Anisotropic Mesh Adaptation for the Finite Element Solution of Anisotropic Diffusion Problems

By

Xianping Li

Submitted to the graduate degree program in the Department of Mathematics and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Chairperson, Weizhang Huang

Committee members

Tyrone Duncan

Jie Han

Erik Van Vleck

Hongguo Xu

Date Defended: 08/15/2011

The Dissertation Committee for Xianping Li certifies
that this is the approved version of the following dissertation:

Anisotropic Mesh Adaptation for the Finite Element Solution of Anisotropic Diffusion Problems

Chairperson, Weizhang Huang

Tyrone Duncan

Jie Han

Erik Van Vleck

Hongguo Xu

Date approved: 08/15/2011

Abstract

Anisotropic diffusion problems arise in many fields of science and engineering and are modeled by partial differential equations (PDEs) or represented in variational formulations. Standard numerical schemes can produce spurious oscillations when they are used to solve those problems. A common approach is to design a proper numerical scheme or a proper mesh such that the numerical solution satisfies discrete maximum principle (DMP). For problems in variational formulations, numerous research has been done on isotropic mesh adaptation but little work has been done for anisotropic mesh adaptation.

In this dissertation, anisotropic mesh adaptation for the finite element solution of anisotropic diffusion problems is investigated. A brief introduction for the related topics is provided. The anisotropic mesh adaptation based on DMP satisfaction is then discussed. An anisotropic non-obtuse angle condition is developed which guarantees that the linear finite element approximation of the steady state problem satisfies DMP. A metric tensor is derived for use in mesh generation based on the anisotropic non-obtuse angle condition. Then DMP satisfaction and error based mesh adaptation are combined together for the first time.

For problems in variational formulations, two metric tensors for anisotropic mesh adaptation and one for isotropic mesh adaptation are developed. For anisotropic mesh adaptation, one metric tensor (based on Hessian recovery) is semi-a posterior and the

other (based on hierarchical basis error estimator) is completely a posterior. The metric tensor for isotropic mesh adaptation is completely a posterior. All the metric tensors incorporate structural information of the underlying problem into their design and generate meshes that adapt to changes in the structure.

The application of anisotropic diffusion filter in image processing is briefly discussed. Numerical examples demonstrate that anisotropic mesh adaptation can significantly improve computational efficiency while still providing good quality result. More research is needed to investigate DMP satisfaction for parabolic problems.

Acknowledgment

I would like to express sincere appreciation to my advisor Professor Weizhang Huang for his sagacious guidance and inspiring encouragement through out my PhD study in Mathematics. Without his grateful support, I would never have completed this degree.

I greatly appreciate my committee members Professor Tyrone Duncan, Professor Jie Han, Professor Erik Van Vleck and Professor Hongguo Xu for their support for my research and caring about my living. I would also like to thank Professor Margaret Bayer for her support in my teaching.

Thanks should also go to Ms. Debbie Garcia, Ms. Gloria Prothe, Ms. Kerrie Brecheisen and Ms. Erinn Barroso for their help with administrative arrangement and many other things.

I would like to thank my son David Lawrence Li who brings great joy into my life, my wife Hongjuan Zhou who provides unlimited love and support, and my parents-in-law who helped taking care of little David during their visit.

Thanks also go to my friends for their help in many ways. The financial support provided by the National Science Foundation (USA) under Grants DMS-0410545 and DMS-0712935 is also highly appreciated.

Table of Contents

Some Notations

If not defined otherwise, the following notations are used throughout this dissertation.

Ω : domain, a connected polygon or polyhedron;

$\mathscr{T}_h$: a triangulation of Ω;

K : a mesh element;

$|K|$: the volume of element K;

$\mathbb{D}_K$: the diffusion matrix on element K after applying numerical quadrature;

F_K : affine mapping from reference element to K;

M : $= M(\boldsymbol{x})$, the metric tensor used for mesh generation;

$M_K := \dfrac{1}{|K|}\displaystyle\int_K M(\boldsymbol{x})d\boldsymbol{x}$, the average of metric on element K;

$\rho_K := \sqrt{\det(M_K)}$;

N : number of elements in the mesh;

N_v : number of vertices in the mesh;

α_h : regularization parameter to prevent M_K from being singular.

Chapter 1

Introduction

Anisotropic diffusion problems arise in many fields of science and engineering such as plasma physics, petroleum engineering, and image processing. Some problems are modeled by partial differential equations (PDEs) and others are represented naturally in variational formulations.

A common feature of anisotropic diffusion problems is the heterogeneity and anisotropy of the diffusion coefficient, which varies with location (heterogeneity) and direction (anisotropy). When standard numerical methods are used to solve those problems, spurious oscillations may occur in the computed solution and can cause problems in the sequential computation. A common approach to avoiding this difficulty is to design a proper numerical scheme and/or a proper mesh so that the numerical solution validates the discrete counterpart of the maximum principle (DMP) satisfied by the continuous solution. For variational problems, it is desirable for the mesh to retain the properties of the materials or structure of the underlying problem.

On the other hand, mesh adaptation has become an imperative tool for use in numerical solution of PDEs and variational problems. It has been amply demonstrated that anisotropic mesh adaptation can significantly improve computational efficiency over isotropic mesh adaptation, especially for problems with strong anisotropic features.

In this dissertation, we study anisotropic mesh adaptation for the finite element solution of anisotropic diffusion problems. Problems in both PDE and variational forms are considered. A prototype PDE model for the steady state diffusion problems is given by

$$-\nabla\cdot(\mathbb{D}\nabla u)=f, \qquad \text{in} \quad \Omega \tag{1.1}$$

subject to the Dirichlet boundary condition

$$u(\boldsymbol{x})=g(\boldsymbol{x}), \qquad \text{on} \quad \partial\Omega \tag{1.2}$$

where $\Omega\subset\mathbb{R}^d$ $(d=1,2,\text{ or }3)$ is the physical domain, f and g are given functions, and $\mathbb{D}=\mathbb{D}(\boldsymbol{x})$ is the diffusion matrix assumed to be symmetric and strictly positive definite on Ω.

The variational form of (1.1) is

$$I[u]=\int_\Omega\left(\frac{1}{2}\nabla u\cdot\mathbb{D}\nabla u-f\,u\right)d\boldsymbol{x}, \quad \forall u\in U_g \tag{1.3}$$

where U_g is the set of functions satisfying the Dirichlet boundary condition (1.2).

Although only the Dirichlet boundary condition is considered in this dissertation, mixed boundary conditions can be treated without major modification.

The boundary value problem (BVP) (1.1) and (1.2), or (1.3) and (1.2) becomes a heterogeneous anisotropic diffusion problem when $\mathbb{D}$ changes with location and its eigenvalues are not all equal at least on a portion of Ω.

When the anisotropy is significant, spurious oscillation and numerical dissipation may occur in the computed solution. Existing research shows that the alignment of mesh elements along the fast diffusion direction helps reduce the numerical dissipation. However, it is very difficult to align the elements, especially when $\mathbb{D}$ is heterogeneous.

On the other hand, tremendous research has been done to improve the satisfaction of discrete maximum principle (DMP), either by designing a proper numerical scheme or a better mesh. A well-known mesh condition is the "non-obtuse angle condition" by Ciarlet and Raviart [39], which guarantees the linear finite element solution to satisfy DMP for isotropic diffusion problems. For anisotropic diffusion problems, however, none of the existing research can guarantee the satisfaction of DMP.

In this dissertation, we develop the so-called anisotropic non-obtuse angle condition for the finite element solution of heterogeneous anisotropic diffusion problems, which guarantees the DMP satisfaction. The condition is a generalization of the well known non-obtuse angle condition. Several variants of the new condition are obtained for convenience of mesh adaptation.

For variational problems, numerous research has been done on isotropic mesh adaptation for the finite element solution. However, little work has been done on anisotropic mesh adaptation. In this dissertation, a bound for the first variation is derived and a formula for the metric tensor is defined for use in anisotropic mesh adaptation.

In the next few sections, we give a brief introduction of the background for related topics including anisotropic diffusion problems, variational problems, anisotropic mesh adaptation, finite element approximation and discrete maximum principle (DMP). The outline of this dissertation is given in §1.6.

1.1 Anisotropic diffusion problems

Anisotropic diffusion problems arise in various areas of science and engineering including plasma physics in fusion experiments and astrophysics [67, 68, 69, 99, 110, 113], petroleum reservoir simulation [1, 2, 43, 55, 97], and image processing [36, 37, 87,

98, 102, 123]. Brief descriptions for problems in those three areas are given in the following subsections.

1.1.1 Plasma physics

Plasma is a macroscopically neutral gas which consists of free electrons and ions. It is often called the *fourth state of matter*. The positive and negative charges can move independently and respond strongly to electromagnetic fields, which causes the properties of the plasma in the direction parallel to the magnetic field to be different from those perpendicular to it. In this sense, magnetized plasma (plasma with a strong magnetic field) are anisotropic.

Magnetohydrodynamics (MHD) is commonly used to describe the behavior of magnetized plasmas. The set of equations which describe MHD is a combination of the Navier-Stokes equations of fluid dynamics and Maxwell's equations of electromagnetism, which have to be solved simultaneously. But using operator splitting, heat conduction equation can be evolved independently of the MHD equations [100]. Hence, some researchers consider only the heat conduction equation and focus on important issues such as monotonicity and pollution of perpendicular heat flux.

The heat flux in plasma contains contributions from two parts. One part is from electron motions moving primarily along magnetic field lines, and the other part arises due to particle collisions driving cross-field diffusion. Heat flows primarily along the field lines with little conduction in the perpendicular direction. Let $\chi_{\|}$ and $\chi_{\perp}$ denote the heat conductivity coefficient parallel and perpendicular to the magnetic field line, respectively. Then $\chi_{\|}$ is much larger than $\chi_{\perp}$ in magnetized plasmas [113, 99, 110]. In fusion experiments, the ratio $\chi_{\|}/\chi_{\perp}$ can exceed 10^{10} [69].

The general heat conduction equation for the plasma is in the following form:

$$\frac{\partial e}{\partial t} = -\nabla \cdot \boldsymbol{q} + Q, \tag{1.4}$$

$$\boldsymbol{q} = -n\left[\chi_{\parallel}\boldsymbol{b}\boldsymbol{b}^T + \chi_{\perp}(I - \boldsymbol{b}\boldsymbol{b}^T)\right]\nabla T, \tag{1.5}$$

where e is the internal energy per unit volume, $\boldsymbol{q}$ is the heat flux, Q is the heat source, n is the number density (i.e., the average number of particles per unit volume), $T \equiv (\gamma-1)e/n$ is the temperature with $\gamma = 5/3$ as the ratio of specific heats for an ideal gas, and $\boldsymbol{b}$ is the unit vector along the magnetic field line. Generally speaking, n may vary in space and time, but in many studies, n is considered as uniformly distributed and taken as constant (e.g. $n = 1$).

As can be seen from the definition of the heat flux $\boldsymbol{q}$ (1.5), the problem (1.4) is an anisotropic diffusion problem with significant anisotropy due to the large ratio $\chi_{\parallel}/\chi_{\perp}$.

One goal in the plasma physics research is to control the thermonuclear fusion power by magnetic confinement or inertial confinement. In toroidal magnetic confinement, the magnetic coils are arranged to produce a toroidal field, and additional poloidal magnetic field inside the cross section of the torus is applied to suppress instability due to charge separation. The net magnetic field line is helically twisted as shown in Fig. 1.1.

The most advanced toroidal confinement system is the tokamak. The plasma current is driven in pulses along the toroidal direction by magnetic induction using the electric transformer method and produces the poloidal component of the helical magnetic field. The poloidal magnetic field produced by the plasma current inside the plasma ring is stronger than that outside the plasma ring. For the tokamak equilibrium, a vertical field is added to reduce the poloidal field inside the ring and to increase the poloidal field outside the ring. Fig. 1.2 shows a schematic of ITER (International Thermonuclear

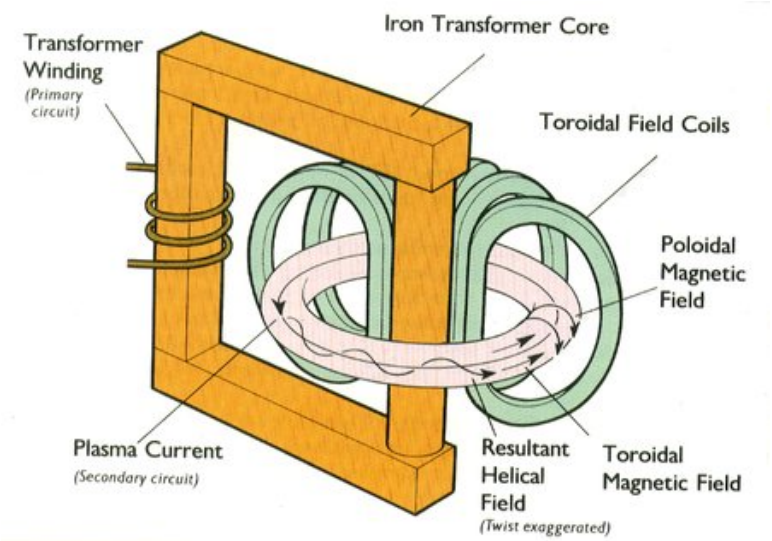


Figure 1.1: Schematic of ITER-8, from http://www.laetusinpraesens.org/iter/iter8.php

Energy Reactor), which is currently under construction and will be the largest tokamak in the world.

The hot plasma is confined and isolated from the relatively cold vessel walls so that the energy can be retained for as long as possible. The vacuum vessel is not in contact with the plasma and acts as the first safety confinement barrier [96]. Fig. 1.3 shows a sketch of the energy flow in a plasma.

In laboratory plasmas, the ratio of parallel and perpendicular conduction coefficients is very high ($\chi_{\parallel}/\chi_{\perp} \sim 10^{10}$). If an improper scheme is used, the numerical dissipation in the perpendicular direction may swamp the true perpendicular diffusion. One way to solve the problem is to align the coordinate along the field direction. However, it is very difficult and also expensive to align the coordinate in computations, especially when the plasma is also heterogeneous. Moreover, spurious solution (e.g., negative temperature) may occur which may lead to imaginary sound speed and cause problems in sequential computations.

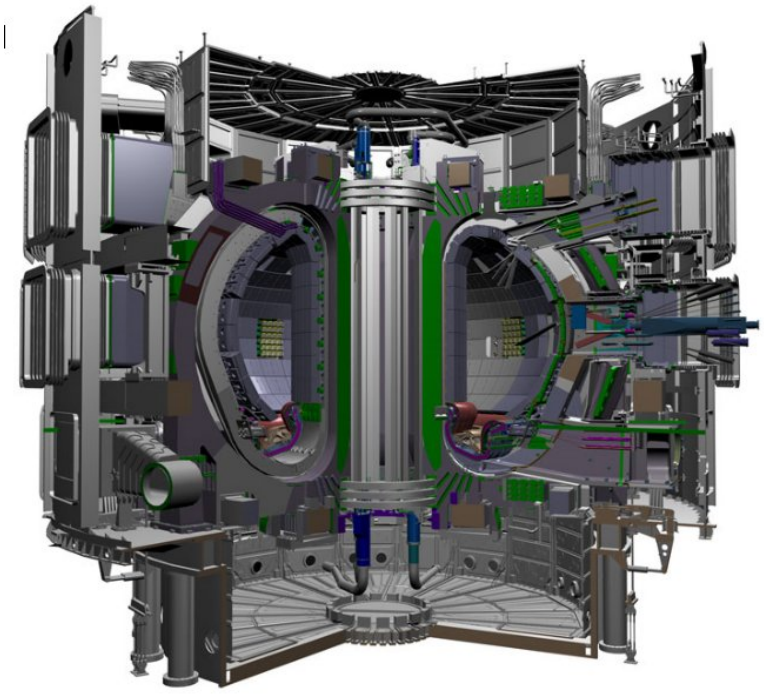


Figure 1.2: Schematic of ITER, from http://www.iter.org

Therefore, the numerical simulation of the heat conduction of plasmas must not only produce a physically meaningful temperature distribution but also avoid excessive numerical dissipation in the directions perpendicular to the magnetic field. Numerous research has been done in these topics. For example, Sharma and Hammett ([110]) showed that standard algorithms for anisotropic diffusion based on centered asymmetric and symmetric differencing do not preserve monotonicity. They applied slope limiters to modify the heat flux to avoid negative temperature. Günter *et al*. applied symmetric finite difference schemes ([69]) and a finite element scheme ([68, 67]) for unaligned coordinates to reduce the pollution of perpendicular numerical diffusion.

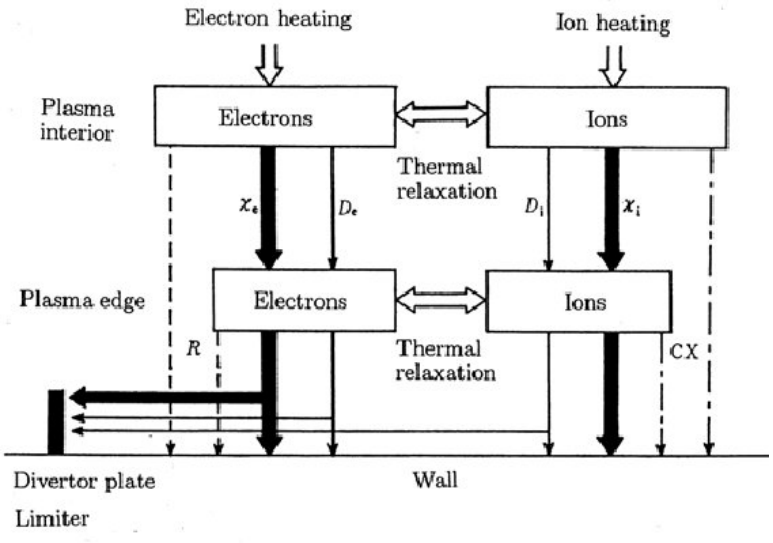


Figure 1.3: Energy flow of ions and electrons in a plasma. Bold arrows: thermal conduction (χ). White arrows: convective loss (D). Dashed arrows: radiation loss (R). Dot-dashed arrows: charge exchange loss (CX). (from [96])

1.1.2 Petroleum engineering

Another important field exhibiting anisotropic diffusion is Petroleum Engineering. A great deal of oil and gas are produced everyday but it still cannot meet the world's demand. Besides reservoirs in simple geologic formations, a large amount of oil and gas reservoirs in complicated geologic formations and deep seabeds have been explored and put in production. A major topic in oil and gas production is on fluid flow in porous media in geologic formations. A porous medium is a material permeated by an interconnected network of pores which are typically filled with a fluid (liquid or gas). Fig. 1.4 shows a sketch of a porous medium.

One of the major tasks of petroleum engineering is to produce crude oil or natural gas which is stored in the porous media in geologic formations. The flow of liquid (water, crude oil, or gas) in the reservoir rock depends on the properties of the rock including porosity, permeability, adhesion and reaction with the fluid.

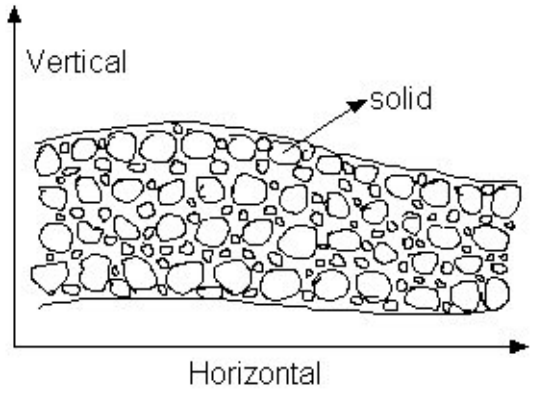


Figure 1.4: Sketch of a porous medium.

Porosity is the ratio of the pore space in a rock to the total volume of the rock. Some of these pores are isolated, while others are interconnected. The ratio of volume of interconnected pores to total volume of the rock is called effective porosity since only the fluid from interconnected pores can be produced out. In this sense, porosity is a measure of the capacity of the reservoir rock to store producible fluids in its pores.

Permeability is the capacity of a porous medium to transmit fluids through its interconnected pores. If the medium is saturated with a single liquid phase, the capacity is called absolute permeability (or permeability). If more than one phase of liquid exist in the medium, the observed permeability by the porous medium to one fluid phase is called the effective permeability to the particular liquid phase. Permeability changes with location, and even at the same location, may depend on the flow direction.

If one of the properties (porosity or permeability) is independent of location, the reservoir rock is called homogeneous in that property; otherwise, the rock is called heterogeneous in that property. If one property is independent of directions, the porous medium is called isotropic in that property; otherwise, it is called anisotropic and the

property shows a directional bias. For example, a reservoir can be homogeneous in porosity but heterogeneous in permeability.

The relation between the flux and pressure gradient for a single phase flow in a porous medium is described by Darcy's law as follows.

$$\boldsymbol{q} = -\frac{\boldsymbol{K}}{\mu}\nabla p,$$

where $\boldsymbol{q}$ is flux, $\boldsymbol{K}$ is the permeability tensor, μ is the viscosity, and ∇p is the pressure gradient.

Combining Darcy's law and the mass-conservation equation yields the single phase flow equation

$$C\frac{\partial p}{\partial t} = \nabla \cdot \left(\frac{\boldsymbol{K}}{\mu}\nabla p\right) + Q, \tag{1.6}$$

where C is the compression coefficient, Q is the source (e.g., oil production rate) and the fluid is assumed to be slightly compressible. In the steady state case, the equation becomes

$$\nabla \cdot \left(\frac{\boldsymbol{K}}{\mu}\nabla p\right) = 0.$$

Ignoring the capillary pressure and gravity force and assuming that water and oil are not miscible and incompressible, a simplified two-phase (crude oil and water) flow under constant temperature is described by

$$\begin{aligned}
\phi\frac{\partial S_o}{\partial t} &= \nabla \cdot \left(\frac{\boldsymbol{K}_0}{\mu_o}\nabla p\right) + Q_o, \\
\phi\frac{\partial S_w}{\partial t} &= \nabla \cdot \left(\frac{\boldsymbol{K}_w}{\mu_w}\nabla p\right) + Q_w, \\
S_o + S_w &= 1,
\end{aligned} \tag{1.7}$$

where ϕ is the porosity, S_o and S_w are the saturations of oil and water, respectively, $\boldsymbol{K}_0$ and $\boldsymbol{K}_w$ are the relative permeabilities of oil and water, respectively, μ_o and μ_w are viscosities of oil and water, respectively; Q_o and Q_w are oil production rate and water injection rate, respectively.

In geological formation, the permeability in vertical direction is usually much less than that in horizontal direction in reservoir rocks. Many oil reservoirs contain natural fractures, which have much higher permeability than general reservoir rocks. Hence, problem (1.6) or (1.7) is typically an isotropic diffusion problem.

1.1.3 Image processing

The third field of anisotropic diffusion is image processing. Image processing is a technique of signal processing for which the input is an image. The technique includes mathematical operations which are applied to the image data in order to improve the visual appearance (enhancement or restoration) or reveal key features and structures (shape recognition or object detection) of the image. Image processing has a variety of field applications [65] such as medicine, biology, astronomy, morphology, light microscopy, and remote sensing.

A traditional method in image processing is accomplished by a simple transformation usually performed by convolution with a filter function, which can hardly satisfy some mathematical axioms required by the procedure. On the other hand, the image processing methods based on partial differential equations (PDEs) are suitable to meet the axioms and provide some advantages. PDE-based image processing techniques are widely used for smoothing, restoration, segmentation, and recognition.

The simplest and best investigated PDE method for smoothing images is to apply a linear diffusion process:

$$\partial_t u = \Delta u, \qquad u(\boldsymbol{x}, 0) = f(\boldsymbol{x}),$$

which is equivalent to the convolution with a Gaussian kernel [123]. It is difficult for the linear model to obtain accurately the locations of the edges at coarse scales. The only way to locate the edges that have been detected at a coarse scale is by tracking across the scale space to their position in the original image. This technique is complicated and expensive.

The first PDE formulation of a nonlinear diffusion method was proposed by Perona and Malik [102] in order to avoid the blurring and localization problems of linear diffusion filtering. An inhomogeneous process is applied to reduce the diffusivity at the locations which are likely to be edges. The Perona-Malik filter is based on the equation

$$\partial_t u = \nabla \cdot (g(|\nabla u|^2)\nabla u), \qquad g(s^2) = \frac{1}{1 + s^2/\lambda^2}, \tag{1.8}$$

for some parameter $\lambda > 0$. This model provides high quality edge detection. However, the diffusion coefficient is a scalar (which varies spatially), so it is in fact an heterogeneous but isotropic model, and the flux $\boldsymbol{j} = -g\nabla u$ is always parallel to ∇u.

In certain applications, it is desirable to bias the flux towards the orientation of interesting features. Then an anisotropic diffusion tensor has to be applied. The general form is

$$\partial_t u = \nabla \cdot (\mathbb{D}\nabla u) + \beta(f - u), \tag{1.9}$$

where $\mathbb{D}(\boldsymbol{x}, u, \nabla u)$ is a diffusion tensor, $\beta \geq 0$ is a parameter and f is the given (or observed) image. The diffusion tensor $\mathbb{D}$ can be taken differently corresponding to

various applications and requirements and is anisotropic in general. Hence the problem (1.9) is an anisotropic diffusion problem.

Nonlinear diffusion filtering is usually performed with explicit schemes, which requires small time step size to be stable. Weickert et al. [124] presented semi-implicit schemes which are stable for all time step sizes. Some finite element schemes are also developed, for example, see [10, 11, 12, 51, 106].

A major application of image processing is magnetic resonance imaging (MRI). The diffusion coefficient is a parameter that directly reflects the molecular diffusion in the tissues. Molecular diffusion refers to the random translational motion of molecules that results from the thermal energy carried by these molecules [16]. Since molecular mobility in tissues may not be the same in all directions, the diffusion is an anisotropic three-dimensional process. The overall effect observed in a diffusion MRI image voxel reflects the displacement distribution of the water molecules present within this voxel. The observation of this displacement distribution may then provide unique clues to the structure and geometric organization of the tissues. Since 1990s, the diffusion MRI has been successfully applied to the analysis of brain tissues for some disease and it provides some patients with the opportunity to receive suitable treatment at a stage when their brain tissues might still be salvageable.

Anisotropic diffusion models work well in many image processing applications. The majority of effort focus on designing new continuous PDE models and discussing their mathematical properties but pay less attention to improving the numerical solutions. If improper numerical schemes are applied, spurious solutions (e.g., image artifacts) may occur, especially when heterogeneity and anisotropy are significant. Hence it is desirable to study the numerical solutions of existing PDE models. Furthermore, mesh adaptation can be applied to improve efficiency, especially for large scale images [11, 12, 106].

1.2 Variational problems

A brief description for variational problems is given in this section. Many anisotropic diffusion problems are modeled naturally in variational form. The diffusion problems introduced in PDE form (1.1) can also be casted into the variational form (1.3). A general variational problem is given by

$$I[u] = \int_{\Omega} F(\boldsymbol{x}, u, \nabla u) d\boldsymbol{x}, \quad \forall u \in U_g \tag{1.10}$$

where $\Omega \subset \mathbb{R}^d$ ($d = 1$, 2, or 3) is the physical domain and U_g is the set of functions satisfying the Dirichlet boundary condition

$$u(\boldsymbol{x}) = g(\boldsymbol{x}) \quad \text{on} \quad \partial\Omega$$

for a given function $g = g(\boldsymbol{x})$.

One common feature of those problems is that they have a natural variational formulation with which the governing equation can be derived through minimization. In most cases, variational problems can be transformed into boundary value problems of PDEs. Methods specially designed for solving PDEs can thus be used for solving many variational problems. Unfortunately, these methods generally do not take structural advantages of variational problems in their design. In the context of mesh adaptation and mesh movement, it has been argued by a number of researchers that the variational formulation should be used as a natural, compelling optimality criterion for the design of computational meshes for variational problems; see, among others, [13, 14, 15, 18, 28, 30, 31, 33, 41, 42, 46, 58, 59, 60, 66, 81, 119].

For example, Felippa [58, 59] proposes a variational principle associated with the governing equations to select optimal finite elements.

Delfour et al. [47] study the optimal triangular meshing for a class of linear second order elliptic problems using the idea of "speed method". They provide explicit expressions for partial derivatives of the associated energy functional with respect to coordinates, and obtain the best positions of the nodes which minimizes the solution error in the natural norm associated with the original problem.

Becker and Rannacher [13, 14] develop a "dual-weighted-residual method" for error control and mesh adaptation for finite element approximations of variational problems. They employ duality techniques and combine it with Galerkin orthogonality to derive a posteriori error estimates, which provide the basis of a feedback process for successively constructing economical meshes. By this approach, information about some local quantities of interest can be obtained, which may not be represented by meshes generated based on global error estimates.

Tourigny and Hülsemann [119] develop a mesh movement strategy with which both the node and approximate solution are updated by sequentially solving local minimization problems derived from the variational formulation of the original problem. A similar idea has been used in studies of the finite element solution for nonconvex variational problems [66, 41, 42] and for a relaxed variational problem [60].

Bochev and Lehoucq [18] present the interplay between the algebraic and variational problems for the pure Neumann problem, and demonstrate that finite element methods for the pure Neumann problem originate from two optimization settings. One requires minimization of a quadratic energy functional on a factor space and leads to singular linear system, and the other involves constrained minimization of a quadratic functional and leads to an equality-constrained quadratic program.

Some research has been focused on first-order system least squares (FOSLS) methods for finite element solution of PDEs. In this method, the second-order elliptic problem $\mathscr{L}u = f$ (together with appropriate boundary conditions) is written as a sys-

tem of first-order PDEs $\mathscr{L}_i u = f_i, i = 1, \cdots, M$. The resulting FOSLS L^2 functional is $\sum_{i=1}^{M} a_i \|\mathscr{L}_i u - f_i\|^2$, $a_i > 0$. For example, Jiang and Carey in [81] and Carey and Pehlivanov in [33] use local residuals as the error indicator to develop mesh adaptation scheme. Berndt, Manteuffel and Mccormick [15] establish a local a posteriori error estimate that is valid for any FOSLS L^2 minimization problem. Cai etc. [28] develops ellipticity estimates and discretization error bounds for FOSLS, and establish optimal convergence of multiplicative and additive multigrid algorithms of the discrete systems in [29]. Later, Cai and Starke study the least-squares methods for linear elasticity (see. [30, 31]).

It should be pointed out that most of the existing work for the adaptive numerical solution of variational problems employ isotropic meshes for which the size of elements is allowed to vary from place to place according to some error estimate or indicator while elements are kept almost equilateral. Although this isotropic mesh adaptation has been successfully applied to numerous application problems, it has a tendency to concentrate too many mesh elements in regions of large solution error. This is especially true for problems whose solutions exhibit strong anisotropic features. For this type of problems, computational efficiency can be significantly improved by using a properly chosen anisotropic mesh for which the size, shape, and orientation of elements all are allowed to vary. Anisotropic mesh adaptation has been successfully applied to the numerical solution of PDEs, e.g. see [5, 6, 21, 34, 45, 49, 74, 73, 88, 107, 111], but little work has been done for variational problems. Hence, it is desirable to study anisotropic mesh adaptation for variational problems.

1.3 Anisotropic mesh adaptation through metric specification

It is well-known that mesh adaptation has become an indispensable tool for use in the numerical solution of partial differential equations and variational problems. Computational efficiency can be significantly improved by placing more mesh elements in regions of larger solution error and less elements in regions of smaller error [8, 26, 32, 64, 80, 118].

When the adaptive finite element solution is concerned, the mesh $\mathscr{T}_h$ is generated according to the behavior of the error in the approximation u_h. In this dissertation, the so-called M-uniform mesh approach is taken [75, 80], with which an adaptive mesh is generated as a uniform mesh in the metric specified by a strictly positive definite tensor $M = M(\boldsymbol{x})$. Such a mesh will hereafter be called *an M-uniform mesh*. In this approach a scalar metric tensor (i.e. the product of a scalar function with the identity matrix) will lead to an isotropic mesh, while a full metric tensor will generally result in an anisotropic mesh. In this sense, the mesh generation procedure is the same for both isotropic and anisotropic mesh generation in this approach.

The key to the approach of mesh adaptation is to specify the metric tensor M and to generate M-uniform meshes for a given M. The second task can be achieved using various meshing strategies. Indeed, a number of algorithms and computer codes have been developed in the last decade for generating M-uniform meshes for a given M. Examples include the Delaunay-type triangulation method [19, 20, 34, 101], the advancing front method [63], the bubble mesh method [126], the method combining local modification with smoothing or node movement [70, 4, 21, 49, 109], and the computer code BAMG (Bidimensional Anisotropic Mesh Generator) developed by Hecht [71] using the Delaunay-type triangulation method [34]. On the other hand, a number of strategies

have been developed for specifying the metric tensor M; e.g. see [19, 20, 34, 61, 74]. Particularly, formulas for M based on the Hessian of the physical solution are used in [19, 20, 34], largely motivated by the results of D'Azevedo [44] and D'Azevedo and Simpson [45] on linear interpolation for quadratic functions on triangles. Several formulas for the metric tensor are developed in [74] based on interpolation error on simplicial elements.

An iterative procedure for solving PDEs is shown in Fig. 1.5. In our computation, each run is stopped after ten iterations. We have found that there is very little improvement in the computed solution after ten iterations for all the examples considered. A new mesh is generated using the computer code BAMG (bidimensional anisotropic mesh generator) developed by Hecht [71] based on a Delaunay-type triangulation method [34]. The code allows the user to supply his/her own metric tensor defined on a background mesh. In our computation, the background mesh has been taken as the most recent mesh available.

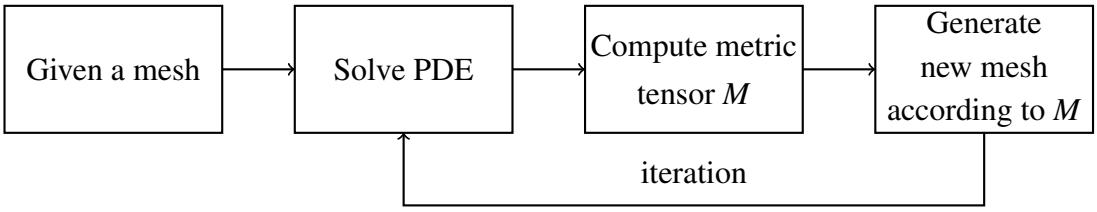


Figure 1.5: An iterative procedure for numerically solving PDE using M-uniform mesh approach.

It is shown in [75, 80] that when the reference element $\hat{K}$ is taken to be equilateral and unitary in volume, a simplicial M-uniform mesh $\mathscr{T}_h$ for a given $M = M(\boldsymbol{x})$ satisfies

$$\rho_K |K| = \frac{\sigma_h}{N}, \quad \forall K \in \mathscr{T}_h \tag{1.11}$$

$$\frac{1}{d} \operatorname{tr}\left((F_K')^T M_K F_K'\right) = \det\left((F_K')^T M_K F_K'\right)^{\frac{1}{d}}, \quad \forall K \in \mathscr{T}_h \tag{1.12}$$

where $|K|$ denotes the volume of K, N is the number of mesh elements, F_K is the affine mapping from $\hat{K}$ to K, F_K' is the Jacobian matrix of F_K, and

$$M_K = \frac{1}{|K|}\int_K M(\boldsymbol{x})d\boldsymbol{x}, \quad \rho_K = \sqrt{\det(M_K)}, \quad \sigma_h = \sum_{K\in\mathcal{T}_h}\rho_K|K|. \tag{1.13}$$

Condition (1.11), referred to as *the equidistribution condition*, determines the size of K from ρ_K. The larger ρ_K is, the smaller $|K|$ is. On the other hand, (1.12), called *the alignment condition*, characterizes the shape and orientation of K in the sense that the principal axes of the circumscribed ellipsoid of K are parallel to the eigenvectors of M_K while their lengths are reciprocally proportional to the square roots of the respective eigenvalues [75].

The focus of this dissertation is to develop and specify metric tensors for finite element approximation of anisotropic diffusion problems and variational problems. More details are discussed in later chapters.

1.4 Finite element approximation

The finite element method is a popular discretization method used to find approximate solutions for differential or integral equations arising from engineering and science. A given domain is decomposed into a collection of subdomains (or elements) and the governing equation is approximated over each subdomain. Then all elements are assembled using the relationships among them to obtain the solution to the whole domain.

Compared to other discretization methods such as the finite difference method and the finite volume method, the finite element method has some advantages including good mathematical foundation and flexible domain approximation (cf. [82, 54, 25, 108, 122]). For example, when the boundary of the geometric domain is complicated, it is

hard to use rectangular grids in the finite difference method to approximate the domain accurately. In this dissertation, finite element method is the only chosen discretization method.

Consider the linear finite element solution of BVP (1.1) and (1.2). Assume that Ω is a connected polygon or polyhedron and an affine family of simplicial triangulations $\{\mathscr{T}_h\}$ is given thereon. Let

$$U_g = \{v \in H^1(\Omega) \mid v|_{\partial\Omega} = g\}.$$

Denote by $U^h_{g_h} \subset U_{g_h}$ the linear finite element space associated with mesh $\mathscr{T}_h$, where g_h is a linear approximation of g. Then a linear finite element solution $\tilde{u}^h \in U^h_{g_h}$ to BVP (1.1) and (1.2) is defined by

$$\int_\Omega (\nabla v^h)^T \, \mathbb{D} \nabla \tilde{u}^h d\boldsymbol{x} = \int_\Omega f v^h d\boldsymbol{x}, \quad \forall v^h \in U^h_0. \tag{1.14}$$

This equation can be rewritten as

$$\sum_{K \in \mathscr{T}_h} \int_K (\nabla v^h)^T \, \mathbb{D} \nabla \tilde{u}^h d\boldsymbol{x} = \sum_{K \in \mathscr{T}_h} \int_K f v^h d\boldsymbol{x}, \quad \forall v^h \in U^h_0. \tag{1.15}$$

Generally speaking, the integrals in (1.15) cannot be carried out analytically, and numerical quadrature is needed. We assume that a quadrature rule has been chosen on the reference element $\hat{K}$ for this purpose,

$$\int_{\hat{K}} v(\xi) d\xi \approx |\hat{K}| \sum_{k=1}^m \hat{w}_k v(\hat{b}_k), \quad \sum_{k=1}^m \hat{w}_k = 1, \tag{1.16}$$

where $\hat{w}_k$'s are the weights and $\hat{b}_k$'s the quadrature nodes. A 2D example of such quadrature rules is given by $\hat{w}_k = \frac{1}{3}$ $(k = 1,2,3)$ and the barycentric coordinates (or

areal coordinates for triangles) of the nodes $(\frac{1}{6},\frac{1}{6},\frac{2}{3})$, $(\frac{1}{6},\frac{2}{3},\frac{1}{6})$, and $(\frac{2}{3},\frac{1}{6},\frac{1}{6})$. And a 3D example is $\hat{w}_i = \frac{1}{4}$ $(i = 1,2,3,4)$ and the barycentric coordinates of the nodes $(a,a,a,1-3a)$, $(a,a,1-3a,a)$, $(a,1-3a,a,a)$, and $(1-3a,a,a,a)$ with $a = \frac{5-\sqrt{5}}{20}$; e.g., see [54].

Let F_K be the affine mapping from $\hat{K}$ to K such that $K = F_K(\hat{K})$, and denote $b_k^K = F_K(\hat{b}_k)$, $k = 1,\cdots,m$. Upon applying (1.16) to the integrals in (1.15) and changing variables, the finite element approximation problem becomes seeking $u^h \in U_{g_h}^h$ such that

$$\sum_{K\in\mathscr{T}_h} |K| \sum_{k=1}^{m} \hat{w}_k \, (\nabla v^h|_K)^T \, \mathbb{D}(b_k^K) \, \nabla u^h|_K = \sum_{K\in\mathscr{T}_h} |K| \sum_{k=1}^{m} \hat{w}_k f(b_k^K) \, v^h(b_k^K), \quad \forall v^h \in U_0^h \tag{1.17}$$

where $\nabla v^h|_K$ and $\nabla u^h|_K$ denote the restriction of ∇v^h and ∇u^h on K, respectively. Note that we have used in (1.17) the fact that $\nabla v^h|_K$ and $\nabla u^h|_K$ are constant. Letting

$$\mathbb{D}_K = \sum_{k=1}^{m} \hat{w}_k \mathbb{D}(b_k^K), \tag{1.18}$$

we can rewrite (1.17) into

$$\sum_{K\in\mathscr{T}_h} |K| \, (\nabla v^h|_K)^T \, \mathbb{D}_K \, \nabla u^h|_K = \sum_{K\in\mathscr{T}_h} |K| \sum_{k=1}^{m} \hat{w}_k f(b_k^K) \, v^h(b_k^K), \quad \forall v^h \in U_0^h. \tag{1.19}$$

We now express (1.19) in a matrix form. Denote the numbers of the elements, vertices, and interior vertices of $\mathscr{T}_h$ by N, N_v, and N_{vi}, respectively. Assume that the vertices are ordered in such a way that the first N_{vi} vertices are the interior vertices. Then U_0^h and u^h can be expressed as

$$U_0^h = \text{span}\{\phi_1,\cdots,\phi_{N_{vi}}\} \tag{1.20}$$

and

$$u^h = \sum_{j=1}^{N_{vi}} u_j \phi_j + \sum_{j=N_{vi}+1}^{N_v} u_j \phi_j, \tag{1.21}$$

where ϕ_j is the linear basis function associated with the j-th vertex, $\boldsymbol{a}_j$. Note that the boundary condition (1.2) can be approximated by

$$u_j = g_j \equiv g(\boldsymbol{a}_j), \quad j = N_{vi}+1, ..., N_v. \tag{1.22}$$

Substituting (1.21) into and taking $v^h = \phi_i$ ($i = 1, ..., N_{vi}$) in (1.19) and combining the resulting equations with (1.22), we obtain the linear algebraic system

$$A\,\boldsymbol{u} = \boldsymbol{f}, \tag{1.23}$$

where

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & I \end{bmatrix}, \tag{1.24}$$

I is the identity matrix of size $(N_v - N_{vi})$, and

$$\boldsymbol{u} = (u_1, ..., u_{N_{vi}}, u_{N_{vi}+1}, ..., u_{N_v})^T,$$

$$\boldsymbol{f} = (f_1, ..., f_{N_{vi}}, g_{N_{vi}+1}, ..., g_{N_v})^T.$$

The entries of the stiffness matrix A and the right-hand-side vector $\boldsymbol{f}$ are given by

$$a_{ij} = \sum_{K \in \mathscr{T}_h} |K| \, (\nabla \phi_i|_K)^T \, \mathbb{D}_K \, \nabla \phi_j|_K, \quad i = 1, ..., N_{vi}, \; j = 1, ..., N_v, \tag{1.25}$$

$$f_i = \sum_{K \in \mathscr{T}_h} |K| \sum_{k=1}^{m} \hat{w}_k f(b_k^K) \, \phi_i(b_k^K), \quad i = 1, ..., N_{vi}. \tag{1.26}$$

Recall that (1.23) and (1.24) have been obtained under the Dirichlet boundary condition (1.2). It is not difficult to show that a linear system in the same form can be obtained for mixed boundary conditions provided that $\Gamma_D \neq \emptyset$, with Γ_D being the part of the boundary where the Dirichlet condition is imposed. Therefore, the mesh conditions developed in this dissertation also work for mixed boundary conditions with $\Gamma_D \neq \emptyset$.

1.5 Discrete maximum principle (DMP)

BVP (1.1) and (1.2) is a representative example of anisotropic diffusion problems which satisfies the (continuous) maximum principle

$$\max_{\boldsymbol{x}\in\Omega\cup\partial\Omega} u(\boldsymbol{x}) \leq \max\{0, \max_{\boldsymbol{s}\in\partial\Omega} g(\boldsymbol{s})\} \tag{1.27}$$

provided that $f(\boldsymbol{x}) \leq 0$ holds for all $\boldsymbol{x} \in \Omega$. When a standard numerical method, such as a finite element, a finite difference, or a finite volume method, is used to solve the problem, spurious oscillations can occur in the computed solution.

The numerical solution of BVP (1.1) and (1.2) has been studied extensively in the past, and a major effort has been made to avoid spurious oscillations in the numerical solution. A common strategy is to develop numerical schemes satisfying the discrete counterpart of (1.27) – the so-called discrete maximum principle (DMP), which are known to produce numerical solutions free of spurious oscillations [38, 121]. The studies can be traced back to early works by Varga [121], Ciarlet [38], Ciarlet and Raviart [39], and Stoyan [114, 115], where a number of sufficient conditions in a general and abstract setting are obtained for a class of linear elliptic partial differential equations (PDEs). For example, denote by $A\boldsymbol{u} = \boldsymbol{f}$ the linear algebraic system resulting from the application of a numerical scheme to a linear elliptic PDE supplemented with a Dirich-

let boundary condition, where A is the $n \times n$ stiffness matrix, $\boldsymbol{u}$ is the unknown vector, and $\boldsymbol{f}$ the right-hand-side vector. Then, a sufficient condition is given as follows.

Lemma 1.5.1. *([115]) If the stiffness matrix A satisfies*

(a) *that A is monotone with $A^{(-)}$ being nonsingular, or singular but irreducible; and* (1.28)

(b) *that $A^{(-)}\boldsymbol{e}^{(n)} \geq 0$,* (1.29)

then the numerical scheme satisfies DMP.

Here, matrix A is said to be monotone if A is nonsingular and $A^{-1} \geq 0$ (i.e., all entries of A^{-1} are non-negative), and $A^{(-)}$ and $\boldsymbol{e}^{(n)}$ are defined as

$$a_{ij}^{(-)} = \begin{cases} a_{ii}, & \text{for } i = j \\ a_{ij}, & \text{for } i \neq j,\ a_{ij} \leq 0\,, \\ 0, & \text{for } i \neq j,\ a_{ij} > 0 \end{cases} \quad \boldsymbol{e}^{(n)} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}. \tag{1.30}$$

Note that condition (1.29) is equivalent to that $A^{(-)}$ has nonnegative row sums. Moreover, $A = A^{(-)}$ and the condition (1.28) holds when A is an M-matrix [120]. From Lemma 1.5.1 we have the following lemma.

Lemma 1.5.2. *If the stiffness matrix A is an M-matrix and has nonnegative row sums, then the numerical scheme satisfies DMP.*

Numerical schemes satisfying DMP have been developed along the line of those sufficient conditions by either designing a proper discretization for the underlying PDE or employing a suitable mesh. To date most success has been made for the isotropic diffusion case where $\mathbb{D}$ is in the scalar matrix form, $\mathbb{D} = a(\boldsymbol{x})I$, with $a(\boldsymbol{x})$ being a scalar function; e.g., see [23, 27, 39, 84, 85, 86, 90, 91, 116, 125]. In particular, it is shown

in [23, 39] that the linear finite element method (FEM) satisfies DMP when the mesh is simplicial and satisfies the so-called non-obtuse angle condition requiring that the dihedral angles of all mesh elements be non-obtuse. In two dimensions this condition can be replaced by a weaker condition (the Delaunay condition) that the sum of any pair of angles opposite a common edge is less than or equal to π [91, 116]. Similar mesh conditions are developed in [84, 85, 86, 90] for elliptic problems with a nonlinear diffusion coefficient in the form $\mathbb{D} = a(\boldsymbol{x}, u, \nabla u)I$ and with mixed boundary conditions. Burman and Ern [27] propose a nonlinear stabilized Galerkin approximation for the Laplace operator and prove that it satisfies DMP on arbitrary meshes and for arbitrary space dimension without resorting to the non-obtuse angle condition.

On the other hand, the anisotropic diffusion case is more difficult and only limited success has been made [43, 50, 67, 68, 69, 89, 103, 104, 105, 93, 94, 95, 97, 110]. For example, Drăgănescu et al. [50] show that the non-obtuse angle condition fails to guarantee DMP satisfaction in the anisotropic diffusion case. The techniques proposed by Liska and Shashkov [95] and Kuzmin et al. [89] to locally modify (or repair) the underlying numerical scheme, by Sharma and Hammett [110] to employ slope limiters in the discretization of the PDE, by Mlacnik and Durlofsky [97] to optimize the mesh for a multipoint flux approximation (MPFA) finite volume method (e.g., see [1, 2] for the method), and by Li et al. [93] to optimize a triangular mesh for the finite element solution, help reduce spurious oscillations. A nonlinear, first order finite volume method developed by Le Potier [103, 104] and further improved by Lipnikov et al. [94] gives rise to a stiffness M-matrix on arbitrary meshes when applied to parabolic PDEs but fails to satisfy DMP when applied to steady-state elliptic problems. A first order finite difference method having similar features is proposed by Le Potier [105].

1.6 Outline

In the above sections of this chapter, we have given a brief introduction of the background of related topics in this dissertation, including anisotropic diffusion problems, variational problems, anisotropic mesh adaptation, finite element approximation and discrete maximum principle (DMP). This section gives an outline of the following chapters.

Chapter 2 discusses the DMP satisfaction for the linear finite element approximation of BVP (1.1) and (1.2). The "anisotropic non-obtuse angle condition" and several variants are derived which guarantee the satisfaction of DMP. The metric tensor based on the anisotropic non-obtuse angle condition is derived to account for DMP satisfaction. Examples are given to demonstrate the successful application of the derived metric tensor.

In Chapter 3, mesh adaptation based on both DMP satisfaction and interpolation error estimate is addressed, and an optimal metric tensor is obtained by minimizing an interpolation error bound. Numerical examples are presented to show the advantage of the combination.

Chapter 4 discusses anisotropic mesh adaptation for variational problems (1.10). A bound for the first variation of a general functional is derived. A formula for the metric tensor for use in anisotropic mesh adaptation is defined such that the bound is minimized on an M-uniform mesh. Numerical examples are given to demonstrate the feature that the resulting mesh also adapts to changes in the structure of the underlying problem.

Chapter 5 discusses the application of anisotropic mesh adaptation in image processing. The advantage of anisotropic mesh adaptation is demonstrated and DMP satisfaction is discussed.

Chapter 6 summarizes the obtained results from this study and discusses further research interests.

This dissertation is mostly based on the following publications.

- W. Huang, L. Kamenski, and X. Li, Anisotropic mesh adaptation for variational problems using error estimation based on hierarchical bases, *Canadian Applied Mathematics Quarterly* (Special issue for the 30th anniversary of CAIMS, in press)([78]).

- X. Li and W. Huang, An anisotropic mesh adaptation method for the finite element solution of heterogeneous anisotropic diffusion problems. *J. Comput. Phys.* 229 (2010), 8072-8094 ([92]).

- W. Huang and X. Li, An anisotropic mesh adaptation method for the finite element solution of variational problems, *Finite Elements in Analysis and Design*, 46 (2010) 61-73 ([79]).

- X. Li, D. Svyatskiy, and M. Shashkov. Mesh adaptation and discrete maximum principle for 2D anisotropic diffusion problems. *LANL technical report*, LA-UR 10-01227, 2007 ([93]).

Chapter 2

An anisotropic non-obtuse angle condition for DMP satisfaction

In this chapter we study the linear finite element solution of BVP (1.1) and (1.2) with a general diffusion matrix $\mathbb{D} = \mathbb{D}(\boldsymbol{x})$.

Firstly, we develop a generalization of the well known non-obtuse angle condition, the so-called anisotropic non-obtuse angle condition, so that the linear finite element approximation satisfies DMP when the mesh is simplicial and satisfies this condition. Then we derive a metric tensor for use in mesh generation based on the anisotropic non-obtuse angle condition. This is done by adopting the so-called M-uniform mesh approach (cf. §1.3) where an anisotropic mesh is generated as a uniform mesh in the metric specified by a tensor M. M-uniform meshes generated with the metric tensor satisfy the anisotropic non-obtuse angle condition and are aligned with the diffusion matrix $\mathbb{D}$ in the sense that the principal axes of the circumscribed ellipsoid of the elements are parallel to the primary diffusion direction of $\mathbb{D}$.

2.1 Anisotropic non-obtuse angle condition

Following the notation in §1.4, we now study under what mesh conditions the scheme (1.23) satisfies DMP. Our basic tool is Lemma 1.5.2, i.e., we show that A is an M-matrix and has non-negative row sums when the mesh satisfies the condition (2.5) below. To this end, we first introduce some notation.

Denote the vertices of K by $\boldsymbol{a}_1^K, \boldsymbol{a}_2^K, \cdots, \boldsymbol{a}_{d+1}^K$. The edge matrix of K is defined as

$$E_K = [\boldsymbol{a}_2^K - \boldsymbol{a}_1^K, \boldsymbol{a}_3^K - \boldsymbol{a}_1^K, \cdots, \boldsymbol{a}_{d+1}^K - \boldsymbol{a}_1^K].$$

From the definition of simplices, E_K is nonsingular [112]. Then, as shown in Fig. 2.1, a set of $\boldsymbol{q}$-vectors can be defined as

$$[\boldsymbol{q}_2^K, \boldsymbol{q}_3^K, \cdots, \boldsymbol{q}_{d+1}^K] = E_K^{-T}, \quad \boldsymbol{q}_1^K = -\sum_{i=2}^{d+1} \boldsymbol{q}_i^K. \tag{2.1}$$

This set of vectors has the following properties.

(i) By definition, it follows that

$$\begin{aligned} \boldsymbol{q}_i^K \cdot (\boldsymbol{a}_j^K - \boldsymbol{a}_1^K) &= \delta_{ij}, \\ \boldsymbol{q}_1^K \cdot (\boldsymbol{a}_j^K - \boldsymbol{a}_i^K) &= \delta_{1j}, \end{aligned} \qquad i = 2, \cdots, d+1;\ j = 1, \cdots, d+1 \tag{2.2}$$

where δ_{ij} is the Kronecker delta function.

(ii) Denote by $\boldsymbol{S}_i^K$ the face opposite to vertex $\boldsymbol{a}_i^K$ (i.e., the face not having $\boldsymbol{a}_i$ as a vertex). Then (2.2) implies that $\boldsymbol{q}_i^K$ is the inward normal to the face $\boldsymbol{S}_i^K$; see Fig. 2.1.

(iii) The dihedral angle, α_{ij}, between any two faces $\boldsymbol{S}_i^K$ and $\boldsymbol{S}_j^K$ $(i \neq j)$ is defined as the supplement of the angle between the inward normals to the faces. It can be

calculated by

$$\cos(\alpha_{ij}) = -\frac{\boldsymbol{q}_i^K \cdot \boldsymbol{q}_j^K}{\|\boldsymbol{q}_i^K\| \, \|\boldsymbol{q}_j^K\|}, \quad i \neq j. \tag{2.3}$$

(iv) It is known [22, 90] that, for any vertex of K with the global and local indices i and i_K, respectively, there holds

$$\nabla \phi_i|_K = \boldsymbol{q}_{i_K}^K. \tag{2.4}$$

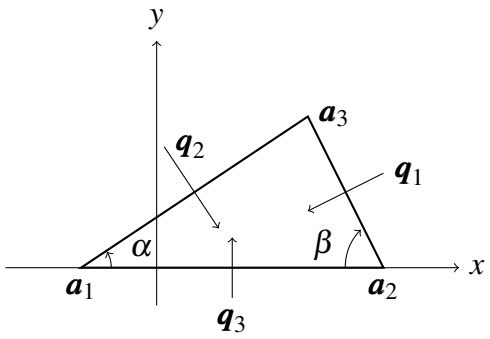


Figure 2.1: A sketch of the $\boldsymbol{q}$ vectors for an arbitrary element. The angles sharing the edge connecting vertices $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ are α and β.

The main result of this section is stated in the following theorem.

Theorem 2.1.1. *If the mesh satisfies the anisotropic non-obtuse angle condition*

$$(\boldsymbol{q}_i^K)^T \, \mathbb{D}_K \, \boldsymbol{q}_j^K \leq 0, \quad \forall i \neq j, \; i, j = 1, 2, ..., d+1, \; \forall K \in \mathscr{T}_h \tag{2.5}$$

then the linear finite element scheme (1.19) for solving BVP (1.1) and (1.2) satisfies DMP.

Proof. We prove this theorem using Lemma 1.5.2. That is, we show that the stiffness matrix A has non-negative row sums and is an M-matrix when the mesh satisfies condition (2.5).

(i) We first show that A has non-negative row sums. From (1.24) we only need to show $\sum_{j=1}^{N_v} a_{ij} \geq 0$ for $i = 1, ..., N_{vi}$. From (1.25) we have

$$\begin{aligned}
\sum_{j=1}^{N_v} a_{ij} &= \sum_{j=1}^{N_v} \sum_{K \in \mathscr{T}_h} |K| \, (\nabla \phi_i|_K)^T \, \mathbb{D}_K \, \nabla \phi_j|_K \\
&= \sum_{K \in \mathscr{T}_h} |K| \, (\nabla \phi_i|_K)^T \, \mathbb{D}_K \, \nabla \left(\sum_{j=1}^{N_v} \phi_j \right) \Bigg|_K \\
&= 0, \qquad (2.6)
\end{aligned}$$

where we have used the fact that $\sum_{j=1}^{N_v} \phi_j(\boldsymbol{x}) \equiv 1$ for any $\boldsymbol{x} \in K$.

(ii) Next we show that

$$a_{ij} \leq 0, \qquad \forall \, i \neq j, \; i, j = 1, ..., N_v \qquad (2.7)$$

$$a_{ii} \geq 0, \qquad \forall \, i = 1, ..., N_v. \qquad (2.8)$$

Let ω_i (or ω_j) be the patch of the elements containing $\boldsymbol{a}_i$ (or $\boldsymbol{a}_j$) as a vertex. Notice that $\nabla \phi_i|_K = 0$ when $K \notin \omega_i$. Denote the local indices of vertices $\boldsymbol{a}_i$ and $\boldsymbol{a}_j$ on K by i_K and j_K, respectively. Then from (1.25), (2.4), and (2.5), we have, for $i \neq j$, $i = 1, ..., N_{vi}$, $j = 1, ..., N_v$,

$$\begin{aligned}
a_{ij} &= \sum_{K \in \omega_i \cap \omega_j} |K| \, (\nabla \phi_i|_K)^T \, \mathbb{D}_K \, \nabla \phi_j|_K \\
&= \sum_{K \in \omega_i \cap \omega_j} |K| (\boldsymbol{q}_{i_K}^K)^T \, \mathbb{D}_K \, \boldsymbol{q}_{j_K}^K \qquad (2.9) \\
&\leq 0. \qquad (2.10)
\end{aligned}$$

From (1.24) it is obvious that $a_{ij} = 0$ for $i \neq j$, $i = N_{vi} + 1, \cdots N_v$, $j = 1, ..., N_v$. Hence, the off-diagonal entries of A are non-positive.

The inequality (2.8) follows immediately from (1.24), (1.25), and the positive definiteness of $\mathbb{D}_K$.

(iii) We now show that A_{11} defined in (1.24) is an M-matrix. Notice that the nonnegativeness of the row sums of A and the properties (2.7) and (2.8) imply that A_{11} is diagonally dominant. In theory, we can show that A_{11} is an M-matrix by proving it is irreducible [120]. However, we will need to assume that any pair of interior vertices is connected at least by an interior edge path [50]. To avoid this additional restriction on the mesh, we instead opt to show A_{11} is symmetric and positive definite, which together with (2.7) and (2.8) implies that A_{11} is an M-matrix [120].

From (1.25) it is obvious that A_{11} is symmetric. It suffices to show A_{11} is positive definite. From the strictly positive definiteness of the diffusion matrix $\mathbb{D}$, there exists a positive constant β such that

$$\mathbb{D}_K \geq \beta \boldsymbol{I}, \quad \forall K \in \mathscr{T}_h.$$

For any vector $\boldsymbol{v} = (v_1, ..., v_{N_{vi}})^T$, we define $v^h = \sum_{i=1}^{N_{vi}} v_i \phi_i \in U_0^h$. From the definition of A_{11} and the fact that $\nabla v^h|_K$ is constant on K, we have

$$\begin{aligned}
\boldsymbol{v}^T A_{11} \boldsymbol{v} &= \sum_{K \in \mathscr{T}_h} |K| \, (\nabla v^h|_K)^T \, \mathbb{D}_K \, \nabla v^h|_K \\
&\geq \beta \sum_{K \in \mathscr{T}_h} |K| \, (\nabla v^h|_K)^T \, \nabla v^h|_K \\
&= \beta \sum_{K \in \mathscr{T}_h} \int_K (\nabla v^h)^T \nabla v^h d\boldsymbol{x} \\
&= \beta \int_\Omega (\nabla v^h)^T \nabla v^h d\boldsymbol{x} \\
&\geq \beta C_p \int_\Omega |v^h|^2 d\boldsymbol{x},
\end{aligned}$$

where in the last step we have used Poincare's inequality and $C_p > 0$ is the associated constant. For any nonzero vector $\boldsymbol{v}$, $v^h = \sum_{i=1}^{N_{vi}} v_i \phi_i \not\equiv 0$ and is piecewise linear and continuous on Ω. Consequently,

$$\boldsymbol{v}^T A_{11} \boldsymbol{v} \geq \beta C_p \int_\Omega |v^h|^2 d\boldsymbol{x} > 0, \quad \forall \boldsymbol{v} \neq 0,$$

which implies that A_{11} is positive definite. Hence, A_{11} is an M-matrix.

(iv) From (1.24) it is easy to verify that the inverse of A is given by

$$A^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1} A_{12} \\ 0 & I \end{bmatrix}.$$

Then (2.7) and the fact $A_{11}^{-1} \geq 0$ imply that $A^{-1} \geq 0$ and therefore A is an M-matrix.

We have shown above that A is an M-matrix and has non-negative row sums. By Lemma 1.5.2 we conclude that the linear FEM satisfies DMP when the simplicial mesh satisfies (2.5). □

Remark 2.1.1. For the isotropic case where $\mathbb{D} = a(\boldsymbol{x})\boldsymbol{I}$ for some scalar function $a(\boldsymbol{x})$, condition (2.5) reduces to the well known non-obtuse angle condition [22, 39]

$$\boldsymbol{q}_i^K \cdot \boldsymbol{q}_j^K \leq 0, \quad \forall i \neq j, \forall K \in \mathscr{T}_h, \tag{2.11}$$

which requires the dihedral angles α_{ij} (cf. (2.3)) of all mesh elements be non-obtuse. Thus, condition (2.5) is a generalization of the non-obtuse angle condition. An alternative interpretation of (2.5) is that the dihedral angles of element K, measured in the Riemannian metric $\mathbb{D}_K$ (piecewise constant), are non-obtuse. □

Remark 2.1.2. It is interesting to point out that an explicit mesh condition similar to (2.5) is obtained by Eigestad et al. [52] for a multipoint flux approximation (MPFA)

finite volume method on triangular meshes for anisotropic homogeneous media (i.e., $\mathbb{D}$ is constant). Moreover, (2.5) reduces to a mesh condition obtained by Li et al. [93] for a similar situation with constant $\mathbb{D}$ and triangular meshes. To see this, let the eigen-decomposition of the constant diffusion matrix $\mathbb{D}$ be

$$\mathbb{D} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} k_1 & 0 \\ 0 & k_2 \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}. \tag{2.12}$$

For an arbitrary triangular element K, denote the angles sharing the edge connecting vertices $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ by α and β; see Fig. 2.1. Then, a mesh condition of [93] is given by

$$\begin{cases} -k_1 \sin\beta \sin\alpha + k_2 \cos\beta \cos\alpha \le 0, \\ -k_2 \cos\beta \le 0, \\ -k_2 \cos\alpha \le 0, \end{cases} \tag{2.13}$$

provided that the edge connecting $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ is parallel to the primary diffusion direction $(\cos\theta, \sin\theta)^T$ (the eigenvector corresponding to the first eigenvalue of $\mathbb{D}$, k_1). We now show that (2.5) reduces to (2.13) for the current situation. Without loss of generality we assume that the primary diffusion direction and the edge connecting $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ are in the direction of the x-axis; cf. Fig. 2.1. (In this case we have $\theta = 0$.) It is not difficult to obtain

$$\boldsymbol{q}_1 = c_1 \begin{bmatrix} -\sin\beta \\ -\cos\beta \end{bmatrix}, \quad \boldsymbol{q}_2 = c_2 \begin{bmatrix} \sin\alpha \\ -\cos\alpha \end{bmatrix}, \quad \boldsymbol{q}_3 = c_3 \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

where c_1, c_2, and c_3 are positive constants. From these and (2.12), (2.5) reduces to

$$\begin{cases} \boldsymbol{q}_1^T \mathbb{D}_K \boldsymbol{q}_2 = \boldsymbol{q}_1^T \mathbb{D} \boldsymbol{q}_2 = c_1 c_2 (-k_1 \sin\alpha \sin\beta + k_2 \cos\alpha \cos\beta) \le 0, \\ \boldsymbol{q}_1^T \mathbb{D}_K \boldsymbol{q}_3 = \boldsymbol{q}_1^T \mathbb{D} \boldsymbol{q}_3 = c_1 c_3 (-k_2 \cos\beta) \le 0, \\ \boldsymbol{q}_2^T \mathbb{D}_K \boldsymbol{q}_3 = \boldsymbol{q}_2^T \mathbb{D} \boldsymbol{q}_3 = c_2 c_3 (-k_2 \cos\alpha) \le 0, \end{cases}$$

which gives (2.13). □

It is often more convenient to express the anisotropic non-obtuse angle condition (2.5) in terms of mapping F_K from $\hat{K}$ to K. Denote the Jacobian matrix of F_K by F_K'. We define the vectors $\hat{\boldsymbol{q}}_k$, $k = 1, ..., d+1$ for the reference element $\hat{K}$ as in (2.1). The chain rule of differentiation implies

$$\nabla \phi_i = (F_K')^{-T} \nabla_\xi \hat{\phi}_i,$$

where $\hat{\phi}_i(\boldsymbol{\xi}) = \phi_i(F_K(\boldsymbol{\xi}))$. From (2.4), we have

$$\boldsymbol{q}_i = (F_K')^{-T} \hat{\boldsymbol{q}}_i.$$

Inserting this into (2.5) we obtain the following theorem.

Theorem 2.1.2. *If the mesh satisfies*

$$\hat{\boldsymbol{q}}_i^T (F_K')^{-1} \mathbb{D}_K (F_K')^{-T} \hat{\boldsymbol{q}}_j \le 0, \quad \forall i \ne j,\ i, j = 1, ..., d+1,\ \forall K \in \mathscr{T}_h \tag{2.14}$$

then the linear finite element scheme (1.19) for solving BVP (1.1) and (1.2) satisfies DMP.

Corollary 2.1.1. *Suppose that the reference element $\hat{K}$ is taken as a simplex with non-obtuse dihedral angles. If the mesh satisfies*

$$(F_K')^{-1}\mathbb{D}_K(F_K')^{-T} = C_K I, \quad \forall K \in \mathscr{T}_h \tag{2.15}$$

where C_K is a positive constant on K and I is the $d \times d$ identity matrix, then the linear finite element scheme (1.19) for solving BVP (1.1) and (1.2) satisfies DMP.

Proof. Since $\hat{K}$ is a simplex with non-obtuse dihedral angles, we have

$$\hat{\boldsymbol{q}}_i^T \hat{\boldsymbol{q}}_j \leq 0,\ i \neq j, \quad i,j = 1,...,d+1.$$

From this it is easy to see that (2.15) is sufficient for (2.14) to hold. □

The mesh condition (2.15) will be used in the next section to develop metric tensor accounting for DMP satisfaction. The metric tensor is needed in anisotropic mesh generation. It is emphasized that (2.15), as well as mesh conditions (2.5) and (2.14), can also be used more directly via direct minimization [93, 97] or variational formulation [72] for optimizing the current mesh to improve DMP satisfaction.

2.2 Metric tensor based on DMP satisfaction

In this section we develop a metric tensor for use in anisotropic mesh generation based on mesh condition (2.15). We adopt the so-called M-uniform mesh approach [74, 75, 80] as introduced in §1.3.

To determine M from mesh condition (2.15), we first notice that the left and right sides of (1.12) represents the arithmetic and geometric means of the eigenvalues of matrix $(F_K')^T M_K F_K'$, respectively. From the arithmetic-mean geometric-mean inequal-

ity, (1.12) implies that all of the eigenvalues are equal to each other. In other words, $(F_K')^T M_K F_K'$ is a scalar matrix, i.e.,

$$(F_K')^T M_K F_K' = \tilde{C}_K I \quad \text{or} \quad (F_K')^{-1} M_K^{-1} (F_K')^{-T} = \tilde{C}_K^{-1} I \tag{2.16}$$

for some constant $\tilde{C}_K$. A direct comparison of (2.16) with (2.15) suggests that the metric tensor M be chosen in the form

$$M_{DMP,K} = \theta_K \mathbb{D}_K^{-1}, \quad \forall K \in \mathscr{T}_h, \tag{2.17}$$

where $\theta = \theta_K > 0$ is an arbitrary piecewise constant function. Thus, any M-uniform mesh associated with a metric tensor in the form (2.17) satisfies condition (2.15). The following theorem follows from Corollary 2.1.1.

Theorem 2.2.1. *Suppose that the reference element $\hat{K}$ is taken to be equilateral and unitary in volume. For an M-uniform mesh associated with any metric tensor in the form (2.17), the linear finite element scheme (1.19) for solving BVP (1.1) and (1.2) satisfies DMP.*

Remark 2.2.1. Since an M-uniform mesh satisfies the alignment condition (1.12), we can conclude that when M is chosen in the form (2.17), a corresponding M-uniform mesh is aligned with the diffusion matrix $\mathbb{D}$ in the sense that the principal axes of the circumscribed ellipsoid of the elements are parallel to the eigenvectors of $\mathbb{D}_K$ while their lengths are proportional to the square roots of the respective eigenvalues. As a consequence, the length of an element is greater in a faster diffusion direction and smaller in a slower diffusion direction. A small length scale of mesh elements in slow diffusion directions helps reduce numerical dissipation in those directions. □

Remark 2.2.2. Note that $\theta = \theta_K$ in (2.17) is arbitrary. Thus, in addition to satisfying DMP, there is a degree of freedom for the mesh to account for other considerations. In the next chapter we shall consider mesh adaptation based on error estimate and choose θ_K to minimize a certain error bound. □

2.3 Numerical results

In this section we present three two-dimensional examples to demonstrate the performance of metric tensors M_{DMP} in (2.17) with $\theta_K = 1$ based on DMP satisfaction. For comparison purpose, we also include numerical results obtained with almost uniform meshes (labelled with M_{unif}). The iterative procedure for solving PDEs has been given in §1.3 (cf. Fig. 1.5).

Example 2.3.1. The first example is to consider BVP (1.1) and (1.2) with

$$f \equiv 0, \quad \Omega = [0,1]^2 \backslash \left[\frac{4}{9}, \frac{5}{9}\right]^2, \quad g = 0 \text{ on } \Gamma_{out}, \quad g = 2 \text{ on } \Gamma_{in},$$

where Γ_{out} and Γ_{in} are the outer and inner boundaries of Ω, respectively; see Fig. 2.2. The diffusion matrix is given by (2.12) with $k_1 = 1000$, $k_2 = 1$, and θ being the angle of the primary diffusion direction (parallel to the first eigenvector of $\mathbb{D}$).

This example satisfies the maximum principle and the solution (whose analytical expression is unavailable) stays between 0 and 2. Our goal is to produce a numerical solution which also satisfies DMP and stays between 0 and 2. This example has been studied in [89, 93].

We first consider the case of constant $\mathbb{D}$ with $\theta = \pi/4$. Fig. 2.3 shows finite element solutions obtained with M_{unif} and M_{DMP}. Meshes and solution contours obtained with those two metric tensors are shown in Figs. 2.4 and 2.5, respectively. No overshoots

in the finite element solutions are observed for all cases. However, undershoots and unphysical minima occur in the solution obtained with M_{unif} ($u_{min} = -0.0602$) (cf. Fig. 2.5)(a)).

The results confirm the theoretical prediction that the solutions obtained with M_{DMP} satisfy DMP and no overshoot/undershoot and no unphysical extremum occur. It should be pointed out that the solution contour obtained with an almost uniform mesh is smooth and the sharp jumps of the solution are smeared; see Figs. 2.3(a) and 2.5(a).

Next we consider a case of variable $\mathbb{D}$ with $\theta = \pi \sin(x)\cos(y)$. The finite element solutions, meshes, and solution contours are shown in Figs. 2.6, 2.7, and 2.8, respectively. Similar observations as for the constant $\mathbb{D}$ case can be made. Especially, undershoots and unphysical extrema occur in the solutions obtained with M_{unif} but not with M_{DMP}. Once again, the results confirm our theoretical predictions in the previous sections.

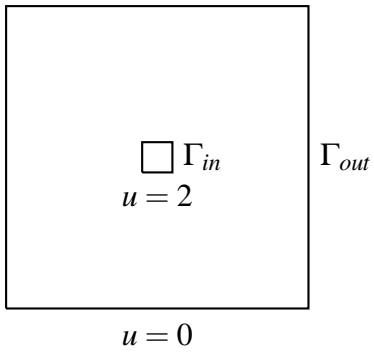
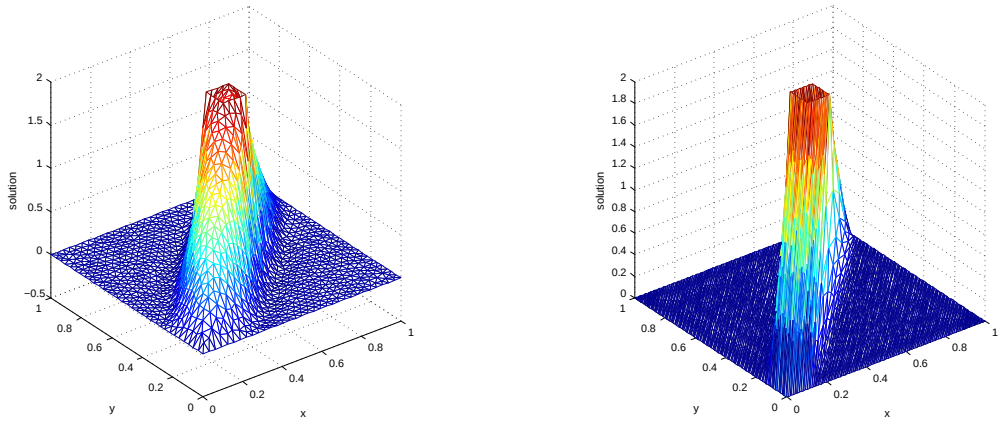


Figure 2.2: The physical domain and boundary conditions for Example 2.3.1.

Example 2.3.2. In this example, we consider BVP (1.1) and (1.2) with

$$f \equiv 0, \quad g(x,0) = g(16,y) = 0,$$



(a): M_{unif} (b): M_{DMP}

Figure 2.3: Example 2.3.1 with constant $\mathbb{D}$. Finite element solutions obtained with (a) M_{unif} and (b) M_{DMP}.

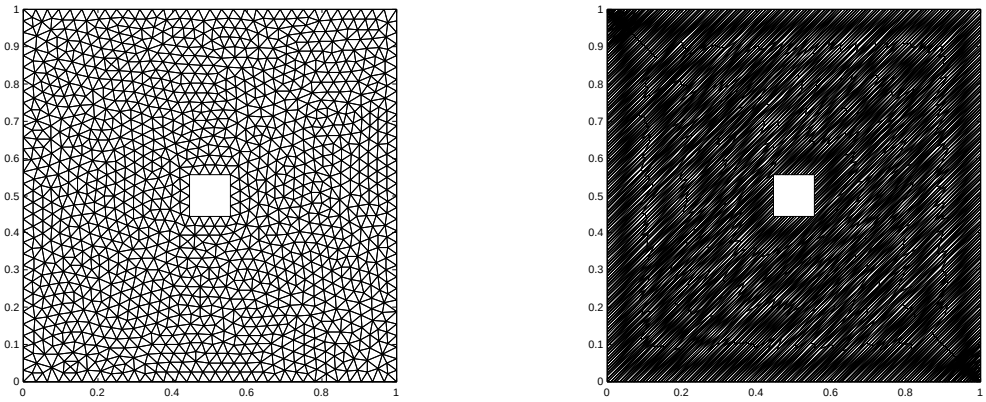
$$g(0,y) = \begin{cases} 0.5y & \text{if } 0 \le y < 2, \\ 1 & \text{if } 2 \le y \le 16, \end{cases} \quad \text{and } g(x,16) = \begin{cases} 1 & \text{if } 0 \le x \le 14, \\ 8 - 0.5x & \text{if } 14 < x \le 16. \end{cases}$$

The diffusion matrix is defined as

$$\mathbb{D}(x,y) = \begin{pmatrix} 500.5 & 499.5 \\ 499.5 & 500.5 \end{pmatrix}.$$

This is a simple example with a constant but anisotropic $\mathbb{D}$ and with a continuous boundary condition. It satisfies the maximum principle and its solution stays between 0 and 1.

Numerical solutions, meshes, and solution contours are shown in Figs. 2.9, 2.10, and 2.11, respectively. For this example, both undershoots and overshoots are observed in the computed solutions with M_{unif} but not with with M_{DMP}. This example demonstrates that a scheme violating DMP can produce unphysical extrema even for a simple



(a): M_{unif}, $N_v = 2460$ (b): M_{DMP}, $N_v = 2530$

Figure 2.4: Example 2.3.1 with constant $\mathbb{D}$. Meshes obtained with (a) M_{unif} and (b) M_{DMP}.

problem with constant diffusion, continuous boundary conditions, and a convex domain.

Example 2.3.3. This example is given by (1.1) and (1.2) with

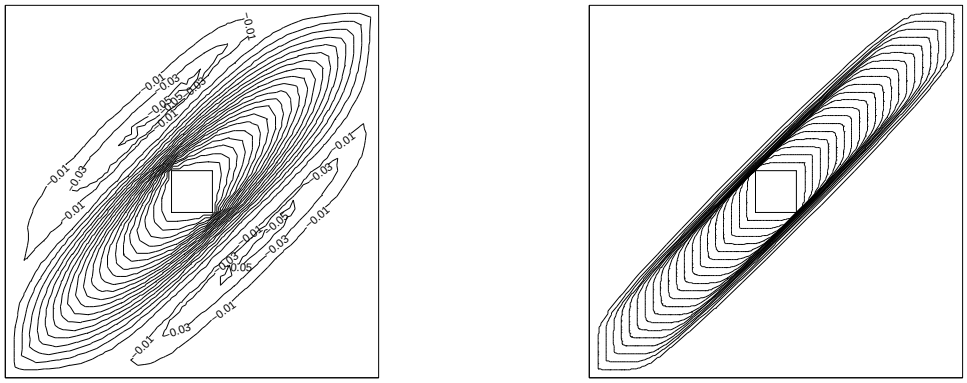
$$\Omega = (0,1)\times(0,1), \quad f(x,y) = \begin{cases} 4.0, & \text{if } x < 0.5 \\ -5.6, & \text{if } x > 0.5 \end{cases}, \quad u = u_{exact} \text{ on } \partial\Omega,$$

$$\mathbb{D}(x,y) = \begin{cases} D_1, & \text{if } x < 0.5, \\ D_2, & \text{if } x > 0.5, \end{cases} \quad D_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad D_2 = \begin{pmatrix} 10 & 3 \\ 3 & 1 \end{pmatrix}.$$

The problem has the exact solution

$$u(x,y) = \begin{cases} 1 - 2y^2 + 4xy + 2y + 6x, & \text{if } x \le 0.5 \\ -2y^2 + 1.6xy - 0.6x + 3.2y + 4.3, & \text{if } x > 0.5. \end{cases} \tag{2.18}$$

Note that the value and primary diffusion direction of the diffusion matrix change across the line $x = 0.5$. This example has been studied in [89].



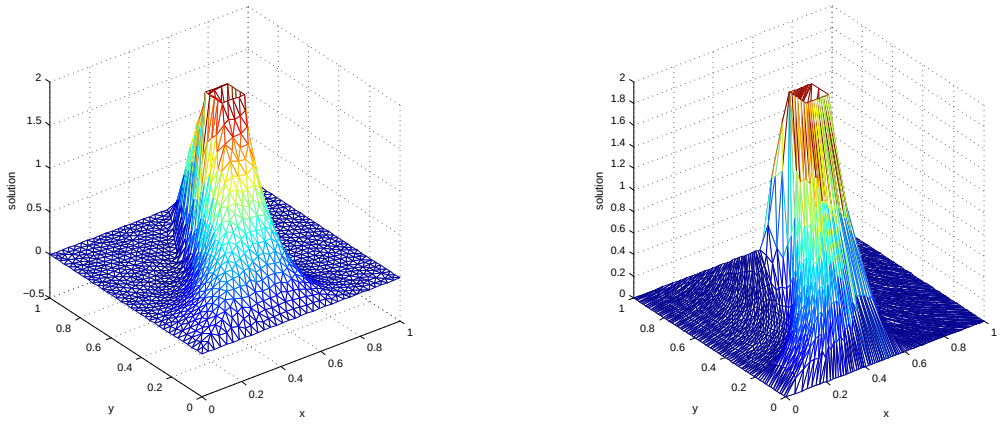
(a): M_{unif}, $u_{min} = -0.0602$ (b): M_{DMP}, $u_{min} = 0$

Figure 2.5: Example 2.3.1 with constant $\mathbb{D}$. Contours of the finite element solutions obtained with (a) M_{unif} and (b) M_{DMP}.

Solutions and meshes obtained with M_{unif} and M_{DMP} are shown in Fig. 2.12. For this example, no overshoots and undershoots is observed for all numerical solutions. The meshes obtained with M_{DMP} show a better alignment with the primary diffusion direction than that obtained with M_{unif}. The results are consistent with what is expected from the construction of the metric tensors. The errors were computed and compared in §3.2 with other results.

2.4 Conclusions and comments

In this chapter we have developed a mesh condition (2.5) under which the linear finite element approximation of anisotropic diffusion problem (1.1) and (1.2) validates the discrete counterpart of the maximum principle satisfied by the continuous problem. The condition is a generalization of the well known non-obtuse angle condition developed for isotropic diffusion problems and requires that the dihedral angles of mesh elements measured in a metric depending only on the diffusion matrix be non-obtuse.



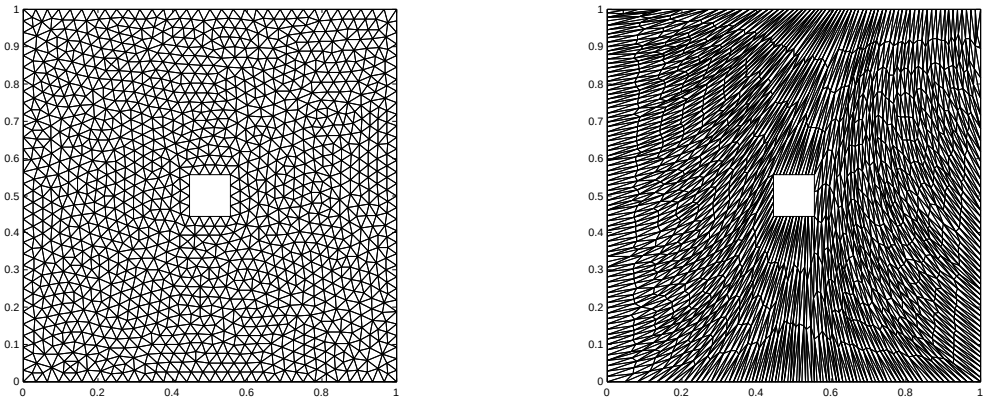
(a): M_{unif} (b): M_{DMP}

Figure 2.6: Example 2.3.1 with variable $\mathbb{D}$. Finite element solutions obtained with (a) M_{unif} and (b) M_{DMP}.

We have also developed two variants of the anisotropic non-obtuse angle condition, (2.14) and (2.15), which can be more convenient to use in actual mesh generation. Indeed, metric tensor (2.17) for use in anisotropic mesh generation is derived based on (2.15) for accounting for DMP satisfaction. Features of these metric tensors are illustrated in numerical examples.

It is worth pointing out that condition (2.5) has been derived based on the local stiffness matrix on a mesh element. Like the non-obtuse angle condition for isotropic diffusion problems, (2.5) can be relaxed by considering the global stiffness matrix as a whole in two dimensions [76]. Moreover, we have restricted our attention to linear PDE (1.1) and Dirichlet boundary condition (1.2). But the procedure developed in this work can be extended to problems with nonlinear diffusion $\mathbb{D} = \mathbb{D}(\boldsymbol{x}, u, \nabla u)$ and mixed boundary conditions (e.g., see [84, 85, 86, 90]) without major modification.

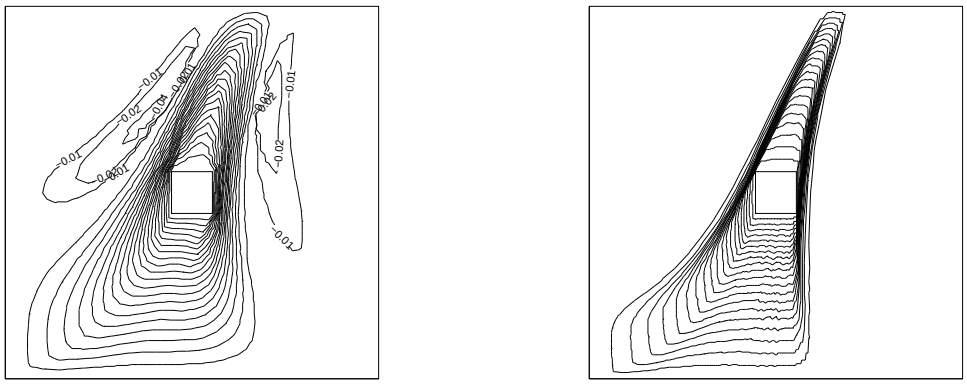
Although the numerical examples have been presented in 2D, the anisotropic non-obtuse angle condition (2.5) and the corresponding metric tensor formula (2.17) are d-dimensional ($d = 1, 2, 3$). In 3D, a Delaunay triangulation may not guarantee the sat-



(a): M_{unif}, $N_v = 2460$ (b): M_{DMP}, $N_v = 2510$

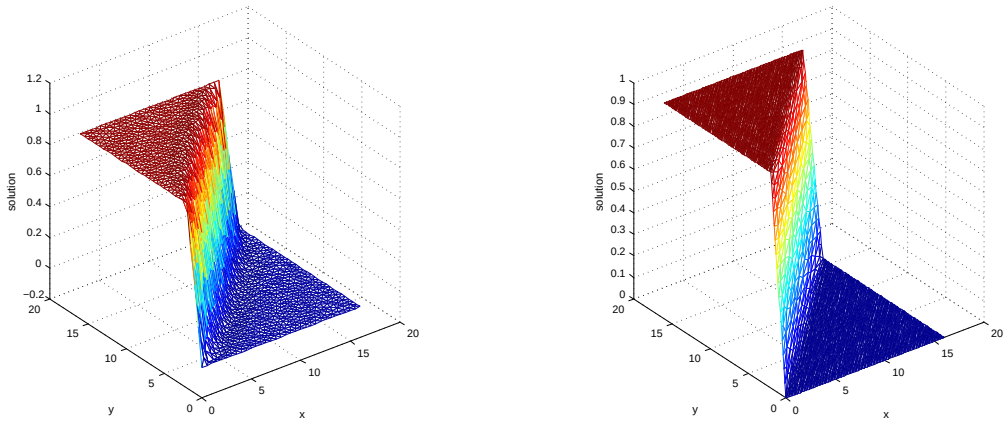
Figure 2.7: Example 2.3.1 with variable $\mathbb{D}$. Meshes obtained with (a) M_{unif} and (b) M_{DMP}.

isfaction of DMP [91]. Nevertheless, polyhedrons can be decomposed into tetrahedra satisfying the non-obtuse angle condition (2.11) and therefore the numerical solution satisfies DMP; e.g., see [24]. It is interesting to know that if it also works for the anisotropic non-obtuse angle condition (2.5) for a given metric tensor M and if a 3D triangulation can be generated to (approximately) satisfy the M-uniform mesh conditions (1.11) and (1.12). Those will be the topics to investigate in the future.



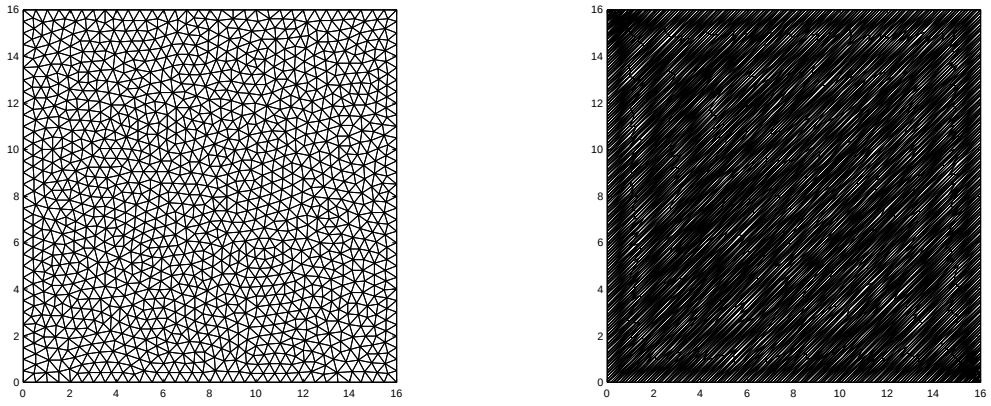
(a): M_{unif}, $u_{min} = -0.0506$ (b): M_{DMP}, $u_{min} = 0$

Figure 2.8: Example 2.3.1 with variable $\mathbb{D}$. Contours of the finite element solutions obtained with (a) M_{unif} and (b) M_{DMP}.



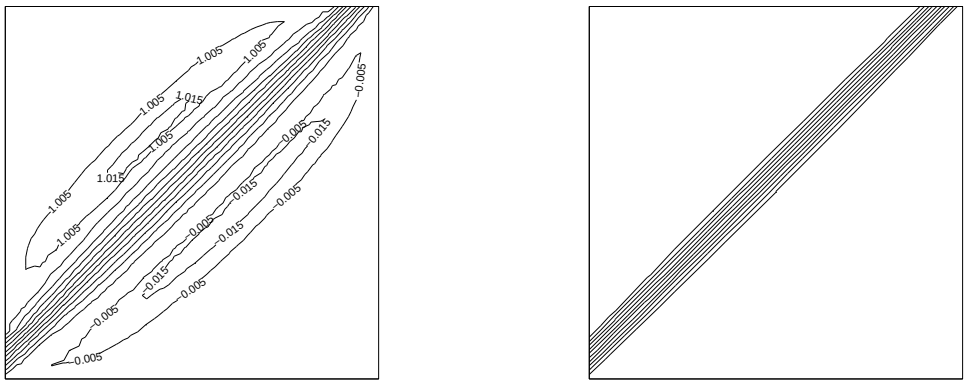
(a): M_{unif}, $u_{min} = -0.0208$ (b): M_{DMP}, $u_{min} = 0$

Figure 2.9: Example 2.3.2. Finite element solutions obtained with (a) M_{unif} and (b) M_{DMP}.



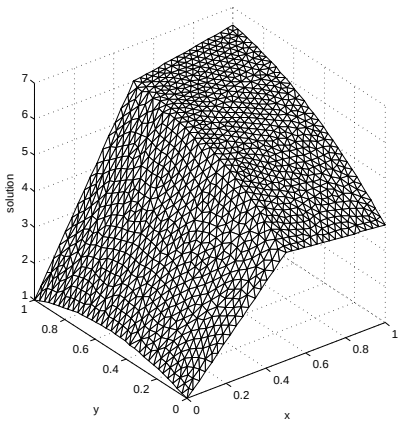
(a): M_{unif}, $N_v = 2480$ (b): M_{DMP}, $N_v = 2754$

Figure 2.10: Example 2.3.2. The adaptive meshes obtained with (a) M_{unif} and (b) M_{DMP}.

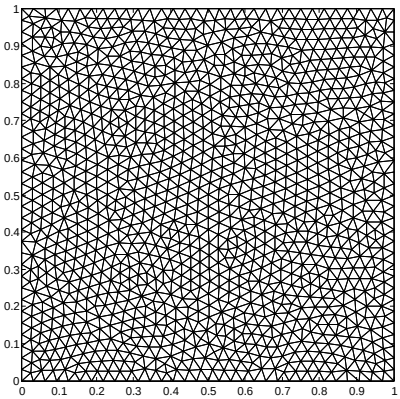


(a): M_{unif}, $u_{min} = -0.0208$, $u_{max} = 1.0209$ (b): M_{DMP}, $u_{min} = 0$, $u_{max} = 1.0$

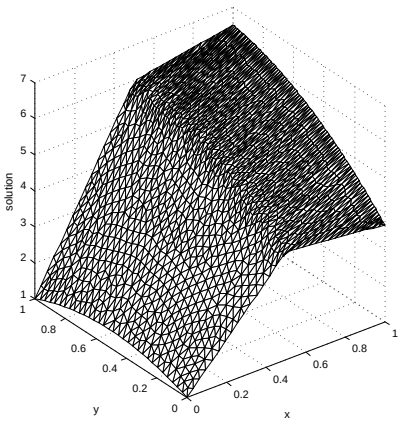
Figure 2.11: Example 2.3.2. Contours of the finite element solutions obtained with (a) M_{unif} and (b) M_{DMP}.



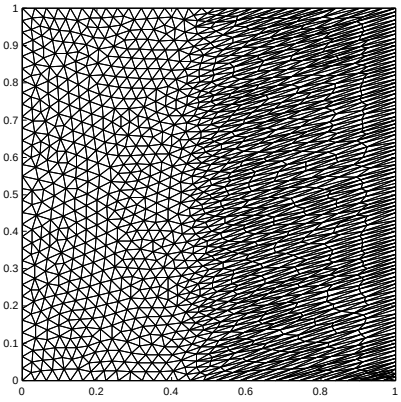
(a): M_{unif}, numerical solution, $u_{min} = 0$



(b): M_{unif}, mesh, $N_v = 2350$



(c): M_{DMP}, numerical solution, $u_{min} = 0$



(d): M_{DMP}, mesh, $N_v = 2415$

Figure 2.12: Example 2.3.3. Numerical solutions and meshes obtained with M_{unif} and M_{DMP}.

Chapter 3

DMP satisfaction and mesh adaptivity

In the previous chapter, we have discussed anisotropic mesh adaptation based on DMP satisfaction. If the mesh satisfies the anisotropic non-obtuse angle condition (2.5), the linear finite element solution is guaranteed to satisfy DMP.

On the other hand, mesh adaptation based on error estimate is important for improving the efficiency and accuracy of the computation as mentioned in §1.3. Moreover, for anisotropic diffusion problems (cf. §1.1), the numerical dissipation along slow diffusion directions can be reduced via both the alignment of the mesh elements along the fast diffusion direction and the small spacing of mesh elements in the slow diffusion direction. Hence, both DMP satisfaction and error based mesh adaptation are important for the numerical solution of anisotropic diffusion problems. So far, existing research has focused on either DMP satisfaction or mesh adaptation. But none has combined both DMP satisfaction and error based mesh adaptation for the numerical solution of anisotropic diffusion problems.

For simplicity, we refer to mesh adaptation based on error estimates as "mesh adaptivity", and the adaptation based on DMP satisfaction as "DMP satisfaction". In this chapter, the combination of DMP satisfaction and mesh adaptivity is investigated. An optimal metric tensor accounting for both considerations is obtained by minimizing an

interpolation error bound, and advantages of using adaptive, DMP-satisfied meshes are demonstrated in numerical examples. To our best knowledge, this is the first effort that mesh adaptivity and DMP satisfaction are combined in the numerical solution of isotropic/anisotropic diffusion problems.

3.1 Metric tensor based on both DMP satisfaction and mesh adaptivity

In Chapter 2, we have developed a metric tensor based on DMP satisfaction in the form of (2.17). As mentioned in Remark 2.2.2, the parameter θ_K can be taken to account for mesh adaptivity. In this section we develop a metric tensor taking both the satisfaction of DMP and mesh adaptivity into consideration with the scalar function $\theta = \theta_K$ in (2.17) being determined to minimize an interpolation error bound. For simplicity, we consider here an error bound for linear Lagrange interpolation. Other interpolation error bounds (e.g., see [75]) can be considered without major modification.

Lemma 3.1.1. *([75]) Let $K \subset \mathbb{R}^d$ be a simplicial element and Π_h be the linear Lagrange interpolation operator. Then,*

$$|v - \Pi_h v|_{H^1(K)} \leq C \|(F_K')^{-1}\| \left[\int_K \left[tr\left((F_K')^T |H(v)| F_K' \right) \right]^2 d\boldsymbol{x} \right]^{\frac{1}{2}}, \quad \forall v \in H^2(K) \quad (3.1)$$

where $\|\cdot\|$ denotes the l_2 matrix norm, $H(v)$ is the Hessian of v, and $|H(v)| = \sqrt{H(v)^2}$.

Lemma 3.1.2. *For any given $d \times d$ symmetric matrix S, there holds that*

$$|\, tr(A^T S A)| \leq\ tr(A^T A)\ \|S\|, \quad \forall A \in \mathbb{R}^{d \times d}. \quad (3.2)$$

If S is further positive definite, then

$$\|S\|^{-1}\ tr(A^TSA) \leq\ tr(A^TA) \leq\ tr(A^TSA)\ \|S^{-1}\|. \tag{3.3}$$

Proof. Denote the eigen-decomposition of S by

$$S = Q\Sigma Q^T,$$

where Q is an orthogonal matrix, $\Sigma = \text{ diag}(\lambda_1, ..., \lambda_d)$, and $\lambda_i,\ i = 1, ..., d$ are the eigenvalues of S. Write

$$A^TQ = [\boldsymbol{v}_1, ..., \boldsymbol{v}_d].$$

Then

$$A^TSA = (A^TQ)\Sigma(Q^TA) = [\boldsymbol{v}_1, ..., \boldsymbol{v}_d]\Sigma[\boldsymbol{v}_1, ..., \boldsymbol{v}_d]^T = \sum_i \lambda_i \boldsymbol{v}_i \boldsymbol{v}_i^T.$$

It follows that

$$\begin{aligned} |\ \text{tr}(A^TSA)| &= |\sum_i \lambda_i\ \text{tr}(\boldsymbol{v}_i\boldsymbol{v}_i^T)| \\ &= |\sum_i \lambda_i \|\boldsymbol{v}_i\|^2| \\ &\leq \sum_i \|\boldsymbol{v}_i\|^2 \cdot |\lambda|_{max} \\ &= \text{tr}(A^TA)\ \|S\|, \end{aligned}$$

which gives (3.2). Inequality (3.3) follows from (3.2) and that

$$\text{tr}(A^TA) = \text{ tr}(A^TS^{\frac{1}{2}}S^{-1}S^{\frac{1}{2}}A) \leq \text{ tr}(A^TSA)\ \|S^{-1}\|. \tag{3.4}$$

□

The scalar function $\theta = \theta_K$ in (2.17) is then determined based on interpolation error bound (3.1). From the definition of the Frobenius matrix norm, we have

$$\|A\| \leq \|A\|_F = \sqrt{\operatorname{tr}(A^T A)} = \sqrt{\operatorname{tr}(AA^T)}, \quad \forall A \in \mathbb{R}^{d\times d}.$$

Using this, taking squares of both sides of (3.1), and summing the result over all elements of $\mathscr{T}_h$, we have

$$\begin{aligned}
|u-\Pi_h u|^2_{H^1(\Omega)} &= \sum_{K\in\mathscr{T}_h} |u-\Pi_h u|^2_{H^1(K)} \\
&\leq C \sum_{K\in\mathscr{T}_h} \|(F_K')^{-1}\|^2 \int_K \left[\operatorname{tr}\left((F_K')^T |H(u)| F_K'\right)\right]^2 d\boldsymbol{x} \\
&\leq C \sum_{K\in\mathscr{T}_h} \|(F_K')^{-1}\|_F^2 \int_K \left[\operatorname{tr}\left((F_K')^T |H(u)| F_K'\right)\right]^2 d\boldsymbol{x} \\
&= C \sum_{K\in\mathscr{T}_h} \left[\operatorname{tr}((F_K')^{-1}(F_K')^{-T})\right] \int_K \left[\operatorname{tr}\left((F_K')^T |H(u)| F_K'\right)\right]^2 d\boldsymbol{x}.
\end{aligned}$$

From Lemma 3.1.2 it follows that

$$\begin{aligned}
&|u-\Pi_h u|^2_{H^1(\Omega)} \\
&\leq C \sum_{K\in\mathscr{T}_h} \left[\operatorname{tr}((F_K')^{-1}\mathbb{D}_K(F_K')^{-T})\right] \cdot \|\mathbb{D}_K^{-1}\| \cdot \int_K \left[\operatorname{tr}((F_K')^T\mathbb{D}_K^{-1}(F_K'))\right]^2 \|\mathbb{D}_K |H(u)|\|^2 d\boldsymbol{x} \\
&= C \sum_{K\in\mathscr{T}_h} |K| \cdot \left[\operatorname{tr}((F_K')^{-1}\mathbb{D}_K(F_K')^{-T})\right] \cdot \left[\operatorname{tr}((F_K')^T\mathbb{D}_K^{-1}(F_K'))\right]^2 \\
&\qquad \times \|\mathbb{D}_K^{-1}\| \cdot \frac{1}{|K|} \int_K \|\mathbb{D}_K |H(u)|\|^2 d\boldsymbol{x}.
\end{aligned} \tag{3.5}$$

Consider an M-uniform mesh $\mathscr{T}_h$ corresponding to a metric tensor M_K in the form (2.17). Then, alignment condition (1.12) reduces to

$$\frac{1}{d} \operatorname{tr}\left((F_K')^T\mathbb{D}_K^{-1}F_K'\right) = \det\left((F_K')^T\mathbb{D}_K^{-1}F_K'\right)^{\frac{1}{d}}. \tag{3.6}$$

From the arithmetic-mean geometric-mean inequality, (3.6) implies that all of the eigenvalues of matrix $(F'_K)^T \mathbb{D}_K^{-1} F'_K$ are equal to each other. As a consequence, all of the eigenvalues of the inverse of $(F'_K)^T \mathbb{D}_K^{-1} F'_K$ are equal to each other, which in turn implies

$$\frac{1}{d} \operatorname{tr}\left((F'_K)^{-1} \mathbb{D}_K (F'_K)^{-T}\right) = \det\left((F'_K)^{-1} \mathbb{D}_K (F'_K)^{-T}\right)^{\frac{1}{d}}. \tag{3.7}$$

Inserting (3.6) and (3.7) into (3.5) and noticing

$$\det\left((F'_K)^T \mathbb{D}_K^{-1} F'_K\right) = |K|^2 \det(\mathbb{D}_K)^{-1}, \quad \det\left((F'_K)^{-1} \mathbb{D}_K (F'_K)^{-T}\right) = |K|^{-2} \det(\mathbb{D}_K),$$

we have

$$|u - \Pi_h u|^2_{H^1(\Omega)} \leq C \sum_{K \in \mathscr{T}_h} |K|^{\frac{d+2}{d}} \det(\mathbb{D}_K)^{-\frac{1}{d}} \|\mathbb{D}_K^{-1}\| \cdot \frac{1}{|K|} \int_K \|\mathbb{D}_K |H(u)|\|^2 d\boldsymbol{x}. \tag{3.8}$$

Rewrite this bound as

$$|u - \Pi_h u|^2_{H^1(\Omega)} \leq C \sum_{K \in \mathscr{T}_h} |K|^{\frac{d+2}{d}} B_K, \tag{3.9}$$

where

$$B_K = \det(\mathbb{D}_K)^{-\frac{1}{d}} \|\mathbb{D}_K^{-1}\| \cdot \frac{1}{|K|} \int_K \|\mathbb{D}_K |H(u)|\|^2 d\boldsymbol{x}. \tag{3.10}$$

Notice that $\int_K \|\mathbb{D}_K |H(u)|\|^2 d\boldsymbol{x}$ and therefore B_K can vanish locally. To ensure the positive definiteness of the metric tensor to be defined, we regularize the above bound with a parameter $\alpha_h > 0$ as

$$|u - \Pi_h u|^2_{H^1(\Omega)} \leq C \sum_{K \in \mathscr{T}_h} |K|^{\frac{d+2}{d}} [\alpha_h + B_K] = C\alpha_h \sum_{K \in \mathscr{T}_h} |K|^{\frac{d+2}{d}} \left[1 + \frac{1}{\alpha_h} B_K\right]. \tag{3.11}$$

From Hölder's inequality, we have

$$\begin{aligned}\sum_{K\in\mathscr{T}_h}|K|^{\frac{d+2}{d}}\left[1+\frac{1}{\alpha_h}B_K\right] &= \sum_{K\in\mathscr{T}_h}\left(|K|\left[1+\frac{1}{\alpha_h}B_K\right]^{\frac{d}{d+2}}\right)^{\frac{d+2}{d}} \\ &\geq N^{-\frac{2}{d}}\left(\sum_{K\in\mathscr{T}_h}|K|\left[1+\frac{1}{\alpha_h}B_K\right]^{\frac{d}{d+2}}\right)^{\frac{d+2}{d}}, \quad (3.12)\end{aligned}$$

with equality in the last step if and only if

$$|K|\left[1+\frac{1}{\alpha_h}B_K\right]^{\frac{d}{d+2}} = \text{ constant}, \quad \forall K\in\mathscr{T}_h. \quad (3.13)$$

A direct comparison of this with equidistribution condition (1.11) suggests that the optimal ρ_K be defined as

$$\rho_K = \left[1+\frac{1}{\alpha_h}B_K\right]^{\frac{d}{d+2}}. \quad (3.14)$$

From the relation $\rho_K = \sqrt{\det(M_K)}$, we find the optimal θ_K and M_K as

$$\theta_K = \rho_K^{\frac{2}{d}}\det(\mathbb{D}_K)^{\frac{1}{d}} = \left[1+\frac{1}{\alpha_h}B_K\right]^{\frac{2}{d+2}}\det(\mathbb{D}_K)^{\frac{1}{d}}, \quad (3.15)$$

$$M_{DMP+adap,K} = \left[1+\frac{1}{\alpha_h}B_K\right]^{\frac{2}{d+2}}\det(\mathbb{D}_K)^{\frac{1}{d}}\,\mathbb{D}_K^{-1}, \quad (3.16)$$

where B_K is defined in (3.10). With the so-defined metric tensor, the error bound can be obtained from (3.11) and (3.12) for a corresponding M-uniform mesh as

$$|u-\Pi_h u|_{H^1(\Omega)} \leq CN^{-\frac{1}{d}}\sqrt{\alpha_h}\sigma_h^{\frac{d+2}{2d}}. \quad (3.17)$$

To complete the definition, we need to determine the regularization parameter α_h. We follow [74] to define α_h such that

$$\sigma_h \equiv \sum_{K\in\mathscr{T}_h} \rho_K |K| \leq 2|\Omega|, \tag{3.18}$$

with which roughly 50% of the mesh points are concentrated in regions of large ρ_K. From (3.14) and Jensen's inequality, we have

$$\begin{aligned}
\sigma_h &= \sum_{K\in\mathscr{T}_h} |K| \left[1 + \frac{1}{\alpha_h} B_K\right]^{\frac{d}{d+2}} \\
&\leq \sum_{K\in\mathscr{T}_h} |K| \left[1 + \alpha_h^{-\frac{d}{d+2}} B_K^{\frac{d}{d+2}}\right] \\
&= |\Omega| + \alpha_h^{-\frac{d}{d+2}} \sum_{K\in\mathscr{T}_h} |K| B_K^{\frac{d}{d+2}}.
\end{aligned} \tag{3.19}$$

By requiring the above bound to be less than or equal to $2|\Omega|$, we obtain

$$\alpha_h = \left(\frac{1}{|\Omega|} \sum_{K\in\mathscr{T}_h} |K| B_K^{\frac{d}{d+2}}\right)^{\frac{d+2}{d}}. \tag{3.20}$$

Combining (3.17) with (3.18) and (3.20) and summarizing the above derivation, we have the following theorem.

Theorem 3.1.1. *Suppose that the reference element $\hat{K}$ is chosen to be equilateral and unitary in volume. For any M-uniform simplicial mesh corresponding to the metric tensor (3.16), the linear finite element scheme (1.19) for solving BVP (1.1) and (1.2) satisfies DMP and the interpolation error for the exact solution u is bounded by*

$$|u - \Pi_h u|_{H^1(\Omega)} \leq C N^{-\frac{1}{d}} \left(\sum_{K\in\mathscr{T}_h} |K| B_K^{\frac{d}{d+2}}\right)^{\frac{d+2}{2d}}, \tag{3.21}$$

where B_K is defined in (3.10).

It is remarked that the metric tensor (3.16) (cf. (3.10)) depends on the second derivatives of the exact solution u which is what we are seeking/approximating. In actual computation, the second derivatives are replaced with approximations obtained with a Hessian recovery technique such as the one of using piecewise quadratic polynomials fitting in least-squares sense to nodal values of the computed solution (e.g., see [74]). A hierarchical basis error estimator can also be used to approximate the Hessian of the exact solution. It is shown in [77] that the least-squares fitting and the hierarchical basis methods work comparably for all considered cases except for one where the diffusion coefficient is discontinuous and the interfaces are predefined in the mesh. In this case, the latter works better than the former since hierarchical basis estimation does not over-concentrate mesh elements near the interfaces. Since our main goal is to study DMP satisfaction instead of the discontinuity of the diffusion coefficient, we choose to use the least squares fitting method for Hessian recovery in our computation due to its simplicity and problem independent feature.

It is interesting to note that the term in the bracket in (3.21) can be viewed as a Riemann sum of an integral, i.e.,

$$\sum_{K\in\mathscr{T}_h} |K| B_K^{\frac{d}{d+2}} \sim \int_\Omega \det(\mathbb{D})^{-\frac{1}{d+2}} \|\mathbb{D}^{-1}\|^{\frac{d}{d+2}} \cdot \|\mathbb{D}|H(u)|\|^{\frac{2d}{d+2}} d\boldsymbol{x}.$$

Thus, the interpolation error has an asymptotic bound as

$$\begin{aligned} |u-\Pi_h u|_{H^1(\Omega)} &\leq CN^{-\frac{1}{d}} \left(\sum_{K\in\mathscr{T}_h} |K| B_K^{\frac{d}{d+2}} \right)^{\frac{d+2}{2d}} \\ &\sim CN^{-\frac{1}{d}} \left(\int_\Omega \det(\mathbb{D})^{-\frac{1}{d+2}} \|\mathbb{D}^{-1}\|^{\frac{d}{d+2}} \cdot \|\mathbb{D}|H(u)|\|^{\frac{2d}{d+2}} d\boldsymbol{x} \right)^{\frac{d+2}{2d}}. \end{aligned} \tag{3.22}$$

We emphasize that both the satisfaction of DMP and mesh adaptation (through minimization of an error bound) are taken into account in the definition of metric tensor (3.16). An interesting question is what the interpolation error bound looks like if mesh adaptation is not taken into consideration. For example, we consider a case $\theta_K = 1$ in (2.17). This gives the metric tensor

$$M_K = \mathbb{D}_K^{-1}. \tag{3.23}$$

Recall that the interpolation error is bounded in (3.9), i.e.,

$$|u - \Pi_h u|_{H^1(\Omega)} \leq C \left(\sum_{K \in \mathscr{T}_h} |K|^{\frac{d+2}{d}} B_K \right)^{\frac{1}{2}}, \tag{3.24}$$

where B_K is defined in (3.10). Moreover, for an M-uniform mesh corresponding to this metric tensor the equidistribution condition (1.11) reduces to

$$\det(\mathbb{D}_K)^{-\frac{1}{2}} |K| = \frac{\sigma_h}{N}, \tag{3.25}$$

where $\sigma_h = \sum_{K \in \mathscr{T}_h} \det(\mathbb{D}_K)^{-\frac{1}{2}} |K|$. Inserting (3.25) into (3.24), we have

$$\begin{aligned}
|u - \Pi_h u|_{H^1(\Omega)} &\leq C \left(\sum_{K \in \mathscr{T}_h} |K| \left(\det(\mathbb{D}_K)^{\frac{1}{2}} \frac{\sigma_h}{N} \right)^{\frac{2}{d}} B_K \right)^{\frac{1}{2}} \\
&= C N^{-\frac{1}{d}} \sigma_h^{\frac{1}{d}} \left(\sum_{K \in \mathscr{T}_h} |K| \det(\mathbb{D}_K)^{\frac{1}{d}} B_K \right)^{\frac{1}{2}} \\
&= C N^{-\frac{1}{d}} \left(\sum_{K \in \mathscr{T}_h} \det(\mathbb{D}_K)^{-\frac{1}{2}} |K| \right)^{\frac{1}{d}} \left(\sum_{K \in \mathscr{T}_h} |K| \det(\mathbb{D}_K)^{\frac{1}{d}} B_K \right)^{\frac{1}{2}}.
\end{aligned}$$

Thus,

$$|u-\Pi_h u|_{H^1(\Omega)} \leq CN^{-\frac{1}{d}} \left(\sum_{K\in\mathscr{T}_h} \det(\mathbb{D}_K)^{-\frac{1}{2}}|K| \right)^{\frac{1}{d}} \left(\sum_{K\in\mathscr{T}_h} |K|\det(\mathbb{D}_K)^{\frac{1}{d}} B_K \right)^{\frac{1}{2}} \tag{3.26}$$

$$\sim CN^{-\frac{1}{d}} \left(\int_\Omega \det(\mathbb{D})^{-\frac{1}{2}} d\boldsymbol{x} \right)^{\frac{1}{d}} \left(\int_\Omega \|\mathbb{D}^{-1}\| \cdot \|\mathbb{D}\,|H(u)|\|^2 d\boldsymbol{x} \right)^{\frac{1}{2}}. \tag{3.27}$$

This is the interpolation error bound for an M-uniform mesh corresponding to metric tensor (3.23).

From Hölder's inequality, it follows that

$$\left(\sum_{K\in\mathscr{T}_h} |K| B_K^{\frac{d}{d+2}} \right)^{\frac{d+2}{2d}} \leq \left(\sum_{K\in\mathscr{T}_h} \det(\mathbb{D}_K)^{-\frac{1}{2}}|K| \right)^{\frac{1}{d}} \left(\sum_{K\in\mathscr{T}_h} |K|\det(\mathbb{D}_K)^{\frac{1}{d}} B_K \right)^{\frac{1}{2}}.$$

Thus, the solution-dependent factor of bound (3.21) is small than or equal to that of bound (3.26). In this sense, $M_{DMP+adap}$ defined in (3.16) leads to a more accurate interpolant than M_{DMP} defined in (3.23) (or (2.17) with $\theta_K = 1$).

Moreover, from the standard interpolation theory we recall that the interpolation error for a uniform mesh is bounded by

$$|u-\Pi_h u|_{H^1(\Omega)} \leq CN^{-\frac{1}{d}} \left(\int_\Omega \|\nabla^2 u\|^2 d\boldsymbol{x} \right)^{\frac{1}{2}}. \tag{3.28}$$

It is easy to see that the solution dependent factor in error bound (3.22) for $M_{DMP+adap}$ is in the order of $|\nabla^2 u|_{L^{\frac{2d}{d+2}}(\Omega)}$ and those in (3.27) for M_{DMP} and (3.28) for a uniform mesh are in the order of $|\nabla^2 u|_{L^2(\Omega)}$. Thus, (3.22) has the smallest solution dependent factor, an indication of the advantage of using adaptive meshes. On the other hand, the error bounds (3.22) and (3.27) depend on the determinant and norm of the diffusion matrix $\mathbb{D}$ and its inverse. This indicates that DMP satisfaction may sacrifice accuracy.

Indeed, as we shall see in §3.2, the solution error for DMP-bound meshes can sometimes be larger than that for a uniform mesh.

3.2 Numerical results

In this section, we take the same examples as shown in §2.3 but for $M_{DMP+adap}$ in (3.16) combining DMP satisfaction and mesh adaptivity. The results are compared with those obtained with M_{DMP} in §2.3. For comparison purpose, we also include numerical results obtained with a metric tensor M_{adap} based on minimization of a bound on the H^1 semi-norm of linear interpolation error [74]:

$$M_{adap,K} = \rho_K^{\frac{2}{d}} \det\left(I + \frac{1}{\alpha_h}|H_K(u)|\right)^{-\frac{1}{d}} \left[I + \frac{1}{\alpha_h}|H_K(u)|\right], \tag{3.29}$$

where

$$\rho_K = \left\|I + \frac{1}{\alpha_h}|H_K(u)|\right\|_F^{\frac{d}{d+2}} \det\left(I + \frac{1}{\alpha_h}|H_K(u)|\right)^{\frac{1}{d+2}},$$

and α_h is defined implicitly through

$$\sum_{K\in\mathscr{T}_h} \rho_K|K| = 2|\Omega|.$$

As mentioned before, the second derivatives of the exact solution are replaced with approximations obtained with a Hessian recovery technique in actual computation. An iterative procedure for solving PDEs is given in §1.3 (cf. Fig. 1.5).

Example 2.3.1. For both cases with a constant and a variable θ we consider, the exact solution has sharp jumps near the inner boundary (cf. Figs. 2.3 and 2.6) so mesh adaptation is needed for a proper resolution of them.

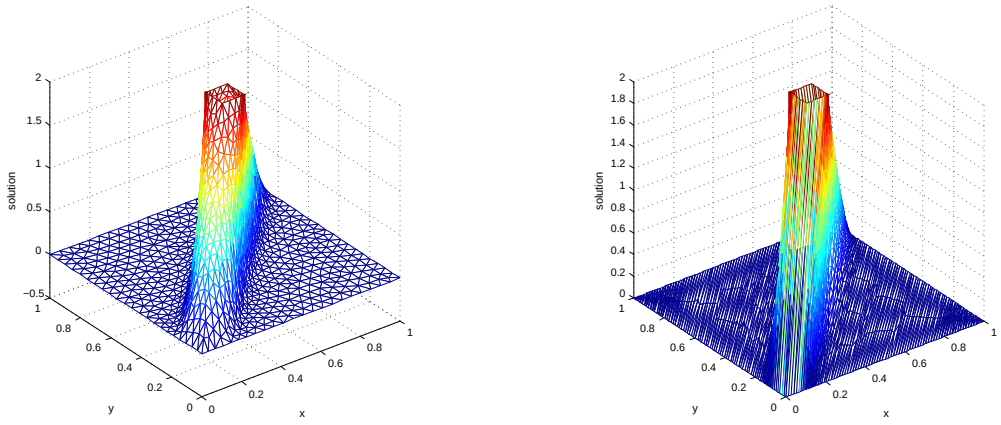
Fig. 3.1 shows finite element solutions obtained with M_{adap} and $M_{DMP+adap}$ for the case of constant $\mathbb{D}$ with $\theta = \pi/4$. The meshes and solution contours are shown in Figs. 3.2 and 3.3, respectively. No overshoots in the finite element solutions are observed for all cases. However, undershoots and unphysical minima occur in he solutions obtained with M_{adap} ($u_{min} = -0.0039$) (cf. Fig. 3.1 (a), 3.3(a)). Fig. 3.4 shows the decrease of $-u_{min}$ as the mesh is refined. For the range of the number of mesh elements considered, the undershooting improves at a rate of $-u_{min} = O(N^{-0.5})$ for both M_{unif} and M_{adap}.

The results confirm the theoretical prediction that the solutions obtained with $M_{DMP+adap}$ satisfy DMP and no overshoot/undershoot and no unphysical extremum occur. The solution contour obtained with $M_{DMP+adap}$ is comparable to the one obtained with M_{DMP} (cf. Fig. 2.5(b)). Furthermore, the mesh obtained with $M_{DMP+adap}$ (cf. Fig. 3.2 (b)) distributes more elements than the one obtained with M_{DMP} (cf. Fig. 2.4 (b)) near the inner boundary where the sharp jump of the solution occurs, which helps to improve the accuracy and reduce the numerical dissipation.

For the case of variable $\mathbb{D}$ with $\theta = \pi \sin(x)\cos(y)$, the finite element solutions, meshes, and solution contours are shown in Figs. 3.5, 3.6, and 3.7, respectively. Similar observations as for the constant $\mathbb{D}$ case can be made. Again, more mesh elements are distributed near the inner boundary for $M_{DMP+adap}$ (cf. Fig. 3.6(b)) than for M_{DMP} (cf. Fig. 2.7(b)).

Example 2.3.2. Numerical solutions, meshes and solution contours are shown in Figs. 3.8, 3.9, and 3.10, respectively. For this example, both undershoots and overshoots are observed in the computed solutions with M_{adap} but not with $M_{DMP+adap}$. Comparing Fig. 3.9(b) and 2.10(b), $M_{DMP+adap}$ gives better mesh adaptation in the place where sharp jump of solution occurs than M_{DMP}.

Example 2.3.3. Solutions and meshes obtained with M_{adap} and $M_{DMP+adap}$ are shown in Fig. 3.11. For this example, no overshoots and undershoots are observed

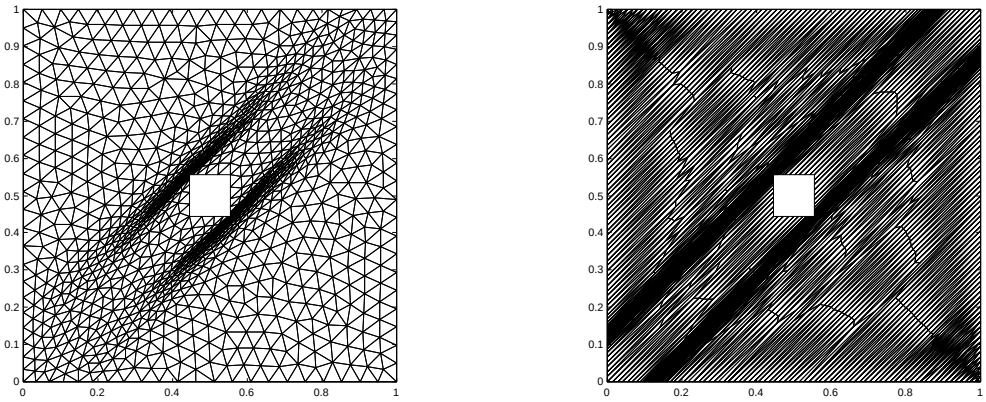


(a): M_{adap} (b): $M_{DMP+adap}$

Figure 3.1: Example 2.3.1 with constant $\mathbb{D}$. Finite element solutions obtained with (a) M_{adap} and (b) $M_{DMP+adap}$.

for all numerical solutions. The meshes obtained with $M_{DMP+adap}$ show a better alignment with the primary diffusion direction than that obtained with M_{adap}. Elements are concentrated along the line $x = 0.5$ for the meshes obtained with $M_{DMP+adap}$ (cf. Fig. 3.11(d)) whereas there is no concentration in the mesh for M_{DMP} (cf. Fig. 2.12)(d). The results are consistent with what is expected from the construction of the metric tensors.

When the interface ($x = 0.5$) is not predefined in the mesh, the H^1 semi-norm and L^2 norm of the error are shown in Fig. 3.12 as functions of the number of mesh elements. Metric tensor M_{adap} leads to far more accurate results than the other three metric tensors, which produce comparable results for the considered range of N. Moreover, M_{adap} and $M_{DMP+adap}$ give the same convergence rate, i.e., $|e^h|_{H^1(\Omega)} = O(N^{-0.5})$ and $\|e^h\|_{L^2(\Omega)} = O(N^{-1})$, while M_{unif} and M_{DMP} result in a slower convergence rate, $|e^h|_{H^1(\Omega)} = O(N^{-0.25})$ and $\|e^h\|_{L^2(\Omega)} = O(N^{-0.5})$. This demonstrates the advantage of using adaptive meshes. Interestingly, the results in [89] (Table 4) obtained with a slope-limited scheme for triangular meshes also show a similar slow convergence.



(a): M_{adap}, $N_v = 2583$ (b): $M_{DMP+adap}$, $N_v = 2474$

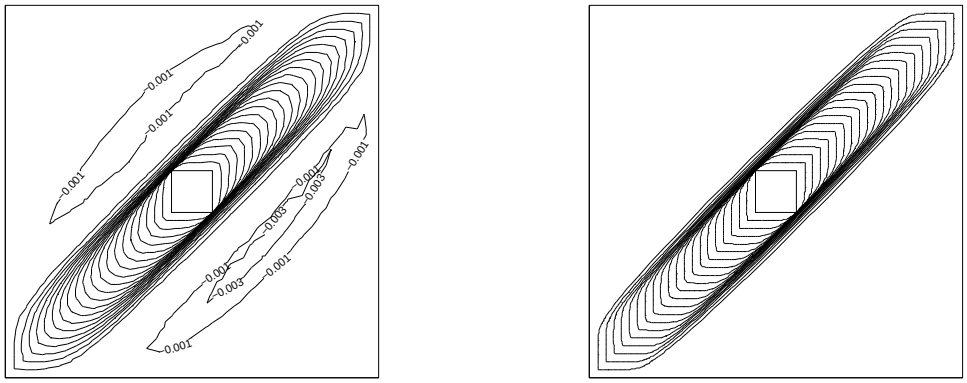
Figure 3.2: Example 2.3.1 with constant $\mathbb{D}$. Meshes obtained with (a) M_{adap} and (b) $M_{DMP+adap}$.

If the interface is predefined in the mesh, then the solution (2.18) can be approximated accurately in the linear finite element space. As shown in Fig. 3.13, all metric tensors produce comparable solutions and the same convergence rate $|e^h|_{H^1(\Omega)} = O(N^{-0.5})$ and $\|e^h\|_{L^2(\Omega)} = O(N^{-1})$.

3.3 Conclusions and comments

In this chapter, we have developed an optimal metric tensor $M_{DMP+adap}$ (3.16) accounting for both DMP satisfaction and mesh adaptivity (based on error estimate) for the linear finite element approximation of anisotropic diffusion problem (1.1) and (1.2). The metric tensor is obtained from (2.17) (which satisfies (2.15)) by minimizing an interpolation error bound.

In terms of DMP satisfaction, the results are comparable with those obtained with metric tensor M_{DMP} (2.17) discussed in Chapter 2. The numerical solutions obtained with M_{DMP} and $M_{DMP+adap}$ both satisfy the discrete maximum principle (DMP). More-



(a): M_{adap}, $u_{min} = -0.0039$ (b): $M_{DMP+adap}$, $u_{min} = 0$

Figure 3.3: Example 2.3.1 with constant $\mathbb{D}$. Contours of the finite element solutions obtained with (a) M_{adap} and (b) $M_{DMP+adap}$.

over, $M_{DMP+adap}$ provides mesh adaptation based on error estimate in addition to DMP satisfaction, which helps improving the efficiency and accuracy and reducing the numerical dissipation in the computations.

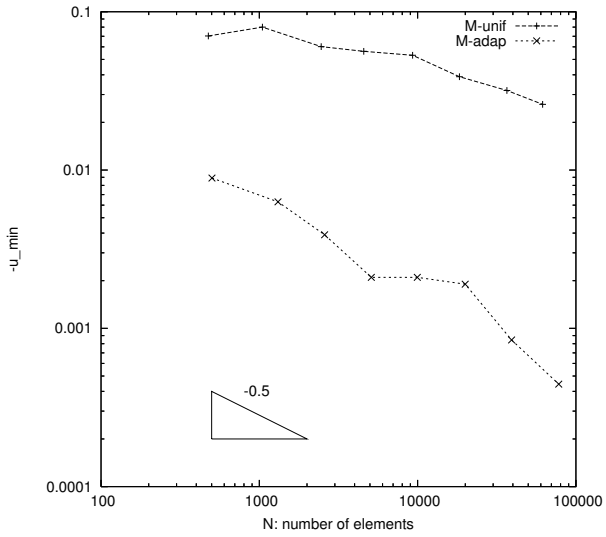
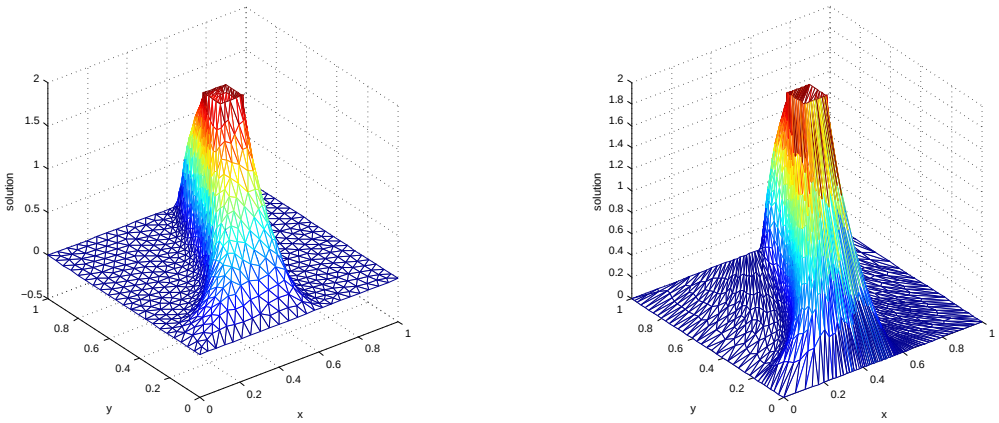
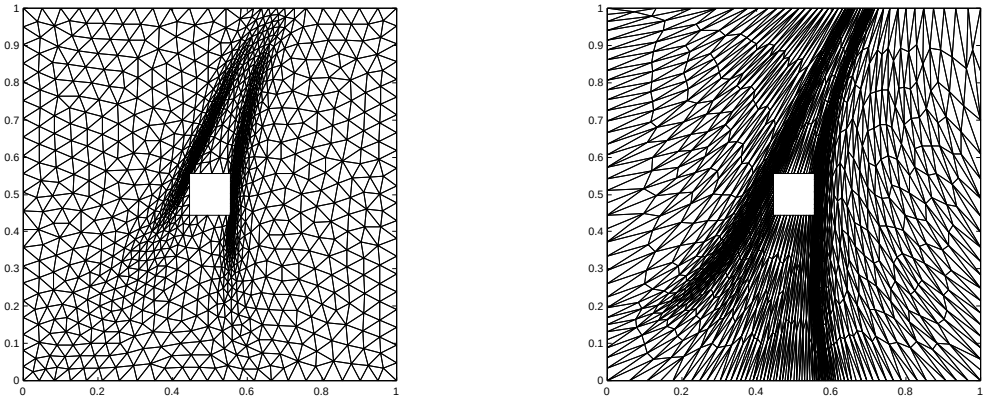


Figure 3.4: Example 2.3.1 with constant $\mathbb{D}$. The undershoot, $-u_{min}$, is shown as functions of the number of elements.



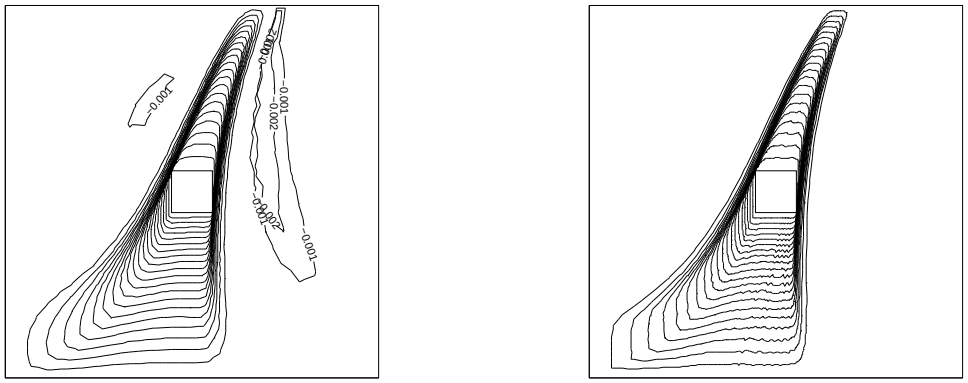
(a): M_{adap} (b): $M_{DMP+adap}$

Figure 3.5: Example 2.3.1 with variable $\mathbb{D}$. Finite element solutions obtained with (a) M_{adap} and (b) $M_{DMP+adap}$.



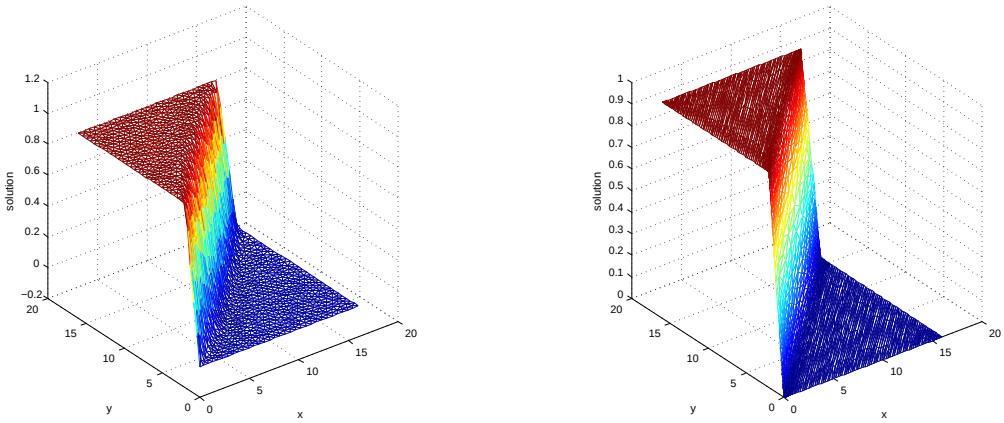
(a): M_{adap}, $N_v = 2568$ (b): $M_{DMP+adap}$, $N_v = 2463$

Figure 3.6: Example 2.3.1 with variable $\mathbb{D}$. Meshes obtained with (a) M_{adap} and (b) $M_{DMP+adap}$.



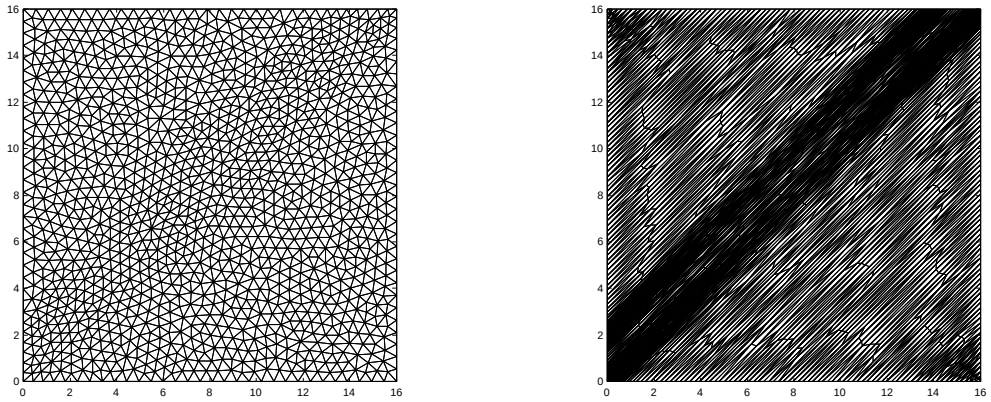
(a): M_{adap}, $u_{min} = -0.0039$ (b): $M_{DMP+adap}$, $u_{min} = 0$

Figure 3.7: Example 2.3.1 with variable $\mathbb{D}$. Contours of the finite element solutions obtained with (a) M_{adap} and (b) $M_{DMP+adap}$.



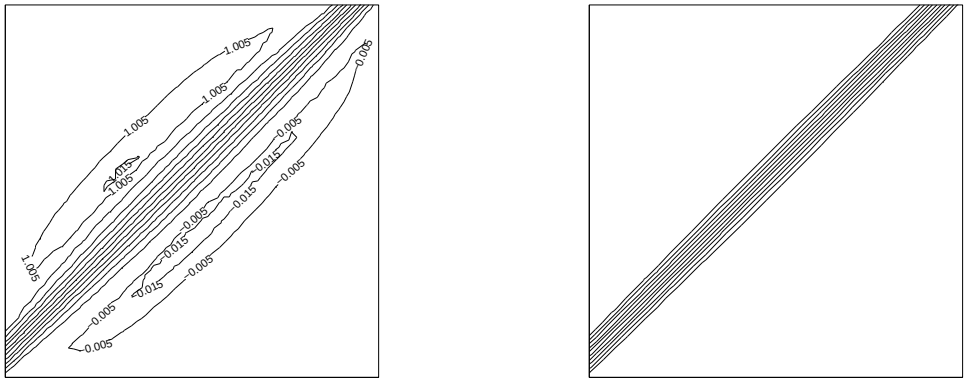
(a): M_{adap}, $u_{min} = -0.0195$ (b): $M_{DMP+adap}$, $u_{min} = 0$

Figure 3.8: Example 2.3.2. Finite element solutions obtained with (a) M_{adap} and (b) $M_{DMP+adap}$.



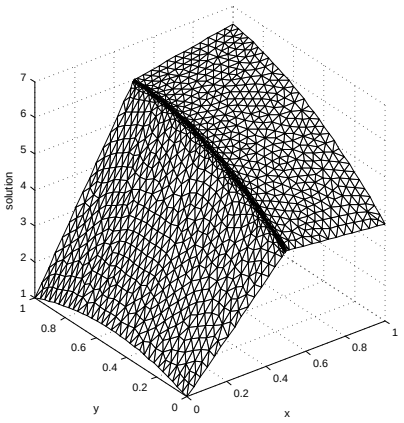
(a): M_{adap}, $N_v = 2480$ (b): $M_{DMP+adap}$, $N_v = 2529$

Figure 3.9: Example 2.3.2. The adaptive meshes obtained with (a) M_{adap} and (b) $M_{DMP+adap}$.

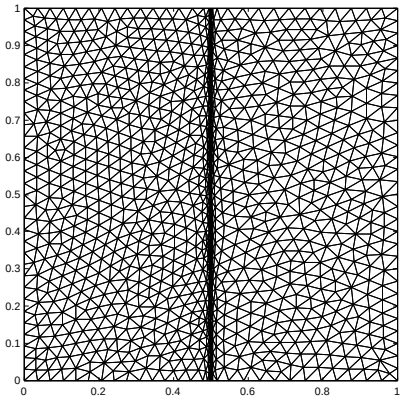


(a): M_{adap}, $u_{min} = -0.0195$, $u_{max} = 1.0160$ (b): $M_{DMP+adap}$, $u_{min} = 0$, $u_{max} = 1.0$

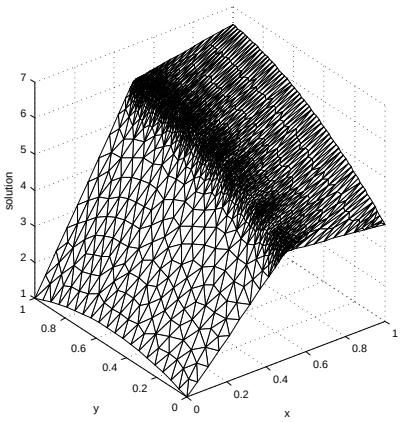
Figure 3.10: Example 2.3.2. Contours of the finite element solutions obtained with (a) M_{adap} and (b) $M_{DMP+adap}$.



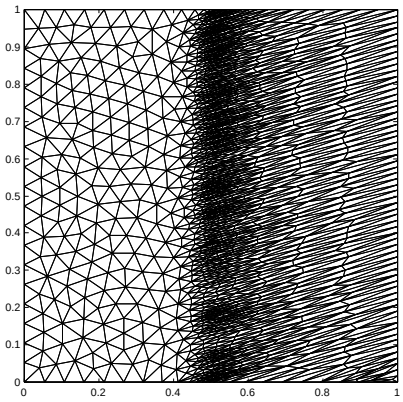
(a): M_{adap}, numerical solution, $u_{min} = 0$



(b): M_{adap}, mesh, $N_v = 2362$

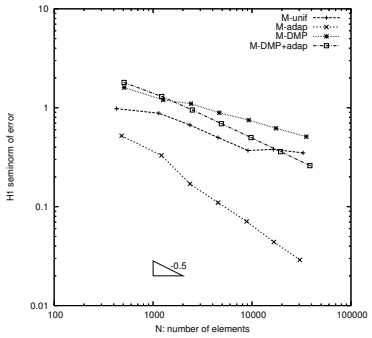


(c): $M_{DMP+adap}$, numerical solution, $u_{min} = 0$

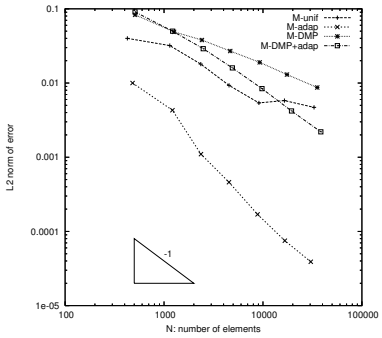


(d): $M_{DMP+adap}$, mesh, $N_v = 2490$

Figure 3.11: Example 2.3.3. Numerical solutions and meshes obtained with M_{adap} and $M_{DMP+adap}$.

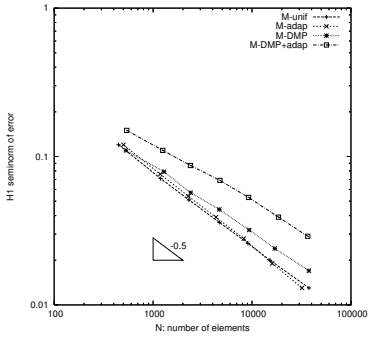


(a): H^1 semi-norm of error

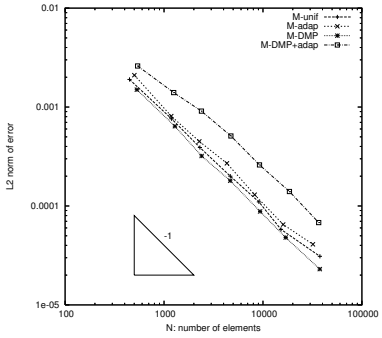


(b): L^2 norm of error

Figure 3.12: Example 2.3.3. The H^1 semi-norm and L^2 norm of solution error are shown as functions of the number of elements for metric tensors M_{unif}, M_{adap}, M_{DMP}, and $M_{DMP+adap}$.



(a): H^1 semi-norm of error



(b): L^2 norm of error

Figure 3.13: Example 2.3.3. The H^1 semi-norm and L^2 norm of solution error are shown as functions of the number of elements for metric tensors M_{unif}, M_{adap}, M_{DMP}, and $M_{DMP+adap}$. The interface ($x = 0.5$) is predefined in the mesh.

Chapter 4

Variational formula for anisotropic mesh adaptation

In Chapters 2 and 3, we have studied anisotropic mesh adaptation for anisotropic diffusion problems in PDE forms. In this chapter, we consider the anisotropic mesh adaptation for problems in the variational form as introduced in §1.2.

To be specific but without loss of generality, we consider the functional of the form

$$I[u] = \int_{\Omega} F(\boldsymbol{x}, u, \nabla u) d\boldsymbol{x}, \quad \forall u \in U_g \tag{4.1}$$

where $\Omega \subset \mathbb{R}^d$ ($d = 1$, 2, or 3) is the physical domain and U_g is the set of functions satisfying the Dirichlet boundary condition

$$u(\boldsymbol{x}) = g(\boldsymbol{x}) \quad \text{on } \partial\Omega \tag{4.2}$$

for a given function $g = g(\boldsymbol{x})$. As mentioned before, we consider only the Dirichlet boundary condition for simplicity, but other types of boundary conditions can be treated without major modification. Note that BVP (1.1) can be written in the form (4.1) with

$$F(\boldsymbol{x}, u, \nabla u) = \frac{1}{2} \nabla u^T (\mathbb{D} \nabla u).$$

The variational problem associated with (4.1) is to seek a minimizer $u \in U_g$ such that

$$I[u] = \min_{v \in U_g} I[v]. \tag{4.3}$$

A necessary condition for u to be a minimizer is that the first variation of the functional vanishes. This leads to the Galerkin formulation

$$\delta I[u;v] \equiv \int_\Omega (F_u(\boldsymbol{x}, u, \nabla u)v + F_{\nabla u}(\boldsymbol{x}, u, \nabla u) \cdot \nabla v)\, d\boldsymbol{x} = 0, \quad \forall v \in U_0 \tag{4.4}$$

where F_u and $F_{\nabla u}$ are the partial derivatives of F with respect to u and ∇u, respectively, and $U_0 = U_g$ with $g = 0$.

We consider the linear finite element approximation for the variational problem. Assume that Ω is a polygon or a polyhedron and an affine family of triangulations $\{\mathscr{T}_h\}$ is given for Ω. Denote by $U^h_{g_h}$ the linear finite element space associated with $\mathscr{T}_h$. Then, a linear finite element approximation $u_h \in U^h_{g_h}$ can be sought either through direct minimization

$$I[u_h] = \min_{v_h \in U^h_{g_h}} I[v_h] \tag{4.5}$$

or by solving the Galerkin formulation

$$\delta I[u_h, v_h] = \int_\Omega (F_u(\boldsymbol{x}, u_h, \nabla u_h)v_h + F_{\nabla u}(\boldsymbol{x}, u_h, \nabla u_h) \cdot \nabla v_h)\, d\boldsymbol{x} = 0, \quad \forall v_h \in U^h_0. \tag{4.6}$$

4.1 Metric tensor for anisotropic mesh adaptation

In this section we define the metric tensor M for use in anisotropic mesh adaptation for the variational problem (4.3). The procedure is to first derive a bound on the variation

$\delta I[u_h, v]$ and then to define M such that the obtained bound is minimized on corresponding M-uniform meshes.

The use of the variation is motivated by the following observation. For a uniformly convex quadratic functional [56] in the form (4.1), there exists a positive constant β such that

$$\|e_h'\|^2_{L^2(\Omega)} \leq \beta \, |\delta I[e_h, e_h]| = \beta \, |\delta I[u_h, e_h]|, \tag{4.7}$$

where $e_h = u - u_h$ is the error and $\|e_h'\|_{L^2(\Omega)}$ is the L^2 norm of the gradient of the error. Thus, the quantities

$$|\delta I[u_h, e_h]| \quad \text{and} \quad \left(\frac{|\delta I[u_h, e_h]|}{\|e_h'\|_{L^2(\Omega)}} \right)^2 \tag{4.8}$$

are equivalent to $\|e_h'\|^2_{L^2(\Omega)}$ and minimizing their bounds is equivalent to minimizing error bounds. Consequently, it is reasonable to define M based on bounds for these quantities. This idea is extended to general functionals in this and next sections. Specifically, a bound on the first quantity in (4.8) is used in this section for defining M for anisotropic mesh adaptation while a bound on the second quantity is employed in the next section for isotropic mesh adaptation.

It is worth pointing out that the so-obtained metric tensor is semi-a posteriori in the sense that it is defined in terms of residuals, edge jumps, and the Hessian of the exact solution. This is in contrast to most previous work (e.g. [19, 20, 34, 61, 74]) where M is defined based on the Hessian of the exact solution and thus completely a priori. As mentioned in §3.1, in the actual computation the Hessian of the exact solution is replaced with approximations obtained through a recovery process from the computed solution. In this section, least squares fitting is used for Hessian recovery, while a hierarchical basis error estimator is discussed in §4.3.

We begin with deriving a bound for $\delta I[u_h, e_h]$. Denote by Π_h the operator of linear finite element interpolation on mesh $\mathscr{T}_h$. Recall that for any $v \in U_0$, we have the orthogonality condition, $\delta I[u_h, \Pi_h v] = 0$. Then we have

$$\begin{aligned}
\delta I[u_h, v] &= \delta I[u_h, v - \Pi_h v] \\
&= \int_\Omega [F_u(\boldsymbol{x}, u_h, \nabla u_h)(v - \Pi_h v) + F_{\nabla u}(\boldsymbol{x}, u_h, \nabla u_h) \cdot \nabla (v - \Pi_h v)] \, d\boldsymbol{x} \\
&= \sum_{K \in \mathscr{T}_h} \int_K [F_u(\boldsymbol{x}, u_h, \nabla u_h)(v - \Pi_h v) + F_{\nabla u}(\boldsymbol{x}, u_h, \nabla u_h) \cdot \nabla (v - \Pi_h v)] \, d\boldsymbol{x} \\
&= \sum_{K \in \mathscr{T}_h} \int_K [F_u(\boldsymbol{x})(v - \Pi_h v) + F_{\nabla u}(\boldsymbol{x}) \cdot \nabla (v - \Pi_h v)] \, d\boldsymbol{x}, \qquad (4.9)
\end{aligned}$$

where, for simplicity, we have used the notation

$$F_u(\boldsymbol{x}) = F_u(\boldsymbol{x}, u_h(\boldsymbol{x}), \nabla u_h(\boldsymbol{x})), \quad F_{\nabla u}(\boldsymbol{x}) = F_{\nabla u}(\boldsymbol{x}, u_h(\boldsymbol{x}), \nabla u_h(\boldsymbol{x})).$$

From the divergence theorem, we can rewrite (4.9) as

$$\begin{aligned}
\delta I[u_h, v] &= \sum_{K \in \mathscr{T}_h} \int_K [F_u(\boldsymbol{x}) - \nabla \cdot F_{\nabla u}(\boldsymbol{x})] \, (v - \Pi_h v) d\boldsymbol{x} \\
&+ \sum_{K \in \mathscr{T}_h} \int_{\partial K} (v - \Pi_h v) F_{\nabla u}(\boldsymbol{x}) \cdot \overrightarrow{n} \, ds, \qquad (4.10)
\end{aligned}$$

where $\overrightarrow{n}$ denotes the outward normal to the face ∂K.

Let $\partial \mathscr{T}_h$ be the collection of all faces of mesh $\mathscr{T}_h$ and K and K' be the two elements sharing the common face γ. We define the residual r_h and the edge jump R_h as

$$r_h(\boldsymbol{x}) = F_u(\boldsymbol{x}) - \nabla \cdot F_{\nabla u}(\boldsymbol{x}), \quad \forall \boldsymbol{x} \in K, \quad \forall K \in \mathscr{T}_h \qquad (4.11)$$

$$R_h(\boldsymbol{x}) = \begin{cases} (F_{\nabla u}(\boldsymbol{x}) \cdot \overrightarrow{n}_\gamma)|_K + (F_{\nabla u}(\boldsymbol{x}) \cdot \overrightarrow{n}_\gamma)|_{K'}, & \boldsymbol{x} \in \gamma, \quad \forall \gamma \in \partial \mathscr{T}_h \backslash \partial \Omega \\ 0, & \boldsymbol{x} \in \gamma, \quad \forall \gamma \in \partial \Omega. \end{cases} \qquad (4.12)$$

Then (4.10) becomes

$$\delta I[u_h, v] = \sum_{K\in\mathscr{T}_h}\int_K r_h(\boldsymbol{x})(v-\Pi_h v)d\boldsymbol{x} + \sum_{\gamma\in\partial\mathscr{T}_h}\int_\gamma (v-\Pi_h v)R_h(\boldsymbol{x})ds. \tag{4.13}$$

Taking $v = e_h \equiv u - u_h$ in the above equation and from Schwarz's inequality, we obtain

$$\begin{aligned}
&|\delta I[u_h, e_h]| \\
\leq\ & \sum_{K\in\mathscr{T}_h}\int_K |r_h(\boldsymbol{x})(e_h-\Pi_h e_h)|d\boldsymbol{x} + \sum_{\gamma\in\partial\mathscr{T}_h}\int_\gamma |(e_h-\Pi_h e_h)R_h(\boldsymbol{x})|ds \\
\leq\ & \sum_{K\in\mathscr{T}_h} \|r_h\|_{L^2(K)}\|e_h-\Pi_h e_h\|_{L^2(K)} + \sum_{\gamma\in\partial\mathscr{T}_h}\|R_h\|_{L^2(\gamma)}\|e_h-\Pi_h e_h\|_{L^2(\gamma)} \qquad (4.14)\\
=\ & \sum_{K\in\mathscr{T}_h}\left[\|r_h\|_{L^2(K)}\|e_h-\Pi_h e_h\|_{L^2(K)} + \frac{1}{2}\sum_{\gamma\in\partial K}\|R_h\|_{L^2(\gamma)}\|e_h-\Pi_h e_h\|_{L^2(\gamma)}\right]. \qquad (4.15)
\end{aligned}$$

For further derivation, we need to estimate $\|e_h-\Pi_h e_h\|_{L^2(K)}$ and $\|e_h-\Pi_h e_h\|_{L^2(\gamma)}$. To this end, we recall that $\{\mathscr{T}_h\}$ is assumed to be an affine family of triangulations on Ω. As a consequence, for any element $K\in\mathscr{T}_h$, there exists an invertible affine mapping $F_K : \hat{K}\to K$ such that $K = F_K(\hat{K})$, where $\hat{K}$ is the reference element chosen here to be equilateral and have the unitary volume. The Jacobian matrix of the mapping F_K is denoted by $F_K^{'}$. Note that it is a constant matrix on K. The following Lemma gives anisotropic bounds on interpolation error on elements and element faces.

Lemma 4.1.1. *Assume that $u\in H^2(\Omega)$. Then for any $K\in\mathscr{T}_h$,*

$$\|e_h-\Pi_h e_h\|_{L^2(K)} \leq C\left[\int_K\Big(\ tr\Big((F_K^{'})^T|H(u)|F_K^{'}\Big)\Big)^2 d\boldsymbol{x}\right]^{\frac{1}{2}}, \tag{4.16}$$

$$\left[\sum_{\gamma\in\partial K}\frac{1}{|\gamma|}\|e_h-\Pi_h e_h\|^2_{L^2(\gamma)}\right]^{\frac{1}{2}} \leq C\left[\frac{1}{|K|}\int_K\Big(\ tr\Big((F_K^{'})^T|H(u)|F_K^{'}\Big)\Big)^2 d\boldsymbol{x}\right]^{\frac{1}{2}}, \tag{4.17}$$

where $tr(\cdot)$ denote the trace of a matrix, $H(u)$ is the Hessian of the solution u, $|H(u)| = \sqrt{H^2(u)}$, $|K|$ denotes the volume of K, and C is a constant independent of $\mathscr{T}_h$ and u.

The inequality (4.16) is proved in [74] and (4.17) is a consequence of the trace theorem and the change of variables.

We continue to develop the bound on $\delta I[u_h, e_h]$. From (4.16) we have

$$\begin{aligned} & \|r_h\|_{L^2(K)} \|e_h - \Pi_h e_h\|_{L^2(K)} \\ \leq\ & C \|r_h\|_{L^2(K)} \left[\int_K \left(\operatorname{tr}\left((F_K')^T |H(u)| F_K' \right) \right)^2 d\boldsymbol{x} \right]^{\frac{1}{2}} \\ =\ & C \|r_h\|_{L^2(K)} |K|^{\frac{1}{2}} \left[\frac{1}{|K|} \int_K \left(\operatorname{tr}\left((F_K')^T |H(u)| F_K' \right) \right)^2 d\boldsymbol{x} \right]^{\frac{1}{2}}. \end{aligned} \tag{4.18}$$

On the other hand,

$$\begin{aligned} & \sum_{\gamma \in \partial K} \|R_h\|_{L^2(\gamma)} \|e_h - \Pi_h e_h\|_{L^2(\gamma)} \\ =\ & \sum_{\gamma \in \partial K} |\gamma|^{\frac{1}{2}} \|R_h\|_{L^2(\gamma)} |\gamma|^{-\frac{1}{2}} \|e_h - \Pi_h e_h\|_{L^2(\gamma)} \\ \leq\ & \left(\sum_{\gamma \in \partial K} |\gamma| \|R_h\|^2_{L^2(\gamma)} \right)^{\frac{1}{2}} \left(\sum_{\gamma \in \partial K} |\gamma|^{-1} \|e_h - \Pi_h e_h\|^2_{L^2(\gamma)} \right)^{\frac{1}{2}} \\ \leq\ & \left(\sum_{\gamma \in \partial K} |\gamma|^{\frac{1}{2}} \|R_h\|_{L^2(\gamma)} \right) \left(\sum_{\gamma \in \partial K} |\gamma|^{-1} \|e_h - \Pi_h e_h\|^2_{L^2(\gamma)} \right)^{\frac{1}{2}} \\ \leq\ & C \left(\sum_{\gamma \in \partial K} |\gamma|^{\frac{1}{2}} \|R_h\|_{L^2(\gamma)} \right) \left[\frac{1}{|K|} \int_K \left(\operatorname{tr}\left((F_K')^T |H(u)| F_K' \right) \right)^2 d\boldsymbol{x} \right]^{\frac{1}{2}}. \end{aligned} \tag{4.19}$$

Substituting (4.18) and (4.19) into (4.15), we have

$$
\begin{aligned}
& |\delta I[u_h, e_h]| \\
\leq\ & C \sum_{K\in\mathscr{T}_h} \left[\|r_h\|_{L^2(K)} |K|^{\frac{1}{2}} + \sum_{\gamma\in\partial K} |\gamma|^{\frac{1}{2}} \|R_h\|_{L^2(\gamma)} \right] \\
& \qquad \cdot \left[\frac{1}{|K|} \int_K \Big(\operatorname{tr}\Big((F_K^{'})^T |H(u)| F_K^{'} \Big) \Big)^2 d\boldsymbol{x} \right]^{\frac{1}{2}} \\
=\ & C \sum_{K\in\mathscr{T}_h} \left[\frac{\|r_h\|_{L^2(K)}}{|K|^{\frac{1}{2}}} + \frac{1}{|K|} \sum_{\gamma\in\partial K} |\gamma|^{\frac{1}{2}} \|R_h\|_{L^2(\gamma)} \right] \\
& \qquad \cdot |K| \left[\frac{1}{|K|} \int_K \Big(\operatorname{tr}\Big((F_K^{'})^T |H(u)| F_K^{'} \Big) \Big)^2 d\boldsymbol{x} \right]^{\frac{1}{2}} \\
\approx\ & C \sum_{K\in\mathscr{T}_h} \left[\frac{\|r_h\|_{L^2(K)}}{|K|^{\frac{1}{2}}} + \frac{1}{|K|} \sum_{\gamma\in\partial K} |\gamma|^{\frac{1}{2}} \|R_h\|_{L^2(\gamma)} \right] \\
& \qquad \cdot |K| \ \operatorname{tr}\Big((F_K^{'})^T |H_K(u)| F_K^{'} \Big), \qquad (4.20)
\end{aligned}
$$

where $H_K(u)$ denotes the value of $H(u)$ at the center of element K and in the last step we have used

$$
\left[\frac{1}{|K|} \int_K \Big(\operatorname{tr}\Big((F_K^{'})^T |H(u)| F_K^{'} \Big) \Big)^2 d\boldsymbol{x} \right]^{\frac{1}{2}} \approx \operatorname{tr}\Big((F_K^{'})^T |H_K(u)| F_K^{'} \Big).
$$

Denoting

$$
\langle r_h \rangle_{L^2(K)} = \frac{\|r_h\|_{L^2(K)}}{|K|^{\frac{1}{2}}} = \left[\frac{1}{|K|} \int_K r_h^2 d\boldsymbol{x} \right]^{\frac{1}{2}}, \quad \langle R_h \rangle_{L^2(\gamma)} = \frac{\|R_h\|_{L^2(\gamma)}}{|\gamma|^{\frac{1}{2}}} = \left[\frac{1}{|\gamma|} \int_\gamma R_h^2 ds \right]^{\frac{1}{2}}, \qquad (4.21)
$$

then (4.20) becomes

$$
|\delta I[u_h, e_h]| \lesssim C \sum_{K\in\mathscr{T}_h} \left[\langle r_h \rangle_{L^2(K)} + \frac{1}{|K|} \sum_{\gamma\in\partial K} |\gamma| \langle R_h \rangle_{L^2(\gamma)} \right] |K| \ \operatorname{tr}\Big((F_K^{'})^T |H_K(u)| F_K^{'} \Big). \qquad (4.22)
$$

We now use this bound to define the metric tensor M. To ensure that M is strictly positive definite, we need to regularize the bound (4.22). For a positive constant α_h (which is to be determined), we have

$$\begin{aligned}
& |\delta I[u_h, e_h]| \\
\lesssim\ & C \sum_{K\in\mathscr{T}_h} \left[\alpha_h + \langle r_h \rangle_{L^2(K)} + \frac{1}{|K|} \sum_{\gamma\in\partial K} |\gamma| \langle R_h \rangle_{L^2(\gamma)} \right] \\
& \qquad \cdot |K|\ \mathrm{tr}\left((F_K^{'})^T \left(\alpha_h I + |H_K(u)| \right) F_K^{'} \right) \\
=\ & C\,\alpha_h^2 \sum_{K\in\mathscr{T}_h} \left[1 + \frac{1}{\alpha_h} \langle r_h \rangle_{L^2(K)} + \frac{1}{\alpha_h |K|} \sum_{\gamma\in\partial K} |\gamma| \langle R_h \rangle_{L^2(\gamma)} \right] \\
& \qquad \cdot |K|\ \mathrm{tr}\left((F_K^{'})^T \left(I + \frac{1}{\alpha_h} |H_K(u)| \right) F_K^{'} \right) \\
=\ & C\,\alpha_h^2 \sum_{K\in\mathscr{T}_h} \left[1 + \frac{1}{\alpha_h} \langle r_h \rangle_{L^2(K)} + \frac{1}{\alpha_h |K|} \sum_{\gamma\in\partial K} |\gamma| \langle R_h \rangle_{L^2(\gamma)} \right] \\
& \qquad \cdot |K|\ \mathrm{tr}\left((F_K^{'})^T H_{K,\alpha} F_K^{'} \right), \qquad (4.23)
\end{aligned}$$

where

$$H_{K,\alpha} = I + \frac{1}{\alpha_h} |H_K(u)|. \qquad (4.24)$$

The metric tensor M is determined so that the bound (4.23) is minimized for M-uniform meshes. As mentioned in §1.3, an M-uniform mesh satisfies the alignment and equidistribution conditions (see [73])

$$\frac{1}{d}\ \mathrm{tr}\left((F_K^{'})^T M_K F_K^{'} \right) = \det\left((F_K^{'})^T M_K F_K^{'} \right)^{\frac{1}{d}}, \qquad (4.25)$$

$$\rho_K |K| = \frac{\sigma_h}{N}, \qquad (4.26)$$

where M_K is an average of M over K, $\rho_K = \sqrt{\det(M_K)}$, $\det(\cdot)$ denote the determinant of a matrix, N is the number of elements of $\mathscr{T}_h$, and

$$\sigma_h = \sum_{K \in \mathscr{T}_h} \rho_K |K|. \tag{4.27}$$

Comparing (4.23) with (4.25) suggests that M be chosen in the form

$$M_K = \theta_K H_{K,\alpha} \quad \forall K \in \mathscr{T}_h \tag{4.28}$$

for some scalar function $\theta = \theta_K$. Inserting this into (4.25), the alignment condition yields

$$\frac{1}{d} \operatorname{tr}\left((F_K^{'})^T H_{K,\alpha} F_K^{'}\right) = \det\left((F_K^{'})^T H_{K,\alpha} F_K^{'}\right)^{\frac{1}{d}} = |K|^{\frac{2}{d}} \det\left(H_{K,\alpha}\right)^{\frac{1}{d}}, \tag{4.29}$$

where we have used $\det(F_K^{'}) = |K|$. Substituting (4.29) into (4.23), we have

$$\begin{aligned} & |\delta I[u_h, e_h]| \\ \lesssim\ & C\alpha_h^2 \sum_{K \in \mathscr{T}_h} \left[1 + \frac{1}{\alpha_h} \langle r_h \rangle_{L^2(K)} + \frac{1}{\alpha_h |K|} \sum_{\gamma \in \partial K} |\gamma| \langle R_h \rangle_{L^2(\gamma)}\right] \\ & \qquad \cdot |K|^{\frac{d+2}{d}} \det\left(H_{K,\alpha}\right)^{\frac{1}{d}}. \end{aligned} \tag{4.30}$$

We now proceed to determine $\theta = \theta_K$ in (4.28) using the equidistribution condition (4.26). From Hölder's inequality, we have

$$\left[\frac{1}{N} \sum_K \left(|K| \rho_K\right)^{\frac{d+2}{d}}\right]^{\frac{d}{d+2}} \geq \frac{1}{N} \sum_K |K| \rho_K = \frac{\sigma_h}{N}$$

or

$$\sum_K (|K|\rho_K)^{\frac{d+2}{d}} \geq (\sigma_h)^{\frac{d+2}{d}} N^{-\frac{2}{d}}, \tag{4.31}$$

with the lower bound being attained for a mesh satisfying (4.26). This suggest, by comparing the left-hand side of (4.31) with the right-hand side of (4.30), that $\rho = \rho_K$ be defined as

$$\rho_K = \left[1 + \frac{1}{\alpha_h}\langle r_h\rangle_{L^2(K)} + \frac{1}{\alpha_h|K|}\sum_{\gamma\in\partial K}|\gamma|\langle R_h\rangle_{L^2(\gamma)}\right]^{\frac{d}{d+2}} \det(H_{K,\alpha})^{\frac{1}{d+2}},$$

or

$$\rho_K = \left[1 + \frac{1}{\alpha_h}\langle r_h\rangle_{L^2(K)} + \frac{1}{\alpha_h|K|}\sum_{\gamma\in\partial K}|\gamma|\langle R_h\rangle_{L^2(\gamma)}\right]^{\frac{d}{d+2}} \det\left(I + \frac{1}{\alpha_h}|H_K(u)|\right)^{\frac{1}{d+2}}. \tag{4.32}$$

With the so-defined ρ, the right-hand side of (4.30) attains its lower bound for a mesh satisfying the equidistribution condition (4.26). Now the variation of the functional (4.1) has a bound as

$$|\delta I[u_h, e_h]| \lesssim C\,\alpha_h^2\,(\sigma_h)^{\frac{d+2}{d}} N^{-\frac{2}{d}} \tag{4.33}$$

for an M-uniform mesh. Recall that $|\delta I[u_h, e_h]|$ is equivalent to $\|e_h'\|^2_{L^2(\Omega)}$ for uniformly convex quadratic functionals. In this case, (4.33) implies that, when α_h and σ_h are bounded, $\|e_h'\|_{L^2(\Omega)} = O(N^{-\frac{1}{d}})$ or the H^1 semi-norm of the error has a first order convergence as $N \to \infty$.

From the relation $\rho = \sqrt{\det(M)}$ and the form (4.28), we obtain the metric tensor as

$$M_K = \rho_K^{\frac{2}{d}} \det\left(I + \frac{1}{\alpha_h}|H_K(u)|\right)^{-\frac{1}{d}} \left[I + \frac{1}{\alpha_h}|H_K(u)|\right]. \tag{4.34}$$

Following [74], we choose the parameter α_h such that

$$\sigma_h \equiv \sum_{K\in\mathscr{T}_h} \rho_K |K| = 2|\Omega|$$

or

$$\sum_{K\in\mathscr{T}_h} |K| \left[1 + \frac{1}{\alpha_h}\langle r_h\rangle_{L^2(K)} + \frac{1}{\alpha_h|K|}\sum_{\gamma\in\partial K}|\gamma|\langle R_h\rangle_{L^2(\gamma)}\right]^{\frac{d}{d+2}} \det\left(I + \frac{1}{\alpha_h}|H_K(u)|\right)^{\frac{1}{d+2}} = 2|\Omega|. \tag{4.35}$$

With this choice, roughly fifty percents of the mesh elements will be concentrated in the regions of large ρ [74]. Moreover, mesh concentration is invariant under the scaling transformation of the solution u. It is easy to show that (4.35) has a solution $\alpha_h > 0$. Once α_h is computed, the adaptation function ρ and the metric tensor M can be determined by (4.32) and (4.34), respectively. A new mesh can then be generated based on the metric tensor.

From (4.32) and (4.34), one can see that this definition of the metric tensor involves the residual r_h, the edge jump R_h, and the Hessian of the solution, $H_K(u)$. The former two are computable because they are based on the computed solution, whereas $H_K(u)$ is not since the exact solution is generally unknown. In this sense, (4.34) is semi-a posteriori. In actual computation, the Hessian of the exact solution is typically approximated through recovery from the computed solution.

4.2 Metric tensor for isotropic mesh adaptation

Interestingly, a completely a posteriori formula can be obtained for isotropic mesh adaptation in a similar procedure as that used in the previous section. To see this,

we first recall some isotropic bounds for interpolation error on elements and element faces in the following lemma. The reader is referred to [3] for its proof.

Lemma 4.2.1. *Assume that $u \in H^2(\Omega)$ and $\{\mathscr{T}_h\}$ is an affine family of regular triangulations for Ω. Then, for any $K \in \mathscr{T}_h$,*

$$\|e_h - \Pi_h e_h\|_{L^2(K)} \leq C h_K \|e_h'\|_{L^2(\tilde{K})}, \tag{4.36}$$

$$\|e_h - \Pi_h e_h\|_{L^2(\gamma)} \leq C h_K^{\frac{1}{2}} \|e_h'\|_{L^2(\tilde{K})}, \tag{4.37}$$

where h_K is the diameter of the element K and $\tilde{K}$ is the subdomain consisting of elements sharing a common face with element K.

Substituting (4.36) and (4.37) into (4.14) and applying Schwarz's inequality, we have

$$|\delta I[u_h, e_h]| \leq C \|e_h'\|_{L^2(\Omega)} \left[\sum_{K \in \mathscr{T}_h} h_K^2 \|r_h\|_{L^2(K)}^2 + \sum_{\gamma \in \partial \mathscr{T}_h} h_K \|R_h\|_{L^2(\gamma)}^2 \right]^{\frac{1}{2}}. \tag{4.38}$$

Thus, using the notation (4.21) we have

$$\begin{aligned} \left(\frac{|\delta I[u_h, e_h]|}{\|e_h'\|_{L^2(\Omega)}} \right)^2 &\leq C \left[\sum_{K \in \mathscr{T}_h} h_K^2 \|r_h\|_{L^2(K)}^2 + \sum_{\gamma \in \mathscr{T}_h} h_K \|R_h\|_{L^2(\gamma)}^2 \right] \\ &\leq C \sum_{K \in \mathscr{T}_h} \left[h_K^2 \|r_h\|_{L^2(K)}^2 + \frac{1}{2} \sum_{\gamma \in \partial K} h_K \|R_h\|_{L^2(\gamma)}^2 \right] \\ &\leq C \sum_{K \in \mathscr{T}_h} |K|^{1+\frac{2}{d}} \left[\langle r_h \rangle_{L^2(K)}^2 + \frac{1}{|K|^{\frac{2}{d}}} \sum_{\gamma \in \partial K} \langle R_h \rangle_{L^2(\gamma)}^2 \right], \end{aligned} \tag{4.39}$$

where in the last step we have used $|K| \sim h_K^d$ and $|\gamma| \sim h_K^{d-1}$ and denotation of $\langle r_h \rangle$ and $\langle R_h \rangle$ as in (4.21). As pointed out in the previous section (cf. (4.8)), the left-hand side of the above inequality is equivalent to $\|e_h'\|_{L^2(\Omega)}^2$ when the functional is uniformly

convex and quadratic. In the same time, the right hand-side term is of the same form as a standard a posteriori error estimate for elliptic PDEs; e.g., see [3].

The bound in (4.39) is used to define M in such a way that it is minimized for M-uniform meshes. The same procedure employed in the previous section for anisotropic mesh adaptation can be used for this purpose. Indeed, we obtain

$$M_K = \rho_K^{\frac{2}{d}} I \tag{4.40}$$

and

$$\rho_K = \left[1 + \frac{1}{\alpha_h} \langle r_h \rangle^2_{L^2(K)} + \frac{1}{\alpha_h |K|^{\frac{2}{d}}} \sum_{\gamma \in \partial K} \langle R_h \rangle^2_{L^2(\gamma)} \right]^{\frac{d}{d+2}}, \tag{4.41}$$

where α_h is defined as

$$\alpha_h = \left\{ \frac{1}{|\Omega|} \sum_{K \in \mathscr{T}_h} |K| \left[\langle r_h \rangle^{\frac{2d}{d+2}}_{L^2(K)} + \left(\frac{1}{|K|^{\frac{2}{d}}} \sum_{\gamma \in \partial K} \langle R_h \rangle^2_{L^2(\gamma)} \right)^{\frac{d}{d+2}} \right] \right\}^{\frac{d+2}{d}}. \tag{4.42}$$

It is remarked that for this choice of α_h, we have

$$\sigma_h = \sum_{K \in \mathscr{T}_h} \rho_K |K| \leq 2|\Omega|. \tag{4.43}$$

Moreover, for the corresponding M-uniform mesh it holds

$$\left(\frac{|\delta I[u_h, e_h]|}{\|e_h'\|_{L^2(\Omega)}} \right)^2 \leq C \alpha_h N^{-\frac{2}{d}}. \tag{4.44}$$

For the case of uniformly convex quadratic functionals, the left-hand side is equivalent to $\|e_h'\|^2_{L^2(\Omega)}$ and thus (4.44) implies $\|e_h'\|_{L^2(\Omega)} = O(N^{-\frac{1}{d}})$, meaning that the error has a first order convergence (which corresponds to $O(h)$ for a uniform mesh) as $N \to \infty$.

Furthermore, it is obvious that the metric tensor defined in (4.40) involves only residual r_h and edge jump R_h and thus is completely a posteriori.

4.3 Error estimation based on hierarchical bases

In §1.2, we developed a metric tensor for use in anisotropic mesh adaptation for variational problems. The metric tensor developed is a semi-posterior in the sense that it involves the Hessian of the exact solution. As mentioned before, the Hessian is approximated by least squares fitting in the actual computation.

In this section, we investigate the use of a global hierarchical basis error estimator (HBEE) developed in [77] for the development of an anisotropic metric tensor for variational problems. The new metric tensor is completely a posteriori and based on residual, edge jumps and the hierarchical basis error estimator. Numerical results show that it performs comparable with existing metric tensors based on Hessian recovery. A few sweeps of the symmetric Gauss-Seidel iteration for solving the global error problem prove sufficient to provide directional information necessary for successful mesh adaptation.

The procedure is similar to the one described in §1.2. A posterior hierarchical basis error estimator (HBEE) is employed on elements and element faces in the bound (4.15). The computation of the error estimator is based on a general framework, details on which can be found among others in the work of Bank and Smith [9] or Deuflhard et al. [48]. The approach is briefly explained as follows.

Recall that $u_h \in U_g^h$ is a linear finite element solution of the Galerkin formulation (4.6) and the error is $e_h = u - u_h$. Let

$$\bar{U}_g^h = U_g^h \oplus W^h,$$

where W^h is the linear span of the edge bubble functions. Obviously, $\bar{U}_g^h$ is a subspace of piecewise quadratic functions. Recall also that

$$\delta I[u_h + e_h, v] = 0 \qquad \forall v \in U_0.$$

Denote by $B_h[u_h; \cdot, \cdot]$ a bilinear form resulting from a linearization of $\delta I[\cdot, \cdot]$ about u_h with respect to the first argument. The error estimate z_h is then defined as the solution of the approximate linear error problem

$$\begin{cases} \text{Find } z_h \in W^h \text{ such that} \\ \delta I[u_h, w_h] + B_h[u_h; z_h, w_h] = 0 \quad \forall w_h \in W^h. \end{cases}$$

The estimate z_h can be viewed as a projection of the true error onto the subspace W^h. Note that $\Pi_h z_h = 0$ by construction, and thus $z_h = z_h - \Pi_h z_h$.

This definition of the error estimate is global and its solution can be costly. To avoid the expensive exact solution in numerical computation, we employ only a few sweeps of the symmetric Gauss-Seidel iteration for the resulting linear system, which proves to be sufficient for the purpose of mesh adaptation [77].

We now assume that z_h provides a reliable local estimate on $e_h - \Pi_h e_h$, i.e., there exist constants $C_1 > 0$ and $C_2 > 0$ such that

$$\|e_h - \Pi_h e_h\|_{L^2(K)} \leq C_1 \|z_h\|_{L^2(K)} \quad \text{and} \quad \|e_h - \Pi_h e_h\|_{L^2(\gamma)} \leq C_2 \|z_h\|_{L^2(\gamma)}.$$

Then we can replace $e_h - \Pi_h e_h$ with z_h in (4.15) and develop the bound on $\delta I[u_h, e_h]$ in terms of z_h.

Recalling that $\Pi_h z_h = 0$ and using element-wise interpolation error estimates in [74], we have

$$\begin{aligned}
\|z_h\|_{L^2(K)} &= \|z_h - \Pi_h z_h\|_{L^2(K)} \\
&\leq C \left[\int_K \left(\operatorname{tr}\left((F_K')^T \, |H(z_h)| \, F_K' \right) \right)^2 d\boldsymbol{x} \right]^{\frac{1}{2}} \\
&= C \, |K|^{\frac{1}{2}} \, \operatorname{tr}\left((F_K')^T \, |H_K(z_h)| \, F_K' \right), \qquad (4.45)
\end{aligned}$$

and

$$\begin{aligned}
\left[\sum_{\gamma \in \partial K} \frac{1}{|\gamma|} \|z_h\|^2_{L^2(\gamma)} \right]^{\frac{1}{2}} &= \left[\sum_{\gamma \in \partial K} \frac{1}{|\gamma|} \|z_h - \Pi_h z_h\|^2_{L^2(\gamma)} \right]^{\frac{1}{2}} \\
&\leq C \left[\frac{1}{|K|} \int_K \left(\operatorname{tr}\left((F_K')^T \, |H(z_h)| \, F_K' \right) \right)^2 d\boldsymbol{x} \right]^{\frac{1}{2}} \\
&= C \operatorname{tr}\left((F_K')^T \, |H_K(z_h)| \, F_K' \right),
\end{aligned}$$

where $\operatorname{tr}(\cdot)$ denote the trace of a matrix, $H_K(z_h)$ is the Hessian of z_h on K, $|H_K(z_h)| = \sqrt{H_K^2(z_h)}$, $|K|$ is the volume of K, F_K is a mapping from the reference element $\hat{K}$ to element K, and C is a constant independent of $\mathscr{T}_h$ and z_h. Thus,

$$\begin{aligned}
\sum_{\gamma \in \partial K} \|R_h\|_{L^2(\gamma)} \|z_h\|_{L^2(\gamma)} &= \sum_{\gamma \in \partial K} |\gamma|^{\frac{1}{2}} \|R_h\|_{L^2(\gamma)} \, |\gamma|^{-\frac{1}{2}} \|z_h\|_{L^2(\gamma)} \\
&\leq \left(\sum_{\gamma \in \partial K} |\gamma| \, \|R_h\|^2_{L^2(\gamma)} \right)^{\frac{1}{2}} \left(\sum_{\gamma \in \partial K} \frac{1}{|\gamma|} \|z_h\|^2_{L^2(\gamma)} \right)^{\frac{1}{2}} \\
&\leq \left(\sum_{\gamma \in \partial K} |\gamma|^{\frac{1}{2}} \, \|R_h\|_{L^2(\gamma)} \right) \left(\sum_{\gamma \in \partial K} \frac{1}{|\gamma|} \|z_h\|^2_{L^2(\gamma)} \right)^{\frac{1}{2}} \\
&\leq C \left(\sum_{\gamma \in \partial K} |\gamma|^{\frac{1}{2}} \, \|R_h\|_{L^2(\gamma)} \right) \operatorname{tr}\left((F_K')^T \, |H_K(z_h)| \, F_K' \right). \qquad (4.46)
\end{aligned}$$

Substituting (4.45) and (4.46) into (4.15) leads to

$$
\begin{aligned}
|\delta I[u_h, e_h]| &\leq C \sum_{K \in \mathscr{T}_h} \left[|K|^{\frac{1}{2}} \|r_h\|_{L^2(K)} + \sum_{\gamma \in \partial K} |\gamma|^{\frac{1}{2}} \|R_h\|_{L^2(\gamma)} \right] \ \mathrm{tr}\left((F_K')^T \, |H_K(z_h)| \, F_K' \right) \\
&= C \sum_{K \in \mathscr{T}_h} \left[\frac{\|r_h\|_{L^2(K)}}{|K|^{\frac{1}{2}}} + \frac{1}{|K|} \sum_{\gamma \in \partial K} |\gamma| \frac{\|R_h\|_{L^2(\gamma)}}{|\gamma|^{\frac{1}{2}}} \right] |K| \ \mathrm{tr}\left((F_K')^T \, |H_K(z_h)| \, F_K' \right) \\
&= C \sum_{K \in \mathscr{T}_h} \left[\langle r_h \rangle_{L^2(K)} + \frac{1}{|K|} \sum_{\gamma \in \partial K} |\gamma| \langle R_h \rangle_{L^2(\gamma)} \right] |K| \ \mathrm{tr}\left((F_K')^T \, |H_K(z_h)| \, F_K' \right),
\end{aligned}
\tag{4.47}
$$

where $\langle r_h \rangle$ and $\langle R_h \rangle$ are defined in (4.21).

We now use bound (4.47) to define the metric tensor M. To ensure that M is strictly positive definite, we first regularize the bound with a positive constant α_h (to be determined), i.e.,

$$
\begin{aligned}
|\delta I[u_h, e_h]| \leq C \sum_{K \in \mathscr{T}_h} & \left[\alpha_h + \langle r_h \rangle_{L^2(K)} + \frac{1}{|K|} \sum_{\gamma \in \partial K} |\gamma| \langle R_h \rangle_{L^2(\gamma)} \right] \\
& \cdot |K| \ \mathrm{tr}\left((F_K')^T \, (\alpha_h I + |H_K(z_h)|) \, F_K' \right) \\
= C \alpha_h^2 \sum_{K \in \mathscr{T}_h} & \left[1 + \frac{1}{\alpha_h} \langle r_h \rangle_{L^2(K)} + \frac{1}{\alpha_h |K|} \sum_{\gamma \in \partial K} |\gamma| \langle R_h \rangle_{L^2(\gamma)} \right] \\
& \cdot |K| \ \mathrm{tr}\left((F_K')^T H_{K,\alpha}(z_h) F_K' \right),
\end{aligned}
\tag{4.48}
$$

where

$$
H_{K,\alpha}(z_h) = I + \frac{1}{\alpha_h} |H_K(z_h)| .
$$

The optimal metric tensor is obtained by minimizing bound (4.48) for M-uniform meshes. As mentioned in §1.3, an M-uniform mesh satisfies the alignment condition (1.12) and the equidistribution condition (1.11). The conditions are repeated below for

convenience. The alignment condition is

$$\frac{1}{d}\,\mathrm{tr}\left((F_K')^T M_K F_K'\right) = \det\left((F_K')^T M_K F_K'\right)^{\frac{1}{d}} \tag{4.49}$$

and the equidistribution condition is

$$\rho_K\,|K| = \frac{\sigma_h}{N}, \tag{4.50}$$

We now pay our attention to the $\mathrm{tr}(\cdot)$ factor in (4.48). Notice that in general,

$$\frac{1}{d}\,\mathrm{tr}\left((F_K')^T H_{K,\alpha}(z_h) F_K'\right) \geq \det\left((F_K')^T H_{K,\alpha}(z_h) F_K'\right)^{\frac{1}{d}}.$$

From (4.49) we can see that the equality in the above inequalities holds if we choose $M = M_K$ in the form

$$M_K = \theta_K H_{K,\alpha}(z_h) \qquad \forall K \in \mathscr{T}_h \tag{4.51}$$

for some scalar function $\theta = \theta_K$. Indeed, with this choice of M_K the alignment condition (4.49) reads as

$$\frac{1}{d}\,\mathrm{tr}\left((F_K')^T H_{K,\alpha} F_K'\right) = \det\left((F_K')^T H_{K,\alpha} F_K'\right)^{\frac{1}{d}} = |K|^{\frac{2}{d}} \det\left(H_{K,\alpha}(z_h)\right)^{\frac{1}{d}}, \tag{4.52}$$

where we have used $|\det(F_K')| = |K|$ and assumed $|\hat{K}| = 1$. Substituting (4.52) into (4.48) yields

$$\begin{aligned} |\delta I[u_h, e_h]| \leq C\alpha_h^2 \sum_{K\in\mathscr{T}_h} \Bigg[1 + \frac{1}{\alpha_h}\langle r_h\rangle_{L^2(K)} + \frac{1}{\alpha_h |K|}\sum_{\gamma\in\partial K} |\gamma|\,\langle R_h\rangle_{L^2(\gamma)}\Bigg] \\ \cdot |K|^{\frac{d+2}{d}} \det\left(H_{K,\alpha}(z_h)\right)^{\frac{1}{d}}. \end{aligned} \tag{4.53}$$

Next, we use the equidistribution condition (4.50) to determine $\theta = \theta_K$ in (4.51). From Hölder's inequality, we have

$$\left[\frac{1}{N}\sum_K \left(|K|\,\rho_K\right)^{\frac{d+2}{d}}\right]^{\frac{d}{d+2}} \geq \frac{1}{N}\sum_K |K|\,\rho_K = \frac{\sigma_h}{N}$$

or

$$\sum_K \left(|K|\rho_K\right)^{\frac{d+2}{d}} \geq \left(\sigma_h\right)^{\frac{d+2}{d}} N^{-\frac{2}{d}}, \tag{4.54}$$

with the lower bound being attained for a mesh satisfying (4.50). Comparing the left-hand side of (4.54) with the right-hand side of (4.53) suggests that $\rho = \rho_K$ be defined as

$$\begin{aligned}\rho_K &= \left[1 + \frac{1}{\alpha_h}\left\langle r_h\right\rangle_{L^2(K)} + \frac{1}{\alpha_h\,|K|}\sum_{\gamma\in\partial K}|\gamma|\left\langle R_h\right\rangle_{L^2(\gamma)}\right]^{\frac{d}{d+2}} \det\left(H_{K,\alpha}(z_h)\right)^{\frac{1}{d+2}} \\ &= \left[1 + \frac{1}{\alpha_h}\left\langle r_h\right\rangle_{L^2(K)} + \frac{1}{\alpha_h\,|K|}\sum_{\gamma\in\partial K}|\gamma|\left\langle R_h\right\rangle_{L^2(\gamma)}\right]^{\frac{d}{d+2}} \det\left(I + \frac{1}{\alpha_h}\left|H_K(z_h)\right|\right)^{\frac{1}{d+2}}.\end{aligned} \tag{4.55}$$

From relations $\rho_K = \sqrt{\det(M_K)}$ and $M_K = \theta_K H_{K,\alpha}(z_h)$ we can obtain θ_K. The metric tensor M_K is then given by

$$M_K = \rho_K^{\frac{2}{d}} \det\left(I + \frac{1}{\alpha_h}\left|H_K(z_h)\right|\right)^{-\frac{1}{d}}\left[I + \frac{1}{\alpha_h}\left|H_K(z_h)\right|\right]. \tag{4.56}$$

With this choice of ρ_K (and M_K), the right-hand side of (4.53) attains its lower bound for a mesh satisfying the equidistribution condition (4.50). Then, the variation

of the functional (4.1) has an upper bound as

$$|\delta I[u_h, e_h]| \leq C\alpha_h^2 (\sigma_h)^{\frac{d+2}{d}} N^{-\frac{2}{d}}.$$

To complete the definition of the metric tensor, we need to choose the regularity parameter α_h. Following [74], we choose it such that

$$\sigma_h \equiv \sum_{K \in \mathscr{T}_h} \rho_K |K| = 2|\Omega|$$

or

$$\sum_{K \in \mathscr{T}_h} |K| \left[1 + \frac{1}{\alpha_h} \langle r_h \rangle_{L^2(K)} + \frac{1}{\alpha_h |K|} \sum_{\gamma \in \partial K} |\gamma| \langle R_h \rangle_{L^2(\gamma)} \right]^{\frac{d}{d+2}} \det \left(I + \frac{1}{\alpha_h} |H_K(z_h)| \right)^{\frac{1}{d+2}}$$
$$= 2|\Omega|. \tag{4.57}$$

With this choice, roughly fifty percents of the mesh elements will be concentrated in the regions of large ρ [74]. It is easy to show that (4.57) has a unique solution since its left-hand side is monotonically decreasing with α_h increasing and tends to $|\Omega|$ as $\alpha_h \to \infty$ and to $+\infty$ as $\alpha_h \to 0$. Moreover, it can be solved using a simple iteration scheme such as the bisection method.

Once α_h is computed, the adaptation function ρ and the metric tensor M can be determined by (4.55) and (4.56), respectively. A new mesh can then be generated based on the metric tensor.

This definition of the metric tensor involves the residual r_h, the edge jump R_h, and the HBEE z_h. All these quantities are based on the computed solution; in this sense, (4.56) is a posteriori.

4.4 Numerical results

In this section we present numerical results obtained for a selection of two-dimensional problems. The metric tensors (4.34), (4.40) and (4.56) are used with the mesh adaptation procedure outlined in §1.3. The corresponding results are labelled with "VPaniso", "VPiso" and "VPhb", respectively.

For comparison purpose, we also include the results obtained with a formula for the metric tensor developed in [74] based on interpolation error. For completeness, we record the formula as

$$M_K = \rho_K^{\frac{2}{d}} \det\left(I + \frac{1}{\alpha_h}|H_K(u)|\right)^{-\frac{1}{d}} \left[I + \frac{1}{\alpha_h}|H_K(u)|\right],$$

$$\rho_K = \left\|I + \frac{1}{\alpha_h}|H_K(u)|\right\|_F^{\frac{d}{d+2}} \det\left(I + \frac{1}{\alpha_h}|H_K(u)|\right)^{\frac{1}{d+2}},$$

where $\|\cdot\|_F$ is the Frobenius matrix norm. Results obtained with this metric tensor are labelled with "IEaniso".

Since formulas (4.34), (4.40) and (4.56) are developed based on the bounds on the quantities in (4.8) that are equivalent to the H^1 seminorm of the error for uniformly convex quadratic functionals, we measure the error in the H^1 seminorm in our computation and comparison. We first focus on comparison among "IEaniso", "VPiso" and "VPaniso". Then "VPhb" is considered and compared with "VPiso" and "VPaniso".

Example 4.4.1. The first example is

$$I[u] = \int_\Omega \left(\frac{1}{2}|\nabla u|^2 - uf\right) dxdy, \tag{4.58}$$

where Ω is the unit square. The function f is chosen such that the exact solution is

$$u_{exact} = \frac{1}{1+e^{\frac{x+y-1.25}{2\varepsilon}}}. \tag{4.59}$$

The parameter is taken as $\varepsilon = 0.005$ in the computation. Note that the problem is equivalent to the boundary value problem

$$\begin{cases} -\nabla\cdot(\nabla u) = f & \text{in} \quad \Omega \\ u = u_{exact} & \text{on} \quad \partial\Omega. \end{cases} \tag{4.60}$$

Fig. 4.1 shows adaptive meshes corresponding to "IEaniso", "VPiso" and "VPaniso". One can see that they all have correct mesh concentration. Moreover, the meshes obtained with IEaniso and VPaniso are obviously anisotropic. They have a much better alignment with the solution and lead to smaller error than the isotropic mesh obtained with VPiso. In Fig. 4.2, the H^1 seminorm of the finite element error is plotted as functions of the number of elements (N). It can be seen that the error has a first order convergence, i.e., $\|e_h'\|_{L^2(\Omega)} = O(N^{-\frac{1}{2}})$, for all three formulas of the metric tensor. The result confirms the theoretical prediction in (4.33) and (4.44). It also shows that anisotropic meshes produce a significantly smaller error than an isotropic one, a known fact in the context of anisotropic mesh adaptation for numerical solution of PDEs.

It is emphasized that the formulas of the metric tensor based on the variational formulation, VPiso and VPaniso, work well for solving the current variational problem. Moreover, VPaniso is comparable in performance with IEaniso, an existing formula based on interpolation error.

To give some sense of the estimates (4.22) and (4.39), we let

$$\eta_{h,aniso} \equiv \left\{ \sum_{K\in\mathscr{T}_h} \left[\langle r_h \rangle_{L^2(K)} + \frac{1}{|K|} \sum_{\gamma\in\partial K} |\gamma| \langle R_h \rangle_{L^2(\gamma)} \right] |K| \ \mathrm{tr}\left((F_K^{'})^T \, |H_K(u)| \, F_K^{'} \right) \right\}^{\frac{1}{2}},$$

$$\eta_{h,iso} \equiv \left\{ \sum_{K\in\mathscr{T}_h} |K|^{1+\frac{2}{d}} \left[\langle r_h \rangle^2_{L^2(K)} + \frac{1}{|K|^{\frac{2}{d}}} \sum_{\gamma\in\partial K} \langle R_h \rangle^2_{L^2(\gamma)} \right] \right\}^{\frac{1}{2}},$$

and plot them as functions of N in Fig. 4.3. In the current situation, each of $\eta_{h,aniso}$ and $\eta_{h,iso}$, up to a multiplicative constant, defines a bound for the error $\|e_h^{'}\|_{L^2(\Omega)}$. Particularly, $\eta_{h,iso}$ has the same form as a standard residual-based a posteriori error estimate. It should be pointed out that they cannot be directly compared against each other because they are not asymptotically exact and the multiplicative constants can have a different value. On the other hand, Fig. 4.3 does show that the quantities decrease at the rate of $N^{-0.5}$, confirming the theoretical predictions (4.33) and (4.44).

Example 4.4.2. Our second example is the functional

$$I[u] = \int_\Omega \left(\frac{1}{2} \nabla u \cdot \mathbb{D} \nabla u - fu \right) dxdy, \tag{4.61}$$

where Ω is the unit square and

$$\mathbb{D} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 1000 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{4.62}$$

with constant θ. The function f is chosen such that the exact solution is given by

$$u_{exact} = 2\cos(\pi x)\sin(2\pi y) + 2. \tag{4.63}$$

The solution is shown in Fig. 4.4.

We consider two cases of θ: $\theta = \pi/6$ and $\theta = 5\pi/6$. Note that when θ changes from $\pi/6$ to $5\pi/6$, the structure of the problem changes while the exact solution remains the same. The numerical solution is shown in Fig. 4.4. By construction, we expect that the mesh concentration changes for methods VPiso and VPaniso but stays the same for IEaniso. This is confirmed in Fig. 4.5 where adaptive meshes are shown for the three formulas of the metric tensor and for the two values of θ.

The error in the H^1 seminorm is plotted in Fig. 4.6 as function of the number of elements. Once again, the first convergence order of the error can be observed from the figure.

Example 4.4.3. The next example is an anisotropic variational problem [17] defined by the functional

$$I[u] = \int_\Omega \left(\left(1 + |\nabla u|^2\right)^{\frac{3}{4}} + 1000 \left| \frac{\partial u}{\partial y} \right|^2 \right) dxdy, \tag{4.64}$$

where Ω is the unit square. The boundary condition is given as

$$\begin{cases} u = 1 & \text{on } x = 0 \text{ or } x = 1, \\ u = 2 & \text{on } y = 0 \text{ or } y = 1. \end{cases} \tag{4.65}$$

No analytical solution is available for this example. Unlike the previous two examples, the functional (4.64) is not quadratic. Hence, the quantities in (4.8) are not equivalent to $\|e_h'\|^2_{L^2(\Omega)}$, and this example is a test for the formulas of the metric tensor based on the variational formulation.

A computed solution is shown in Fig. 4.7, from which we can see that the solution has sharp layers near boundaries $x = 0$ and $x = 1$. Fig. 4.8 shows adaptive meshes obtained corresponding to IEaniso, VPiso and VPaniso. One can see that they have

correct mesh concentration. The result shows that VPaniso is comparable to IEaniso even for the variational problem with a non-quadratic functional.

Adaptive meshes obtained with the three different metric tensors (VPiso, VPaniso, and VPhb) are given in Fig. 4.9. They all have correct mesh concentration, but the anisotropic metric tensors (Figs. 4.9(b) and 4.9)(c) provide a much better alignment with the boundary layers. Again, both anisotropic meshes are comparable, although mesh elements near the boundary layer in the HBEE-based adaptive mesh ("VPhb") have a larger aspect ratio than elements of the mesh obtained by means of the Hessian recovery ("VPaniso"). This could be due to the smoothing nature of the Hessian recovery: usually, it operates on a larger patch, thus introducing an additional smoothing effect, which affects the grading of the elements' size and orientation. The global hierarchical basis error estimator does not have this handicap and, in this example, the mesh obtained by means of HBEE is slightly better aligned with the steep boundary layers.

Example 4.4.4. This example is an energy functional used in image processing with observed image $p(x,y)$ and reconstructed image $u(x,y)$ [7, 37]:

$$I[u] = \int_{\Omega} \left(p(x,y) - (Ru)(x,y)\right)^2 + \alpha\,\phi(|\nabla u|)\big)\,dxdy, \tag{4.66}$$

where Ω is the unit square, R is a linear operator of $L^2(\Omega) \to L^2(\Omega)$, ϕ is a function $\mathbb{R}^+ \to \mathbb{R}^+$, $\alpha \in \mathbb{R}^+$ is a parameter, and $0 \le p(x,y) \le 1$.

In our computation, we choose $R = I$, $\phi(t) = \sqrt{1+t^2}$, $\alpha = 1$ and $p = \frac{1}{1+e^{1000(x+y-1.25)}}$ together with the boundary condition

$$u = p(x,y) \qquad \text{on } \partial\Omega. \tag{4.67}$$

The functional (4.66) becomes

$$I[u] = \int_\Omega \left(\left(1 + |\nabla u|^2\right)^{\frac{1}{2}} + (p-u)^2 \right) dxdy, \tag{4.68}$$

which, once again, is not quadratic. No analytical solution is available for this example. A computed solution and adaptive meshes are shown in Fig. 4.10. The results demonstrate that VPaniso is comparable to IEaniso.

The adaptive meshes obtained based on VPiso, VPaniso and VPhb are shown together in Fig. 4.11. As in the previous example, the anisotropic metric tensors are comparable and provide a better mesh adaptation than the isotropic one. Again, the HBEE-based mesh has a slightly larger maximum aspect ratio.

Fig. 4.12 shows a result for $p = 0.5$ together with the boundary condition

$$\begin{cases} u = 0 & \text{on } x = 0 \text{ or } x = 1, \\ u = 1 & \text{on } y = 0 \text{ or } y = 1. \end{cases} \tag{4.69}$$

Again, the results demonstrate that VPaniso is comparable to IEaniso.

Example 4.4.5. Lastly, we compare the results obtained from VPaniso (4.34) based on variational formulation and IEaniso (or (3.29) based on PDE formulation for the anisotropic diffusion problems discussed in chapters 2 and 3. We consider again the Example 2.3.1. For convenience, we restate the problem here.

The PDE form of the example is

$$-\nabla \cdot (\mathbb{D}\nabla u) = 0, \qquad \text{in} \quad \Omega,$$

and the variational form is

$$I[u] = \int_\Omega \left(\frac{1}{2} \nabla u \cdot \mathbb{D} \nabla u \right) d\boldsymbol{x}, \quad \forall u \in U_g.$$

The domain and boundary conditions are

$$\Omega = [0,1]^2 \backslash \left[\frac{4}{9}, \frac{5}{9} \right]^2, \quad u = 0 \text{ on } \Gamma_{out}, \quad u = 2 \text{ on } \Gamma_{in},$$

where Γ_{out} and Γ_{in} are the outer and inner boundaries of Ω, respectively; see Fig. 2.2. The diffusion matrix is given by

$$\mathbb{D} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1000 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}.$$

with θ being the angle of the primary diffusion direction (parallel to the first eigenvector of $\mathbb{D}$). This problem satisfies the maximum principle and the solution stays between 0 and 2.

We first consider the case of constant $\mathbb{D}$ with $\theta = \pi/4$. Fig. 4.13 shows finite element solutions obtained with VPaniso and IEaniso (or M_{adap}). Meshes and solution contours obtained with those two metric tensors are shown in Figs. 4.14 and 4.15, respectively. Undershoots and unphysical minima are observed in both the solution obtained with IEaniso ($u_{min} = -0.0039$) and VPaniso ($u_{min} = -0.0119$). The undershoot is more significant with VPaniso than that with IEaniso.

The results obtained from both VPaniso and IEaniso (or M_{adap}) violate discrete maximum principle, which is reasonable since no DMP-satisfied condition is applied in the derivation of the corresponding metric tensors. The mesh obtained with IEaniso has better alignment with the solution than the one obtained with VPaniso. On the

other hand, the mesh obtained with VPaniso may preserve more structure information for the underlying problem than IEaniso. However, we cannot see much benefit of using VPaniso based on the examples we have studied. More studies are needed to demonstrate the adavantage of using VPaniso.

Next we consider a case of variable $\mathbb{D}$ with $\theta = \pi \sin(x)\cos(y)$. The finite element solutions, meshes, and solution contours are shown in Figs. 4.16, 4.17, and 4.18, respectively. Similar observations as for the constant $\mathbb{D}$ case can be made.

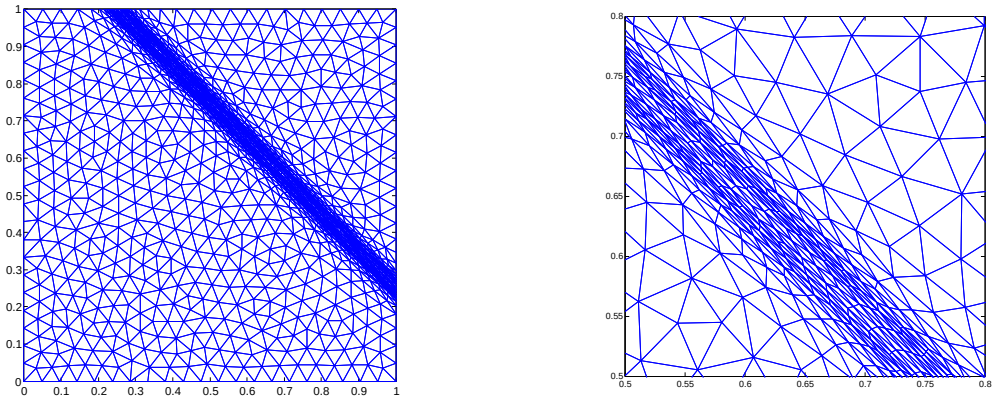
4.5 Conclusions and comments

In this chapter, the metric tensor for use in the adaptive finite element solution of variational problems has been developed based on the underlying variational formulation. Three formulas, (4.34), (4.40), and (4.56) have been obtained. The first one (4.34) is semi-a posteriori in the sense that it involves the residual r_h and the edge jump R_h, both dependent on the computed solution, and the Hessian of the exact solution which is approximated using least squares fitting in actual computation. The second one (4.40) is for isotropic mesh adaptation and is completely a posteriori, involving only r_h and R_h. The third one (4.56) is for anisotropic mesh adaptation but also a posterior based on residual, edge jumps and the hierarchical basis error estimator.

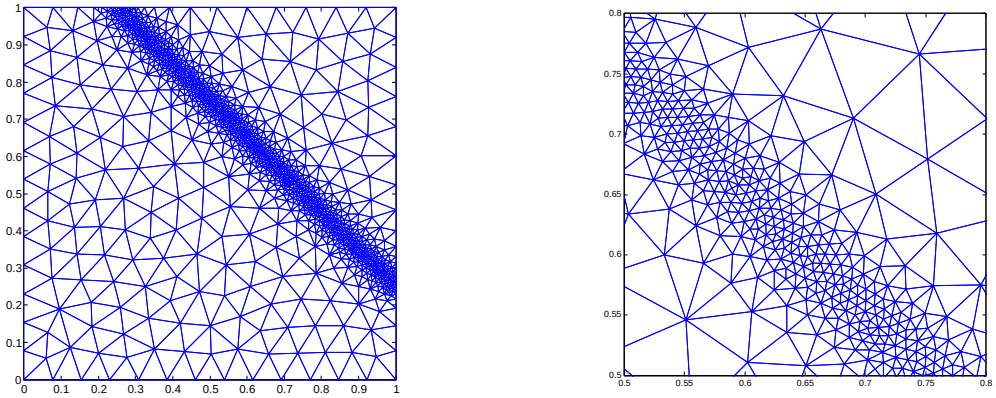
Unlike the existing ones, the new formulas incorporate structural information of the underlying problem into their design and generate meshes which adapt to changes in the structure of the underlying problem. This work is motivated by the argument that the underlying variational formulation should naturally be used for the design of computational meshes for the numerical solution of variational problems. This idea has been used and advocated in the past by a number of researchers such as in [13, 14, 15, 18, 28, 30, 31, 33, 41, 42, 46, 58, 59, 60, 66, 81, 119]. The numerical results have

shown that the approach is feasible and can be used for anisotropic mesh adaptation for variational problems. More studies are needed to demonstrate the advantages of the approach over interpolation error based methods or other methods without using variational structures.

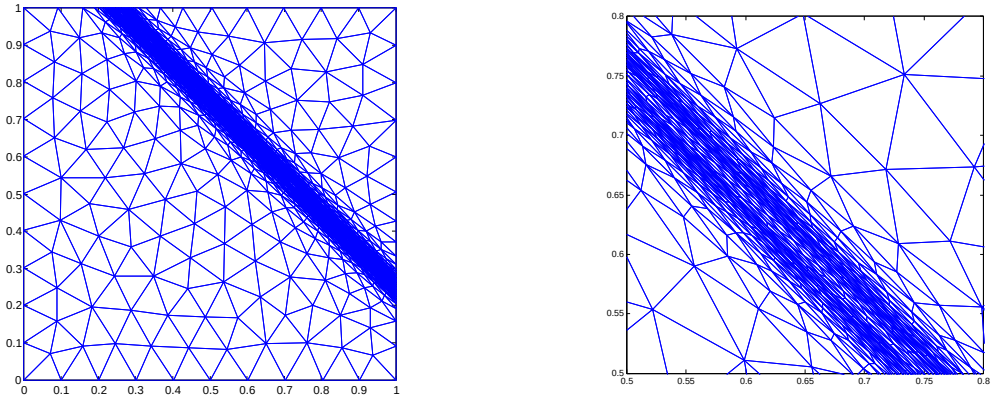
Numerical results confirm the conclusion of [77] that a global HBEE can be a successful alternative to Hessian recovery in mesh adaptation; a fast approximate solution of the global error problem is sufficient to provide directional information for anisotropic mesh adaptation. They also confirm the conjecture that good mesh adaptation does not require a convergent Hessian recovery or an accurate error estimator, but rather some additional information of global nature, although it is still unclear which information exactly is necessary for successful anisotropic adaptation.



(a) IEaniso, $N_v = 2387$, $|e|_{H^1} = 0.28$, and a close-up at (0.65,0.65).



(b) VPiso, $N_v = 2390$, $|e|_{H^1} = 0.79$, and a close-up at (0.65,0.65).



(c) VPaniso, $N_v = 2353$, $|e|_{H^1} = 0.17$, and a close-up at (0.65,0.65).

Figure 4.1: Example 4.4.1. Adaptive meshes obtained with different formulas of the metric tensor.

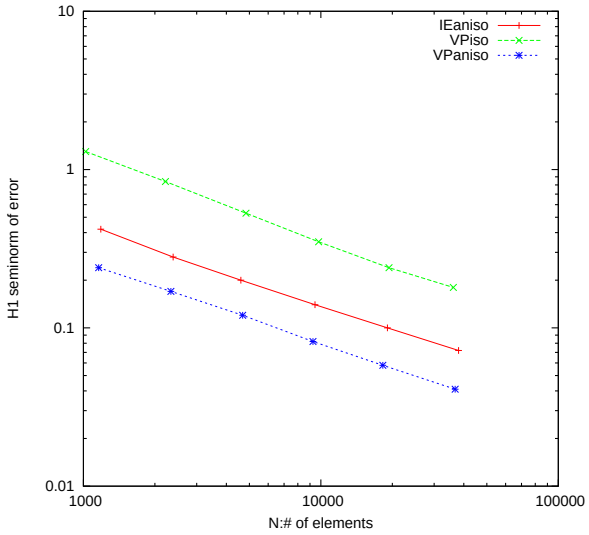
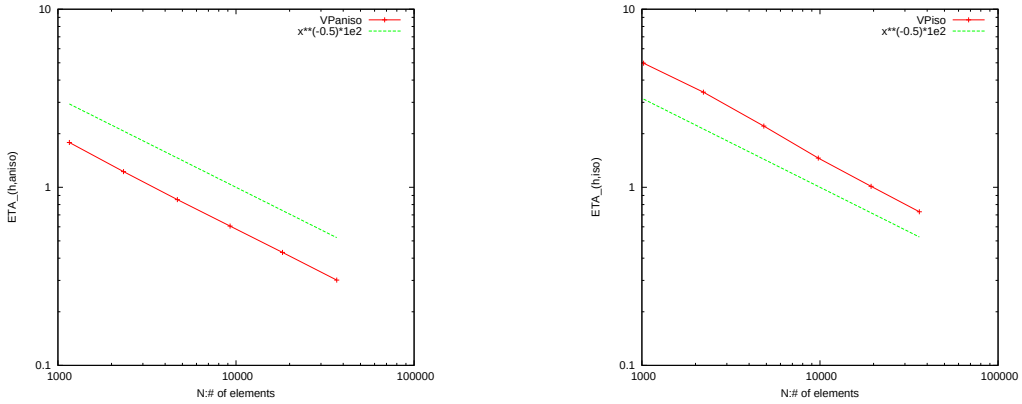


Figure 4.2: Example 4.4.1. A comparison of H^1 seminorm of the error for the finite element solutions obtained using different mesh adaptation strategies, $\varepsilon = 0.005$.



(a): Anisotropic Error Estimate

(b): Isotropic Error Estimate

Figure 4.3: Example 4.4.1. Error estimates associated with VPiso and VPaniso are plotted as functions of the number of elements.

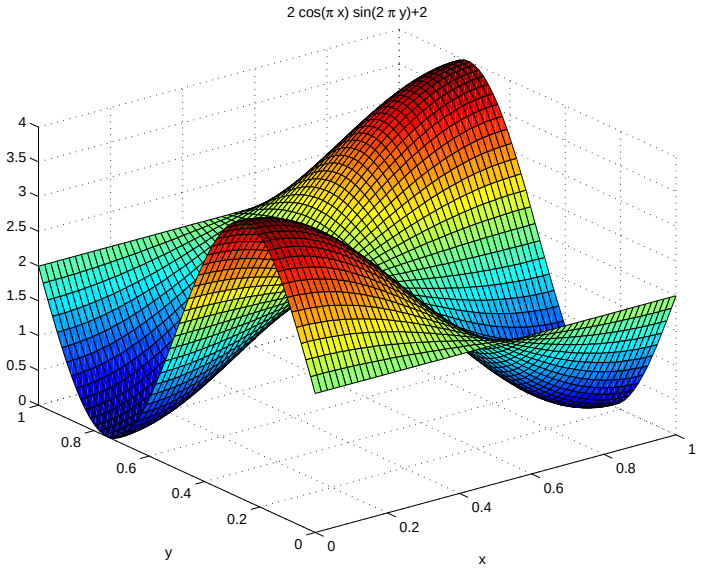
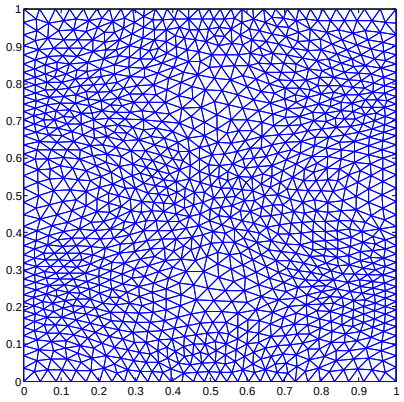
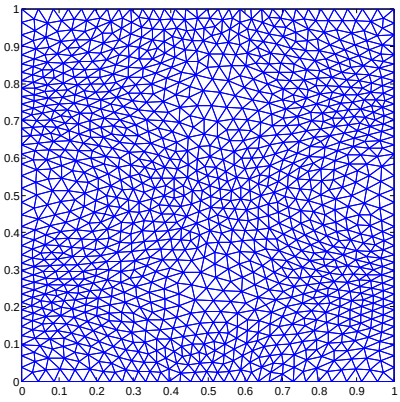


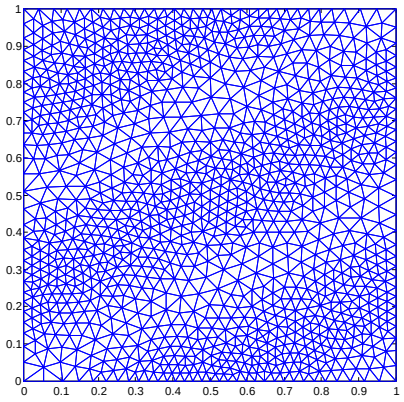
Figure 4.4: The solution of Example 4.4.2.



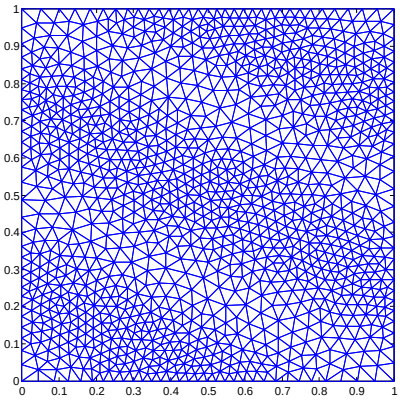
(a): IEaniso, $\theta = \pi/6$



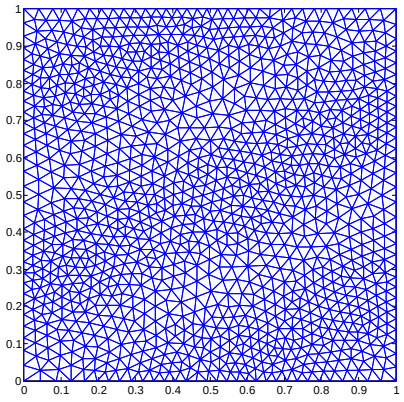
(b): IEaniso, $\theta = 5\pi/6$



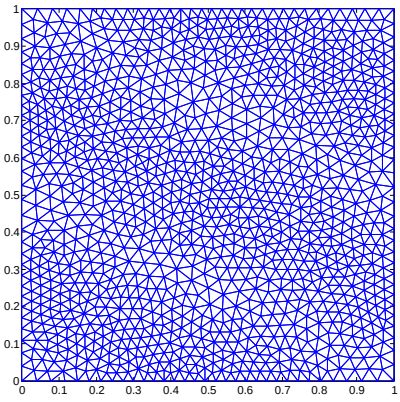
(c): VPiso, $\theta = \pi/6$



(d): VPiso, $\theta = 5\pi/6$

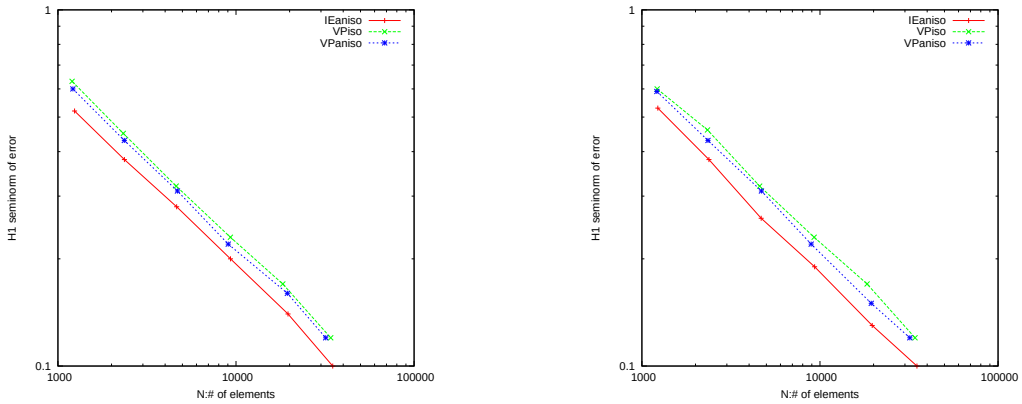


(e): VPaniso, $\theta = \pi/6$



(f): VPaniso, $\theta = 5\pi/6$

Figure 4.5: Example 4.4.2. Adaptive meshes obtained with different formulas of the metric tensor for $\theta = \pi/6$ and $\theta = 5\pi/6$.



(a): $\theta = \pi/6$ (b): $\theta = 5\pi/6$

Figure 4.6: Example 4.4.2. The H^1 seminorm of the error is plotted as function of the number of elements for $\theta = \pi/6$ and $\theta = 5\pi/6$.

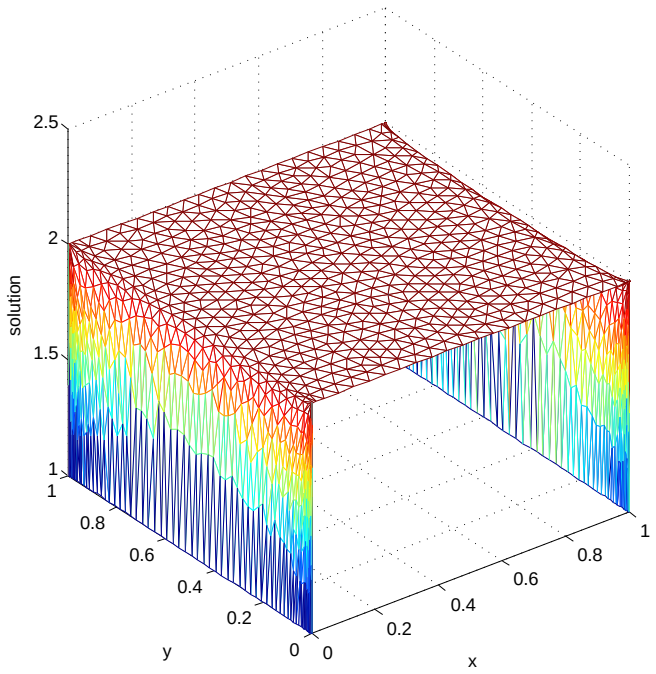
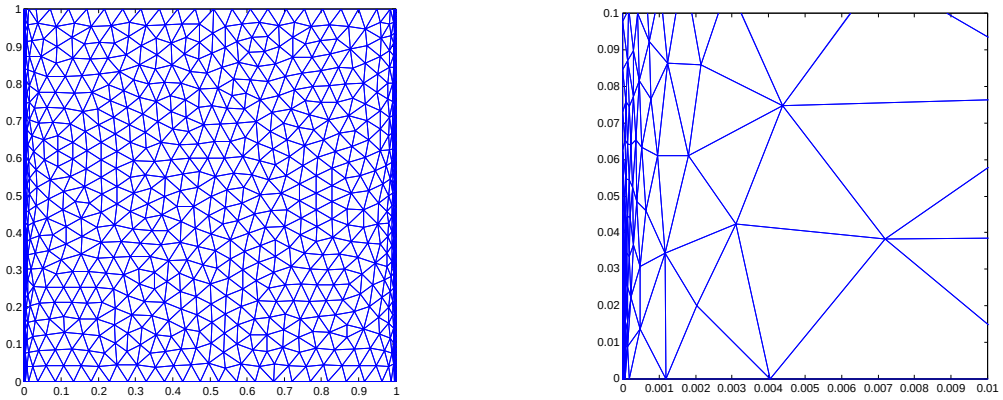
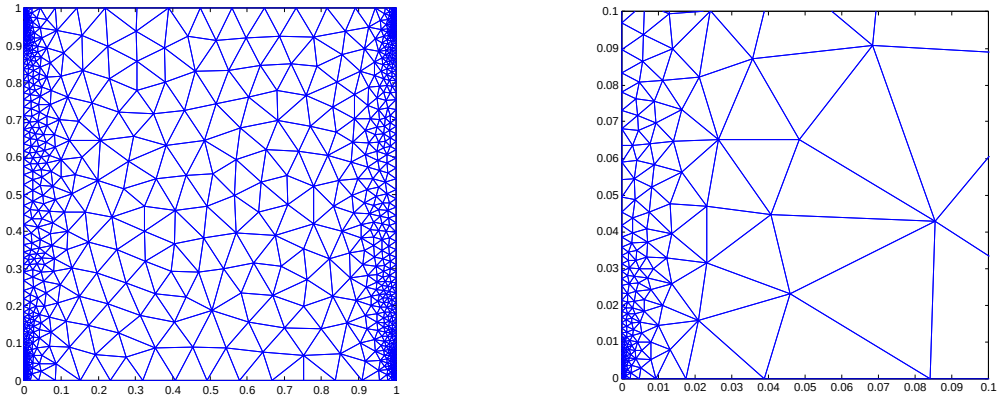


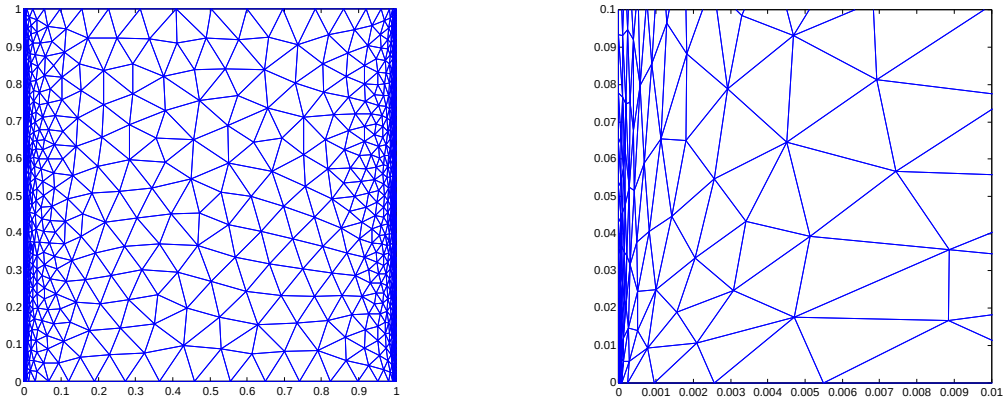
Figure 4.7: A computed solution for Example 4.4.3.



(a) IEaniso, $N_v = 2676$, and a close-up at (0,0).

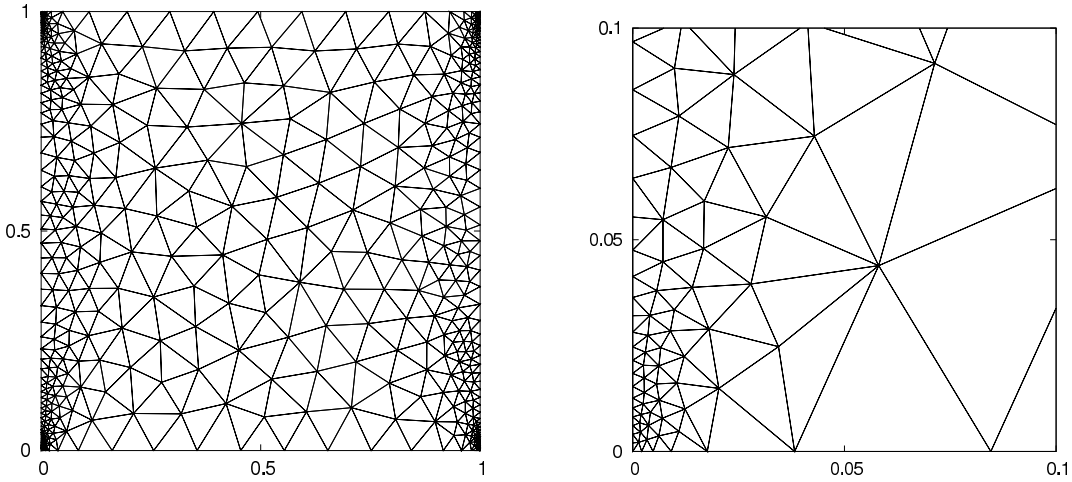


(b) VPiso, $N_v = 2643$, and a close-up at (0,0).

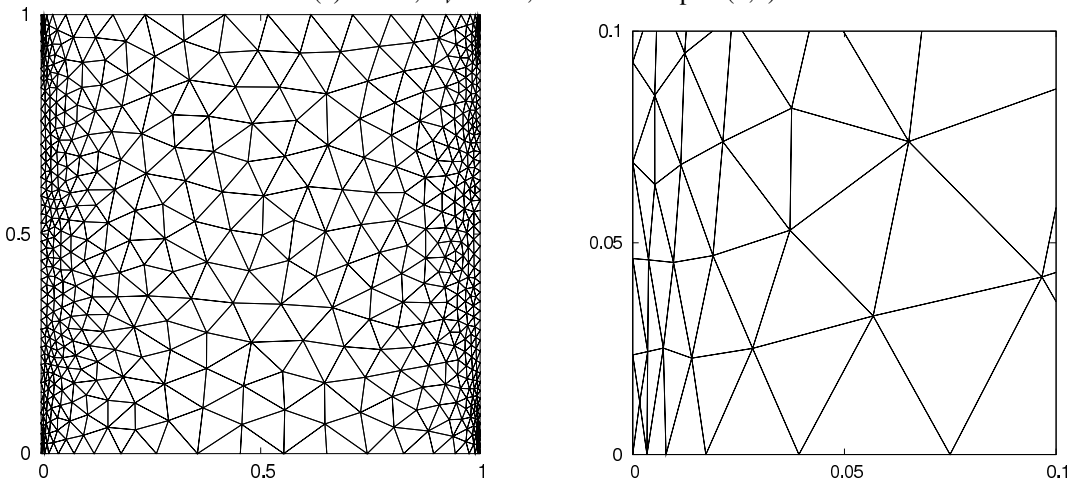


(c) VPaniso, $N_v = 2666$, and a close-up at (0,0).

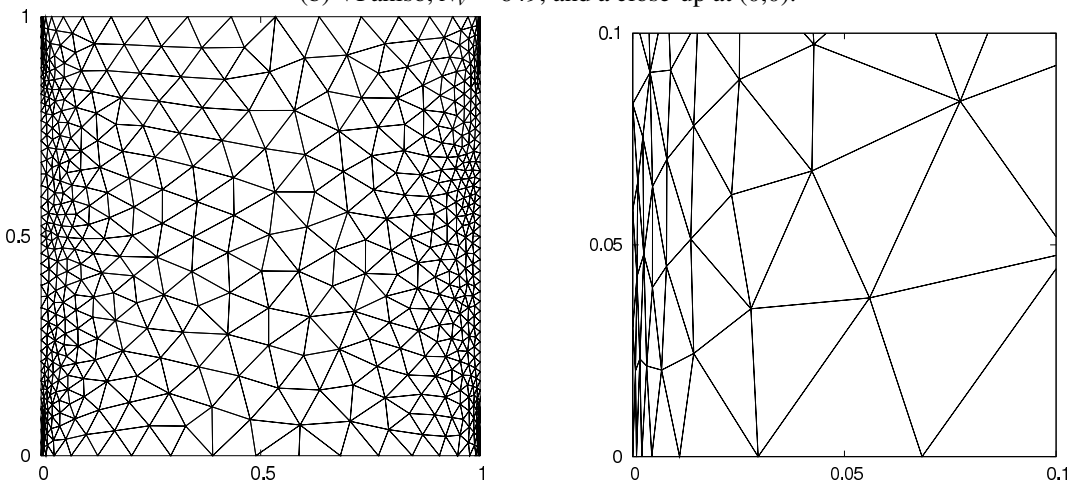
Figure 4.8: Adaptive meshes obtained with different formulas of the metric tensor for Example 4.4.3.



(a) VPiso, $N_v = 644$, and a close-up at (0,0).

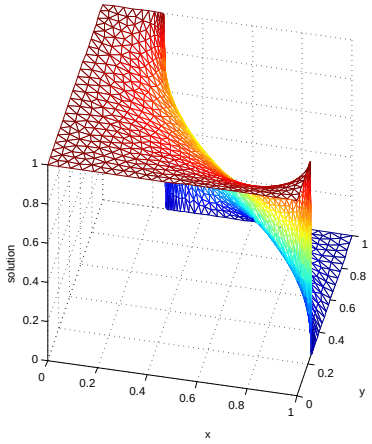


(b) VPaniso, $N_v = 649$, and a close-up at (0,0).

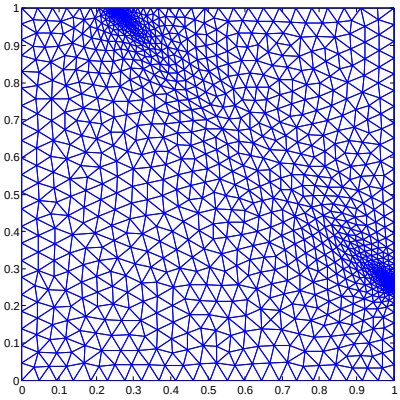


(c) VPhb, $N_v = 639$, and a close-up at (0,0).

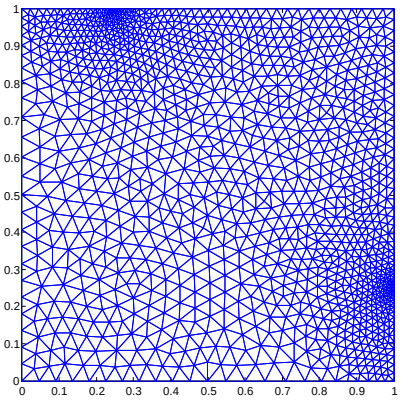
Figure 4.9: Example 4.4.3: adaptive meshes obtained based on VPiso, VPaniso and VPhb.



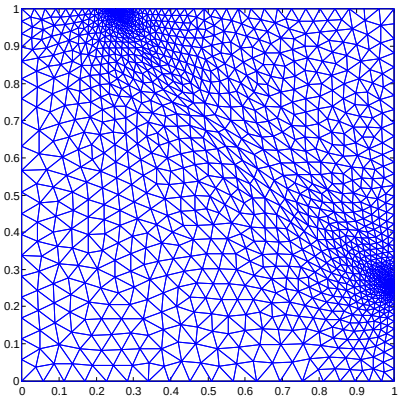
(a): Computed solution



(b): IEaniso Mesh, $N_v = 2545$

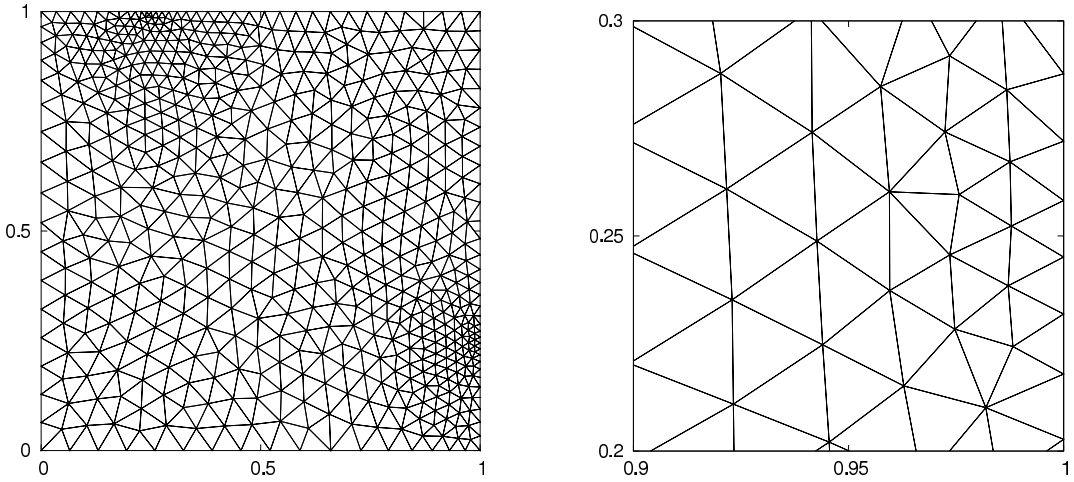


(c): VPiso Mesh, $N_v = 2518$

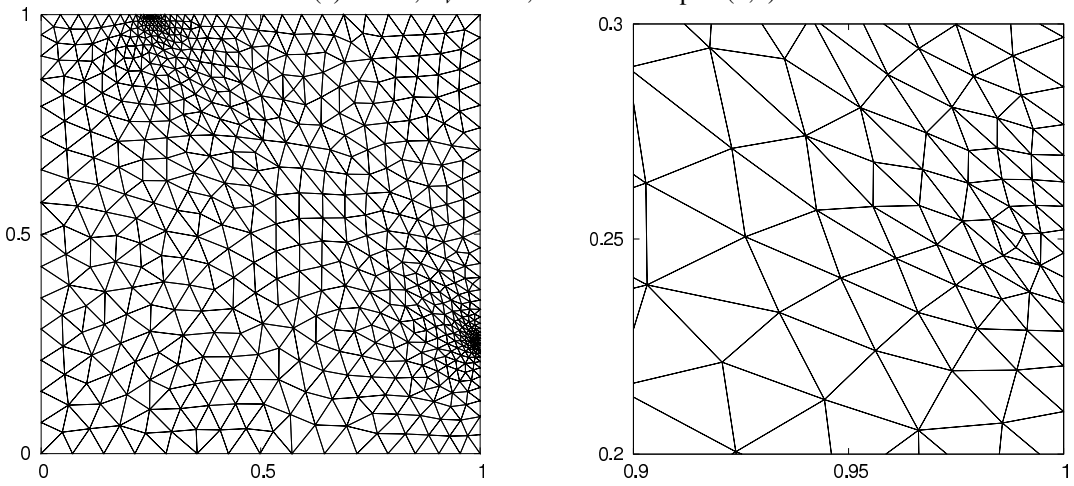


(d): VPaniso Mesh, $N_v = 2527$

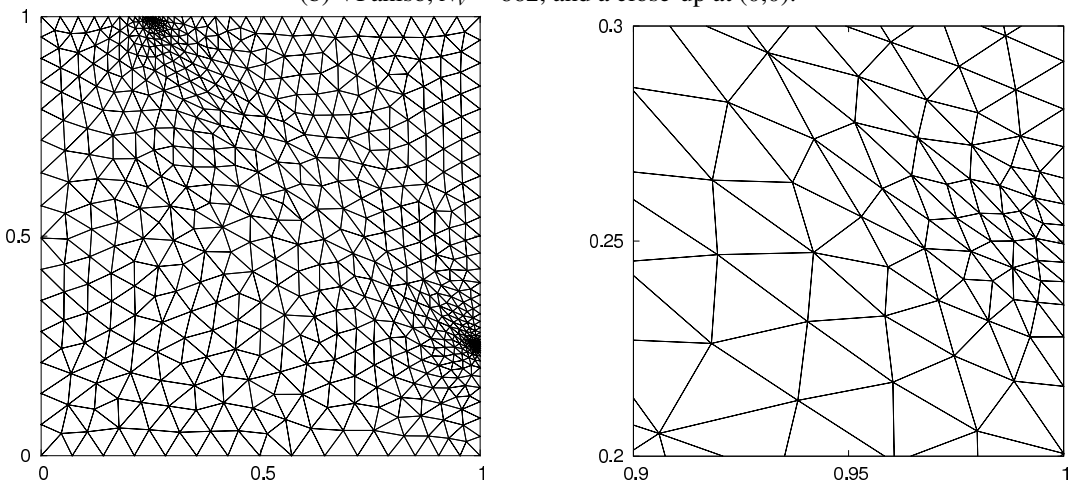
Figure 4.10: Example 4.4.4. A computed solution and adaptive meshes obtained based on IEaniso, VPiso and VPaniso.



(a) VPiso, $N_v = 657$, and a close-up at (0,0).

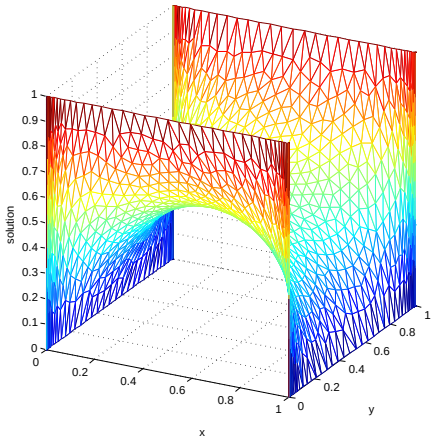


(b) VPaniso, $N_v = 662$, and a close-up at (0,0).

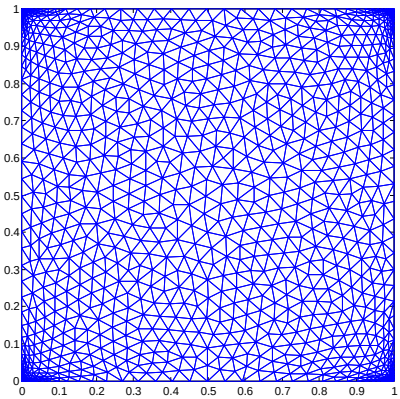


(c) VPhb, $N_v = 656$, and a close-up at (0,0).

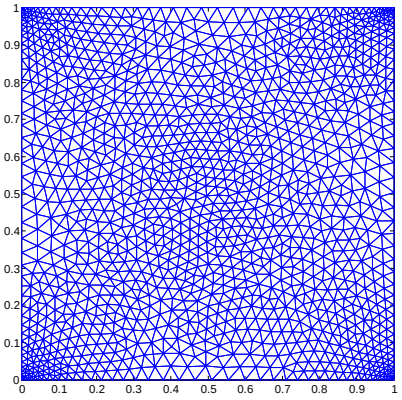
Figure 4.11: Example 4.4.4. Adaptive meshes obtained based on VPiso, VPaniso and VPhb.



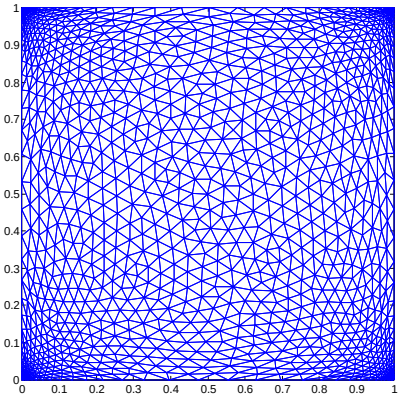
(a): Computed solution



(b): IEaniso Mesh, $N_v = 2667$

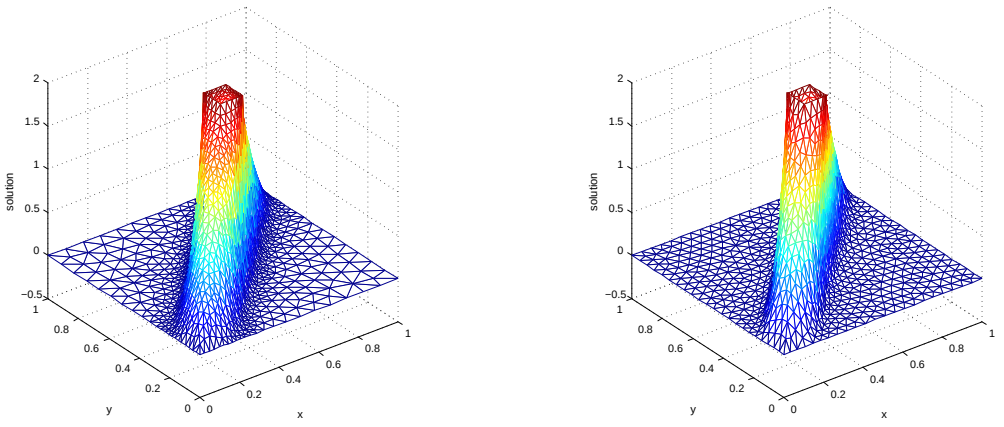


(c): VPiso Mesh, $N_v = 2656$



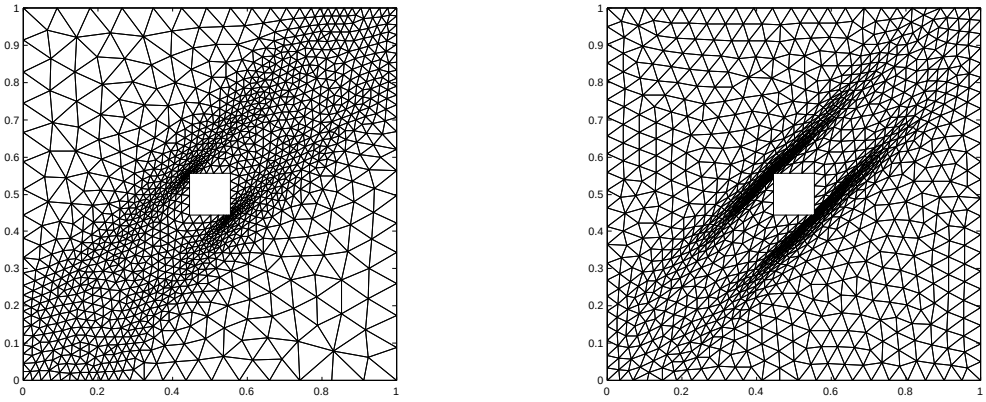
(d): VPaniso Mesh, $N_v = 2666$

Figure 4.12: A computed solution and adaptive meshes obtained with different formulas of the metric tensor for Example 4.4.4 with $p = 0.5$ and boundary condition (4.69).



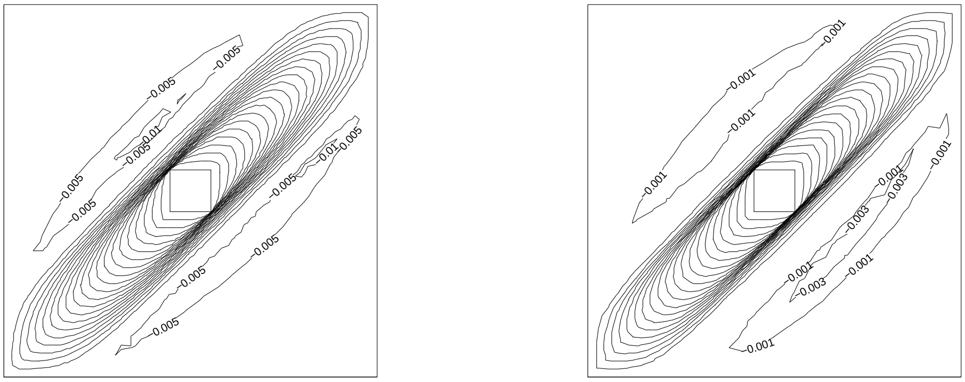
(a): VPaniso (b): IEaniso

Figure 4.13: Example 4.4.5 with constant $\mathbb{D}$. Finite element solutions obtained with (a) VPaniso and (b) IEaniso.



(a): VPaniso, $N_v = 2584$ (b): IEaniso, $N_v = 2583$

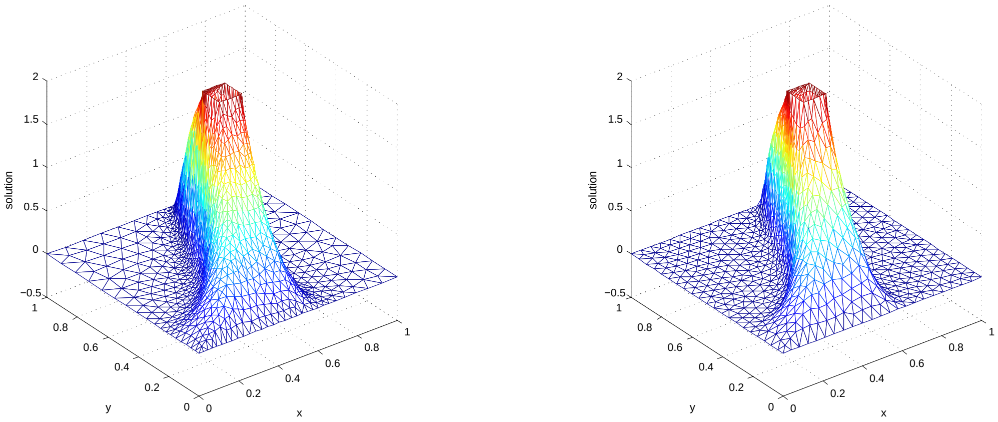
Figure 4.14: Example 4.4.5 with constant $\mathbb{D}$. Meshes obtained with (a) VPaniso and (b) IEaniso.



(a): VPaniso, $u_{min} = -0.0119$

(b): IEaniso, $u_{min} = -0.0039$

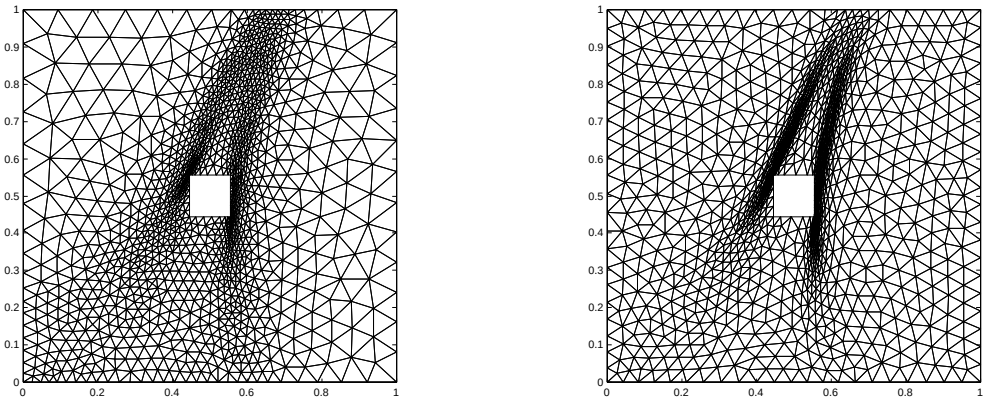
Figure 4.15: Example 4.4.5 with constant $\mathbb{D}$. Contours of the finite element solutions obtained with (a) VPaniso and (b) IEaniso.



(a): VPaniso

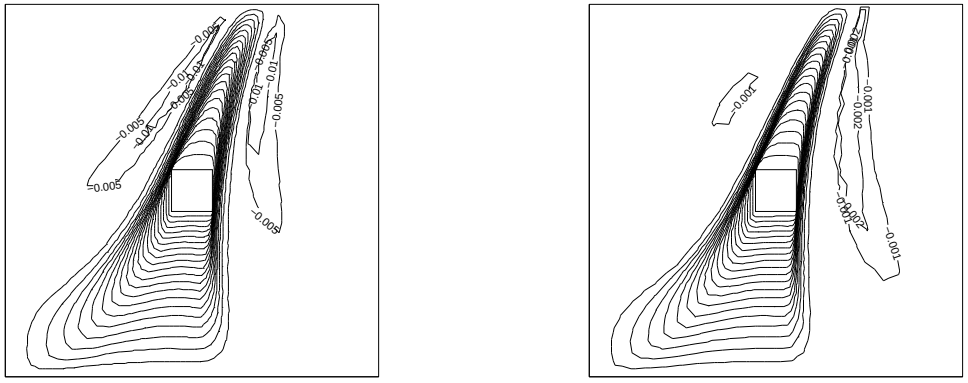
(b): IEaniso

Figure 4.16: Example 4.4.5 with variable $\mathbb{D}$. Finite element solutions obtained with (a) VPaniso and (b) IEaniso.



(a): VPaniso, $N_v = 2593$ (b): IEaniso, $N_v = 2568$

Figure 4.17: Example 4.4.5 with variable $\mathbb{D}$. Meshes obtained with (a) VPaniso and (b) IEaniso.



(a): VPaniso, $u_{min} = -0.0175$ (b): IEaniso, $u_{min} = -0.0039$

Figure 4.18: Example 4.4.5 with variable $\mathbb{D}$. Contours of the finite element solutions obtained with (a) VPaniso and (b) IEaniso.

Chapter 5

Application in Image Processing

In this chapter, we study the application of anisotropic mesh adaptation in image processing. Background of image processing was introduced in §1.1.3. The application of anisotropic diffusion in image processing is addressed in the next section. The application of anisotropic mesh adaptation is discussed in §5.2.

5.1 Anisotropic diffusion in image processing

As mentioned in §1.1.3, anisotropic diffusion filters have become a popular tool in image processing; e.g. see [102, 123, 16, 57, 117, 62, 51]. They have many applications including noise removal, edge detection, texture enhancement or segmentation. Since the focus of this dissertation is on mesh adaptation, we restrict our attention to noise removal and edge detection. A basic model, as mentioned in §1.1.3, is the Perona-Malik filter (1.8). In this section, we consider a general model based on [53].

$$
\begin{aligned}
u_t &= \nabla \cdot (\mathbb{D}\nabla u) + \beta(u^0 - u), \qquad \text{in} \quad \Omega_T = \Omega \times (0,T] && (5.1)\\
\partial u &= 0, \qquad \text{on} \quad \partial\Omega \times [0,T] && (5.2)\\
u &= u^0, \qquad \text{on} \quad \Omega \times \{t = 0\}, && (5.3)
\end{aligned}
$$

where $\Omega = [0,1] \times [0,1]$, u^0 is the given (or observed) image called "initial image", $\beta \geq 0$ is a parameter, and $\mathbb{D} = \mathbb{D}(\boldsymbol{x})$ is the diffusion matrix assumed to be symmetric and strictly positive definite on Ω.

The model (5.1) is called the Generalized Perona-Malik model (GPM) if the diffusion tensor $\mathbb{D}$ is chosen as

$$
\mathbb{D} = \frac{1}{\left(1 + \frac{|\nabla u|^2}{\lambda}\right)^{\alpha}} \boldsymbol{I}_2, \tag{5.4}
$$

where $\boldsymbol{I}_2$ is the identity matrix of dimension 2 and $\alpha > 0$, $\lambda > 0$ are parameters. The model becomes the well-known Perona-Malik model (PM) when choosing $\beta = 0$ in (5.1) and $\alpha = 1$ in (5.4).

If choosing $\beta = 0$ in (5.1) and $\alpha = 0.5$, $\lambda = 1$ in (5.4), the model represents a regularized version of the Total Variation model (TV), which is given by

$$
\mathbb{D} = \frac{1}{|\nabla u|} \boldsymbol{I}_2. \tag{5.5}
$$

For the examples we studied, we choose $\lambda = 100$ in TV model which provides better denoising effect than $\lambda = 1$.

As can be seen from (5.4) and (5.5), PM model and TV model are in fact isotropic models. For anisotropic model, the diffusion tensor $\mathbb{D}$ should be chosen anisotropic, i.e., its eigenvalues should not be all equal at least on a portion of Ω.

5.2 Anisotropic mesh adaptation in image processing

This section discusses anisotropic mesh adaptation for finite element solution of the General Perona-Malik model (5.1). Conventionally, anisotropic diffusion models in image processing are solved using finite difference method which provides a natural discretization on a fixed rectangular grid. Explicit schemes are commonly used although they are restricted to small time step sizes to ensure stability. Semi-implicit schemes which posses better stability properties are considered in [35].

On the other hand, finite element techniques in image processing are considered in [10, 106, 11, 12, 51]. One of the advantages of using the finite element method is the ease to apply mesh adaptation to improve the computational efficiency. For example, considering a gray scale image with size 256×256. The number of degree of freedom (DOF) using the finite difference method will be $65,536$, and the number of vertices will also be $65,536$ for a regular triangle mesh using the finite element method. With an adaptive mesh, only a small amount of elements (or DOF) is sufficient to provide comparable image quality. Another advantage of anisotropic mesh adaptation is that some important features of the original image can be preserved by the concentration of mesh elements in the corresponding area.

Since PM model and TV model do not show much difference for the examples we examined, only the results from PM model are included in this dissertation. For computational convenience, we just consider gray scale images of size 256×256. Colorful images with large sizes can be treated similarly.

The results obtained from three meshing strategies are compared. One strategy is to use fixed mesh with $65,536$ vertices, the results is labeled with "M_{fixed}" where no adaptation is applied for the mesh. Another strategy is to use the metric tensor M_{adap} (3.29) based on minimization of an interpolation error bound [74]. The third strategy is

to use the metric tensor M_{DMP} (2.17) developed in §2.2. Note that M_{DMP} is developed for steady state problem (1.1) and (1.2). Although it guarantees DMP satisfaction for the finite element solution of (1.1) and (1.2), it does not guarantee DMP satisfaction for parabolic PDEs.

5.3 Numerical results

In this section, we present the results of anisotropic mesh adaption applied in image processing. In all of the computations except addressed otherwise, a fixed time step $\Delta t = 10^{-5}$ is used and the computation stops after 10 time evolutions, i.e., at $t = 10^{-4}$. For the Perona-Malik (PM) model, we choose $\beta = 0$ in (5.1), $\lambda = 100$ and $\alpha = 1$ in (5.4).

The noisy images (J) are obtained by applying Gaussian noise (with mean 0 and variance 0.01) to the original images (I) using the following MATLAB commands:

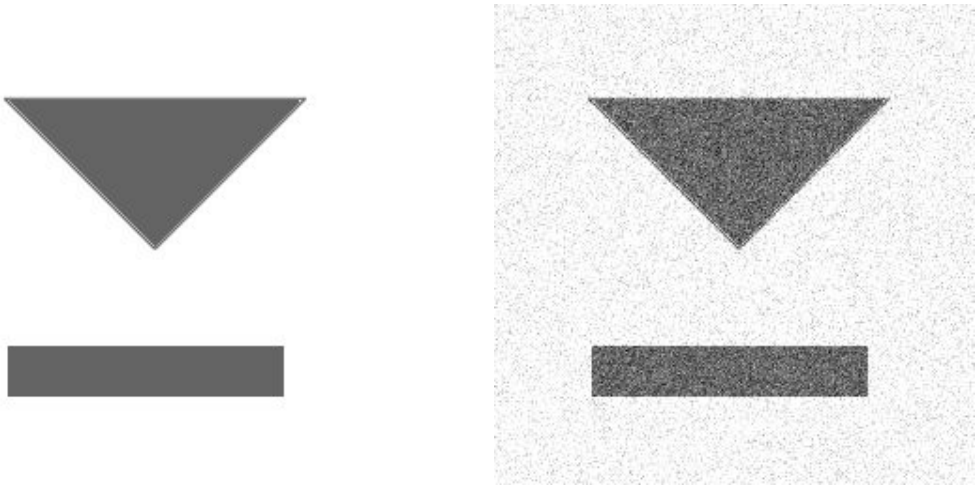
```
% read original image
I = imread('image name');
% apply Gaussian noise to I
J = imnoise(I, 'gaussian', 0, 0.01);
```

Finite element discretization for space and backward Euler method for time are applied to solve the parabolic problem. The discrete linear system is solved via the UMFPACK Direct Solver. The computations are performed on a computer with 2.1 GHz Intel Core 2 Duo processor and 3 GB 667 MHz DDR2 SDRAM memory.

Example 5.3.1. Our first example is an noise removal or edge detection problem for image shown in Fig. 5.1. The results are shown in Fig. 5.2. The goal is to remove the noise around the triangle and rectangle. Adaptive meshes using M_{adap} and M_{DMP}

with about $5,000$ vertices provide good quality of final image comparing to fixed mesh using M_{fixed} with $65,536$ vertices. It is noted that both M_{adap} and M_{DMP} provide results that are comparable to [10] and better than [12].

For the computation of 10 time evolutions, it takes 377 seconds using M_{fixed} ($N_v = 65536$) and 28 seconds using either M_{adap} with $N_v = 5118$ or M_{DMP} with $N_v = 5108$. As can be seen, the computational efficiency is improved significantly.



(a): Original image (b): Noisy image

Figure 5.1: Example 5.3.1. Image before processing.

Example 5.3.2. Our second example is similar to Example 5.3.1 except the edge is on a circle. The results are shown in Fig. 5.3 and are consistent with the observations made in Example 5.3.1.

Example 5.3.3. The third example is denosing of a picture of "famous" Lena in Fig. 5.4. The results are shown in Fig. 5.5. As can be seen from 5.5(d) and (f), the concentration of mesh elements preserves some features of the image. However, the processed images, 5.5(c) based on M_{adap} and (e) based on M_{DMP} are not as clear as the one obtained using fixed full image mesh 5.5(a). This is reasonable since a mesh with

just $5,500$ vertices cannot preserve all the information in the initial image with $65,536$ vertices. By using more mesh elements during anisotropic mesh adaptation, the results are getting better, see Fig. 5.6 and 5.7.

Fig. 5.7 demonstrate that M_{adap} and M_{DMP} are comparable to M_{fixed} for a complicated image when the numbers of elements are comparable. In this case, however, the advantage of using anisotropic mesh adaptation is not obvious. For the 10 time evolutions, it takes 366 seconds for M_{fixed} with $N_v = 65536$, 457 seconds for M_{adap} with $N_v = 68657$ and 312 seconds for M_{DMP} with $N_v = 49803$. To improve computational efficiency, the number of elements should be reduced. Therefore, anisotropic mesh adaptation is helpful to preserve key features with less elements.

Example 5.3.4. The last example is an image of human lung. The results are shown in Fig. 5.8. For image data, all the solution values should be between 0 and 1. However, the minimum value of the solution obtained based on the three meshing strategies are negative. Specifically, $u_{min} = -1.9e-7$ for M_{fixed} (5.8(b)), $u_{min} = -1.4e-4$ for M_{adap} (5.8(c)), and $u_{min} = -1.9e-3$ for M_{DMP} (5.8(d)). The negative solution violates DMP and provides artifact in the final image.

After 30 time evolutions, the artifacts become significant and observable, see 5.8(e) and (f). Even using more mesh elements ($N_v = 13,000$) in computation, artifact still exists. This example demonstrates that DMP satisfaction is really an important topic in image processing.

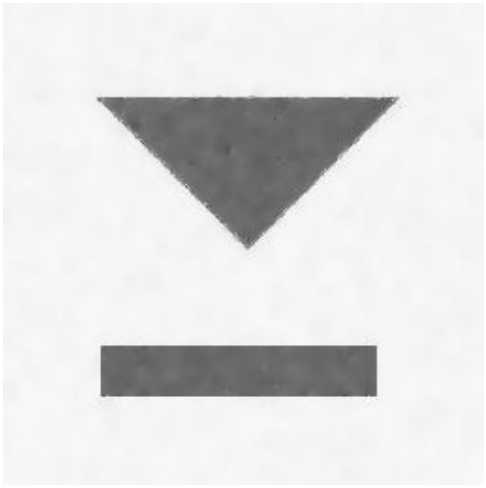
Mass lumping technique [40, 83] is commonly used in the finite element method to preserve monotonicity of the numerical solution. Fig. 5.9 shows the results while mass lumping is applied for the computation. The solution is improved significantly (in terms of DMP) but still violates DMP. Hence, it is desirable to investigate DMP satisfaction for parabolic problems.

5.4 Conclusions and comments

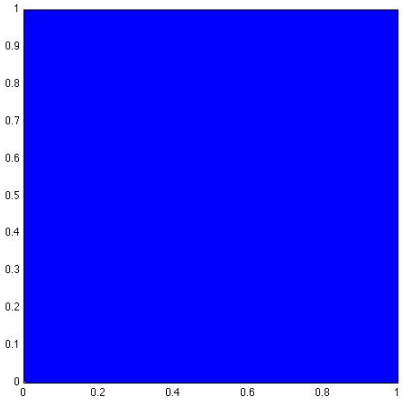
Anisotropic mesh adaptation based on M_{adap} (3.29) and M_{DMP} (2.17) are applied in image processing. The results are compared with those obtained using a fixed full image mesh. The results demonstrate that anisotropic mesh adaptation can improve computational efficiency while keeping comparable quality of the final solution. The concentration of mesh elements also preserves some features of the initial image. Our results are comparable to those in [10] and [106], and better than those in [12].

For computations using M_{adap} and M_{DMP}, there exists interpolation error from image data to initial adaptive mesh. One way to resolve this problem is to use the full image mesh with $65,536$ vertices as the initial mesh for the computation. By this way, the quality of the final image can be improved to some extent. However, the initial computation using full image mesh is costly, especially when the image size is large. For computational efficiency, we prefer starting with a small mesh while accepting the initial interpolation error.

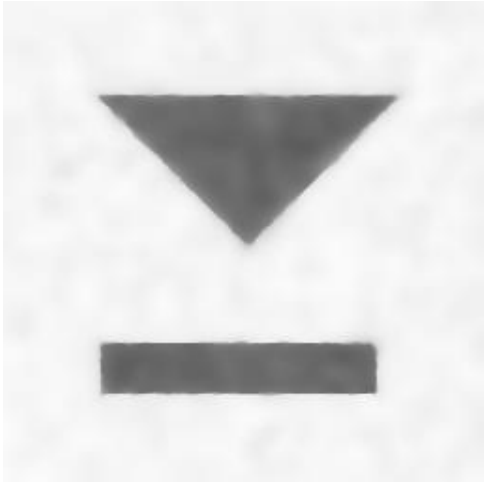
It is worth mentioning that DMP satisfaction is also important during image processing. Improper schemes or meshes may lead to the occurrence of artifact in the final solution (see Example 5.3.4). Hence, it is desirable to utilize a DMP-satisfied mesh in the image processing. This will be a future research topic to develop DMP mesh conditions for time dependent anisotropic diffusion problems.



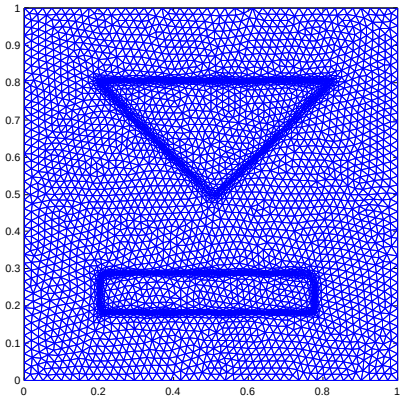
(a): Solution using M_{fixed}



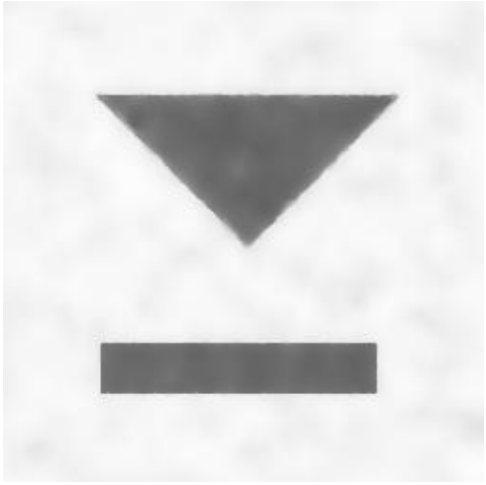
(b): Mesh using M_{fixed}, $N_v = 65536$



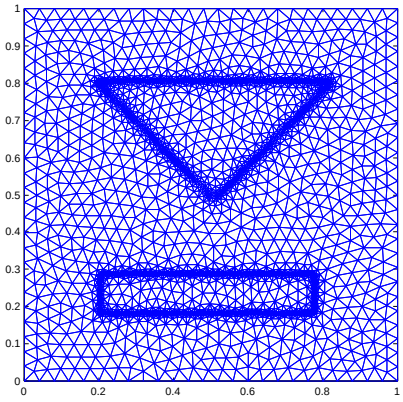
(c): Solution using M_{adap}



(d): Mesh using M_{adap}, $N_v = 5118$

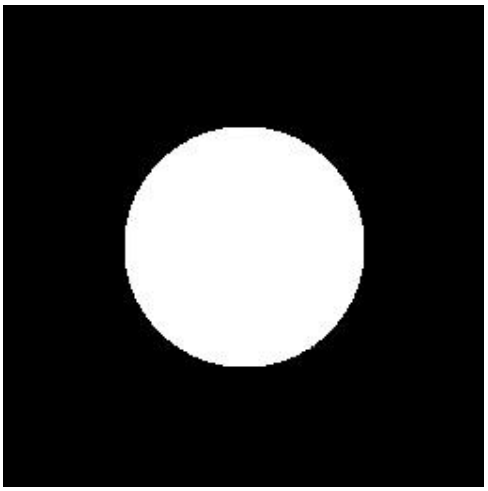


(e): Solution using M_{DMP}

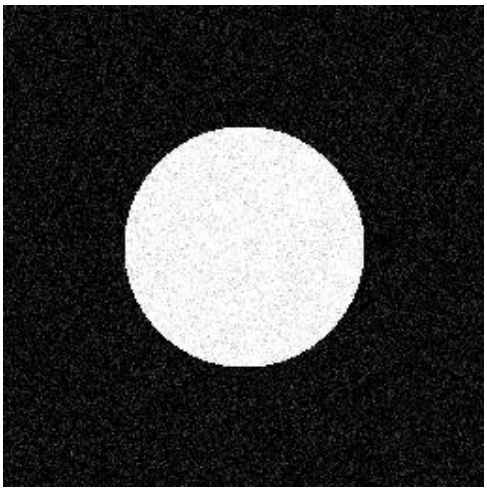


(f): Mesh using M_{DMP}, $N_v = 5108$

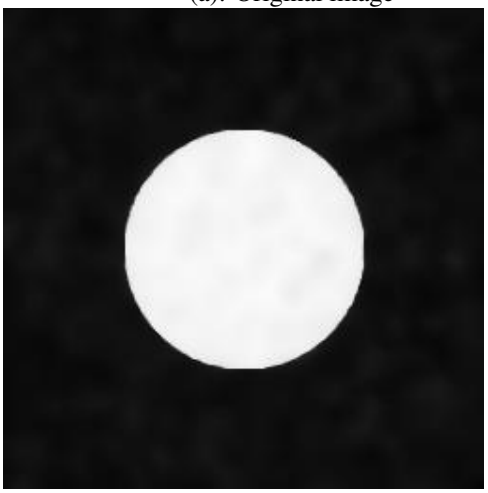
Figure 5.2: Example 5.3.1. Image processing using PM model.



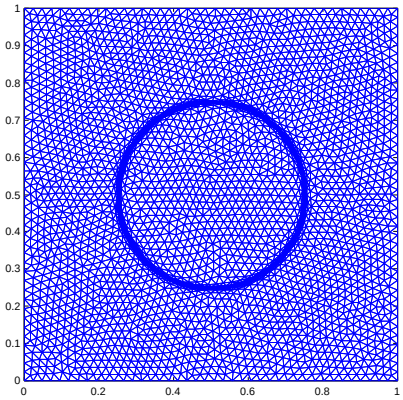
(a): Original image



(b): Noisy image

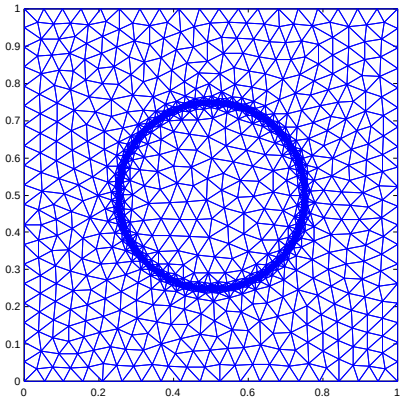


(c): Solution using M_{adap}



(d): Mesh using M_{adap}, $N_v = 5171$

(e): Solution using M_{DMP}



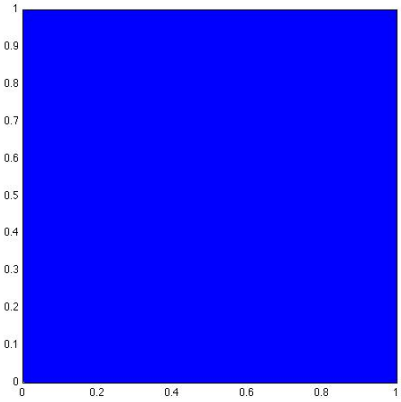
(f): Mesh using M_{DMP}, $N_v = 4899$

Figure 5.3: Example 5.3.2. Image processing using PM model.

(a): Original image (b): Noisy image

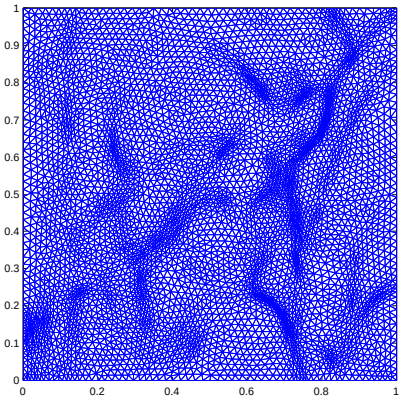
Figure 5.4: Example 5.3.3. Image before processing.

(a): Solution using M_{fixed}



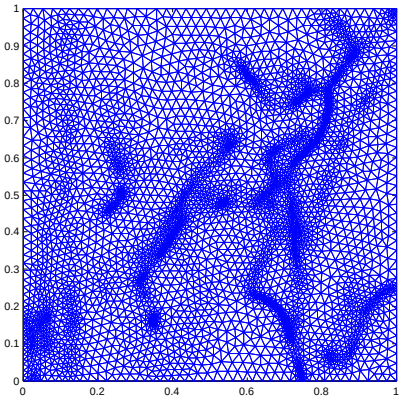
(b): Mesh using M_{fixed}, $N_v = 65536$

(c): Solution using M_{adap}



(d): Mesh using M_{adap}, $N_v = 5477$

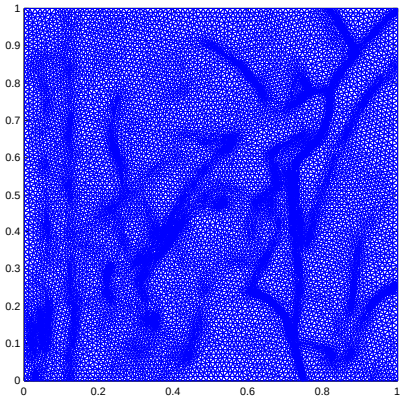
(e): Solution using M_{DMP}



(f): Mesh using M_{DMP}, $N_v = 5425$

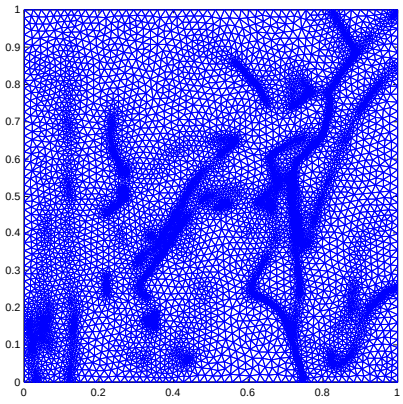
Figure 5.5: Example 5.3.3. Image processing using PM model.

(a): Solution using M_{adap}



(b): Mesh using M_{adap}, $N_v = 12581$

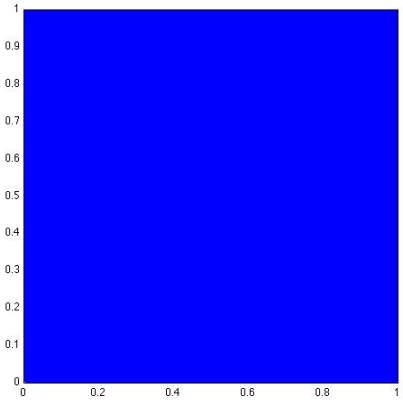
(c): Solution using M_{DMP}



(d): Mesh using M_{DMP}, $N_v = 10626$

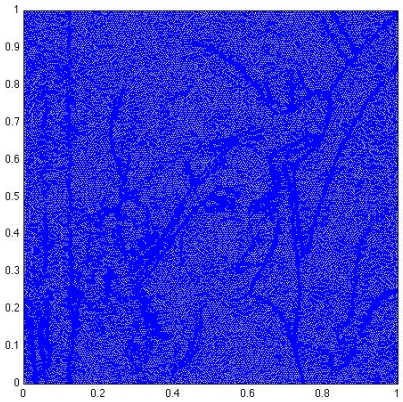
Figure 5.6: Example 5.3.3. Image processing using PM model with more mesh elements.

(a): Solution using M_{fixed}



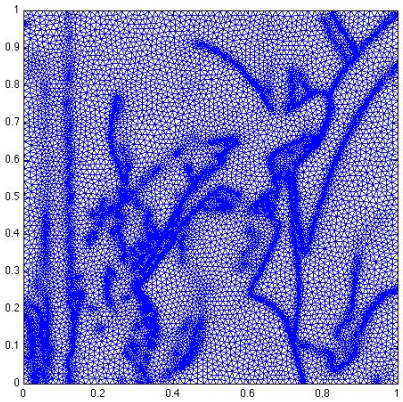
(b): Mesh using M_{fixed}, $N_v = 65536$

(c): Solution using M_{adap}



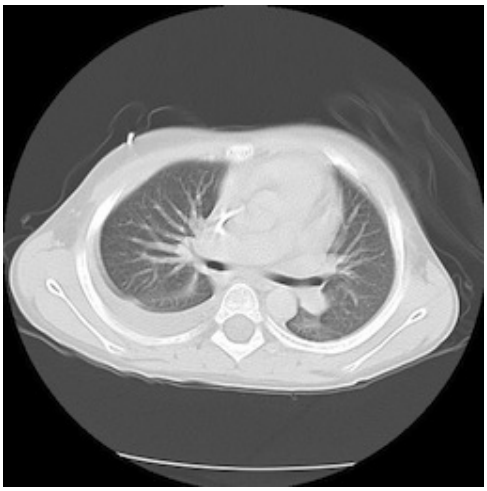
(d): Mesh using M_{adap}, $N_v = 68657$

(e): Solution using M_{DMP}

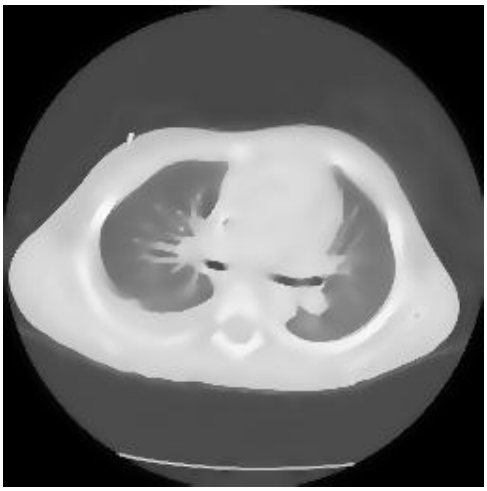


(f): Mesh using M_{DMP}, $N_v = 49803$

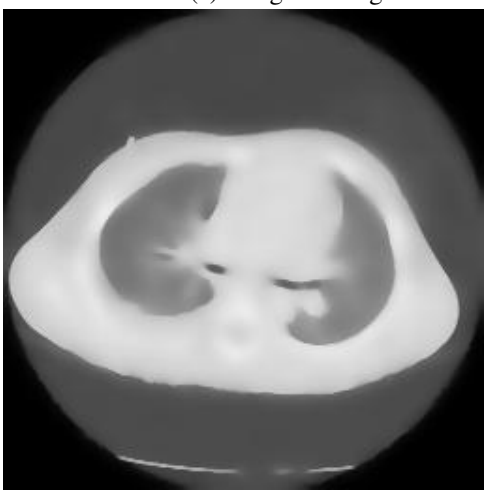
Figure 5.7: Example 5.3.3. Image processing using PM model.



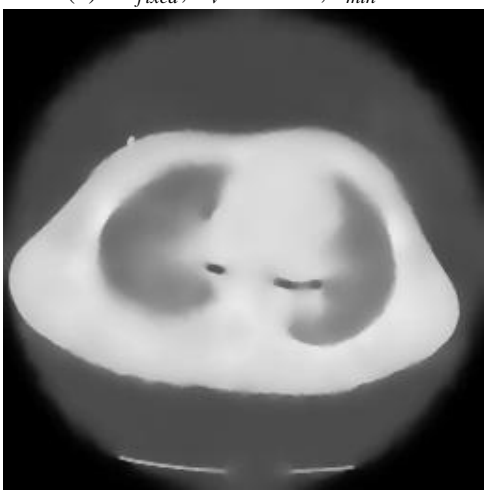
(a): Original image



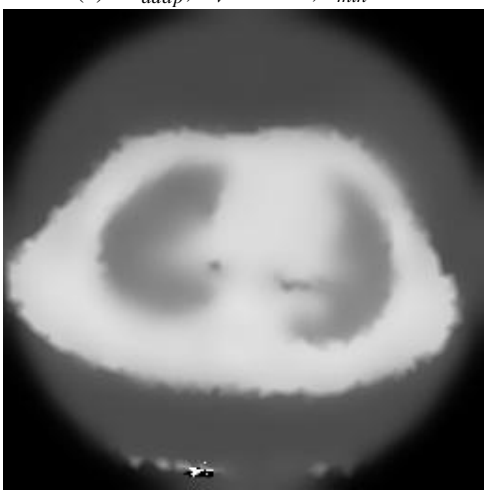
(b): M_{fixed}, $N_v = 65536$, $u_{min} = -1.9e-7$



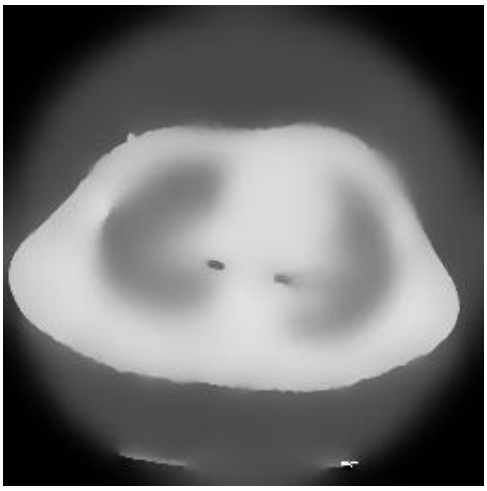
(c): M_{adap}, $N_v = 6539$, $u_{min} = -1.4e-4$



(d): M_{DMP}, $N_v = 5221$, $u_{min} = -1.9e-3$

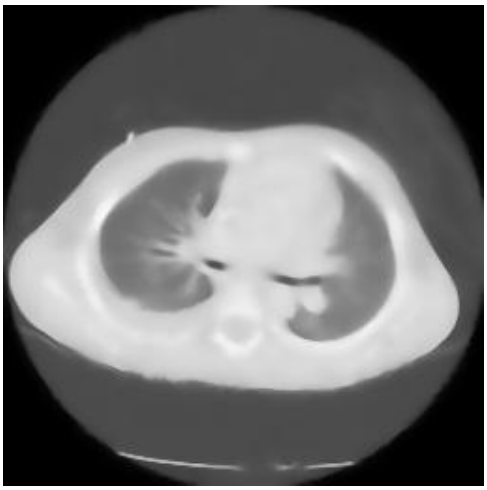


(e): M_{adap}, 30 time evolutions,
$N_v = 7345$, $u_{min} = -86.7$

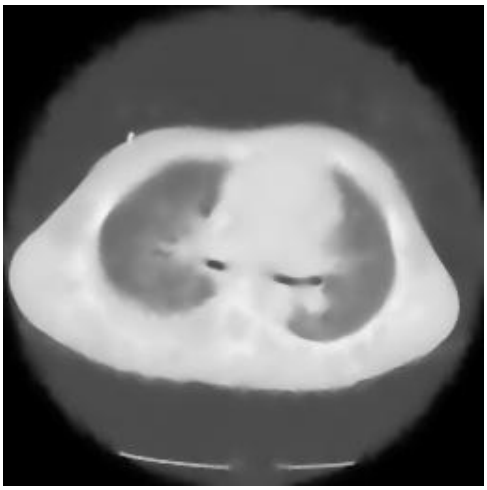


(f): M_{DMP}, 30 time evolutions,
$N_v = 5437$, $u_{min} = -11.4$

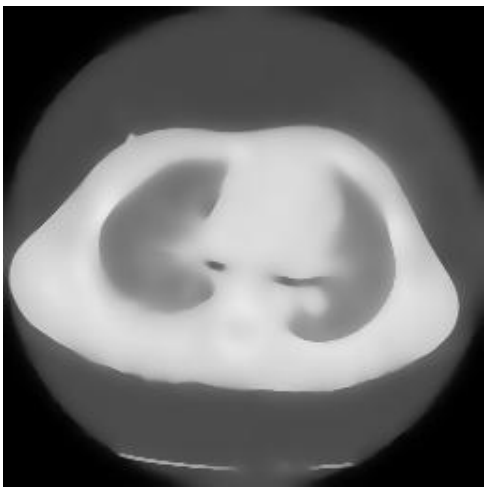
Figure 5.8: Example 5.3.4. Image processing using PM model.



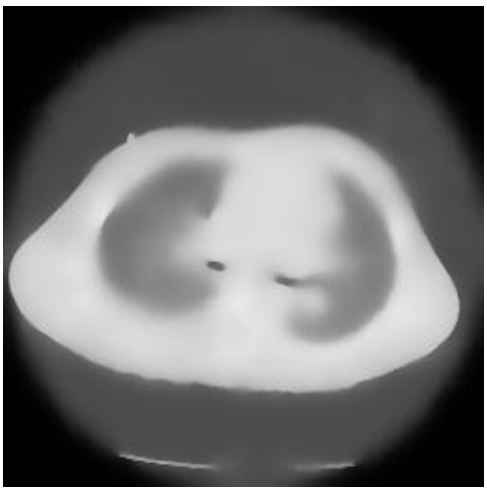
(a): M_{adap}, 10 time evolutions, $N_v = 6467$, $u_{min} = -9.8e-5$



(b): M_{DMP}, 10 time evolutions, $N_v = 5309$, $u_{min} = -2.5e-4$



(c): M_{adap}, 30 time evolutions, $N_v = 6668$, $u_{min} = -2.9e-6$



(d): M_{DMP}, 30 time evolutions, $N_v = 5227$, $u_{min} = -6.1e-5$

Figure 5.9: Example 5.3.4. Image processing using PM model and mass lumping technique.

Chapter 6

Summary and future research topics

This section summarizes the obtained results in this dissertation, and discusses future research topics in this area.

6.1 Summary

In this study, anisotropic mesh adaptation for the finite element solution of anisotropic diffusion problems has been considered. The background of related topics is introduced in Chapter 1. For the boundary value problem (1.1) and (1.2), the discrete maximum principle (DMP) is discussed in Chapters 2 and 3. In particular, an anisotropic non-obtuse angle condition (2.5) is derived in Chapter 2, which is a generalization of the well known non-obtuse angle condition developed for isotropic diffusion problems. Two variants of the anisotropic non-obtuse angle condition, (2.14) and (2.15), are developed for actual mesh generation. A metric tensor (2.17) is derived based on (2.15) for accounting for DMP satisfaction. Chapter 3 discusses the combination of DMP satisfaction and error based mesh adaptation, and a metric tensor accounting for both (3.16) is developed. This is the first time to combine the error based anisotropic mesh adaptation and DMP satisfaction for isotropic/anisotropic diffusion problems.

Chapter 4 discusses the mesh adaptation for the linear finite element solution for variational problem (4.1). Three formulas, (4.34), (4.40), and (4.56) have been developed based on the underlying variational formulation. The first one (4.34) is semi-a posteriori in the sense that it involves the residual r_h and the edge jump R_h, both dependent on the computed solution, and the Hessian of the exact solution which is approximated using least squares fitting in actual computation. The second one (4.40) is for isotropic mesh adaptation and is completely a posteriori, involving only r_h and R_h. The third one (4.56) is for anisotropic mesh adaptation and also is a posterior based on residual, edge jumps and the hierarchical basis error estimator. All the formulas incorporate structural information of the underlying problem into their design and generate meshes which adapt to changes in the structure of the underlying problem.

Chapter 5 discusses the application of anisotropic mesh adaptation in image processing. Examples demonstrate that anisotropic mesh adaptation can significantly improve computational efficiency while still providing good quality processing. Metric tensors M_{adap} (3.29) based on interpolation error estimate and M_{DMP} (2.17) based on DMP satisfaction for elliptic problem (1.1) are applied in image processing and provide results comparable to published literature. More research is needed to investigate DMP satisfaction for parabolic problems to avoid artifacts in the final image.

6.2 Future research topics

Violating discrete maximum principle (DMP) during computations may cause severe consequences. For instance, in laboratory plasmas, the ratio of parallel and perpendicular conduction coefficients is very high. Improper schemes may produce spurious solutions (such as negative temperature) which may lead to imaginary sound speed. In image processing, artifacts may occur if DMP is not satisfied.

In this dissertation, DMP is considered only for the particular form of elliptic problem (1.1) in Chapters 2 and 3. It is desirable to study the DMP satisfaction for a general elliptic problem in the form of

$$-\nabla\cdot(\mathbb{D}(\boldsymbol{x},u,\nabla u)\nabla u)+\boldsymbol{b}(\boldsymbol{x})\nabla u+c(\boldsymbol{x})u=f, \qquad \text{in} \quad \Omega, \tag{6.1}$$

and for parabolic problems in the form of

$$u_t-\nabla\cdot(\mathbb{D}(\boldsymbol{x},u,\nabla u)\nabla u)+\boldsymbol{b}(\boldsymbol{x})\nabla u+c(\boldsymbol{x})u=f, \qquad \text{in} \quad \Omega. \tag{6.2}$$

Some research has been done on DMP of particular parabolic problems, which utilizes the concept of M-matrices and provides sufficient conditions which are relatively restrictive. It is worthwhile to extend the anisotropic non-obtuse angle condition and to provide a more relaxed sufficient condition for general elliptic and parabolic problems.

For variational problems, more studies are needed to demonstrate the advantages of the approach over interpolation error based methods or other methods without using variational structures.

Finally, it is desirable to apply schemes that satisfies DMP to real world problems arising from science and engineering. The application to image processing is discussed in Chapter 5. However, only the metric tensor based on steady state problems is used for mesh adaptation, which does not guarantee DMP satisfaction for time dependent problem. Once the condition for DMP satisfaction for parabolic problems is developed, it can be applied to the PDE-based image processing techniques in order to prevent occurrence of artifacts during the computation. Applications to other diffusion problems, for instance, flow in porous media, can also be considered.

Bibliography

[1] I. Aavatsmark, T. Barkve, Ø. Bøe, and T. Mannseth. Discretization on unstructured grids for inhomogeneous, anisotropic media. I. Derivation of the methods. *SIAM J. Sci. Comput.*, 19:1700–1716 (electronic), 1998. Cited on 3, 25

[2] I. Aavatsmark, T. Barkve, Ø. Bøe, and T. Mannseth. Discretization on unstructured grids for inhomogeneous, anisotropic media. II. Discussion and numerical results. *SIAM J. Sci. Comput.*, 19:1717–1736 (electronic), 1998. Cited on 3, 25

[3] M. Ainsworth and J.T. Oden. *A posteriori error estimation in finite element analysis*. John Wiley & Sons, New York, 2000. Pure and Applied Mathematics. Cited on 80, 81

[4] D. Ait-Ali-Yahia, G. Baruzzi, W.G. Habashi, M. Fortin, J. Dompierre, and M.G. Vallet. Anisotropic mesh adaptation: towards user-independent, mesh-independent and solver-independent cfd. part ii: structured grids. *Int. J. Numer. Meth. Fluids*, 39:657 – 673, 2002. Cited on 17

[5] T. Apel. *Anisotropic finite elements: local estimates and applications*. B. Gl Teubner Stuttgart, Leipzig, 1999. Cited on 16

[6] T. Apel and M. Dobrowolski. Anisotropic interpolation with applications to the finite element method. *Computing*, 47:277 – 293, 1992. Cited on 16

[7] G. Aubert and L. Vese. A variational method in image recovery. *SIAM J. Numer. Aanl.*, 34:1948 – 1979, 1997. Cited on 93

[8] W. Bangerth and R. Rannacher. *Adaptive Finite Element Methods for Differential Equations*. Birkhauser Verlag, Basel, 2003. Lectures in Mathematics. Cited on 17

[9] R.E. Bank and R.K. Smith. A posteriori error estimates based on hierarchical bases. *SIAM J. Numer. Anal.*, 30:921–935, 1993. Cited on 82

[10] E. Bänsch and K. Mikula. A coarsening finite element strategy in image selective smoothing. *Comput. Visual. Sci*, 1:53–61, 1997. Cited on 13, 113, 115, 117

[11] E. Bänsch and K. Mikula. Adaptivity in 3d image processing. *Comput. Visual. Sci*, 4:21–30, 2001. Cited on 13, 113

[12] C. Bazán and P. Blomgren. Adaptive finite element method for image processing. *In Proceedings of COMSOL Multiphysics Conference*, pages 337–381, Boston, Massachusetts, 2005. Cited on 13, 113, 115, 117

[13] R. Becker and R. Rannacher. A feed-back approach to error control in finite element methods: basic analysis and examples. *East-West J. Numer. Math.*, 4:237 – 264, 1996. Cited on 14, 15, 96

[14] R. Becker and R. Rannacher. An optimal control approach to a posteriori error estimation in finite element methods. *Acta Numer.*, 10:1 – 102, 2001. Cited on 14, 15, 96

[15] M. Berndt, T.A. Manteuffel, and S.F. Mccormick. Local error estimates and adaptive refinement for first-order system least squares (fosls). *Electr. Trans. Numer. Anal.*, 6:35 – 43, 1997. Cited on 14, 16, 96

[16] D.L. Bihan, J. Mangin, C. Poupon, C.A. Clark, S. Pappata, N. Molko, and H. Chabriat. Diffusion tensor imaging: concepts and applications. *Journal of Magnetic Resonance Imaging*, 13:534 – 546, 2001. Cited on 13, 111

[17] M. Bildhauer. *Convex Variational Problems: Linear, Nearly Linear and Anisotropic Growth Conditions*. Springer, New York, 2003. Cited on 92

[18] P. Bochev and R.B. Lehoucq. On the finite element solution of the pure neumann problem. *SIAM Review*, 47:50 – 66, 2005. Cited on 14, 15, 96

[19] H. Borouchaki, P.L. George, P. Hecht, P. Laug, and E. Saletl. Delaunay mesh generation governed by metric specification: Part I. Algorithms. *Finite Elements in Analysis and Design*, 25:61–83, 1997. Cited on 17, 18, 71

[20] H. Borouchaki, P.L. George, and B. Mohammadi. Delaunay mesh generation governed by metric specification: Part II. applications. *Finite Element in Analysis and Design*, 25:85 – 109, 1997. Cited on 17, 18, 71

[21] F.J. Bossen and P.S. Heckbert. A pliant method for anisotropic mesh generation. In *Proceedings, 5th International Meshing roundtable*, pages 63 – 74, Albuquerque, NM, 1996. Sandia National Laboratories. Sandia Report 96-2301. Cited on 16, 17

[22] J. Brandts, S. Korotov, and M. Křížek. Dissection of the path-simplex in $\mathbb{R}^n$ into n path-subsimplices. *Linear Algebra and its Applications*, 421:382–393, 2007. Cited on 30, 33

[23] J. Brandts, S. Korotov, and M. Křížek. The discrete maximum principle for linear simplicial finite element approximations of a reaction-diffusion problem. *Linear Algebra and its Applications*, 429:2344–2357, 2008. Cited on 24, 25

[24] J. Brandts, S. Korotov, M. Křížek, and J. Šolc. On nonobtuse simplicial partitions. *SIAM Rev.*, 51:317–335, 2009. Cited on 44

[25] S. Brenner and R. Scott. *The Mathematical Theory of Finite Element Methods, 3rd Edition*. Springer, 2007. Cited on 19

[26] C.J. Budd, W. Huang, , and R.D. Russell. Adaptivity with moving grids. *Acta Numerica*, 18:111–241, 2009. Cited on 17

[27] E. Burman and A. Ern. Discrete maximum principle for galerkin approximations of the laplace operator on arbitrary meshes. *C. R. Acad. Sci. Paris*, Ser.I 338:641–646, 2004. Cited on 24, 25

[28] Z. Cai, R. Lazarov, T.A. Manteuffel, and S.F. Mccormick. First-order system least squares for second-order partial differential equations: part i. *SIAM J. Numer. Anal.*, 31:1785 – 1799, 1994. Cited on 14, 16, 96

[29] Z. Cai, T.A. Manteuffel, and S.F. Mccormick. First-order system least squares for second-order partial differential equations: part ii. *SIAM J. Numer. Anal.*, 34:425 – 454, 1997. Cited on 16

[30] Z. Cai and G. Starke. First-order system least squares for the stress-displacement formulation: linear elasticity. *SIAM J. Numer. Anal.*, 41:715 – 730, 2003. Cited on 14, 16, 96

[31] Z. Cai and G. Starke. Least-squares methods for linear elasticity. *SIAM J. Numer. Anal.*, 42:826 – 842, 2004. Cited on 14, 16, 96

[32] G. Carey. *Computational Grids: Generation, Adaptation, and Solution Strategies*. Taylor & Francis, Washington, DC, 1997. Cited on 17

[33] G.F. Carey and A. Pehlivanov. Local error estimation and adaptive remeshing scheme for least-squares mixed finite elements. *Comput. Methods Appl. Mech. Engrg.*, 150:125 – 131, 1997. Cited on 14, 16, 96

[34] M.J. Castro-Díaz, F. Hecht, B. Mohammadi, and O. Pironneau. Anisotropic unstructured mesh adaption for flow simulations. *Int. J. Numer. Meth. Fluids*, 25:475–491, 1997. Cited on 16, 17, 18, 71

[35] F. Catté, P.L. Lions, J.M. Morel, and T. Coll. Image selective smoothing and edge detection by nonlinear diffusion. *SIAM J. Numer. Anal.*, 29:182–193, 1992. Cited on 113

[36] T.F. Chan and J. Shen. Non-texture inpainting by curvature driven diffusions (cdd). *J. Vis. Commun. Image Rep*, 12:436–449, 2000. Cited on 4

[37] T.F. Chan, J. Shen, and L. Vese. Variational pde models in image processing. *Not. AMS J.*, 50:14–26, 2003. Cited on 4, 93

[38] P.G. Ciarlet. Discrete maximum principle for finite difference operators. *Aequationes Math.*, 4:338–352, 1970. Cited on 23

[39] P.G. Ciarlet and P.-A. Raviart. Maximum principle and uniform convergence for the finite element method. *Comput. Methods Appl. Mech. Engrg.*, 2:17–31, 1973. Cited on 3, 23, 24, 25, 33

[40] G. Cohen, P. Joly, J.E. Roberts, and N. Tordjman. Higher order triangular finite elements with mass lumping for the wave equation. *SIAM J. Numer. Anal.*, 201:421–438, 2004. Cited on 116

[41] C. Collins, D. Kinderlehrer, and M. Luskin. Numerical approximation of the solution of variational problem with a double well potential. *SIAM J. Numer. Anal.*, 28:321 – 332, April 1991. Cited on 14, 15, 96

[42] C. Collins and M.Luskin. Optimal order estimates for the finite element approximation of the solution of a nonconvex variational problem. *Math. Comp*, 57:621 – 637, 1991. Cited on 14, 15, 96

[43] P.I. Crumpton, G.J. Shaw, and A.F. Ware. Discretisation and multigrid solution of elliptic equations with mixed derivative terms and strongly discontinuous coefficients. *J. Comput. Phys.*, 116:343–358, 1995. Cited on 3, 25

[44] E.F. D'Azevedo. Optimal triangular mesh generation by coordinate transformation. *SIAM J. Sci. Stat. Comput.*, 12:755 – 786, 1991. Cited on 18

[45] E.F. D'Azevedo and R.B. Simpson. On optimal triangular meshes for minimizing the gradient error. *Numer. Math.*, 59:321 – 348, 1991. Cited on 16, 18

[46] Paula de Oliveira. Existence de maillages optimaux dans les méthodes d'éléments finis. *RAIRO Anal. Numer.*, 14:279 – 290, 1980. Cited on 14, 96

[47] M. Delfour, G. Payre, and J.P. Zolésio. An optimal triangulation for second-order elliptic problems. *Comput. Meth. Appl. Mech. Engrg.*, 50:231 – 261, 1985. Cited on 15

[48] P. Deuflhard, P. Leinen, and H. Yserentant. Concepts of an adaptive hierarchical finite element code. *Impact Comput. Sci. Engrg.*, 1:3 – 35, 1989. Cited on 82

[49] J. Dompierre, M.G. Vallet, Y. Bourgault, M. Fortin, and W.G. Habashi. Anisotropic mesh adaptation: towards user-independent, mesh-independent and solver-independent CFD. Part III: Unstructured meshes. *Int. J. Numer. Meth. Fluids*, 39:675 – 702, 2002. Cited on 16, 17

[50] A. Drăgănescu, T.F. Dupont, and L.R. Scott. Failure of the discrete maximum principle for an elliptic finite element problem. *Math. Comp.*, 74(249):1–23, 2004. Cited on 25, 32

[51] C. Ebmeyer and J. Vogelgesang. Finite element approximation of a forward and backward anisotropic diffusion model in image denoising and form generalization. *Numer. Methods Partial Differential Equations*, 24:646–662, 2008. Cited on 13, 111, 113

[52] G.T. Eigestad, I. Aavatsmark, and M. Espedal. Symmetry and M-matrix issues for the O-method on an unstructured grid. *Comput. Geosci.*, 6:381–404, 2002. Cited on 33

[53] C.M. Elliott, B.Gawron, S. Maier-paape, and E.S. Van Vleck. Discrete dynamics for convex and non-convex smoothing functionals in pde based image restoration. *Communications on Pure and Applied Analysis*, 5:181–200, 2006. Cited on 111

[54] A. Ern and J.L. Guermond. *Theory and practice of finite elements*. Sprigner-Verlag New York, LLC, 2004. Cited on 19, 21

[55] T. Ertekin, J.H. Abou-Kassem, and G.R. King. *Basic applied reservoir simulation*. SPE textbook series, Vol. 7, Richardson, Texas, 2001. Cited on 3

[56] L.C. Evans. *Partial differential equations*. American Mathematical Society, Providence, Rhode Island, 1998. Graduate Studies in Mathematics, Volume 19. Cited on 71

[57] G.C. Fehmers and C. Hocker. Fast structure interpretation with structure-oriented filtering. *Geophysics*, 68:1286–1293, 2003. Cited on 111

[58] C.A. Felippa. Optimization of finite element grids by direct energy search. *Appl. Math. Modeling*, 1:93 – 96, 1976. Cited on 14, 96

[59] C.A. Felippa. Numerical experiments in finite element grid optimization by direct energy search. *Appl. Math. Modeling*, 1:239 – 244, 1977. Cited on 14, 96

[60] D.A. French. On the convergence of finite-element approximations of a relaxed variational problem. *SIAM J. Numer. Anal.*, 27:419 – 436, 1990. Cited on 14, 15, 96

[61] P. Frey and P.L. George. *Mesh Generation: Application to Finite Elements*. Hermes Science, Oxford and Paris, 2000. Cited on 18, 71

[62] S. Fu, Q. Ruan, W. Wang, and Y. Li. Adaptive anisotropic diffusion for ultrasonic image denoising and edge enhancement. *International Journal of Information Technology*, 2:2 – 4, 2006. Cited on 111

[63] R.V. Garimella and M.S. Shephard. Boundary layer meshing for viscous flows in complex domain. In *Proceedings, 7th International Meshing Roundtable*, pages 107–118, Albuquerque, NM, 1998. Sandia National Laboratories. Cited on 17

[64] P.L. George. Automatic mesh generation and finite element computation. In P.G. Ciarlet and J.L. Lions, editors, *Handbook of Numerical Analysis, Vol. IV*, pages 69 – 190, New York, 1996. Elsevier Science B.V. Cited on 17

[65] R.C. Gonzalez and R.E. Woods. *Digital image processing*. Pearson Education. Inc., New Jersey, 2008. third edition. Cited on 11

[66] P.A. Gremaud. Numerical analysis of a nonconvex variational problem related to solid-solid phase transitions. *SIAM J. Numer. Anal.*, 31:111 – 127, 1994. Cited on 14, 15, 96

[67] S. Günter and K. Lackner. A mixed implicit-explicit finite difference scheme for heat transport in magnetised plasmas. *J. Comput. Phys.*, 228:282–293, 2009. Cited on 3, 7, 25

[68] S. Günter, K. Lackner, and C. Tichmann. Finite element and higher order difference formulations for modelling heat transport in magnetised plasmas. *J. Comput. Phys.*, 226:2306–2316, 2007. Cited on 3, 7, 25

[69] S. Günter, Q. Yu, J. Kruger, and K. Lackner. Modelling of heat transport in magnetised plasmas using non-aligned coordinates. *J. Comput. Phys.*, 209:354–370, 2005. Cited on 3, 4, 7, 25

[70] W.G. Habashi, J. Dompierre, Y. Bourgault, D. Ait-Ali-Yahia, M. Fortin, and M.G. Vallet. Anisotropic mesh adaptation: towards user-independent, mesh-independent and solver-independent CFD. Part I: General principles. *Int. J. Numer. Meth. Fluids*, 32:725–744, 2000. Cited on 17

[71] F. Hecht. Bidimensional anisotropic mesh generator. Technical report, INRIA, Rocquencourt, 1997. Source code: http://www.ann.jussieu.fr/ hecht/ftp/bamg/. Cited on 17, 18

[72] W. Huang. Variational mesh adaptation: isotropy and equidistribution. *J. Comput. Phys.*, 174:903 – 924, 2001. Cited on 36

[73] W. Huang. Measuring mesh qualities and application to variational mesh adaptation. *SIAM J. Sci. Comput.*, 26:1643 – 1666, 2005. Cited on 16, 76

[74] W. Huang. Metric tensors for anisotropic mesh generation. *J. Comput. Phys.*, 204:633–665, 2005. Cited on 16, 18, 36, 54, 55, 58, 71, 74, 79, 84, 88, 89, 113

[75] W. Huang. Mathematical principles of anisotropic mesh adaptation. *Comm. Comput. Phys.*, 1:276 – 310, 2006. Cited on 17, 18, 19, 36, 49

[76] W. Huang. Discrete maximum principle and a delaunay-type mesh condition for linear finite element approximations of two-dimensional anisotropic diffusion problems. *Numerical Mathematics: Theory, Methods and Applications*, 4:319 – 334, 2011. Cited on 43

[77] W. Huang, L. Kamenski, and J. Lang. A new anisotropic mesh adaptation method based upon hierarchical a posteriori error estimates. *J. Comput. Phys.*, 229:2179–2198, 2010. Cited on 55, 82, 83, 97

[78] W. Huang, L. Kamenski, and X. Li. Anisotropic mesh adaptation for variational problems using error estimation based on hierarchical bases. *Canadian Applied Mathematics Quarterly*, Special issue for the 30th anniversary of CAIMS, accepted. Cited on 27

[79] W. Huang and X. Li. An anisotropic mesh adaptation method for the finite element solution of variational problems. *Fin. Elem. Anal. Des.*, 46:61–73, 2010. Cited on 27

[80] W. Huang and R.D. Russell. *Adaptive Moving Mesh Methods*. Springer, New York, 2011. Cited on 17, 18, 36

[81] B.N. Jiang and G.F. Carey. Adaptive refinement of least-squares finite elements with element-by-elment conjugate gradient solution. *Int. J. Numer. Meth. Engr.*, 24:569 – 580, 1987. Cited on 14, 16, 96

[82] C. Johnson. *Numerical Solution of Partial Differential Equations by the Finite Element Method*. Cambridge University Press, 1988. Cited on 19

[83] S. Jund and S. Salmon. Arbitrary high-order finite element schemes and high-order mass lumping. *Int. J. Appl. Math. Comput. Sci.*, 17:375–393, 2007. Cited on 116

[84] J. Karátson and S. Korotov. Discrete maximum principles for finite element solutions of nonlinear elliptic problems with mixed boundary conditions. *Numer. Math.*, 99:669–698, 2005. Cited on 24, 25, 43

[85] J. Karátson and S. Korotov. Discrete maximum principles for finite element solutions of some mixed nonlinear elliptic problems using quadratures. *J. Comput. Appl. Math.*, 192:75–88, 2006. Cited on 24, 25, 43

[86] J. Karátson, S. Korotov, and M. Křížek. On discrete maximum principles for nonlinear elliptic problems. *Mathematics and Computers in Simulation*, 76:99–108, 2007. Cited on 24, 25, 43

[87] D.A. Karras and G.B. Mertzios. New pde-based methods for image enhancement using som and bayesian inference in various discretization schemes. *Meas. Sci. Technol.*, 20:104012, 2009. Cited on 4

[88] G. Kunert. A posteriori error estimation for anisotropic tetrahedral and triangular finite element meshes. Technical report, TU Chemnitz, 1999. Ph. D. Thesis. Cited on 16

[89] D. Kuzmin, M.J. Shashkov, and D. Svyatskiy. A constrained finite element method satisfying the discrete maximum principle for anisotropic diffusion problems. *J. Comput. Phys.*, 228:3448–3463, 2009. Cited on 25, 38, 41, 60

[90] M. Křížek and Q. Lin. On diagonal dominance of stiffness matrices in 3d. *East-West J. Numer. Math.*, 3:59–69, 1995. Cited on 24, 25, 30, 43

[91] F.W. Letniowski. Three-dimensional delaunay triangulations for finite element approximations to a second-order diffusion operator. *SIAM J. Sci. Stat. Comput.*, 13:765–770, 1992. Cited on 24, 25, 44

[92] X. Li and W. Huang. An anisotropic mesh adaptation method for the finite element solution of heterogeneous anisotropic diffusion problems. *J. Comput. Phys.*, 229:8072–8094, 2010. Cited on 27

[93] X. Li, D. Svyatskiy, and M. Shashkov. Mesh adaptation and discrete maximum principle for 2D anisotropic diffusion problems. *LANL technical report*, LA-UR 10-01227, 2007. Cited on 25, 27, 34, 36, 38

[94] K. Lipnikov, M. Shashkov, D. Svyatskiy, and Y. Vassilevski. Monotone finite volume schemes for diffusion equations on unstructured triangular and shape-regular polygonal meshes. *J. Comput. Phys.*, 227:492–512, 2007. Cited on 25

[95] R. Liska and M. Shashkov. Enforcing the discrete maximum principle for linear finite element solutions of second-order elliptic problems. *Communications in Computational Physics*, 3:852–877, 2008. Cited on 25

[96] K. Miyamoto. *Plasma Physics and Controlled Nuclear Fusion*. Springer-Verlag Berlin Heidelberg, New York, 2005. Cited on 6, 8

[97] M.J. Mlacnik and L.J. Durlofsky. Unstructured grid optimization for improved monotonicity of discrete solutions of elliptic equations with highly anisotropic coefficients. *J. Comput. Phys.*, 216:337–361, 2006. Cited on 3, 25, 36

[98] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Commun. Pure Appl. Math*, 42:577–685, 1989. Cited on 4

[99] K. Nishikawa and M. Wakatani. *Plasma Physics*. Springer-Verlag Berlin Heidelberg, New York, 2000. Cited on 3, 4

[100] I.J. Parrish and J.M. Stone. Nonlinear evolution of the magnetothermal instability in two dimensions. *Astrophys. J.*, 633:334–348, 2005. Cited on 4

[101] J. Peraire, M. Vahdati, K. Morgan, and O.C. Zienkiewicz. Adaptive remeshing for compressible flow computations. *J. Comput. Phys.*, 72:449–466, 1987. Cited on 17

[102] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12:629–639, 1990. Cited on 4, 12, 111

[103] C. Le Potier. Schéma volumes finis monotone pour des opérateurs de diffusion fortement anisotropes sur des maillages de triangles non structurés. *C. R. Math. Acad. Sci. Paris*, 341:787–792, 2005. Cited on 25

[104] C. Le Potier. A nonlinear finite volume scheme satisfying maximum and minimum principles for diffusion operators. *Int. J. Finite Vol.*, 6:20, 2009. Cited on 25

[105] C. Le Potier. Un schéma linéaire vérifiant le principe du maximum pour des opérateurs de diffusion très anisotropes sur des maillages déformés. *C. R. Math. Acad. Sci. Paris*, 347:105–110, 2009. Cited on 25

[106] T. Preusser and M. Rumpf. An adaptive finite element method for large scale image processing. *Journal of Visual Communication and Image Representation*, 11:183–195, 2000. Cited on 13, 113, 117

[107] W. Rachowicz. An anisotropic *h*-type mesh refinement strategy. *Comput. Meth. Appl. Mech. Engrg.*, 109:169–181, 1993. Cited on 16

[108] J.N. Reddy. *An Introduction to the Finite Element Method, Third edition*. Tata McGraw-Hill Publishing Company Limited, New York, reprint 2008. Cited on 19

[109] J. Remacle, X. Li, M.S. Shephard, and J.E. Flaherty. Anisotropic adaptive simulation of transient flows using discontinuous Galerkin methods. *Int. J. Numer. Meth. Engr.*, 62:899–923, 2003. Cited on 17

[110] P. Sharma and G.W. Hammett. Preserving monotonicity in anisotropic diffusion. *Journal of Computational Physics*, 227:123–142, 2007. Cited on 3, 4, 7, 25

[111] K.G. Siebert. An a posteriori error estimator for anisotropic refinement. *Numer. Math.*, 73:373 – 398, 1996. Cited on 16

[112] D.M.Y. Sommerville. *An introduction to the geometry of n dimensions*. Methuen & Co. LTD., London, 1929. Cited on 29

[113] T.H. Stix. *Waves in Plasmas*. Amer. Inst. of Phys., New York, 1992. Cited on 3, 4

[114] G. Stoyan. On a maximum principle for matrices, and on conservation of monotonicity. With applications to discretization methods. *Z. Angew. Math. Mech.*, 62:375–381, 1982. Cited on 23

[115] G. Stoyan. On maximum principles for monotone matrices. *Linear algebra and its applications*, 78:147–161, 1986. Cited on 23, 24

[116] G. Strang and G.J. Fix. *An Analysis of the Finite Element Method*. Prentice Hall, Englewood Cliffs, NJ, 1973. Cited on 24, 25

[117] Q. Sun, J.A. Hossack, J. Tang, and S.T. Acton. Speckle reducing anisotropic diffusion for 3d ultrasound images. *Computerized Medical Imaging and Graphics*, 28:461–470, 2004. Cited on 111

[118] J.F. Thompson, B.K. Soni, and N.P. Weatherill. *Handbook of Grid Generation*. CRC Press, Boca Ration and London, 1999. Cited on 17

[119] Y. Tourigny and F. Hülsemann. A new moving mesh algorithm for the finite element solution of variational problems. *SIAM J. Numer. Anal.*, 35:1416 – 1438, 1998. Cited on 14, 15, 96

[120] R.S. Varga. *Matrix iterative analysis*. Prentice-Hall, New Jersey, 1962. Cited on 24, 32

[121] R.S. Varga. On a discrete maximum principle. *SIAM J. Numer. Anal.*, 3:355–359, 1966. Cited on 23

[122] Pavel Šolín. *Partial Differential Equations and the Finite Element Method*. Wiley-Interscience, 2005. Cited on 19

[123] J. Weickert. *Anisotropic diffusion in image processing*. Teubner-Verlag, Stuttgart, Germany, 1998. Cited on 4, 12, 111

[124] J. Weickert, B. Romeny, and M. Viergever. Efficient and reliable schemes for nonlinear diffusion filtering. *IEEE Transactions on Image Processing*, 7:398–410, 1998. Cited on 13

[125] J. Xu and L. Zikatanov. A monotone finite element scheme for convection-diffusion equations. *Mathematics of Computation*, 69:1429–1446, 1999. Cited on 24

[126] S. Yamakawa and K. Shimada. High quality anisotropic tetrahedral mesh generation via ellipsoidal bubble packing. In *Proceedings, 9th International Meshing Roundtable*, pages 263–273, Albuquerque, NM, 2000. Sandia National Laboratories. Sandia Report 2000-2207. Cited on 17