KUNGL
TEKNISKA
HÖGSKOLAN

Adaptive Algorithms for Deterministic and Stochastic Differential Equations

Kyoung-Sook Moon

Stockholm 2003

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

Akademisk avhandling som med tillstånd av Kungl Tekniska Högskolan framläges till offentlig granskning för avläggande av teknisk doktorsexamen fredagen den 19 september 2003 kl 10.15 i salongen, biblioteket, Osquars backe 31, Kungl Tekniska Högskolan, Stockholm, Sweden.

ISBN 91-7283-553-2

TRITA-NA-0315
ISSN 0348-2952
ISRN KTH/NA/R--03/15--SE

Universitetsservice US AB, Stockholm 2003

Abstract

Adaptive methods for numerical solutions of differential equations are useful to combine the two goals of good accuracy and efficiency. This thesis contributes to the theoretical understanding of optimal convergence rates of adaptive algorithms. In particular, this work studies stopping, accuracy and efficiency behavior of adaptive algorithms for solving ordinary differential equations (ODEs), partial differential equations (PDEs) and Itô stochastic differential equations (SDEs).

The main ingredient of the adaptive algorithms is an error expansion of the form "Global error $= \sum$ local error $\cdot$ weight + higher order error", with computable leading order terms. The algorithm uses additional computational work to obtain the error estimate because of the weights. However, the approximation of the weights is required to inform where to refine the mesh to achieve optimal efficiency. Based on the a posteriori error expansions with "error indicator := |local error $\cdot$ weight|", the adaptive algorithm subdivides the time steps or elements if the error indicators are greater than TOL/N, and stops if all N time steps or elements have sufficiently small error indicators. Similar algorithms are derived with either stochastic or deterministic time steps for weak approximations of SDEs including stopped diffusions using Monte Carlo Euler method.

There are two main results on efficiency and accuracy of the adaptive algorithms. For accuracy, the approximation errors are asymptotically bounded by the specified error tolerance times a problem independent factor as the tolerance parameter tends to zero. For efficiency, the algorithms decrease the maximal error indicator with a factor, less than 1, or stop with asymptotically optimal number of final time steps or elements. Note that the optimal here refers to error densities of one sign, i.e. possible cancellation of the errors is not taken into account. For a p-th order accurate method, the $L^{\frac{1}{p+1}}$ quasi-norm of the error density is a good measure of the convergence rate for the adaptive approximation, while the number of uniform steps is measured by the larger L^1-norm of the error density.

This analysis of convergence rates of the adaptive algorithms is based on the convergence of an error density, which measures the approximation error for each time step or element. In particular, the error density is proven to converge pointwise on structured adaptive meshes allowing hanging nodes for tensor finite elements and to converge *almost surely* for SDEs as the error tolerance tends to zero.

Finally, numerical experiments illustrate the behavior of the adaptive methods and show that adaptive methods can be more efficient than methods with uniform step sizes.

2000 Mathematics Subject Classification. Primary 65L50, 65C30, 65N50

Key words and phrases. adaptive methods, mesh refinement algorithm, a posteriori error estimate, dual solution, computational complexity, stopped diffusion, finite element method, Monte Carlo method

ISBN 91-7283-553-2 • TRITA-NA-0315 • ISSN 0348-2952 • ISRN KTH/NA/R--03/15--SE

Preface

The work included in this thesis has been carried out from April 1999 to September 2003 at the department of Numerical Analysis and Computer Science (NADA), Royal Institute of Technology (KTH), in Stockholm, Sweden.

First of all, I would like to express my deepest gratitude to my dear advisor, Prof. Anders Szepessy, for being my academic supervisor with enlightening guidance, for continuous encouragement with trust and for all the support throughout this work.

I would also like to thank Raúl Tempone and Georgios E. Zouraris for many valuable discussions about mathematics as well as life. I am grateful to Erik von Schwerin for nice company at different scientific meetings and his comments on the early draft. I thank Mattias Sanberg, Jesper Carlsson and Larisa Beilina for fruitful discussions.

It has been wonderful time for me to cooperate and discuss with all my present and former colleagues at NADA and the Parallel and Scientific Computing Institute (PSCI). In particular, I would like to thank to Prof. Germund Dahlquist, Prof. Axel Ruhe, Prof. Björn Engquist and Prof. Jesper Oppelstrup for finding time for scientific discussions and helping other things.

Last but not least, I wish to thank my dear husband, Dr. Sang-Mo Koo, and our families for all their love, patience and support over the years.

Financial supports from PSCI, NADA and the Swedish National Board for Industrial and Technical Development (NUTEK) are gratefully acknowledged.

Stockholm, September 2003
Kyoung-Sook Moon

Contents

List of Appended Papers

Paper I: K.-S. Moon, A. Szepessy, R. Tempone and G.E. Zouraris, *A variational principle for adaptive approximation of ordinary differential equations*, Numerische Mathematik; DOI 10.1007/s00211-003-0467-8 (2003).

The author participated in discussing the idea of the error expansions for ODEs and contributed to writing the manuscript.

Paper II: K.-S. Moon, A. Szepessy, R. Tempone and G.E. Zouraris, *Convergence rates for adaptive approximation of ordinary differential equations*, Numerische Mathematik; DOI 10.1007/s00211-003-0466-9 (2003).

The author contributed to proving the convergence rate of the adaptive algorithm for ODEs and performed all the numerical simulations.

Paper III: K.-S. Moon, E. Schwerin, A. Szepessy and R. Tempone, *Convergence rates for an adaptive dual weighted residual finite element algorithm*, Technical report, NADA, Royal Institute of Technology, (2003).
Presented in "Multiscale Modelling, Multiresolution and Adaptivity" in Newton Inst., Cambridge, UK, April 7 - 11, 2003.

The author contributed to proving the convergence of the error density for PDEs and writing.

Paper IV: K.-S. Moon, A. Szepessy, R. Tempone and G.E. Zouraris, *Convergence rates for adaptive weak approximation of stochastic differential equations*, Technical report, NADA, Royal Institute of Technology, (2003).
Presented in "First SIAM-EMS conference Applied Mathematics in our Changing World" (Berlin), September, 2001.

The author contributed to proving the convergence rates of the adaptive algorithms for SDEs and writing.

Paper V: K.-S. Moon, *Adaptive Monte Carlo algorithms for stopped diffusion*, Technical report, NADA, Royal Institute of Technology, (2003).
Presented in "SDEs and SPDEs : Numerical Methods and Applications", Edinburgh, UK, 31 March - 4 April 2003.

The author contributed to proving the error expansion for stopped diffusion, performed the numerical simulations, and wrote the manuscript.

Chapter 1

Introduction

Since the time of Leibniz (1646-1719) and Newton (1642-1727), many mathematical models in engineering, physics and finance have been formulated using deterministic or stochastic differential equations. The rapid development of high speed computers over the last decades accelerates the possibilities of efficiently utilizing these models. To use mathematical models on computers we need numerical approximations of solutions to differential equations. As for all reliable measurements, a useful computation requires an estimate of its accuracy. In general, higher accuracy can only be achieved by increasing the computational work. Adaptive methods are useful to combine the two goals of controlling the approximation error and minimizing the computational work. There are numerous adaptive algorithms for differential equations, however the theoretical understanding of optimal convergence rates of adaptive algorithms is not as well developed. This work studies adaptive algorithms and their convergence rates for *ordinary differential equations* (ODEs), *partial differential equations* (PDEs) and for weak approximation of *Itô stochastic differential equations* (SDEs).

The main ingredient of the adaptive algorithms here is an error expansion of the form

$$\text{Global error} = \sum \text{local error} \cdot \text{weight} + \text{higher order error}, \tag{1.1}$$

with computable leading order terms. Here by computable we mean that the leading order error depends only on approximate solution, in other words, the leading order terms are in *a posteriori* form, in contrast to an *a priori* error estimate based on the unknown exact solution. For a deterministic case, such error expansion for differential equations can be derived by three different methods. Firstly, the classical error equation based on linearization derives the error expansion of the form (1.1), see (Henrici, 1962; Harrier et al., 1993; Dahlquist and Björk, 1974; Dahlquist and Björk, 2003). The Galerkin orthogonality using either local problems or the residual also leads to (1.1), see the survey papers (Eriksson et al., 1995; Ainsworth

and Oden, 1997; Becker and Rannacher, 2001). Finally, the error expansion (1.1) can be derived by the variational principle following (Alekseev, 1961) and (Gröbner, 1967). For weak approximations of SDEs, the work (Talay and Tubaro, 1990) develops an error expansion in a priori form under assumptions of sufficient smoothness by the Euler method. Then (Bally and Talay, 1996) extends this error expansion to non-smooth functions using the Malliavin calculus. (Gobet, 2000; Gobet, 2001) extends the results of Talay and Tubaro to killed and reflected diffusion. Based on stochastic flows and discrete dual backward problems, (Szepessy et al., 2001) derives new expansions of the computational error with computable leading order term in a posteriori form. General construction and analysis of the convergence order for approximations of SDEs can be found in (Kloeden and Platen, 1992).

Based on these error estimates, the goal of adaptive algorithms for differential equations is to solve problems with an adapted mesh using as little computational work as possible, for a given level of accuracy. Based on the a posteriori error expansions derived in **Paper I-V**, the global error can be asymptotically approximated by

$$\text{Global error} \approx \sum_{\mathcal{K}} \text{error indicator}, \tag{1.2}$$

where $\mathcal{K}$ is a set of time steps or elements. Then a typical adaptive algorithm does two things iteratively :

(1) if the error indicators satisfy an accuracy condition, then it stops; otherwise

(2) the algorithm chooses where to refine the mesh and recomputes the error indicators then makes an iterative step to **(1)**.

Therefore the indicators not only estimate the localization of the global error but also give information on how to refine the mesh in order to achieve optimal efficiency.

An approximation of the weight function in (1.1), plays an important role for step **(2)**. The weight function solves a certain linear backward dual problem obtained by linearizing a given forward problem around the approximate solution. Therefore, the algorithm uses additional computational work to obtain the error estimate, which also informs where to refine the mesh to achieve optimal efficiency. The use of dual functions is well known for adaptive mesh control for ODEs and PDEs, see (Babuška and Miller, 1984a; Babuška and Miller, 1984b; Babuška and Miller, 1984c; Johnson, 1988; Johnson and Szepessy, 1995; Estep, 1995; Becker and Rannacher, 1996).

Despite the wide use of adaptive algorithms and the well developed theory of a posteriori error estimates, less is known theoretically on the behavior of adaptive mesh refinement algorithms. However, there are recent important contributions; The work (DeVore, 1998) studies the efficiency of nonlinear approximation of functions, including wavelet expansions, based on smoothness conditions in Besov space. Inspired by this approximation result, (Cohen et al., 2000) proves that a wavelet-based adaptive N-term approximation produces an approximation with asymptotically optimal error in the energy norm for linear coercive elliptic problems. Then

the work (Binev et al., 2002) and (Stevenson, 2003) extend the ideas of (Morin et al., 2000) to prove similar optimal error estimates in the energy norm for piecewise linear finite elements applied to the Poisson equation, see also (Dörfler, 1996). The modification includes a coarsening step in the adaptive algorithm to obtain bounds on the computational work. For solving ODEs, (Stuart, 1997) and (Lamba and Stuart, 1998) prove the convergence of `ODE23` of `MATLAB` version 4.2. For strong approximation of SDEs, the work (Hofmann et al., 2000; Hofmann et al., 2001) and (Müller-Gronbach, 2000) prove optimal convergence rates of adaptive and uniform step size control. These topics can be linked to information based complexity theory, see (Bakhvalov, 1971; Werschulz, 1991; Novak, 1996; Traub and Werschulz, 1998).

The main results of this thesis are the construction and analysis of adaptive algorithms for ODEs, PDEs and weak approximation of SDEs. In particular, the key results are

- an analysis of a posteriori error expansions useful for adaptive mesh refinement methods based on local errors and dual solutions,
- development of general adaptive algorithms for solving ODEs, PDEs and SDEs, and
- a rigorous and simple analysis of convergence rates of the adaptive algorithms, including the convergence of an error density, which measures the approximation error for each time step or element.

In this thesis, Chapter 2 describes an adaptive algorithm for ODEs derived in **Paper II** and the error expansion proved in **Paper I**. Chapter 3 presents an adaptive finite element algorithm for a second order elliptic PDE based on the dual weighted residual error expansion using piecewise linear tensor elements analyzed in **Paper III**. Chapter 4 introduces SDEs and the adaptive Monte Carlo algorithms with stochastic and deterministic time steps based on *almost sure* convergence of the error density analyzed in **Paper IV**. The extension of this adaptive Monte Carlo algorithm to stopped diffusion problems, derived in **Paper V**, is also described in Chapter 4. When we refer to the sections and theorems in the papers, we use the paper number as a prefix, i.e. Theorem I.2.1 denotes Theorem 2.1 in **Paper I**.

Chapter 2

Adaptive Algorithm

Let us consider an ordinary differential equation (ODE) of the form

$$\frac{dX(t)}{dt} = a(t, X(t)), \quad 0 < t \leq T, \tag{2.1}$$

with an initial value $X(0) = X_0 \in \mathbb{R}^d$ and a given flux $a : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$. Let us first discretize the time interval $[0, T]$ into N subintervals $0 = t_0 < t_1 < \ldots < t_N = T$ and let $\overline{X}$ be an approximation of X in (2.1) by any p-th order accurate numerical method, satisfying the same initial condition $\overline{X}(0) = X(0) = X_0$.

What we are interested in is to compute a function value $g(X(T))$ for a given general function $g : \mathbb{R}^d \to \mathbb{R}$, which may represent the quantities of physical interest. Using the variational principle, we show in **Paper I** that the global error has the expansion

$$g(X(T)) - g(\overline{X}(T)) = \sum_{n=1}^{N} \left(\bar{e}(t_n), \bar{\Psi}(t_n)\right) + \int_0^T o(\Delta t^p(t))dt, \tag{2.2}$$

where $(\cdot, \cdot)$ is the standard scalar product on $\mathbb{R}^d$. Here the approximate local error is defined by $\bar{e}(t_n) := \gamma(\bar{\overline{X}}(t_n) - \overline{X}(t_n))$, where γ is a Richardson extrapolation constant and the approximate local exact solution $\bar{\overline{X}}$ is computed with smaller time steps or a higher order method than $\overline{X}$. The weight $\bar{\Psi}$ is an approximation of Ψ, which solves the dual equation

$$\begin{aligned} -\frac{d\Psi(s)}{ds} &= (a')^*(s, X(s))\Psi(s), \quad s < T \\ \Psi(T) &= g'(X(T)), \end{aligned} \tag{2.3}$$

where $(a')^*(s, x)$ is the transpose of the Jacobian matrix.

Based on the error expansion in (2.2) and a given error tolerance TOL, the number of time steps are minimized by choosing, for all time steps,

$$\text{error indicator} := |\text{local error} \cdot \text{weight}| = \text{constant}, \tag{2.4}$$

i.e. for $n = 1, \ldots, N$,

$$\bar{r}_n := |\left(\bar{e}(t_n), \bar{\Psi}(t_n)\right)| = \frac{\text{TOL}}{N}.$$

This motivates that the adaptive algorithm, `MSTZ` described in **Paper II**, starts with an initial partition $\Delta t[1]$ and then specifies iteratively a new partition $\Delta t[k+1]$, from $\Delta t[k]$, using the following refinement strategy: for $n = 1, 2, \ldots, N[k]$,

$$\begin{aligned} &\textbf{if } \bar{r}_n[k] > s_1 \frac{\text{TOL}}{N[k]} \textbf{ then} \text{ divide } \Delta t_n[k] \text{ into two sub steps} \\ &\textbf{else} \text{ let the new step be the same as the old} \\ &\textbf{endif} \end{aligned} \tag{2.5}$$

until it satisfies the following stopping criteria:

$$\textbf{if } \left(\max_{1 \le n \le N[k]} \bar{r}_n[k] \le S_1 \frac{\text{TOL}}{N[k]} \right) \textbf{ then } \text{ stop.} \tag{2.6}$$

The choice of the parameters $1 \simeq s_1 < S_1$ is designed to avoid the slow reduction of the maximum error indicator when almost all $\bar{r}_n$ satisfy $\bar{r}_n \le s_1 \text{TOL}/N$. Here the time steps are chosen to be the same for all components of the solution, however the error estimate (2.2) would also be applicable to multi adaptive time steps used in (Logg, 2001).

For a p-th order accurate method, the number N^{uni} of uniform steps to reach a given level of accuracy TOL turns out to be proportional to the p-th root of the L^1-norm of the error density, $\bar{\rho}(t) := \left(\bar{e}(t_n), \bar{\Psi}(t_n)\right)/\Delta t_n^{p+1}$, for $t \in (t_{n-1}, t_n]$, i.e

$$N^{\text{uni}} = \frac{T}{\text{TOL}^{\frac{1}{p}}} \|\bar{\rho}\|_{L^1}^{\frac{1}{p}}. \tag{2.7}$$

On the other hand, the optimal number N^{opt} of adaptive steps is proportional to the p-th root of the smaller $L^{\frac{1}{p+1}}$ quasi-norm of the error density, i.e

$$N^{\text{opt}} = \frac{1}{\text{TOL}^{\frac{1}{p}}} \|\bar{\rho}\|_{L^{\frac{1}{p+1}}}^{\frac{1}{p}}. \tag{2.8}$$

Therefore an optimal adaptive method uses less time steps than a uniform method, due to Jensen's inequality $\|f\|_{L^{\frac{1}{p+1}}} \le T^p \|f\|_{L^1}$. Note that the optimal here refers to error densities of one sign, i.e. possible cancellation of the error is not taken into account here.

Table 2.1 shows an illustrative comparison between the adaptive algorithm `MSTZ`, `Uniform` using uniform step size and the built in ODE solver `ODE45` in `MATLAB` for

	MSTZ	Uniform	ODE45
Error	0.01	0.06	0.04
N	6×10^3	10×10^3	34×10^3
N^{tot}	2×10^4	1.9×10^4	9.2×10^4

Table 2.1. The comparisons of the final number of time steps N and the total number of time steps N^{tot} with the global error Error using a 5-th order explicit Runge-Kutta method with adaptive steps for MSTZ or ODE45 and uniform steps for Uniform, see Example II.3.1.

the Lorenz problem introduced in (Lorenz, 1963). For this example, the algorithm MSTZ achieves higher accuracy with around half the final number of time steps compared to Uniform and with one fifth of the final number of steps compared to ODE45. On the other hand, by taking the total work and storage into account the comparison is less favorable for MSTZ and leaves possibilities for more efficient future implementations.

2.1 Overview of Paper I

Paper I derives an error expansion of the form (2.2) for general approximation of functions of solutions to ODEs. Based on a variational principle following (Alekseev, 1961) and (Gröbner, 1967), the error analysis using local errors gives a simple and precise derivation of the fundamental property that the global error is a weighted sum of the local errors in Theorems I.2.1. Then Theorem I.2.3 transfers the error representation in Theorems I.2.1 to an asymptotic expansion (2.2) with computable approximations of the local errors and the weights, using standard estimates. Theorem I.2.4 derives an alternative error expansion using Richardson extrapolation and Section I.2.4 studies the roundoff errors. In particular, Theorem I.2.5 proves a uniqueness result of desirable error representations for adaptive refinements. Finally, Section I.3 shows that the computable leading order terms of the error expansions are the same whether derived by the variational principle, by the error equation or by the Galerkin orthogonality. However, only the variational principle provides that the local error is a factor in the higher order error terms.

2.2 Overview of Paper II

Based on the error expansion (2.2) derived in **Paper I**, **Paper II** constructs an adaptive algorithm for ODEs and analyzes its asymptotic behavior as the error tolerance parameter tends to zero. For a p-th order accurate method, the $L^{\frac{1}{p+1}}$ quasi-norm of the error density in (2.8) is a good measure of the convergence rate for the adaptive approximation of ODEs, while the number of uniform steps is measured by the larger L^1-norm of the error density in (2.7).

Section II.2 constructs an algorithm which subdivides the time steps with error indicators following the refinement (2.5) and stopping (2.6). Then we analyze the adaptive algorithm in three theorems. Firstly, Theorems II.2.1 (stopping) proves that each refinement level of this adaptive algorithm decreases the maximal error indicator with a factor, less than 1, or stops. Secondly, Theorem II.2.4 (accuracy) shows that the computational error is asymptotically bounded by a given error tolerance. Thirdly, Theorem II.2.5 (efficiency) proves that the algorithm stops with asymptotically optimal number of time steps N in the finest mesh, up to a problem independent factor.

Theorem II.2.7 shows that the total number of steps is bounded by $\mathcal{O}(N)$ for decreasing tolerance, where N is the number of time steps in the finest mesh. However, varying tolerance has the drawback that the final stopping tolerance is not a priori known. For the constant tolerance, the total number of steps including all levels is bounded by the larger $\mathcal{O}(N \log N)$. The adaptive algorithm and the analysis of its convergence rate can also be generalized by including the merging of time steps in Remark II.2.8, see also (Moon, 2001). Finally, Section II.3 presents numerical experiments, which show that the adaptive algorithm needs less number of time steps than `ODE45` in `MATLAB` and `Uniform` to achieve the same level of accuracy for the Lorenz and a singular problem.

Chapter 3

Adaptive Finite Element Algorithm

Let us consider a second order elliptic partial differential equation (PDE) of the form

$$-\mathrm{div}(a\nabla u) = f \tag{3.1}$$

in an open bounded domain $\Omega \subset \mathbb{R}^d$ with boundary condition $u|_{\partial\Omega} = 0$. The goal is to compute a linear functional

$$(u, F) := \int_\Omega uF dx \tag{3.2}$$

for a given function $F \in L^2(\Omega)$. The functional value (3.2) can be a quantity of physical interest such as stress values, mean fluxes and drag and lift coefficients, see the survey paper (Becker and Rannacher, 2001).

Let us first discretize the domain Ω into quadrilaterals K and let h_K be the local mesh size, i.e. the length of the longest edge of K. Define the residual $\mathcal{R}(v) = -\mathrm{div}(a\nabla v) - f$ as a distribution in $H^{-1}(\Omega)$ for $v \in H_0^1(\Omega)$ and let $(\cdot,\cdot)$ denote the duality pairing in $H^{-1} \times H_0^1$, which reduces to the usual inner product in $L^2(\Omega)$ on $L^2 \times L^2$. Here the Sobolev space $H_0^1(\Omega)$ is the usual Hilbert space of functions on Ω, vanishing on $\partial\Omega$, with bounded first derivatives in $L^2(\Omega)$. Let u_h be the finite element approximation of the solution u of (3.1) based on the standard variational formulation, i.e.

$$\begin{aligned} (\mathcal{R}(u), v) &= 0, \quad \forall v \in H_0^1(\Omega), \\ (\mathcal{R}(u_h), v) &= 0, \quad \forall v \in V_h, \end{aligned}$$

where V_h is the set of continuous piecewise bilinear functions in $H_0^1(\Omega)$. Using the dual function $\varphi \in H_0^1(\Omega)$ defined by

$$(a\nabla v, \nabla\varphi) = (F, v), \quad \forall v \in H_0^1(\Omega), \tag{3.3}$$

the global error has the representation

$$(u - u_h, F) = (\mathcal{R}(u_h), \pi\varphi - \varphi)\,, \tag{3.4}$$

where π is the nodal interpolant on V_h.

An introduction to the finite element method is given, e.g., in the books (Johnson, 1987; Brenner and Scott, 1994; Bangerth and Rannacher, 2003). The *a posteriori* error analysis for finite element methods was initiated by the work (Babuška and Rheinboldt, 1978) and has been further developed in (Ladeveze and Leguillon, 1983; Bank and Weiser, 1985; Ainsworth and Oden, 1993). For further references, see the survey paper (Ainsworth and Oden, 1997; Becker and Rannacher, 2001). The use of duality arguments in (3.4) goes back to ideas of (Babuška and Miller, 1984a; Babuška and Miller, 1984b; Babuška and Miller, 1984c) and has been used successfully, see the work (Johnson, 1990; Eriksson and Johnson, 1988; Johnson and Szepessy, 1995; Eriksson et al., 1995; Becker and Rannacher, 1996). The error analysis for PDEs with stochastic coefficients or stochastic load is summarized in (Tempone, 2002). See also the work (Deb et al., 2001; Babuška et al., 2002; Schwab and Todor, 2002).

Inspired by the work (Eriksson et al., 1995; Becker and Rannacher, 2001), **Paper III** derives a computable approximation of (3.4) of the form

$$(u - u_h, F) \simeq \sum_K \bar{\rho}_K h_K^{2+d} \tag{3.5}$$

where the error density $\bar{\rho}$ is essentially independent of the mesh size, since the asymptotic error density is used to find the optimal mesh. Then, following the idea in **Paper II**, we derive and analyze the convergence rate of an adaptive algorithm based on the error expansion (3.5). The extension of the analysis of the adaptive algorithm to PDEs is straightforward except the pointwise convergence of the error density and a hanging node constraint. To prove this convergence, we require structured adapted meshes and analysis specific to tensor finite elements.

Figure 3.1 shows an illustrative adaptive mesh based on the error expansion (3.5). We compute the linear functional

$$(u, F) = \int_\Omega \partial_{x_1} u(x) \exp(-|x - x^*|^2/\epsilon^2) dx/(\pi\epsilon^2),$$

where $x^* = (1/\sqrt{2}, 1/\sqrt{3})$ and the solution $u = sin(\pi x_1)sin(\pi x_2)$ of PDE (3.1) with $a \equiv 1, f = 2\pi^2 \sin(\pi x_1)\sin(\pi x_2)$ and $\Omega = (0,1)^2$ using TOL $= 10^{-3}$ and $\epsilon = 0.03$. The error density has the bound $\bar{\rho}(x) = \mathcal{O}(1)/(|x - x^*| + \epsilon)^3$, which yields $\lim_{\mathrm{TOL}\to 0+}(\mathrm{TOL}N) = \mathcal{O}(1)$, where N is the number of adaptive elements, while the method using uniform elements has the larger number $\mathcal{O}(\epsilon^{-1})/\mathrm{TOL}$ of elements, by a factor ϵ^{-1}.

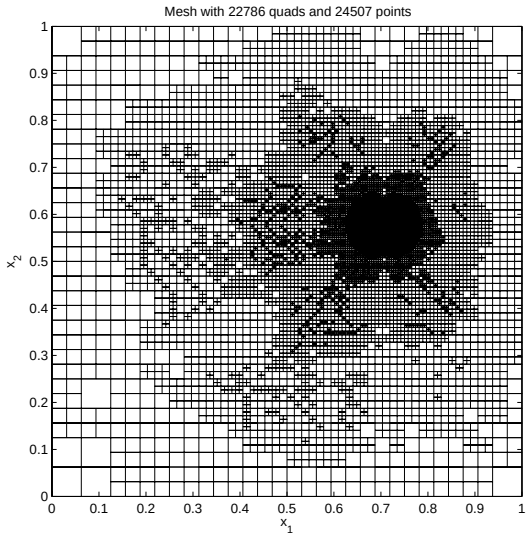

Figure 3.1. An illustrative adaptive mesh based on the error expansion using the dual weighted residual method described in **Paper III**.

3.1 Overview of Paper III

Paper III studies convergence properties of an adaptive algorithm based on the dual weighted residual finite element method. The algorithm computes the piecewise linear tensor approximation of functionals of solutions to second order elliptic partial differential equations in an open bounded domain of $\mathbb{R}^d$.

Section III.2 derives an asymptotic error expansion (3.5) using the dual weighted residual method. This error density in (3.5) is proven to converge pointwise on structured adaptive meshes allowing hanging nodes for tensor finite elements. Note that such pointwise convergence of the error density, based on second order difference quotients, is well known for structured uniform meshes. However, Remark III.2.2 shows by an example that second order difference quotients of smooth functions do not in general converge pointwise for adapted meshes. To prove convergence of the second order difference quotients, in the error density, our proof instead uses localized averages and a symmetrizer.

Section III.3 constructs an algorithm which subdivides the elements with error indicators, $\bar{\rho}_K h_K^{2+d}$, greater than TOL/N and stops if all N elements have sufficiently small error indicators. In particular the algorithm has no coarsening step. Theorems III.3.1, III.3.4 and III.3.5 prove that each refinement level of this adaptive algorithm decreases the maximal error indicator with a factor, less than 1, or stops with an error asymptotically bounded by TOL and with asymptotically optimal number of elements, N, in the finest mesh, up to a problem independent

factor. Note that the optimal here refers to error densities of one sign, i.e. possible cancellation of the error is not taken into account.

For the second order accurate piecewise linear finite elements on a uniform mesh, the number of elements to reach a given approximation error turns out to be proportional to the $d/2$ power of the L^1-norm of the error density. This work shows that the smallest number of isotropic elements in an adaptive mesh is proportional to the $d/2$ power of the smaller $L^{\frac{d}{d+2}}$ quasi-norm of the same error density. These norms of the error density are therefore good measures of the convergence rates and define our optimal number of elements. The total number of elements, including all refinement levels, can be bounded by $\mathcal{O}(N)$, provided that the tolerance in each refinement level decreases by a constant factor, see Theorem III.3.8. On the other hand, with constant tolerance, the total number of elements including all levels is bounded by the larger $\mathcal{O}(N \log N)$.

Chapter 4

Adaptive Monte Carlo Algorithms

Let us consider an Itô Stochastic differential equation (SDE) of the form

$$dX(t) = a(t, X(t))dt + b(t, X(t))dW(t) \tag{4.1}$$

for $t \in [0, T]$ with initial value $X(0) = X_0 \in \mathbb{R}$. The stochastic process $X = \{X(t), 0 \leq t \leq T\}$ is a unique solution of the SDE (4.1) with a drift coefficient $a(t, X(t))$ and a diffusive coefficient $b(t, X(t))$. The one dimensional *Wiener process* $W = \{W(t), 0 \leq t \leq T\}$, also known as *Brownian motion*, depends continuously on t with $W(0) = 0$ with probability 1. For any partition $0 = t_0 < t_1 < \ldots < t_N = T$, the Wiener increments $W(t_{n+1}) - W(t_n)$ for $n = 0, \ldots, N-1$ are independent and normally distributed with

$$E[W(t_{n+1}) - W(t_n)] = 0, \quad E[(W(t_{n+1}) - W(t_n))^2] = t_{n+1} - t_n. \tag{4.2}$$

The corresponding Itô integral of (4.1) is defined to be the limit of the forward Euler method, while the trapezoidal method yields, the so called *Stratonovich integral*. An introduction to numerical approximation of SDEs and an extensive review of the literature can be found in the inspiring book (Kloeden and Platen, 1992) and a survey article (Platen, 1999). See also (Karatzas and Shreve, 1988; Milstein, 1995; Øksendal, 1998).

Let us first discretize the time interval $[0, T]$ into N subintervals $0 = t_0 < t_1 < \ldots < t_N = T$ and let $\overline{X}$ be the Euler approximation of the process X, i.e. start with $\overline{X}(0) = X_0$ and compute $\overline{X}_{n+1}$ by

$$\overline{X}(t_{n+1}) = \overline{X}(t_n) + a(t_n, \overline{X}(t_n))\Delta t_n + b(t_n, \overline{X}(t_n))\Delta W_n, \tag{4.3}$$

where $\Delta t_n := t_{n+1} - t_n$ denotes the time increments and and $\Delta W_n := W(t_{n+1}) - W(t_n)$ denotes the Wiener increments for $n = 0, 1, \ldots, N-1$. The equation (4.1)

with $X(t) \in \mathbb{R}$ can directly be extended to higher dimension $X(t) \in \mathbb{R}^d$ with a multidimensional Wiener process $W = (W_1, \ldots, W_l)$ consisting of independent one dimensional Wiener processes W_i for $i = 1, \ldots, l$.

The applications of stochastic process are constantly growing including finance, signal and image processing, stochastic climate prediction, wave propagation in random media and biology, see, for instance, (Wilmott et al., 1995; Jouini et al., 2001; Steele, 2001; Majda et al., 2001; Abdullaev et al., 2000; Ewens and Grant, 2001). Depending on the problems, one could be interested in either strong convergence

$$\|X(T) - \overline{X}(T)\|_{L^2} = \sqrt{E[(X(T) - \overline{X}(T))^2]} = \mathcal{O}(\sqrt{\Delta t}) \tag{4.4}$$

or weak convergence for a given general function $g : \mathbb{R} \to \mathbb{R}$,

$$E[g(X(T)) - g(\overline{X}(T))] = \mathcal{O}(\Delta t) \tag{4.5}$$

under appropriate assumptions on smoothness of the functions g, a and b. The former approximates the solution pathwise which is useful for instance in the direct simulation of stochastic dynamical system or testing parameter estimators. On the other hand, in many practical problems we only need to approximate the corresponding probability such as the moments of a functional, for instance an option price g on a stock price X or a trend of storm track in climate prediction.

To compute the expected value $E[g(X(T))]$ for a given function g, we apply the Euler approximation (4.3) to get $E[g(\overline{X}(T))]$. But the expected value $E[g(\overline{X}(T))]$ is not computable, thus we use the Monte Carlo method to approximate the expected value by the sample average. Therefore the global error can be split into a time discretization error and a statistical error

$$\begin{aligned} & E[g(X(T))] - \frac{1}{M}\sum_{j=1}^{M} g(\overline{X}(T;\omega_j)) \\ &= \big(E[g(X(T)) - g(\overline{X}(T))]\big) + \left(E[g(\overline{X}(T))] - \frac{1}{M}\sum_{j=1}^{M} g(\overline{X}(T;\omega_j))\right) \\ &=: \mathcal{E}_T + \mathcal{E}_S \end{aligned} \tag{4.6}$$

where M is the number of realizations.

The goal of this work is to develop a numerical method for weak approximations of Itô SDEs using as few time steps N and realizations M as possible, with an approximation error bounded by a given error tolerance TOL, i.e. the event

$$|\mathcal{E}_T + \mathcal{E}_S| \leq \mathrm{TOL} \tag{4.7}$$

has a probability close to one. The statistical error $\mathcal{E}_S$ in (4.6) is asymptotically bounded by

$$|\mathcal{E}_S| \leq c_0 \frac{\overline{\sigma}}{\sqrt{M}} =: \mathtt{E}_S \tag{4.8}$$

using the Central Limit Theorem, see for instance (Durett, 1964; Goodman et al., 2002). Here $\overline{\sigma}$ is the sample average of the standard deviation of $g(\overline{X}(T))$ and c_0

is a positive constant corresponding to confidence interval. This statistical error is independent of the dimension of the problem, which makes the Monte Carlo method useful for high dimensional problems. On the other hand the Monte Carlo approximation converges very slowly as we can see in (4.8). One cure for this slow convergence is variance reduction methods which accelerate the convergence by reducing the standard deviation $\overline{\sigma}$, see (Calflisch, 1998; Fishman, 1996).

Talay and Tubaro (1990) expanded the time discretization error $\mathcal{E}_T$ in (4.6) in powers of N^{-1} in a priori form. Szepessy, Tempone and Zouraris (2001) developed a posteriori error expansions with computable leading order terms based on stochastic flows and discrete dual backward problems; the time discretization error $\mathcal{E}_T$ is then approximated by either

$$\mathcal{E}_T \simeq E\left[\sum_{n=0}^{N-1} \rho(t_n,\omega)(\Delta t_n)^2\right] \tag{4.9}$$

with an error density function $\rho(t_n,\omega)$ for stochastic time stepping or

$$\mathcal{E}_T \simeq \sum_{n=0}^{N-1} E[\rho(t_n,\omega)](\Delta t_n)^2 \tag{4.10}$$

for deterministic time stepping. The stochastic time stepping algorithm based on (4.9) may use different meshes for each realization, while the deterministic time stepping algorithm uses the same mesh for all realizations.

The algorithms choose both the number M of realizations and the number N of time steps adaptively to minimize the computational work subject to an accuracy constraint. The basic structure of the stochastic time stepping algorithm is as follows:

`Algorithm S`

First split a given error tolerance TOL by $\text{TOL} = \text{TOL}_S + \text{TOL}_T$ where TOL_S is a statistical error tolerance and TOL_T is a time discretization error tolerance. Set $m = 1$ and $\texttt{E}_S[m] = +\infty$ and choose $M[m] = M_0$.

Do while $(\texttt{E}_S[m] > \text{TOL}_S)$

For each realization, compute $g(\overline{X}(T;\omega_j))$ for $j = 1,\ldots,M[m]$, until a stopping criterion of a time refinement strategy using TOL_T is satisfied. Compute a new approximation $\texttt{E}_S[m+1]$ and a new number $M[m+1]$ of realizations. Increase m by 1.

end-do.

Accept the approximate solution $\frac{1}{M[m]}\sum_{j=1}^{M[m]} g(\overline{X}(T;\omega_j))$, since it satisfies both the stopping criteria of the statistical and of the time discretization error.

The deterministic time stepping algorithm (`Algorithm D`) first determines the mesh, using a smaller number of realizations than for the stochastic time stepping algorithm, then using this fixed mesh, it computes the sample average of $g(\overline{X}(T;\omega_j))$ for $j = 1, 2, \ldots, M$ by increasing the number M of realizations adaptively.

Table 4.1 shows a comparison of the number of time steps between `Algorithm S`, `Algorithm D` and `Uniform`, which use the Monte Carlo Euler method with a constant time step size. We compute $E[g(X(T))]$ with TOL $= 0.025$ where the process X solves

$$dX(t) = \mathbf{1}_{t\in[0,\alpha]} X(t)dW(t) + \mathbf{1}_{t\in(\alpha,T]} \left(\frac{X(t)dt}{2\sqrt{t-\alpha}} + X(t)dW(t) \right)$$

with the initial value $X(0) = 1$. Here $\mathbf{1}_{y\in A}$ denotes the indicator function, i.e.$\mathbf{1}_{y\in A} = 1$ if $y \in A$, otherwise $\mathbf{1}_{y\in A} = 0$. Table 4.1 shows that for deterministic α, both `Algorithm S` and `Algorithm D` have a remarkable advantage over a computation with uniform time steps. On the other hand, for a uniformly distributed random variable α, the deterministic time stepping `Algorithm D` behaves like the uniform time stepping method `Uniform` which needs $\mathcal{O}(\text{TOL}^{-2})$ time steps to achieve the given level of accuracy. Thus, `Algorithm S` has a clear advantage in this case with the expected number of steps of order $\mathcal{O}(\text{TOL}^{-1})$.

	`Algorithm S`	`Algorithm D`	`Uniform`
$\alpha = T/3$	2.3×10^2	4.6×10^2	1.7×10^5
$\alpha \sim U(1/22, T - 1/22)$	1.9×10^2	1.3×10^5	1.3×10^5

Table 4.1. The comparison of the number of time steps for a singular problem with constant $\alpha = T/3$ or uniformly distributed random variable $\alpha \sim U(1/22, T - 1/22)$.

Stopped diffusion is an example where adaptive methods are very useful. The main difficulty in the approximation of the stopped (or killed) diffusion on the boundary ∂D is that a continuous sample path may exit the given domain D even though a discrete approximate solution does not cross the boundary of D, see Figure 4.1. This hitting of the boundary makes the time discretization error of order $N^{-1/2}$ for the Monte Carlo Euler method with N uniform time steps, see (Gobet, 2000), while the discretization error is of order N^{-1} without boundary in $\mathbb{R}^d \times [0, T]$. To overcome this slow convergence, the work (Mannella, 1999; Baldi et al., 1999; Jansons and Lythe, 2000) propose to generate a uniformly distributed random variable in $(0, 1)$ for each time step and compare it with a known exit probability to decide if the continuous path exits the domain during this time interval. Gobet (2001) proved that a similar method with N uniform time steps in a domain with smooth boundary converges with the rate N^{-1} under some appropriate assumptions. Buchmann (2003) compared different Monte Carlo methods computationally for stopped diffusions.

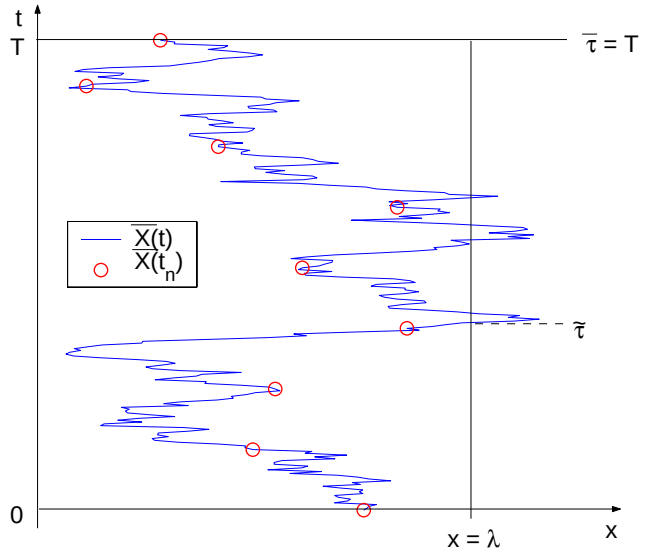

Figure 4.1. An illustrative Monte Carlo Euler trajectory when the first exit time $\widetilde{\tau}$ of the continuous Euler path $\overline{X}(t)$ is much smaller than the first exit time $\overline{\tau}$ of the discrete Euler path $\overline{X}(t_n)$.

Inspired by (Petersen and Buchmann, 2002), we extend the a posteriori error expansion and the stochastic time stepping algorithm, `Algorithm S`, introduced in (Szepessy et al., 2001) to a stopped diffusion problem in **Paper V**. The extended algorithm reduces the computational error by choosing adaptively the size of the time steps near the boundary and recover the convergence rate of order N^{-1} with N adaptive time steps. Figure 4.2 shows an illustrative histogram of the step sizes depending on the distance from the boundary. We compute $E[g(X(\tau),\tau)]$ for $g(x,t) = x^3 e^{-t}$ where the process X solves

$$dX(t) = \frac{11}{36}X(t)dt + \frac{1}{6}X(t)dW(t)$$

with the initial value $X(0) = 1.6$ in a domain $(x,t) \in D \times [0,T]$, where $D = (-\infty, 2)$ and $T = 2$. Here $\tau := \inf\{0 < t < T : X(t) \notin D\}$ is the first exit time from a given domain and we use the extended algorithm of `Algorithm S` with TOL $= 0.05$. To compensate the large error near the boundary, relatively small step sizes are used close to the boundary compared to further away from the boundary as is clear from Figure 4.2.

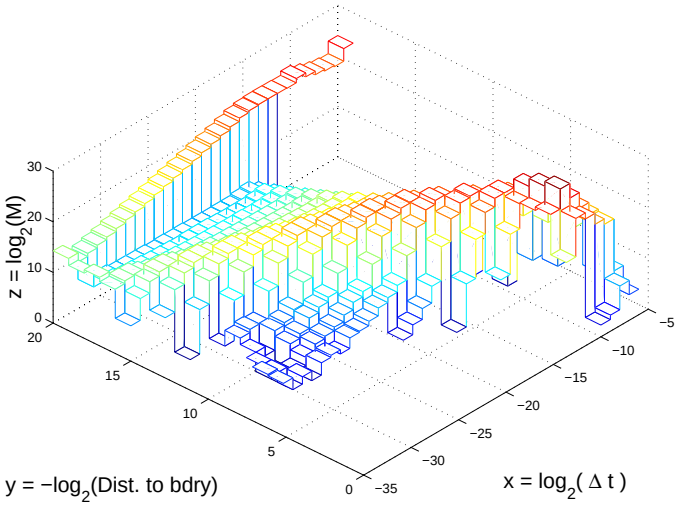

Figure 4.2. The histogram of the step sizes depending on the distance from the boundary using the extended algorithm based on `Algorithm S`. Relatively small step sizes are used close to the boundary to achieve the discretization error of order N^{-1} with N adaptive time steps.

4.1 Overview of Paper IV

Paper IV proves convergence rates of adaptive algorithms for weak approximation of Itô SDEs using the Monte Carlo Euler method. Two adaptive algorithms based either on optimal stochastic time steps or optimal deterministic time steps are studied.

There are two main results on efficiency and accuracy of the adaptive algorithms described in Section IV.3. For accuracy, with probability close to one, the approximation errors in (4.7) are asymptotically bounded by the specified error tolerance times a problem independent factor as the tolerance parameter tends to zero. For efficiency, both the algorithms with stochastic steps and deterministic steps stop with asymptotically optimal expected number of final time steps and optimal number of final time steps respectively, up to a problem independent factor. Note that the optimal here refers to error densities of one sign, i.e. possible cancellation of the error is not taken into account.

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 (4.9) or (4.10), converges *almost surely* as the error tolerance tends to zero. This convergence can be understood by the *almost sure* convergence of the approximate solution, $\overline{X}$, as the maximal step size tends to zero. Although the time steps are not adapted to the standard filtration generated by only W for the stochastic time stepping algorithm, the work (Szepessy et al., 2001) proves 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 IV.4.2. Therefore Theorem IV.4.1 and IV.4.4 use Doob's inequality to prove the *a.s.* convergence of $\overline{X}$. Similar results of pointwise convergence with constant step sizes, adapted to the standard filtration, are surveyed in (Talay, 1995).

Finally numerical examples illustrate the behavior of the adaptive algorithms, motivating when stochastic and deterministic adaptive time steps are more efficient than constant time steps and when adaptive stochastic steps are more efficient than adaptive deterministic steps.

4.2 Overview of Paper V

Paper V presents adaptive algorithms for weak approximation of stopped diffusion using the Monte Carlo Euler method. The goal is to compute an expected value $E[g(X(\tau), \tau)]$ of a given function g depending on the solution X of an Itô SDE and on the first exit time $\tau := \inf\{0 < t < T : X(t) \notin D\}$ from a given domain D.

This work derives adaptive algorithms to reduce the computational error by choosing adaptively the size of the time steps near the boundary. This has the advantage that the exit probability does not need to be computed accurately.

In Section V.2, we approximate the time discretization error, $\mathcal{E}_T$ in (4.6), in a posteriori form by extending the error estimate in (Szepessy et al., 2001) to weak approximation of stopped diffusion. As in (Talay and Tubaro, 1990), the first step to derive an error estimate is to introduce a continuous Euler path $\overline{X}(t)$ for $t \in [t_n, t_{n+1}]$ and $n = 0, \ldots, N-1$ and its first exit time $\widetilde{\tau}$. Then the error $\mathcal{E}_C := E[g(X(\tau), \tau) - g(\overline{X}(\widetilde{\tau}), \widetilde{\tau})]$ between the exact and continuous Euler path is approximated using stochastic flows and dual backward solutions in Section V.2.3. The main idea in this extension is to use difference quotients to replace the stochastic flows that do not exist at the boundary. The approximate error $\mathcal{E}_D := E[g(\overline{X}(\widetilde{\tau}), \widetilde{\tau}) - g(\overline{X}(\overline{\tau}), \overline{\tau})]$ between the continuous and discrete Euler path is derived by a conditional probability using Brownian bridges in Section V.2.2. Note that the exit probability is used here only to decide the time steps, not to approximate the expected values directly. Therefore the accuracy of the approximation of the exit probability is not crucial.

In Section V.3, we propose two stochastic time stepping adaptive algorithms, one controls the time discretization error $\mathcal{E}_T = \mathcal{E}_C + \mathcal{E}_D$ together and the other considers the time discretization errors $\mathcal{E}_C$ and $\mathcal{E}_D$ separately. Numerical results show that both adaptive algorithms achieve the time discretization error of order N^{-1} with N adaptive time steps, while the error is of order $N^{-1/2}$ for a method with N uniform time steps.

Bibliography

Abdullaev, F., Bronski, J., and Papanicolaou, G. (2000). Soliton perturbations and the random kepler problem. *Phys. D.*, 135:369–386.

Ainsworth, M. and Oden, J. (1993). A unified approach to a posteriori error estimation using element residual methods. *Numer. Math.*, 65:23–50.

Ainsworth, M. and Oden, J. (1997). A posteriori error estimation in finite element analysis. *Comput. Methods Appl. Mech. Engrg.*, 142:1–88.

Alekseev, V. (1961). An estimate for the perturbations of the solutions of ordinary differential equations ii. *Vestnik Moskov. Univ. Ser. I Mat. Mech*, 3:3–10.

Babuška, I. and Miller, A. (1984a). The post-processing approach in the finite element method, i: calculations of displacements, stresses and other higher derivatives of the displacements. *Int. J. Numer. Meth. Eng.*, 20:1085–1109.

Babuška, I. and Miller, A. (1984b). The post-processing approach in the finite element method, ii: the calculation of stress intensity factors. *Int. J. Numer. Meth. Eng.*, 20:1111–1129.

Babuška, I. and Miller, A. (1984c). The post-processing approach in the finite element method, iii: a posteriori error estimation and adptive mesh selection. *Int. J. Numer. Meth. Eng.*, 20:2311–2324.

Babuška, I. and Rheinboldt, W. (1978). Error estimates for adaptive finite element computations. *SIAM J. Numer. Anal.*, 15:736–754.

Babuška, I., Tempone, R., and Zouraris, G. (2002). Galerkin finite element approximations of stochastic elliptic partial differential equations. Technical report, Univ. of Texas at Austin, USA.

Bakhvalov, N. (1971). On the optimality of linear methods for operator approximation in convex classes of functions. *USSR Comput. Math. and Math. Phys.*, 11:244–249.

Baldi, P., Caramellino, L., and Iovino, M. (1999). Pricing general barrier options: a numerical approach using sharp large deviations. *Math. Finance*, 9(4):293–322.

Bally, V. and Talay, D. (1996). The law of the euler scheme for stochastic differential equations i: Convergence rate of the distribution function. *Probab. Theory Related Fields*, 104(1):43–60.

Bangerth, W. and Rannacher, R. (2003). *Adaptive finite element methods for differential equations*. ISBN 3-7643-7009-2, Birkhäuser–Verlag.

Bank, R. and Weiser, A. (1985). Some a posteriori error estimators for elliptic partial differential equations. *Math. Comp.*, 44:283–301.

Becker, K. and Rannacher, R. (1996). A feed-back approach to error control in finite element methods: basic analysis and examples. *East-West J. Numer. Math.*, 4(4):237–264.

Becker, K. and Rannacher, R. (2001). An optimal control approach to a posteriori error estimation in finite element methods. *Acta Numerica*, 10:1–102.

Binev, P., Dahmen, W., and DeVore, R. (2002). Adaptive finite element methods with convergence rates. Technical report, Dept. of Math., Univ. of South Carolina.

Brenner, S. and Scott, L. (1994). *The mathematical theory of finite element methods*. Springer-Verlag New York.

Buchmann, F. (2003). Computing exit times with the euler scheme. Technical report, ETH, Zurich, Switzerland.

Calflisch, R. (1998). Monte carlo and quasi-monte carlo methods. *Acta Numerica*, pages 1–49.

Cohen, A., Dahmen, W., and DeVore, R. (2000). Adaptive wavelet method for elliptic operator equations:convergence rates. *Math. Comp.*, 70(233):27–75.

Dahlquist, G. and Björk, Å. (1974). *Numerical Method*. Prentice-Hall, New-York.

Dahlquist, G. and Björk, Å. (2003). *Numerical Mathematics*. SIAM, Philadelphia, http://www.mai.liu.se/∼akbjo/NMbook.html.

Dahmen, W. (1997). Wavelet and multiscale methods for operator equations. *Acta Numerica*, 6:55–228.

Deb, M.-K., Babuška, I., and Oden, J. (2001). Solution of stochastic partial differential equations using galerkin finite element technique. *Comput. Methods Appl. Mech. Engrg.*, 190:6359–6372.

DeVore, R. (1998). Nonlinear approximation. *Acta Numerica*, pages 51–150.

Dörfler, W. (1996). A convergent adaptive algorithm for poisson's equation. *SIAM J. Numer. Anal.*, 33(3):1106–1124.

Durett, R. (1964). *Probability: theory and examples*. Duxbury Press, Belmont CA.

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

Eriksson, K. and Johnson, C. (1988). An adaptive finite element method for linear elliptic problems. *Math. Comp.*, 50:361–383.

Estep, D. (1995). A posteriori error bounds and global error control for approximation of ordinary differential equations. *SIAM J. Numer. Anal.*, 32:1–48.

Evans, L. (1998). *Partial differential equations*. Graduate studies in Math. 19. American Math. Society, USA.

Ewens, W. and Grant, G. (2001). *Statistical methods in bioinformatics*. ISBN 0-387-95229. Springer.

Fishman, G. S. (1996). *Monte Carlo: Concepts, Algorithms, Applications*. Springer-Verlag New York.

Gobet, E. (2000). Weak approximation of killed diffusion using euler schemes. *Stochastic Process. Appl.*, 87(2):167–197.

Gobet, E. (2001). Euler schemes and half-space approximation for the simulation of diffusion in a domain. *ESAIM Probab. Statist.*, 5:261–297.

Goodman, J., Moon, K.-S., Szepessy, A., Tempone, R., and Zouraris, G. E. (2002). *Stochastic and partial differential equations with adaptive numerics*. Lecture note. Royal Institute of Technology, Sweden, http://www.nada.kth.se/kurser/kth/2D1269/aktuellt.html.

Gröbner, W. (1967). *Die Lie-Reihen und ihre Anwendungen*. VEB Deutscher Verlag der Wissenschaften, Berlin.

Harrier, E., Nørsett, S., and Wanner, G. (1993). *Solving Ordinary Differential Equations I*. Springer-Verlag.

Henrici, P. (1962). *Discrete Variable Methods in Ordinary Differential Equations*. John Wiley & Sons Inc.

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

Hofmann, N., Müller-Gronbach, T., and Ritter, K. (2001). The optimal discretization of stochastic differential equations. *J. Complexity*, 17:117–153.

Jansons, K. and Lythe, G. (2000). Efficient numerical solution of stochastic differential equations using exponential timestepping. *J. Stat. Phys.*, 100(5/6):1097–1109.

Johnson, C. (1987). *Numerical solution of partial differential equations by the finite element method*. Studentlitteratur, Lund, Sweden.

Johnson, C. (1988). 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.

Johnson, C. (1990). Adaptive finite element methods for diffusion and convection problems. *Comput. Methods Appl. Mech. Eng.*, 82:301–322.

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

Jouini, E., Cvitanić, J., and Musiela, M. (2001). *Option pricing, interest rates and risk management*. Cambridge University Press.

Karatzas, I. and Shreve, S. (1988). *Brownian motion and stochastic calculus*. Springer-Verlag New York.

Kloeden, P. and Platen, E. (1992). *Numerical solution of stochastic differential equations*. Applications of mathematics 23, Springer–Verlag.

Ladeveze, P. and Leguillon, D. (1983). Error estimate procedure in the finite element method and applications. *SIAM J. Numer. Anal.*, 20:485–509.

Lamba, H. and Stuart, A. M. (1998). Convergence results for the matlab ode23 routine. *BIT*, 38(4):751–780.

Logg, A. (2001). Multi-adaptive galerkin methods for odes. Licentiate thesis, Chalmers University of Technology, Sweden. ISSN 0347-2809.

Lorenz, E. N. (1963). Deterministic non-periodic flows. *J. Atmos. Sci.*, 20:130–141.

Majda, A., Timofeyev, I., and Eijnden, E. V. (2001). A mathematical framework for stochastic climate models. *Comm. Pure Appl. Math.*, 54(8):891–974.

Mannella, R. (1999). Absorbing boundaries and optimal stopping in a stochastic differential equation. *Phys. Lett. A*, 254(5):257–262.

Milstein, G. (1995). *Numerical integration of stochastic differential equations*. Mathematics and its Applications 313. Kluwer Academic Publishers, Dordrecht.

Moon, K.-S. (2001). Convergence rates of adaptive algorithms for deterministic and stochastic differential equations. Licentiate thesis, Royal Institute of Technology, Sweden. TRITA-NA-0134, ISBN 91-7283-196-0.

Moon, K.-S., Szepessy, A., Tempone, R., and Zouraris, G. E. (2001). *Hyperbolic differential equations and adaptive numerics*. Springer–Verlag. in Theory and numerics of differential equations, Durham 2000, Eds. J.F. Blowey, J.P. Coleman and A.W. Craig.

Morin, P., Nochetto, R. H., and Siebert, K. G. (2000). Data oscillation and convergence of adaptive fem. *SIAM J. Numer. Anal.*, 38(2):466–488.

Müller-Gronbach, T. (2000). The optimal uniform approximation of systems of stochastic differential equations. Technical report, Freie Universität Berlin. Nr. A-8-2000.

Novak, E. (1996). On the power of adaption. *J. Complexity*, 12:199–237.

Øksendal, B. (1998). *Stochastic differential equations*. Springer–Verlag, Berlin.

Petersen, W. P. and Buchmann, F. M. (2002). Solving dirichlet problems numerically using the feynman-kac representation. Technical report, ETH, Zurich, Switzerland. Research report no. 2002-01.

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

Schwab, C. and Todor, R. (2002). Finite element method for elliptic problems with stochastic data. Technical report, ETH, Zurich, Switzerland. Research report 2002-05.

Steele, J. (2001). *Stochastic calculus and financial applications*. Appl. of Math. 45. Springer-Verlag New York.

Stevenson, R. (2003). An optimal adaptive finite element method. Technical report, Univ. of Utrecht. preprint no. 1271.

Strauss, W. (1992). *Partial differential equations: An introduction*. John Wiley & Sons, Inc.

Stuart, A. (1997). Probabilistic and detereministic convergence proofs for software for initial value problems. *Numerical Algorithms*, 14:227–260.

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

Talay, D. (1995). *Simulation of Stochastic Differential Systems.* Lecture notes in physics 451, Springer. in Probabilistic Method in Applied Physics (P. Krée ; W. Wedig Eds.).

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

Tempone, R. (2002). *Numerical complexity analysis of weak approximation of stochastic differential equations.* PhD thesis, Royal Institute of Technology, Sweden.

Traub, J. and Werschulz, A. (1998). *Complexity and Information.* Cambridge University Press, Cambridge.

Werschulz, A. (1991). *The Computational Complexity of Differential and Integral Equations, An Information-Based Approach.* The Clarendon Press, Oxford University Press, New York.

Wilmott, P., Howison, S., and Dewynne, J. (1995). *The mathematics of financial derivatives.* Cambridge University Press.