SV79] D.W. Strook and S.R.S. Varadhan. *Multidimensional Diffusion Processes*. Grundlehren der Mathematischen Wissenschaften 233. Springer–Verlag, 1979.

[Tal95] D. Talay. *Probabilistic Methods in Applied Physics*, chapter Simulation of Stochastic Differential Systems, pages 54–96. Number 451 in Lecture notes in physics. Springer, 1995. P. Krée ; W. Wedig Eds.

[The00] T. G. Theting. Solving Wick-stochastic boundary value problems using a finite element method. *Stochastics Stochastics Rep.*, 70(3-4):241–270, 2000.

[TT90] D. Talay and L. Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Anal. Appl.*, 8:483–509, 1990.

[TW98] J. F. Traub and A. G. Werschulz. *Complexity and information*. Cambridge University Press, Cambridge, 1998.

[Våg98] G. Våge. Variational methods for PDEs applied to stochastic partial differential equations. *Math. Scand.*, 82(1):113–137, 1998.

[WHD95] P. Wilmott, S.D. Howinson, and J.N Dewynne. *The Mathematics of Financial Derivatives: A Student Introduction*. Cambridge University Press, 1995.

[Wil98] P. Wilmott. *Derivatives : The Theory and Practice of Financial Engineering*. John Wiley & Sons, November 1998.

[WO97] P. Wilmott and A. Oztukel. Uncertain parameters, an empirical stochastic volatility model and confidence limits. *International Journal of Theoretical and Applied Finance*, 1(1):175–189, 1997.

[Yag87a] A. M. Yaglom. *Correlation theory of stationary and related random functions. Vol. I.* Springer-Verlag, New York, 1987. Basic results.

[Yag87b] A. M. Yaglom. *Correlation theory of stationary and related random functions. Vol. II.* Springer-Verlag, New York, 1987. Supplementary notes and references.