Convergence Rates of Adaptive Algorithms for Stochastic and Partial Differential Equations

ERIK VON SCHWERIN

Licentiate Thesis
Stockholm, Sweden 2005

Nada är en gemensam institution mellan
Kungliga Tekniska högskolan och Stockholms universitet.

Convergence Rates of Adaptive Algorithms for Stochastic and Partial Differential Equations

ERIK VON SCHWERIN

Avhandling som med tillstånd av Kungliga Tekniska Högskolan
framlägges till offentlig granskning för avläggande av teknologie licentiatexamen
onsdagen den 9 februari 2005 kl 10.00
i sal D31, Huvudbyggnaden, Lindstedtsvägen 3,
Kungliga Tekniska högskolan, Stockholm.

ISBN 91-7283-959-7
TRITA-NA-0501
ISSN 0348-2952
ISRN KTH/NA/R--05/01--SE

Akademisk avhandling som med tillstånd av Kungl Tekniska Högskolan framlägges till offentlig granskning för avläggande av teknisk licentiatexamen onsdagen den 9 februari 2005 kl 10.00 i sal D31, Huvudbyggnaden, Kungl Tekniska Högskolan, Lindstedtsvägen 3, Stockholm.

ISBN 91-7283-959-7
TRITA-NA-0501
ISSN 0348-2952
ISRN KTH/NA/R--05/01--SE

Abstract

This work concerns the convergence of adaptive algorithms, based on a posteriori expansions of global errors, for numerical solution of differential equations, where the goal is to compute a functional of the solution. An adaptive algorithm aims to minimise the number of degrees of freedom to make the error in the functional less than a given tolerance. The number of degrees of freedom provides the convergence rate of the adaptive algorithm as the tolerance tends to zero. Provided that the computational work is proportional to the degrees of freedom this gives an estimate of the efficiency of the algorithm. The thesis consists of two papers, the first on the numerical approximation of partial differential equations, and the second on weak approximation of stochastic differential equations with barriers.

The first paper considers approximation of functionals of solutions to second order elliptic partial differential equations in bounded domains of $\mathbb{R}^d$, using isoparametric d-linear quadrilateral finite elements. For an adaptive algorithm, an error expansion with computable leading order term is derived. The computable error density, based on the dual weighted residual error representation,

$$\text{global error} = \sum_{\text{elements}} \text{error density} \cdot \text{mesh size}^{2+d},$$

using localised averages of second order difference quotients of the primal and dual finite element solutions, is proved to converge uniformly as the mesh size tends to zero. The proof splits the error into one part from elements with no edges on the initial mesh and without hanging nodes, and the remaining part, with hanging nodes and edges on the initial mesh, which is asymptotically negligible as the mesh size tends to zero. For each element an error indicator is defined by the computed error density multiplying the local mesh size to the power of $2 + d$. It is proved, using the uniform convergence of the error density, that the adaptive algorithm, based on successive subdivisions of elements, reduces the maximal error indicator with a factor or stops with the error asymptotically bounded by the tolerance using the optimal number of elements for an adaptive isotropic mesh, up to a problem independent factor. Here the optimal number of elements is proportional to the $d/2$ power of the $L^{\frac{d}{d+2}}$ quasi-norm of the error density, whereas a uniform mesh requires a number of elements proportional to the $d/2$ power of the larger L^1 norm of the same error density to obtain the same accuracy. For problems with multiple scales, in particular, these convergence rates may differ much, even though the convergence order may be the same. Numerical experiments for an elasticity problem with a crack and different variants of the averages show that the algorithm is useful in practice also for relatively large tolerances, much larger than the small tolerances needed to theoretically guarantee that the algorithm works well.

The second paper presents an adaptive algorithm for Monte Carlo Euler approximation of the expected value $E[g(X(\tau), \tau)]$ of a given function g depending on the solution X of an Itô stochastic differential equation and on the first exit time τ from a given domain. An error expansion with computable leading order term, for the approximation of $E[g(X(T))]$ with a fixed final time $T > 0$ in [Szepessy, Tempone and Zouraris, Comm. Pure and Appl. Math., 54, 1169-1214, 2001] is extended to the case with stopped diffusion. In the extension conditional probabilities are used to estimate the first exit time error, and difference quotients are used to approximate the initial data of the dual solutions. For the stopped diffusion problem the time discretisation error is of order $N^{-1/2}$ for a method with N uniform time steps. Numerical results show that the adaptive algorithm improve the time discretisation error to the order N^{-1} with N adaptive time steps.

ISBN 91-7283-959-7 • TRITA-NA-0501 • ISSN 0348-2952 • ISRN KTH/NA/R--05/01--SE

Acknowledgments

First of all I wish to thank Prof. Anders Szepessy for being an outstanding advisor: enthusiastic, encouraging and supportive.

To all my colleagues, present and former, at KTH I am grateful for creating a friendly and stimulating working place. Especially, I would like to thank Kyoung-Sook Moon and Raúl Tempone for guiding me into the research as a new PhD student. I also thank Daniel Appelö for reading a draft of the introduction and suggesting improvements.

Financial support has been provided by the Swedish Foundation for Strategic Research (SSF) and is gratefully acknowledged.

Contents

List of Papers

Paper I: Kyoung-Sook Moon, Erik von Schwerin, Anders Szepessy and Raul Tempone, *Convergence Rates for an Adaptive Dual Weighted Residual Finite Element Algorithm*, preprint (2004)

The author of this thesis contributed to the proof of convergence of the error density, performed the numerical computations, which influenced the proof, and contributed to the writing of parts of the paper, mainly on the numerical section.

Paper II: Anna Dzougoutov, Kyoung-Sook Moon, Erik von Schwerin, Anders Szepessy and Raul Tempone, *Adaptive Monte Carlo algorithms for stopped diffusion*, accepted in Lecture Notes in Computational Science and Engineering, Heidelberg, 2005. Springer Verlag.

The author of this thesis derived the error expansion for stopped diffusion in multi dimensional domains and wrote the paper, extending the one dimensional setting in [15], Paper V.

Chapter 1

Background

Differential equations are important in the formulation of mathematical models in many areas of science and engineering. Such models may be used to get an understanding of global properties of the system being modelled, from analytical solutions to the differential equations, from qualitative analysis of the dependence on model parameters, or from approximate numerical solutions for particular parameter values. However, mathematical models are also commonly used, not primarily to study global behaviour, but to predict the values of one or several scalar quantities of particular importance for the application at hand. Mathematically, such quantities correspond to functionals of the solutions to the differential equations. When the underlying differential equations are solved numerically, with finite computational resources, it is desirable to minimise the computational work for a given accuracy in the functional values. In goal oriented adaptivity, for a fixed numerical method of approximation, the degrees of freedom are adapted to both the differential equation and the functional in an attempt to minimise the work needed to meet the error tolerance in the goal functional. Both articles in this thesis aim at understanding of optimal convergence rates for goal oriented adaptive algorithms; one adaptive algorithm is studied in two different settings, namely that of deterministic elliptic partial differential equations in bounded d-dimensional domains using isoparametric d-linear quadrilateral finite element approximations, and that of Itô stochastic differential equations using the Euler Monte Carlo method.

Adaptive and Non-Adaptive Algorithms Consider the problem of computing an approximate value of $g(f)$ of a functional $g\colon X \to \mathbb{R}$ for $f \in F$, where F is a subset of the normed linear space X. Often a numerical method for this problem is on the form

$$g^n(f) = \phi^n(L_1(f), \ldots, L_n(f)), \tag{1.1}$$

where $L_i\colon X \to \mathbb{R}$ are linear functionals and $\phi^n\colon X \to \mathbb{R}$ is linear or nonlinear. The functionals can for example be function evaluations, $L_i(f) = f(x_i)$. The method g^n is called *non-adaptive* if the functionals L_i are the same for all $f \in F$. It is called *adaptive* if the choice of functionals L_i depends on f through the previously computed values $L_1(f), \ldots, L_{i-1}(f)$.

In information based complexity theory there is a general result by Bakhvalov and Smolyak comparing adaptive and non-adaptive methods for approximation of linear functionals, $g : X \to \mathbb{R}$, on a normed linear function space, X. The result states that for any adaptive method (1.1) using a fixed number of linear functionals L_i to approximate the linear g, defined on a symmetric convex subset F of X, there is a linear non-adaptive method whose

maximal error, on F with the same number of linear functionals, is as small that of the adaptive method. A more detailed formulation can be found the overview article [21] by Novak.

How does the adaptive algorithms for computation of linear functionals of solutions to differential equations which are considered here relate to the result of Bakhvalov and Smolyak? The point of view is different in that a fixed method, for example a finite element method of given order, is considered with the aim to construct an adaptive mesh refinement algorithm for that method. Also, in contrast to keeping the number of steps in the algorithm fixed, the aim here is to create an algorithm where the number of steps, as a function of the the error tolerance is close to optimal as the tolerance tends to zero.

Consider a numerical method based on uniform discretisation of a d-dimensional domain with element size h and with approximation error $\Theta(h^p)$, as $h \to 0$, using the notation that $f = \Theta(g)$ if and only if $f = \mathcal{O}(g)$ and $g = \mathcal{O}(f)$. Making the error less than a tolerance TOL requires $\Theta(\mathrm{TOL}^{-d/p})$ elements. Assuming that the work is proportional to the number of time steps the performance of the method can be expressed in terms of the tolerance, as $\mathrm{TOL} \to 0$. This measure of the efficiency is natural to extend to adaptive algorithms as illustrated in a simple setting in the next example.

Example: Numerical Integration The assumption of a convex domain of definition, F, for the adaptive and non-adaptive methods in the result of Bakhvalov and Smolyak mentioned above is important. To illustrate this and to show how convergence rates for adaptive algorithms are measured here, consider the linear functional given by an integral of a known function, $g(f) = \int_0^T f(t)\,dt$, and let the method of numerical integration be the left point rule (forward Euler). Discretise the time interval $[0, T]$ into N subintervals $0 = t_0 < t_1 < \cdots < t_N = T$ with steps $\Delta t_n := t_{n+1} - t_n$. With $\bar{g}$ denoting the numerical approximation of $g(f)$ the global discretisation error becomes

$$g(f) - \bar{g} = \sum_{n=0}^{N-1} \rho_n (\Delta t_n)^2 + \text{ higher order terms}, \tag{1.2}$$

where the error density function ρ is given by $\rho_n := \frac{df}{dt}(t_n)/2$. As an example of a non-adaptive method consider uniform Δt. Using that the number of time steps is

$$N(\Delta t) = \int_0^T \frac{1}{\Delta t(\tau)}\, d\tau, \tag{1.3}$$

the number N_u of uniform steps to reach a given level of accuracy TOL is asymptotically proportional to TOL^{-1} with the L^1-norm of the function ρ in the proportionality constant,

$$N_u \simeq \frac{T}{\mathrm{TOL}} \|\rho\|_{L^1(0,T)}, \tag{1.4}$$

provided that ρ has constant sign. When the number of steps in (1.3) is minimised with the accuracy constraint that the leading order of (1.2) is TOL, the optimal distribution of time steps is

$$\rho_n \Delta t_n^2 = \text{ constant for all } n.$$

With this choice the number N_a of adaptive steps becomes proportional to TOL^{-1} with the smaller $L^{\frac{1}{2}}$-quasi-norm of the error density as the proportionality constant,

$$N_a \simeq \frac{1}{\mathrm{TOL}} \|\rho\|_{L^{\frac{1}{2}}(0,T)}. \tag{1.5}$$

Since the Euler method uses one function evaluation per step the asymptotical number of steps (1.4) and (1.5) give the convergence rates of the Euler method using uniform and optimal adaptive time steps respectively.

Take for example the integrand $f(t) = 1/\sqrt{t+\epsilon}$ for a small positive parameter $\epsilon \ll T$. Since $\rho(t) = \frac{-1}{4(t+\epsilon)^{3/4}}$ the number of uniform steps becomes

$$N_u \simeq \frac{T/4}{\mathrm{TOL}} \int_0^T \frac{dt}{(t+\epsilon)^{3/2}} \approx \frac{T/4}{\mathrm{TOL}} \frac{1}{\epsilon^{1/2}},$$

while the number of adaptive time steps is smaller,

$$N_a \simeq \frac{1/4}{\mathrm{TOL}} \left(\int_0^T \frac{dt}{(t+\epsilon)^{3/4}} \right)^2 \approx \frac{4\sqrt{T}}{\mathrm{TOL}}.$$

The smaller multiple of 1/TOL with an adaptive approach captures the multiple scales introduced by $\epsilon \ll T$. In this example, the integrand can also be viewed as an approximation of the singular $1/\sqrt{t}$, in which case the parameter must be taken $\epsilon^{1/2} = o(\mathrm{TOL})$, so that $N_a/N_u \to 0$ as $\mathrm{TOL} \to 0$.

If $F = \{f \ : \ ||f'||_{L^{\frac{1}{2}}} < M\}$ for a constant M, then the integrand in the example above is in F for some ϵ depending on M. Note that F is non-convex so that the result by Bakhvalov and Smolyak does not apply to the problem of computing $g(f)$ for $f \in F$. In this class of integrands the choice of uniform steps in the non-adaptive method is motivated by considering integrands $f_s(t) = 1/\sqrt{|t-s|+\epsilon}$ with ϵ just large enough for f_s to be in F for all $s \in [0,T]$. However, it is not always the case that optimal non-adaptive discretisations for a fixed method are uniform, as is illustrated in the next example.

Example: Corner Singularity for an Elliptic Partial Differential Equation Let u, in a domain Ω with a crack as in Figure 1.1, be the solution of the Laplace equation

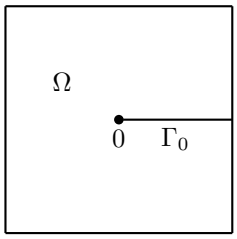

Figure 1.1: Domain with a crack

$$\begin{aligned} -\Delta u &= 0, && \text{in } \Omega, \\ u &= 0, && \text{on } \Gamma_0, \\ u &= f, && \text{on } \partial\Omega \setminus \Gamma_0, \end{aligned}$$

and let $g(f)$, viewed as a linear functional of the Dirichlet boundary values f, be given by the integral of u,

$$(u,1) = \int_\Omega u \, dx.$$

Even if the boundary conditions are taken from a class of smooth functions the solution u will in general have a form like $u(r,\theta) = \sqrt{r}\alpha(\theta) + \beta(r,\theta)$ in polar coordinates with smooth α and β close to the tip of the crack; see for example the textbook [13] by Johnson. For a given solution method of the boundary value problem, for example bilinear finite elements on a grid with square elements and hanging nodes, the a priori information of the singularity of the derivative of the exact solution can be used to construct non-uniform non-adaptive meshes for this particular geometry. On the other hand, in applications where the domain Ω varies adaptive methods allow the mesh to automatically adapt to the geometry without the detailed a priori knowledge of the solution. This is the situation considered in Paper I.

Chapter 2

An Adaptive Algorithm

This chapter describes an adaptive algorithm for computing approximate solutions to problems which can abstractly be stated as:

$$\begin{aligned} &\text{compute the functional } g(u) \\ &\text{where } u \text{ solves an initial or boundary value problem} \\ &\text{for a differential equation in a } d\text{-dimensional domain } \Omega. \end{aligned} \tag{2.1}$$

For a given method of numerical approximation of u, based on discretisation of the domain Ω, the algorithm constructs the final discretisation by iterative refinements of an initial mesh; the algorithm presupposes an expansion of the error in the scalar quantity $g(u)$ of the form

$$\text{Global error} = \sum \text{local error} \cdot \text{weight} + \text{higher order error}, \tag{2.2}$$

depending on the approximation method and on the problem; compare (1.2) in the numerical integration example. The leading order terms must be computable using information on the current mesh. The weight describes the influence of changes in the differential equation on the functional of its solution. The goal of the adaptive algorithm is to, for the given approximation method, approximate $g(u)$ using an adapted mesh with a minimal number of intervals (elements) for error less than a given tolerance.

Concrete formulations of the abstract (2.1) are for example

- the computation of $g(u)$ where u solves an ordinary differential equation

$$\begin{aligned} \frac{du}{dt}(t) &= a(t, u(t)), \qquad 0 < t < T, \\ u(0) &= u_0, \end{aligned} \tag{2.3}$$

  with flux $a\colon [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ and an approximate solution u_h is obtained by any p:th order numerical method using $u_h(0) = u_0$ and $\Omega = [0,T]$ is discretised into $0 = t_0 < t_1 < \cdots < t_N = T$.

- the computation of $g(u)$ where u solves an elliptic partial differential equation in a bounded open domain $\Omega \subset \mathbb{R}^d$ and an approximate solution u_h is obtained using a given finite element method; see the example on page 3.

Equidistribution of Errors Consider (2.1) in a domain of dimension d with a given approximation method of order p. Assume that an asymptotic error expansion (2.2) on the form

$$\text{error} \simeq \sum_n \rho_n h_n^{p+d}$$

is known, where h is the local mesh size, of the non stretched element, and ρ is independent of h. The number of elements that corresponds to a mesh with size h can be determined by

$$N(h) \equiv \int_\Omega \frac{dx}{h^d(x)}. \tag{2.4}$$

If the sign of the error density varies a very small set of elements may give an error in the functional that is close to zero due to cancellation of error contributions of opposite sign. Thus the optimal mesh may consist of very few elements, but it seems difficult to exploit the cancellation of errors when constructing the mesh. Disregarding the possible cancellation by minimising the number of elements N in (2.4) under the constraint

$$\sum_{n=1}^{N} |\rho_n| h_n^{d+p} = \int_\Omega |\rho(x)| h^p dx = \text{TOL},$$

gives the optimum

$$|\rho|(h^*)^{d+p} = \text{constant} \tag{2.5}$$

with corresponding mesh size function

$$h^* \equiv \frac{\text{TOL}^{\frac{1}{p}}}{|\rho|^{\frac{1}{d+p}}} \left(\int_\Omega |\rho(x)|^{\frac{d}{d+p}} dx \right)^{-\frac{1}{p}}. \tag{2.6}$$

This condition is optimal only for density functions ρ with one sign. Moreover, in higher dimension, $d > 1$, it is optimal only for meshes with non stretched elements, that is elements such that each element is described by one element size h.

The adaptive refinement algorithm, described in a generic deterministic form in Algorithm 1 below, is designed to approximate the optimal equidistribution of error contributions (2.5). With $[k]$ denoting quantities on the k:th mesh in the refinement sequence, the accepted mesh k_{stop} ideally fulfils

$$\hat{\rho}_n[k_{\text{stop}}](h_n[k_{\text{stop}}])^{d+p} \approx \frac{\text{TOL}}{N[k_{\text{stop}}]}, \qquad n = 1, 2, \ldots, N[k_{\text{stop}}],$$

where $\hat{\rho}_n$ is a computable approximation of the unsigned error density, $|\rho|$. Thus, after calculating $\hat{\rho}[k]$ from computed approximate primal and dual solutions on level k, the algorithm refines all elements with error indicators $\bar{r}_n[k] \equiv \hat{\rho}_n[k](h_n[k])^{d+p} > s_1 \text{TOL}/N[k]$, where $s_1 \approx 1$ is a constant. The maximal error indicator may reduce slowly when most $\bar{r}_n$ are small, $\bar{r}_n[k] \leq s_1 \text{TOL}/N[k]$, leading to many refinements; to avoid this the refinements stop when all $\bar{r}_n[k] \leq S_1 \text{TOL}/N[k]$ for a constant $S_1 > s_1$. In summary, the new element sizes $h[k+1]$ are obtained from $h[k]$ by:

Algorithm 1: Refinement and stopping

forall intervals (elements) $n = 1, 2, \ldots, N[k]$ **do**
  $\bar{r}_n[k] \equiv \hat{\rho}_n[k](h_n[k])^{d+p}$
  if $\bar{r}_n[k] > s_1 \mathrm{TOL}/N[k]$ **then**
    mark interval (element) n for division
  end
end
if $\max_{1 \le n \le N[k]} \bar{r}_n[k] \le S_1 \mathrm{TOL}/N[k]$ **then**
  stop the refinements
else
  divide every marked interval (element) into 2^d sub intervals (elements)
end

The optimality condition (2.5) was obtained from the assumption of a limit error density, ρ, and the adaptive algorithm constructed to approximate (2.5) using a computed approximate error density $\hat{\rho}$. This is meaningful if $\hat{\rho}$ converges to $|\rho|$ as $\mathrm{TOL} \to 0$. Thus for any particular application the proof of this convergence is crucial for the theoretical analysis of the algorithm. It is possible to analyse the important properties of stopping, accuracy and efficiency of the algorithm in terms of convergence of $\hat{\rho}$.

Stopping of Algorithm 1 Assume the convergence of $\hat{\rho}$ where this positive approximate error density is bounded away from zero by a lower bound δ which tends to 0 with TOL as

$$\delta = \mathrm{TOL}^{\gamma}, \tag{2.7}$$

for a positive parameter γ which depends on the application. Then the change in the density $\hat{\rho}(K)[k]$ in an element K on refinement level k from its value on the parent element on a previous level, $p(K,k)$ can be bounded; it follows from the convergence assumption and (2.7) and an additional assumption (2.12) on the initial mesh size that there exist functions $\hat{c}$ and $\hat{C}$, close to 1 for sufficiently refined meshes, such that

$$\hat{c}(K) \le \frac{\hat{\rho}(K)[p(K,k)]}{\hat{\rho}(K)[k]} \le \hat{C}(K), \tag{2.8a}$$

$$\hat{c}(K) \le \frac{\hat{\rho}(K)[k-1]}{\hat{\rho}(K)[k]} \le \hat{C}(K). \tag{2.8b}$$

The lower bound on the qoutients here can be used, together with the refinement and stopping criteria in Algorithm 1, to prove the following theorem, which shows that the slow reduction of the maximal error indicator is avoided for S_1 chosen suitably larger than s_1.

Theorem (Stopping). *With the adaptive refinement and stopping strategy in Algorithm 1, assume that $\hat{c}$ satisfies* (2.8a)–(2.8b)*, for the elements or time steps corresponding to the maximal error indicator on each refinement level, and that*

$$S_1 \ge \frac{2^d}{\hat{c}} s_1, \quad 1 > \frac{\hat{c}^{-1}}{2^{d+p}}.$$

Then each refinement level either decreases the maximal error indicator with the factor

$$\max_{1 \le n \le N[k+1]} \bar{r}_n[k+1] \le \frac{\hat{c}^{-1}}{2^{d+p}} \max_{1 \le n \le N[k]} \bar{r}_n[k], \tag{2.9}$$

or stops the algorithm.

Accuracy of Algorithm 1 By construction the adaptive algorithm guarantees that the estimate of the global error is bounded by a given error tolerance, TOL. Is also the true global error bounded by TOL asymptotically? The stopping criterion in Algorithm 1 gives an upper bound of the error indicators, which together with the assumed convergence of $\hat{\rho}$ leads to an asymptotical bound of the global error of the kind

$$\limsup_{\mathrm{TOL}\to 0+} \left(\mathrm{TOL}^{-1} \big| g(u) - g(u_h) \big| \right) \leq S_1,$$

where u is the exact solution and u_h the computed approximation. See Theorem 3.2 in Paper I for a precise formulation for the dual weighted residual finite element algorithm considered there for second order elliptic partial differential equations.

Efficiency of Algorithm 1 The goal of the adaptive algorithm is to determine a mesh with a minimal number of elements or time steps, N, for the specified accuracy. The optimality condition (2.6) in the equation (2.4) for N gives the optimal number of adaptive elements

$$N^{\mathrm{opt}} = \int_\Omega \frac{dx}{(h^*(x))^d} = \frac{1}{\mathrm{TOL}^{\frac{d}{p}}} \left(\int_\Omega |\rho[k](x)|^{\frac{d}{d+p}} dx \right)^{\frac{d+p}{p}} = \frac{1}{\mathrm{TOL}^{\frac{d}{p}}} \|\rho\|^{\frac{d}{p}}_{L^{\frac{d}{d+p}}}. \tag{2.10}$$

With a uniform mesh, constant mesh size h, the number of elements, N^{uni}, to achieve $\sum_{i=1}^N |\rho_i| h^{d+p} = \mathrm{TOL}$ becomes instead

$$N^{\mathrm{uni}} = \int_\Omega \frac{dx}{h^d(x)} = \frac{\int_\Omega dx}{\mathrm{TOL}^{\frac{d}{p}}} \left(\int_\Omega |\rho[k](x)| dx \right)^{\frac{d}{p}} = \frac{\int_\Omega dx}{\mathrm{TOL}^{\frac{d}{p}}} \|\rho\|^{\frac{d}{p}}_{L^1}. \tag{2.11}$$

Since, by Jensen's inequality, $\|f\|_{L^{\frac{d}{d+p}}} \leq (\int_\Omega dx)^{\frac{p}{d}} \|f\|_{L^1}$, the asymptotic constant multiplying $1/\mathrm{TOL}^{d/p}$ in the convergence order is smaller for the adaptive method than the uniform element size method. For problems with multiple scale solutions the difference may be significant; compare the integration example in Chapter 1.

From the refinement criterion in Algorithm 1, a lower bound of the error indicators follows for the refined parent error indicator. This, together with the assumption that upper bound of the ratios of the error density (2.8a)–(2.8b) holds for all elements on the final mesh, and an assumption

$$h_K[1] = \Theta(\mathrm{TOL}^s), \tag{2.12}$$

on the initial mesh size to guarantee that, for sufficiently small TOL, all elements on the initial mesh are refined, can be used to show that Algorithm 1 generates a mesh which is optimal, (2.10), up to a multiplicative constant independent of the data,

$$(\mathrm{TOL}^{\frac{d}{p}} N) \leq C \, \|\hat{C}\hat{\rho}\|^{\frac{d}{p}}_{L^{\frac{d}{d+p}}} \leq C \left(\max_{x\in D} \hat{C}(x)^{\frac{d}{p}} \right) \|\hat{\rho}\|^{\frac{d}{p}}_{L^{\frac{d}{d+p}}}, \tag{2.13}$$

with $C \leq (\frac{2^{d+p}}{s_1})^{\frac{d}{p}}$. See Theorem 3.5 in Paper I for a precise formulation in a specific case.

Earlier Works on Algorithm 1 The results presented in this thesis follow previous work on the same adaptive algorithm in other precise settings of (2.1). For an ordinary differential equation(2.3), an error expansion (2.2) is derived by a variational principle in [17] and the convergence properties of the adaptive algorithm are studied in [18].

The works [24, 19] treat the weak approximation of an Itô Stochastic differential equation of the form

$$dX_k(t) = a_k(t, X(t))dt + \sum_{\ell=1}^{\ell_0} b_k^\ell(t, X(t))dW^\ell(t), \quad t > 0, \tag{2.14}$$

where $k = 1, \ldots, d$, and $(X(t;\omega))$ is a stochastic process in $\mathbb{R}^d$, with independent one dimensional Wiener processes $W^\ell(t;\omega)$, $\ell = 1, \ldots, \ell_0$. The functions $a(t,x) \in \mathbb{R}^d$ and $b^\ell(t,x) \in \mathbb{R}^d$, $\ell = 1, \ldots, \ell_0$, are given drift and diffusion fluxes.

Weak approximation of the stochastic differential equation by the Euler Monte Carlo method approximates the expected value $E[g(X(T))]$ of a functional of the solution with a sample average of $g(\overline{X}(T))$, where $\overline{X}(t_n)$ are identically distributed samples of a discrete time approximations of $X(t_n)$ in the times $0 = t_0 < t_1 < \cdots < t_N = T$ using the Euler method,

$$\overline{X}(t_{n+1}) - \overline{X}(t_n) = a(t_n, \overline{X}(t_n))\Delta t_n + \sum_{\ell=1}^{\ell_0} b^\ell(t_n, \overline{X}(t_n))\Delta W_n^\ell, \tag{2.15}$$

for $\Delta t_n \equiv t_{n+1} - t_n$, $\Delta W_n^\ell \equiv W^\ell(t_{n+1}) - W^\ell(t_n)$, $n = 0, 1, 2, \ldots, N-1$. The aim of the adaptive algorithm is to, for a given error tolerance, obtain

$$\left| E[g(X(T))] - \frac{1}{M}\sum_{j=1}^{M} g(\overline{X}(T;\omega_j)) \right| \leq \text{TOL} \tag{2.16}$$

with a probability close to one, and doing this with minimal computational work, proportional to the total number of stochastic time steps N_{ω_j} for the M realisations.

The error in (2.16) splits naturally into two parts,

$$\begin{aligned} &E[g(X(T))] - \frac{1}{M}\sum_{j=1}^{M} g(\overline{X}(T;\omega_j)) \\ &= \big(E[g(X(T)) - g(\overline{X}(T))]\big) + \left(E[g(\overline{X}(T))] - \frac{1}{M}\sum_{j=1}^{M} g(\overline{X}(T)) \right), \end{aligned} \tag{2.17}$$

corresponding to time discretisation error and statistical error.

Talay and Tubaro derived a priori estimates of the error (2.16) in [25]. This is modified to an error expansion with a posteriori computable leading order term in [24] using computable stochastic flows and discrete dual backward problems. In [19] convergence of algorithms based on the error expansion is analysed in terms of stopping, accuracy and efficiency using both stochastic and deterministic time steps in the control of the time discretisation error. With stochastic adaptive time steps Algorithm 1 controls the refinements and the stopping in the computation of each sample path. In those time steps that are marked for refinement the sample value of the Wiener processes in the mid points are simulated using Brownian bridges

$$W^l\left(\frac{t_n + t_{n+1}}{2}\right) = \frac{1}{2}\left(W^l(t_n) + W^l(t_{n+1})\right) + z_n^l, \tag{2.18}$$

where z_n^l are independent normally distributed random variables with mean 0 and variance $(t_{n+1} - t_n)/4$, independent also of previous $W^l(t_j)$.

The statistical error, governed by the Central Limit Theorem, is asymptotically bounded by $c_0\overline{\sigma}/\sqrt{M}$ where $\overline{\sigma}$ is the sample average of the standard deviation of $g(\overline{X}(T))$ and c_0 is a positive constant for a confidence interval.

Paper II extends the earlier work on stochastic differential equations to stopped diffusion problems; see Section 3.4.

Chapter 3

Summary of Papers

3.1 An Adaptive Dual Weighted Residual Finite Element Algorithm

Consider an adaptive finite element algorithm to approximate linear functionals

$$g(u) = (u, F) \equiv \int_\Omega uF \, dx$$

of multi scale solutions, $u : \Omega \to \mathbb{R}$, of the second order elliptic partial differential equation

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

in a given open bounded domain $\Omega \subset \mathbb{R}^d$ with Dirichlet boundary condition $u|_{\partial\Omega} = 0$. The weak form of (3.1) is

$$(a\nabla u, \nabla v) = (f, v), \qquad \forall v \in H_0^1(\Omega),$$

where the Sobolev space $H_0^1(\Omega)$ is the Hilbert space of functions on Ω, vanishing on $\partial\Omega$, such that the first derivatives are in $L^2(\Omega)$. The finite element approximate solution, u_h, solves the corresponding discrete variational form,

$$(a\nabla u_h, \nabla v) = (f, v), \qquad \forall v \in V_h, \tag{3.2}$$

where V_h is a finite dimensional subspace of $H_0^1(\Omega)$; see for example [4] by Brenner and Scott. For the purpose of Paper I, V_h is the set of continuous piecewise isoparametric bilinear quadrilateral functions in $H_0^1(\Omega)$, using an adaptive quadrilateral mesh with hanging nodes. In the dual weighted residual method, see [2] by Becker and Rannacher, a dual function $\varphi \in H_0^1(\Omega)$ defined by

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

is introduced to describe the sensitivity of the functional value on the fluctuations in the solution to the partial differential equation. From the definition of the dual, the error in the functional value is

$$(u - u_h, F) = (a\nabla(u - u_h), \nabla\varphi) = (\mathcal{R}(u_h), -\varphi),$$

with the residual $\mathcal{R}(v) = -\mathrm{div}(a\nabla v) - f$, defined as a distribution on $H^{-1}(\Omega)$ for $v \in H_0^1(\Omega)$. By this and the orthogonality (3.2) applied to $\pi\varphi \in V_h$, where $\pi\varphi$ is the nodal interpolant on V_h, the error in the functional has the dual weighted residual representation

$$(u - u_h, F) = (\mathcal{R}(u_h), \pi\varphi - \varphi)\,. \tag{3.3}$$

Taking inspiration from [9], by Eriksson et.al., and [2] Paper I contains a derivation of a computable approximation $\sum_K \bar{\rho}_K h_K^{d+2}$ of (3.3) for adaptive meshes with at most one hanging node per edge where the refinements of the initial elements are obtained by successive division of elements into 2^d, so that the transformation of each initial element to the reference tensor element maps the corresponding sub mesh to a tensor hanging node mesh.

The new difficulty when elliptic partial differential equations are considered instead of ordinary differential equations is the analysis of the convergence of the error density.

In contrast to the common aproach to derive an a posteriori error estimate, the aim here is to derive a uniformly convergent error density with computable leading order term and formulate an adaptive algorithm with proved convergence rates. The works [1] by Babuška and Vogelius, [8] by Dörfler, and [20] by Morin, Nochetto and Siebert study the convergence of adaptive algorithms for finite element approximations of partial differential equations.

There are also recent work on the convergence rates of adaptive algorithms for numerical solution of elliptic partial differential equations, in terms of the computational work. DeVore [6] shows the efficiency of adaptive approximation of functions, including wavelet expansions. In [5] Cohen, Dahmen and DeVore uses an adaptive N-term wavelet-based approximation algorithm and proves that it produces a solution which is asymptotically optimal in the energy norm error for linear coercive elliptic problems. In [3] by Binev, Dahmen and DeVore and [23] by Stevenson, the ideas in [20] are extended to prove optimal energy norm error estimates using piecewise linear elements for the Poisson equation.

3.2 Paper I: Convergence Rates for an Adaptive Dual Weighted Residual Finite Element Algorithm

This paper establishes basic convergence rates for a dual weighted residual finite element algorithm using isoparametric d-linear quadrilateral finite element approximation to functionals of solutions second order elliptic partial differential equations in open bounded domains of $\mathbb{R}^d$.

Section 2 describes an expansion of the error in the functional, based on (3.3), which is shown in Theorem 2.1 to be uniformly convergent as the mesh size tends to zero for smooth primal and dual solutions. The computable error density using localised averages of second order difference quotients of the primal and dual solutions, gives the leading order term of the error expansion; see Corollary 2.2.

The hanging node constraint implies that the refinement step in Algorithm 1 on page 7 must be modified to include a recursive marking of all neighbours that would otherwise violate the constraint. With that modification, the algorithm is analysed Section 3 following the outline in Chapter 2.

Section 4 presents numerical results for a simplified elasticity problem related to a problem with round corner of small radius introducing a small scale in the solution. The results show that the adaptive algorithm is more efficient for this problem than uniform refinements.

Paper I has entry [16] in the bibliography.

3.3 An Adaptive Algorithm for the Stopped Diffusion Problem

Here the objective is to compute adaptive approximations of an expected value

$$E[g(X(\tau), \tau)] \tag{3.4}$$

of a given function, $g : D \times [0,T] \to \mathbb{R}$, where the stochastic process X solves a stochastic differential equation (2.14) and τ is the first exit time

$$\tau := \inf\{0 < t : (X(t), t) \notin D \times (0,T)\}$$

from a given open domain $D \times (0,T) \subset \mathbb{R}^d \times (0,T)$. These so called barrier problems have applications in physics and finance, for example when pricing barrier options.

The expected value (3.4) is approximated by a sample average of $g(\overline{X}(\overline{\tau}), \overline{\tau})$, where $(\overline{X}, \overline{\tau})$ is an Euler approximation (2.15) of (X, τ) using stochastic adaptive time steps. Like in (2.17) the global error using M realizations, splits into two parts

$$\begin{aligned} &E[g(X(\tau),\tau)] - \frac{1}{M}\sum_{j=1}^{M} g(\overline{X}(\overline{\tau};\omega_j),\overline{\tau}) \\ &= \left(E[g(X(\tau),\tau) - g(\overline{X}(\overline{\tau}),\overline{\tau})]\right) + \left(E[g(\overline{X}(\overline{\tau}),\overline{\tau})] - \frac{1}{M}\sum_{j=1}^{M} g(\overline{X}(\overline{\tau};\omega_j),\overline{\tau})\right), \end{aligned}$$

corresponding to time discretisation error and statistical error.

The main difficulty introduced by the barrier is that the continuous path may exit D even though a discrete approximate solution does not cross the boundary of D. The hitting of the boundary causes the time discretisation error for the Monte Carlo Euler method with N uniform time steps to be of order $N^{-1/2}$ instead of N^{-1} without stopping boundary in $\mathbb{R}^d \times [0,T)$; see [10] by Gobet.

In Mannella [14] and Jansons and Lythe [12] the order N^{-1}, using N uniform time steps is recovered by deciding in each time step whether the continuous path exits a half space domain by simulating a stochastic outcome. In [11] Gobet proves the convergence rate N^{-1} for a similar method, under suitable assumptions including smooth boundary. These methods are efficient when the exit probabilities can be computed accurately, for example when the domain is a half space or has a smooth boundary which can be approximated by tangent planes, but not for a boundary with corners.

3.4 Paper II: Adaptive Monte Carlo Algorithms for Stopped Diffusion

This paper, inspired by Petersen and Buchmann [22], uses an alternative approach to the uniform time step methods of [14], [12] and [11]. The time steps are chosen adaptively for each sample path, decreasing close to the barrier. The advantage of this method is that the exit probability need not be computed accurately, which is difficult for complicated domains D. Section 2 contains a derivation of an expansion of the error with computable leading order term, which is an extension of the corresponding error expansion in [24] for the approximation of $E[g(X(T))]$ for fixed T and $D = \mathbb{R}^d$. The extension uses a conditional probability to estimate the first exit time and it initialises the dual solutions on the barrier with difference approximations of partial derivatives. Section 3 presents an adaptive algorithm based on the estimates in Section 2. Numerical results presented in Section 4 show that the algorithm recovers the time discretisation error of order N^{-1}, for N adaptive time steps.

Bibliography

[1] I. Babuška, and M. Vogelius, *Feedback and adaptive finite element solution of one-dimensional boundary value problems*, Numer. Math. **44** (1984), no. 1, 75–102.

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

[3] P. Binev, W. Dahmen and R. DeVore, *Adaptive Finite Element Methods with convergence rates*, Numer. Math. **97** (2004), no. 2, 219–268.

[4] S.C. Brenner and L.R. Scott, *The mathematical theory of finite element methods*, Texts in Applied Mathematics 15, Springer–Verlag, New York, 1994.

[5] A. Cohen, W. Dahmen and R. DeVore, *Adaptive wavelet methods for elliptic operator equations: convergence rates*, Math. Comp., **70** (2001), no. 233, 25–75

[6] R. A. DeVore, Nonlinear approximation, Acta Numerica, (1998), 51-150.

[7] A. Dzougoutov, K.-S. Moon, E. von Schwerin, A. Szepessy and R. Tempone, *Adaptive Monte Carlo algorithms for stopped diffusion*, Accepted in Lecture Notes in Computational Science and Engineering, Heidelberg, 2005, Springer Verlag, – appended as Paper II

[8] W. Dörfler, A convergent adaptive algorithm for Poisson's equation, SIAM J. Numer. Anal. **33** (1996), no. 3, 1106-1124.

[9] K. Eriksson, D. Estep, P. Hansbo and C. Johnson, *Introduction to Adaptive Methods for Differential Equations*, Acta Numerica, (1995), 105–158.

[10] Gobet E.: Weak approximation of killed diffusion using Euler schemes. Stochastic Process. Appl. **87**, no. 2, 167–197, (2000)

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

[12] Jansons K.M. and Lythe G.D.: Efficient numerical solution of stochastic differential equations using exponential timestepping. J. Stat. Phys. **100**, no. 5/6, 1097–1109, (2000)

[13] C. Johnson, *Numerical solution of partial differential equations by the finite element method*, Studentlitteratur, New York, 1987.

[14] Mannella R.: Absorbing boundaries and optimal stopping in a stochastic differential equation. Phys. Lett. A **254**, no. 5, 257–262, (1999)

[15] K.-S. Moon, *Adaptive Algorithms for Deterministic and Stochastic Differential Equations*, Doctoral thesis, ISBN 91-7283-553-2, Royal Institute of Technology, 2003

[16] K.-S. Moon, E. von Schwerin, A. Szepessy and R. Tempone, *Convergence Rates for an Adaptive Dual Weighted Residual Finite Element Algorithm*, preprint – appended as Paper I, (2004),

[17] K.-S. Moon, A. Szepessy, R. Tempone and G.E. Zouraris, *A variational principle for adaptive approximation of ordinary differential equations*, Numer. Math. **96** (2003), 131–152.

[18] K.-S. Moon, A. Szepessy, R. Tempone and G.E. Zouraris, *Convergence rates for adaptive approximation of ordinary differential equations*, Numer. Math. **96** (2003), 99–129.

[19] K.-S. Moon, A. Szepessy, R. Tempone and G.E. Zouraris: *Convergence rates for adaptive weak approximation of stochastic differential equations*, preprint, (2004)

[20] P. Morin, R. Nochetto and K.G. Siebert, *Data oscillation and convergence of adaptive FEM*, SIAM J. Numer. Anal., **38** (2000), 466–488.

[21] E. Novak, *On the Power of Adaption*, Journal of Complexity, **12** (1996), 199–237.

[22] Petersen W.P. and Buchmann F.M.: Solving Dirichlet problems numerically using the Feynman-Kac representation. BIT **43**, no. 3, 519–540, (2003)

[23] R. Stevenson, *An optimal adaptive finite element method*, Preprint (2004), To appear in SIAM J. Numer. Anal.

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

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