

Mathematics Monograph Series **6**

# **Adaptive Computations: Theory and Algorithms**

Edited by Tao Tang Jinchao Xu

(自适应计算：理论与算法)



SCIENCE PRESS  
Beijing

Mathematics Monograph Series 6

Edited by Tao Tang Jinchao Xu

# **Adaptive Computations: Theory and Algorithms**

(自适应计算：理论与算法)



SCIENCE PRESS  
Beijing

Responsible Editor: Fan Qingkui

Copyright© 2007 by Science Press  
Published by Science Press  
16 Donghuangchenggen North Street  
Beijing 100717, China

Printed in Beijing

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of the copyright owner.

ISBN 978-7-03-018421-4 (Beijing)

# Preface

Adaptive grid methods are among the most important classes of numerical methods for partial differential equations that arise from scientific and engineering computing. The study of this type of methods has been very active in recent years for algorithm design, theoretical analysis and applications to practical computations. This volume contains a number of self-contained articles for adaptive finite element and finite difference methods, which is aimed to provide some introduction materials for graduate students and junior researchers and a collection of references for researchers and practitioners.

These articles mostly grew out of the lectures notes that were given in *Summer Workshops on Adaptive Method, Theory and Application* organized by Tao Tang, Jinchao Xu and Pingwen Zhang in Peking University, China, during June 20 – August 20, 2005. This summer school was aimed to provide a comprehensive and up-to-date presentation of modern theories and practical applications for adaptive computations. The main lecturers of the Summer School include Weizhang Huang of University of Kansas, Natalia Kopteva of University of Limerick, Zhiping Li of Peking University, John Mackenzie of Strathclyde University, Jinchao Xu of Penn State University, Paul Zegeling of Utrecht University, and Zhimin Zhang of Wayne State University. Other lecturers include Tao Tang, Xiaoping Wang of HKUST, Huazhong Tang and Pingwen Zhang (both from Peking University). More detailed information of this summer school can be found in <http://ccse.pku.edu.cn/activities/2005/adaptiveseminar.htm> (which is most in Chinese). This summer school continued in 2006 but with a more focused program (see <http://ccse.pku.edu.cn/06summerschool/school.html>) and it is expected to continue more in the coming years.

The articles in this volume, which are ordered alphabetically by authors, touch upon various aspects of adaptive methods for algorithmic design, theoretical analysis and practical applications. Chapter 1, by Long Chen and Jinchao Xu, summarizes the basic techniques in local adaptive finite element methods and especially the recent advancements on the convergence and complexity of this type of adaptive methods. Chapter 2, again by Long Chen and Jinchao Xu, describes various postprocessing techniques for gradient recovery that include results on patch symmetric grids, mildly structured grids and general unstructured grids. Chapter 3, by Weichang Huang, cov-

ers topics on adaptive algorithms for anisotropic meshes and their theoretical foundations, including basic mathematical principles of mesh adaptation, anisotropic interpolation theory, monitor functions, variational mesh generation, and moving mesh methods. Chapter 4, by Natalia Kopteva, provides a rigorous theoretical analysis on the convergence of moving grid methods for a special class of convection-dominated convection-diffusion problem. Chapter 5, by Zhiping Li, presents a special adaptive grid method, the mesh transformation method, with applications to computation of crystalline and microstructures. Chapter 6, by J.A. Mackenzie and W.R. Mekwi, focuses on moving mesh methods which employ a moving mesh partial differential equation (MMPDE) to evolve the mesh in an appropriate fashion. Chapter 7, by P. A. Zegeling, presents moving grid methods with a summary of many ideas and techniques and applications to reaction-diffusion systems, tumour angiogenesis models, brine transport and magneto-hydrodynamics. Chapter 8, by Zhimin Zhang, discusses two post-processing methods that are instrumental for adaptive finite element methods: the Zienkiewicz-Zhu Superconvergence Patch Recovery (SPR) and recently proposed Polynomial Preserving Recovery (PPR).

It is our great pleasure as editors to thank all of the authors for their hard work to provide us with these high quality articles. We are very grateful for the generous financial supports provided by National Science Foundation of China, The Mathematics Center of Ministry of Education of China, School of Mathematical Sciences of Peking University, and the Joint Research Institute for Applied Mathematics between Peking University and Hong Kong Baptist University. We would like to thank Pingwen Zhang from Peking university for his leadership in helping organize the summer activities in Peking University. We would also like to thank Tammy Lam of Hong Kong Baptist University for the considerable work she put into producing the final layout of the proceedings.

Tao Tang  
Hong Kong Baptist University  
Jinchao Xu  
Pennsylvania State University  
October, 2006

# Contents

<b>Chapter 1</b>	<b>Convergence of Adaptive Finite Element Methods</b>	<b>1</b>
1.1	Introduction	2
1.2	Preliminaries	11
1.3	Residual type error estimator	14
1.4	Convergence of an adaptive finite element method	21
1.5	Optimality of the adaptive finite element method	27
	Bibliography	31
<b>Chapter 2</b>	<b>A Posteriori Error Estimator by Post-processing</b>	<b>34</b>
2.1	Introduction	35
2.2	Linear finite element on patch symmetric grids	39
2.3	Linear finite element on mildly structured grids	45
2.4	Linear finite element on general unstructured grids	53
	Bibliography	63
<b>Chapter 3</b>	<b>Anisotropic Mesh Adaptation and Movement</b>	<b>68</b>
3.1	Introduction	70
3.1.1	Sobolev spaces	71
3.1.2	Mesh terminology	73
3.1.3	Two algebraic inequalities	75
3.2	Basic principles in mesh adaptation	76
3.2.1	Geometric meaning of SVD decomposition	78
3.2.2	Alignment and equidistribution	79
3.2.3	Alignment and equidistribution for finite element meshes	82
3.3	Interpolation theory in Sobolev spaces	83
3.3.1	Finite element terminology	83
3.3.2	Element-wise estimate on interpolation error	83
3.4	Isotropic error estimates	85
3.4.1	Chain rule	85
3.4.2	Isotropic error estimation on a general mesh	87
3.4.3	Error bound on regular triangulations	87
3.5	Anisotropic error estimates	89

3.5.1	An anisotropic error bound . . . . .	89
3.5.2	Anisotropic error estimates independent of coordinate system	90
3.5.3	Bibliographic notes . . . . .	94
3.6	Mesh quality measures and monitor functions . . . . .	95
3.6.1	Mesh quality measures . . . . .	96
3.6.2	The isotropic case . . . . .	100
3.6.3	The anisotropic case: $l = 1$ . . . . .	106
3.6.4	The anisotropic case: $l = 2$ . . . . .	109
3.7	Anisotropic mesh adaptation: Refinement approach . . . . .	111
3.7.1	Metric tensor . . . . .	113
3.7.2	Numerical experiments . . . . .	115
3.8	Anisotropic mesh adaptation: Variational approach . . . . .	120
3.8.1	Functional for mesh alignment . . . . .	121
3.8.2	Functional for equidistribution . . . . .	122
3.8.3	Mesh adaptation functional . . . . .	122
3.8.4	Mesh equation . . . . .	123
3.8.5	Numerical experiments . . . . .	125
3.9	Adaptive moving mesh methods: MMPDE approach . . . . .	130
3.9.1	The MMPDE method . . . . .	132
3.9.2	Numerical examples . . . . .	138
3.10	Adaptive moving mesh methods: GCL approach . . . . .	142
3.10.1	GCL method . . . . .	143
3.10.2	Relation to the Lagrange method and the deformation map method . . . . .	144
3.10.3	Choice of $w$ , $v_{ref}$ , and $\rho$ . . . . .	147
3.10.4	Numerical examples . . . . .	147
3.11	Conclusions . . . . .	150
	Bibliography . . . . .	151
<b>Chapter 4</b>	<b>Convergence Theory of Moving Grid Methods . . . . .</b>	<b>159</b>
4.1	Introduction . . . . .	160
4.2	Maximum norm a posteriori error estimates for a 1d singularly per- turbed convection-diffusion equation . . . . .	160
4.2.1	Model convection-diffusion problem . . . . .	161
4.2.2	Numerical methods and notation . . . . .	162
4.2.3	Stability properties of the differential operator . . . . .	165

4.2.4	First-order error estimates . . . . .	167
4.2.5	Second-order error estimates . . . . .	172
4.3	Full analysis of a robust adaptive method for a 1d convection-diffusion problem . . . . .	176
4.3.1	Upwind difference scheme . . . . .	177
4.3.2	Adaptive mesh movement by equidistribution of the arc-length monitor function . . . . .	178
4.3.3	The algorithm . . . . .	182
4.3.4	The existence theorem . . . . .	185
4.3.5	Accuracy of equidistributed and computed solutions . . . . .	187
4.3.6	How many iterations for $\epsilon$ -uniform accuracy? . . . . .	189
4.3.7	Numerical results . . . . .	203
4.3.8	Possible generalizations . . . . .	207
	Bibliography . . . . .	208
<b>Chapter 5</b>	<b>Computation of Crystalline Microstructures with The Mesh Transformation Method . . . . .</b>	<b>211</b>
5.1	Introduction . . . . .	212
5.1.1	An illustrative example . . . . .	213
5.1.2	The idea of the mesh transformation method (MTM) . . . . .	215
5.2	Mesh transformation and regular re-meshing . . . . .	218
5.2.1	Some numerical examples . . . . .	221
5.3	Regularized mesh transformation methods . . . . .	225
5.3.1	Existence and convergence theorems for RMT . . . . .	226
5.3.2	Regularized periodic relaxation method (RPR) . . . . .	227
5.4	Application in computing non-smooth minimizers . . . . .	232
5.4.1	Discrete relaxation problem . . . . .	233
5.4.2	Numerical results . . . . .	235
	Bibliography . . . . .	240
<b>Chapter 6</b>	<b>On The Use of Moving Mesh Methods to Solve PDEs . . . . .</b>	<b>242</b>
6.1	Introduction . . . . .	243
6.2	Moving mesh partial differential equations (MMPDEs) . . . . .	245
6.2.1	Mesh equidistribution . . . . .	245
6.2.2	Variational formulation . . . . .	246
6.3	Phase change problems . . . . .	248
6.3.1	Convective heat transfer . . . . .	248

6.3.2	Phase-field models . . . . .	256
6.4	A one-dimensional $hr$ -adaptive method . . . . .	261
6.4.1	The $r$ -refinement strategy . . . . .	261
6.4.2	$h$ -refinement strategy . . . . .	263
6.4.3	Monitor function . . . . .	263
6.4.4	Refinement formula . . . . .	265
6.4.5	Outline of $hr$ -adaptive algorithm . . . . .	265
6.4.6	Numerical experiments . . . . .	266
6.5	Conclusions . . . . .	275
	Bibliography . . . . .	276
<b>Chapter 7</b>	<b>Theory and Application of Adaptive Moving Grid Methods . .</b>	<b>279</b>
7.1	Introduction . . . . .	280
7.2	Adaptive moving grids in one dimension . . . . .	283
7.2.1	A simple boundary-value problem . . . . .	283
7.2.2	The equidistribution principle . . . . .	285
7.2.3	Smoothing of the grid in space and time . . . . .	290
7.2.4	Applications . . . . .	298
7.3	The higher-dimensional case . . . . .	307
7.3.1	A tensor-grid approach in 2D . . . . .	307
7.3.2	Smooth adaptive grids based on Winslow's approach . . . . .	313
7.3.3	Three-dimensional adaptive moving grids . . . . .	324
	Bibliography . . . . .	330
<b>Chapter 8</b>	<b>Recovery Techniques in Finite Element Methods . . . . .</b>	<b>333</b>
8.1	Introduction and preliminary . . . . .	334
8.2	Local recovery for 1D . . . . .	337
8.2.1	Motivation-linear element . . . . .	338
8.2.2	Higher order elements . . . . .	340
8.2.3	One dimensional theoretical results . . . . .	357
8.3	Local recovery in higher dimensions . . . . .	364
8.3.1	Methods and examples . . . . .	365
8.3.2	Properties of the gradient recovery operator . . . . .	372
8.3.3	Superconvergence analysis in 2D . . . . .	375
8.4	Quasi-local, semi-local, and global recoveries . . . . .	397
	Bibliography . . . . .	408
<b>Index</b>		<b>413</b>

# Convergence of Adaptive Finite Element Methods

Long Chen and Jinchao Xu  
Department of Mathematics  
Pennsylvania State University  
University Park, PA 16802  
USA

lyc102@gmail.com, xu@math.psu.edu

## Contents

1.1	Introduction . . . . .	2
1.2	Preliminaries . . . . .	11
1.3	Residual type error estimator . . . . .	14
1.4	Convergence of an adaptive finite element method . . . . .	21
1.5	Optimality of the adaptive finite element method . . . . .	27
	Bibliography . . . . .	31

Adaptive methods are now widely used in the scientific computation to achieve better accuracy with minimum degree of freedom. While these methods have been shown to be very successful, the theory ensuring the convergence of the algorithm and the advantages over non-adaptive methods is still under development. In this chapter, we shall survey recent progress on the convergence analysis of adaptive finite element

methods (AFEMs) for second order elliptic partial differential equations [2, 11, 25, 32, 40].

To present the main idea and techniques in their simplest form we restrict ourselves to the linear finite element approximation to the model Poisson equation

$$-\Delta u = f \text{ in } \Omega, \quad \text{and } u = 0 \text{ on } \partial\Omega, \quad (1.0.1)$$

posed on a polygonal domain  $\Omega \subset \mathbb{R}^2$ . We shall prove that there exists an algorithm through the local refinement of triangulations to produce a sequence of approximations of  $u$  in an optimal way.

## 1.1 Introduction

Let  $D \subset \mathbb{R}^n$  be a bounded domain. For an integer  $k \geq 0$ , and  $1 \leq p \leq \infty$ ,  $W^{k,p}(D)$  will denote the usual Sobolev spaces of functions having distributional derivatives up to order  $k$  in  $L^p$ . The norm (or semi-norm) is given for  $1 \leq p < \infty$  by  $\|u\|_{k,p,D} = \left( \sum_{|\alpha| \leq k} \|D^\alpha u\|_p^p \right)^{1/p}$  (or  $|u|_{k,p,D} = \left( \sum_{|\alpha|=k} \|D^\alpha u\|_p^p \right)^{1/p}$ ) with the usual modification for  $p = \infty$ . The subscript  $k$ ,  $p$  or  $D$  will be omitted if  $k = 0$ ,  $p = 2$  or  $D = \Omega$ . The closure of  $C_0^\infty(D)$  with respect to the norm of  $W^{k,p}(D)$  is denoted by  $W_0^{k,p}(D)$ . When  $k < 0$ ,  $W^{k,p}(D)$  is defined as the dual space of  $W_0^{|k|,p}(D)$ . Furthermore  $H^k(D) := W^{k,2}(D)$  and  $H_0^k(D) := W_0^{k,2}(D)$  are Hilbert spaces in the norm  $\|\cdot\|_{k,D}$  or semi-norm  $|\cdot|_{k,D}$ , respectively.

The letter  $C$  or  $c$ , with or without subscript, denote generic constants that may not be the same at different occurrences. To avoid writing these constants repeatedly, by  $x \lesssim y$  we mean that there exist a constant  $C$  such that  $x \leq Cy$ . Obviously  $x \gtrsim y$  is defined as  $y \lesssim x$ , and  $x \simeq y$  as  $x \lesssim y$  and  $x \gtrsim y$ . The letter  $C_i$  with subscript is used to denote specific important constants.

### Finite element methods

The weak formulation of (1.0.1) is: for a given  $f \in H^{-1}$ , find  $u \in H_0^1(\Omega)$  such that

$$a(u, v) = \langle f, v \rangle \quad \forall v \in H_0^1, \quad (1.1.2)$$

where  $a(u, v) = (\nabla u, \nabla v) = \int_\Omega \nabla u \cdot \nabla v$  and  $\langle f, v \rangle$  is the dual pair. In particular when  $f \in L^2$ ,  $\langle f, v \rangle = (f, v) = \int_\Omega f v$ . The Laplace operator can be understood as the linear operator introduced by the bilinear form  $a(\cdot, \cdot)$

$$-\Delta : H_0^1 \mapsto H^{-1}, \quad \text{by } \langle -\Delta u, v \rangle = a(u, v), \quad \forall v \in H_0^1.$$

By the Poincaré inequality, the bilinear form  $a(\cdot, \cdot)$  defines an inner product on  $V$ . The existence and uniqueness of the solution to (1.1.2) follows from Riesz representation theorem. Moreover from

$$|u|_1^2 = a(u, u) = \langle f, u \rangle \leq \|f\|_{-1} |u|_1,$$

we have the stability result

$$|u|_1 \leq \|f\|_{-1}. \quad (1.1.3)$$

Namely the linear operator  $\Delta : H_0^1 \mapsto H^{-1}$  is an isomorphism and  $\Delta^{-1}$  is bounded. Although  $\Delta^{-1}$  is defined on  $H^{-1}$ , we will mainly consider its restriction on  $L^2 \subset H^{-1}$ . In this case by the regularity theory of elliptic partial differential equations (c.f., [28]),  $u = \Delta^{-1}f$  is smoother than  $H^1$  and thus the approximation of  $u$  in  $H^1$  norm with certain rate is possible.

The finite element method is to find an approximation by solving (1.1.2) in finite-dimensional subspaces of  $H_0^1(\Omega)$  based on triangulations of  $\Omega$ . That is, given an  $f \in L^2$ , to find a  $u_h \in V^h \subset H_0^1$  such that

$$a(u_h, v_h) = (f, v_h) \quad \forall v_h \in V^h. \quad (1.1.4)$$

The existence and uniqueness of the solution to (1.1.4) follows from the Riesz representation theorem since  $V^h \subset H_0^1$ .

We now discuss the construction of finite element spaces  $V^h$ . A triangulation  $\mathcal{T}_h$  (also indicated by mesh or grid) of  $\Omega$  is a *conforming* partition of  $\Omega$  into a set of triangles. Conforming means the intersection of any two triangles  $\tau_1$  and  $\tau_2$  in  $\mathcal{T}_h$  either consists of a common vertex  $x_i$ , edge  $E$  or empty. The *interior* node (also indicated by vertices) set and edge set of the triangulation are denoted by  $\mathcal{N}_h$  and  $\mathcal{E}_h$ , respectively. The subscript  $h$  here represents the mesh size  $h := \max_{\tau \in \mathcal{T}_h} \text{diam}(\tau)$  where  $\text{diam}(\tau)$  denotes the diameter of  $\tau$ . It is always assumed that there are a family of triangulations  $\{\mathcal{T}_h\}_{h \in \mathcal{H}}$  with  $h \rightarrow 0$ . Properties of triangulations presented below are assumed to hold uniformly with respect to the whole family.

Throughout this article we shall only consider shape regular triangulations.  $\{\mathcal{T}_h\}_{h \in \mathcal{H}}$  is *shape regular* if

$$\max_{\tau \in \mathcal{T}_h} \frac{\text{diam}(\tau)^2}{|\tau|} \leq \sigma_s$$

where  $|\tau|$  is the area of  $\tau$ . The shape regularity of triangulations assures that angles of the triangulation remains bounded away from 0 and  $\pi$  which is important to con-

control the interpolation error in  $H^1$  norm [1] and the condition number of the stiffness matrix [27].

For the simplicity, we shall only consider linear finite element space. Given a shape regular triangulation  $\mathcal{T}_h$  of  $\Omega$ , we define

$$V^h := \{v \mid v \in C(\bar{\Omega}), v|_{\partial\Omega} = 0 \text{ and } v|_{\tau} \in \mathcal{P}_1(\tau), \forall \tau \in \mathcal{T}_h\},$$

where  $\mathcal{P}_1(\tau)$  denotes the linear polynomial space on  $\tau$ . It is easily to show that  $V^h \subset H_0^1$ .

### Convergence analysis of finite element methods

Classic convergence analysis of finite element methods are most developed for *quasi-uniform* grids.  $\{\mathcal{T}_h\}_{h \in \mathcal{H}}$  is called *quasi-uniform* if it is shape regular and furthermore satisfies the global assumption

$$\frac{\max_{\tau \in \mathcal{T}_h} |\tau|}{\min_{\tau \in \mathcal{T}_h} |\tau|} \leq \sigma_u.$$

Namely all elements of a quasi-uniform grid are around the same size.

For finite element methods there is a natural energy norm defined by the bilinear form  $\|u\|_a^2 := a(u, u)$ . For this model problem  $\|u\|_a = \|\nabla u\| = |u|_1$ . The standard convergence analysis of the finite element methods on quasi-uniform meshes has three main ingredients.

#### 1. Quasi-optimality of the finite element approximation:

$$|u - u_h|_1 \lesssim \inf_{v_h \in V^h} |u - v_h|_1.$$

In particular,  $|u - u_h|_1 \lesssim |u - u_I|_1$ , where  $u_I$  is the nodal interpolation of  $u$  based on  $\mathcal{T}_h$ . Namely  $u_I \in V^h$  and  $u_I(x_i) = u(x_i)$  for all  $x_i \in \mathcal{N}_h$ .

#### 2. Interpolation error estimates for $u_I$ on quasi-uniform triangulations $\mathcal{T}_h$ :

$$|u - u_I|_1 \lesssim h \|u\|_2, \quad \forall u \in H^2.$$

#### 3. Regularity result of elliptic equations: when the domain $\Omega$ is smooth or convex and Lipschitz, $\Delta^{-1} : L^2(\Omega) \mapsto H^2(\Omega) \cap H_0^1(\Omega)$ is a bounded linear operator. That is

$$\|u\|_2 \lesssim \|f\|.$$

Combining 1, 2 and 3, on quasi-uniform triangulations  $\mathcal{T}_h$  of a nice domain  $\Omega$ , we obtain

$$|u - u_h|_1 \lesssim |u - u_I|_1 \lesssim h\|u\|_2 \lesssim h\|f\|.$$

It is easy to see the first order convergence is optimal for general function  $u \in H^2$ . Indeed for  $u = x^2 + y^2$ , one can show the convergence rate cannot be greater than one. Thus in the full regularity case the uniform refinement of the triangulations will produce finite element approximations of  $u$  in an optimal order.

The regularity result, however, does not hold on general Lipschitz domains. To see this, let us give a simple counter example. Given  $\beta \in (0, 1)$ , let  $\Omega = \{(r, \theta) : 0 < r < 1, 0 < \theta < \pi/\beta\}$  and let  $f = r^\beta \sin(\beta\theta)$  and  $u = (1 - r^2)f$ . Being the imaginary part of the complex analytic function  $z^{1/2}$ ,  $f$  is harmonic in  $\Omega$ . A direct calculation shows that

$$-\Delta u = 4(1 + \beta)f \text{ in } \Omega, \quad \text{and} \quad u|_{\partial\Omega} = 0.$$

Note that  $4(1 + \beta)f \in L^\infty \subset L^2$  but  $u \notin H^2$ . Furthermore it is easy to see  $u \in H^s$ , for  $0 < s < 1 + \beta$ . In views of the approximation theory, we cannot expect approximation rate (in  $H^1$  norm) better than  $h^\beta$  if we insist on quasi-uniform grids.

### Equidistribution

To improve the convergence rate, the element size is adapted according to the behavior of the solution. In this case, the element size in areas of the domain where the solution is smooth can stay bounded well away from zero, and thus the global element size is not a good measure of the approximation rate. For this reason, when the optimality of the convergence rate is concerned,  $\#\mathcal{T}$ , the number of triangles, is used to measure the approximation rate in the setting of adaptive methods that involve local refinement.  $\mathcal{T}_N$  is used to denote a triangulation with at most  $N$  triangles which is proportional to the number of degree of freedom. Again by default we are considering a sequence of triangulations  $\{\mathcal{T}_N\}_{N>N_0}$ . Note that  $N = O(h^{-2})$  for quasi-uniform grids  $\mathcal{T}_h$  when  $h \rightarrow 0$ .

The first order convergence of finite element approximation can be recovered through the correct mesh adaptation. We have the following theorem.

**Theorem 1.1.** *Let  $u \in W^{2,1}(\Omega)$ . Suppose a shape regular triangulation  $\mathcal{T}_N$  weakly equidistributes the  $W^{2,1}$  norm of  $u$  in the sense that*

$$|u|_{2,1,\tau} \leq \frac{C}{N} |u|_{2,1,\Omega}, \quad \forall \tau \in \mathcal{T}_N. \quad (1.1.5)$$

Then the finite element approximation  $u_N$  based on  $\mathcal{T}_N$  is of optimal approximation order:

$$|u - u_N|_1 \lesssim N^{-1/2} |u|_{2,1,\Omega}. \quad (1.1.6)$$

*Proof.* Using the embedding:  $W^{2,1}(\Omega) \subset H^1(\Omega) \cap C(\bar{\Omega})$ , we know the nodal interpolation  $u_I$  is well defined and

$$|u - u_I|_{1,\tau} \leq C \|u\|_{2,1,\tau}, \quad \forall \tau \in \mathcal{T}_N.$$

Since the nodal interpolation preserve linear polynomials, by Bramble-Hilbert lemma, we have

$$|u - u_I|_{1,\tau} \leq C |u|_{2,1,\tau}, \quad \forall \tau \in \mathcal{T}_N.$$

We emphasize that constants in the above inequalities do not depend on the element  $\tau$  since those norms are scaling invariant. Squaring and summing over all the elements, we get

$$|u - u_I|_1^2 \lesssim \sum_{\tau \in \mathcal{T}_N} |u|_{2,1,\tau}^2 \lesssim N^{-1} |u|_{2,1,\Omega}^2.$$

In the last step we have used the equidistribution assumption (1.1.5).  $\square$

We would like to stress two important points contained in Theorem 1.1. First we use a weaker norm  $W^{2,1}$  of  $u$  to obtain the optimal convergence order  $N^{-1/2} = \mathcal{O}(h)$ . We still use the second derivative of  $u$  but the norm is shifted from  $L^2$  norm to a weaker one  $L^1$ . In general we may need to work on Besov spaces. We refer the reader to [10, 24, 41] for the definition and properties of Besov space. Here we only remark that such spaces are function spaces suitable for the theory of nonlinear approximation by piecewise polynomials. The regularity theory of elliptic partial differential equation in terms of Besov norms can be found at [3, 4, 20, 21]. It is shown in those works that the solution to the Poisson's equation indeed has typically higher Besov than Sobolev regularity and thus the use of adaptive scheme gives a better asymptotic approximation rate.

Secondly an ideal case of (1.1.5) is the equidistribution principle widely used in the literature:

$$|u|_{2,1,\tau} = \frac{1}{N} |u|_{2,1,\Omega}, \quad \forall \tau \in \mathcal{T}_N.$$

Namely a good mesh adaptation is to equidistribute some local error bound. We would like to elaborate that, in the current setting, equidistribution is indeed a sufficient condition for optimal error, but by no means this has to be a necessary condition. Namely the equidistribution principle can be severely violated but asymptotically optimal error estimates can still be maintained. For example if a bounded number of

elements satisfy

$$|u|_{2,1,\tau} \leq \frac{C}{\sqrt{N}} |u|_{2,1,\Omega},$$

and the rest satisfy (1.1.5), the optimal error estimate (1.1.6) is still valid. This observation leads a better marking strategy discussed later.

We shall show the existence of practical algorithms to generate triangulations satisfying (1.1.5). If  $u$  is known and a little bit smoother than  $W^{2,1}$ , then there are algorithms [5, 11] which can produce such equidistributed triangulation in a nearly optimal way. As an example we present a result below and we refer the reader to [5] for the proof. Recall that a function  $f \in L \log^+ L(\Omega)$  if  $\int_{\Omega} |f \log |f|| < \infty$ .

**Theorem 1.2.** *If  $D^2 u \in L \log^+ L(\Omega)$  and  $|u|_{2,1,\tau}$  can be evaluate exactly, then there exists an algorithm to generate a triangulation  $\mathcal{T}_N$  with  $N$  elements satisfying (1.1.5).*

Those algorithms, however, cannot apply directly to the numerical solution of partial differential equations since the solution  $u$  is unknown.

### Adaptive finite element methods

We now briefly review the main idea of adaptive finite element methods through local refinement. Given an initial triangulation  $\mathcal{T}_0$ , we shall generate a sequence of nested conforming triangulations  $\mathcal{T}_k$  using the following loop

$$\text{SOLVE} \rightarrow \text{ESTIMATE} \rightarrow \text{MARK} \rightarrow \text{REFINE}. \quad (1.1.7)$$

More precisely to get  $\mathcal{T}_{k+1}$  from  $\mathcal{T}_k$  we first solve (1.1.4) to get  $u_k$  on  $\mathcal{T}_k$ . The error is estimated using  $u_k$  and used to mark a set of triangles of  $\mathcal{T}_k$  that are to be refined. Triangles are refined in such a way that the triangulation is still shape regular and conforming.

We shall not discuss the step **SOLVE** which deserves a separate investigation. We assume that the solutions of the finite dimensional problems can be solved to any accuracy efficiently. Examples of such optimal solvers are multigrid method or multigrid-based preconditioned conjugate gradient method [14, 29, 43, 44].

The *a posteriori* error estimators are essential part of the **ESTIMATE** step. To obtain information of the error, we need *a posteriori* error estimator  $\eta = \text{POSTERIORI}(\mathcal{T}_h, u_h, f, a(\cdot, \cdot))$ . Here **POSTERIORI**( $\mathcal{T}_h, u_h, f, a(\cdot, \cdot)$ ) can be thought of as a subroutine which returns a quantity  $\eta$  based on the current triangulation  $\mathcal{T}_h$ , the finite element approximation  $u_h$  on  $\mathcal{T}_h$ , the data  $f$  and possibly the bilinear form  $a(\cdot, \cdot)$ . One may want to approximate the local error bound  $|u|_{2,1,\tau}$  in Theorem 1.1

by  $|u_h|_{2,1,\tau}$ . But for piecewise linear function  $|u_h|_{2,1,\tau} = 0$  for any  $\tau \in \mathcal{T}_h$ . On the other hand, roughly speaking, the derivative of a piecewise constant function is the delta function on edges with the magnitude of the difference of the values in the elements sharing that edge. More precisely for each interior edge  $E \in \mathcal{E}_h$ , we fix a normal vector  $\mathbf{n}_E$ . Let  $\tau$  and  $\tilde{\tau}$  be two triangles sharing the edge  $E$  such that  $\mathbf{n}_E$  is the outward normal vector of  $E$  in  $\tau$ . We define the jump of flux across  $E$  as

$$J_E(u_h) := \left[ \frac{\partial u_h}{\partial \mathbf{n}_E} \right] := \frac{\partial u_h}{\partial \mathbf{n}_E} \Big|_{\tau} - \frac{\partial u_h}{\partial \mathbf{n}_E} \Big|_{\tilde{\tau}}.$$

We shall prove such quantity, read as a discrete version of  $|u|_{2,1,\tau}$ , will be crucial in the *a posteriori* error estimator. The lost information going from  $u$  to  $u_h$  will be captured by the data.

The *a posteriori* error estimators are usually split into local error indicators and they are then employed to make local modifications by dividing the elements whose error indicator is large and possibly coarsening the elements whose error indicator is small. The way we mark these triangles influences the efficiency of the adaptive algorithm. The traditional maximum marking strategy is to mark triangulations  $\tau^*$  such that

$$\eta_{\tau^*} \geq \theta \max_{\tau \in \mathcal{T}_h} \eta_{\tau}, \quad \text{for some } \theta \in (0, 1).$$

This marking strategy is proposed in the pioneering work of Babuška and Vogelius [2]. Such marking strategy is designed to evenly distribute the error. Based our relaxation of the equidistribution principal, we may leave some exceptional elements and focus on the overall amounts of the error. This leads to the bulk criterion firstly proposed by Dörfler [25] in order to prove the convergence of the local refinement strategy. With such strategy, one defines the marking set  $\mathcal{M}$  such that

$$\sum_{\tau \in \mathcal{M}} \eta_{\tau} \geq \theta \sum_{\tau \in \mathcal{T}} \eta_{\tau}, \quad \text{for some } \theta \in (0, 1).$$

We shall always use this marking strategy in the **MARK** step.

In the **REFINE** step, we need to carefully choose the rule for dividing the marked triangles such that the mesh obtained by this dividing rule is still conforming and shape regular. Such refinement rules include red and green refinement [6], longest refinement [36, 37] and newest vertex bisection [39]. In addition we also would like to control the number of elements added to ensure the optimality of the refinement. To this end we shall use newest vertex bisection in this article and briefly mention