# ENRICHED FINITE ELEMENTS METHOD FOR CONVECTION-DIFFUSION-REACTION PROBLEMS

**A Thesis Submitted to**
**the Graduate School of Engineering and Sciences of**
**İzmir Institute of Technology**
**in Partial Fulfillment of the Requirements for the Degree of**

**DOCTOR OF PHILOSOPHY**

**in Mathematics**

**by**
**Ali ŞENDUR**

**November 2012**
**İZMİR**

We approve the thesis of **Ali ŞENDUR**

**Examining Committee Members:**

_____
**Prof. Dr. Oktay PASHAEV**
Department of Mathematics, İzmir Institute of Technology


_____
**Prof. Dr. Turgut ÖZİŞ**
Department of Mathematics, Ege University


_____
**Prof. Dr. Gökmen TAYFUR**
Department of Civil Engineering, İzmir Institute of Technology


_____
**Assoc. Prof. Dr. Gamze TANOĞLU**
Department of Mathematics, İzmir Institute of Technology


_____
**Assoc. Prof. Dr. Burhan PEKTAŞ**
Department of Mathematics and Computer Sciences, İzmir University


**22 November 2012**


_____ | _____
**Prof. Dr. Oktay PASHAEV**
Supervisor, Department of
Mathematics
İzmir Institute of Technology

**Prof. Dr. Ali İhsan NESLİTÜRK**
Co-Supervisor, Department of
Mathematics
İzmir Institute of Technology


_____ | _____
**Prof. Dr. Oğuz YILMAZ**
Head of the Department of
Mathematics

**Prof. Dr. R. Tuğrul SENGER**
Dean of the Graduate School of
Engineering and Sciences

*To My wife: Esra*

# ACKNOWLEDGMENTS

# ABSTRACT

### ENRICHED FINITE ELEMENTS METHOD FOR
### CONVECTION-DIFFUSION-REACTION PROBLEMS

In this thesis, we consider stabilization techniques for linear convection-diffusion-reaction (CDR) problems. The survey begins with two stabilization techniques: streamline upwind Petrov-Galerkin method (SUPG) and Residual-free bubbles method (RFB). We briefly recall the general ideas behind them, trying to underline their potentials and limitations.

Next, we propose a stabilization technique for one-dimensional CDR problems based on the RFB method and particularly designed to treat the most interesting case of small diffusion. We replace the RFB functions by their cheap, yet efficient approximations which retain the same qualitative behavior. The approximate bubbles are computed on a suitable sub-grid, the choice of whose nodes are critical and determined by minimizing the residual of a local problem. The resulting numerical method has similar stability features with the RFB method for the whole range of problem parameters. We also note that the location of the sub-grid nodes suggested by the strategy herein coincides with the one described by Brezzi and his coworkers.

Next, the approach in one-dimensional case is extended to two-dimensional CDR problems. Based on the numerical experiences gained with this work, the pseudo RFBs retain the stability features of RFBs for the whole range of problem parameters.

Finally, a numerical scheme for one-dimensional time-dependent CDR problem is studied. A numerical approximation with the Crank-Nicolson operator for time and a recent method suggested by Neslitürk and his coworkers for the space discretization is constructed. Numerical results confirm the good performance of the method.

# ÖZET

## KONVEKSİYON-DİFÜZYON-REAKSİYON PROBLEMLERİ İÇİN ZENGİNLEŞTİRİLMİŞ SONLU ELEMANLAR YÖNTEMİ

Bu çalışmada difüzyon katsayısı küçük bir parametre olan konveksiyon-difüzyon reaksiyon (KDR) probleminin sayısal çözümleri için dengelenmiş sonlu elemanlar yöntemleri ele alınmaktadır. Konveksiyonun baskın olduğu durumlarda KDR denkleminin tabaka davranışı sergilediği ve standart sonlu elemanlar çözümünün, problemin tam çözümüne aykırı olarak yüksek devinimli salınımlar gösterdiği bilinen bir gerçektir. Bu nedenle, problemin fiziksel gerçekliği ile uyumlu sayısal çözümler üreten algoritmalar geliştirmek önemlidir. Çalışmamıza klasik dengeleme tekniklerinden Streamline-Upwind Petrov Galerkin (SUPG) ve Residual free bubbles (RFB) metotlarını incelemekle başlayacağız.

Daha sonra difüzyon katsayısı küçük bir parametre olan KDR probleminin bir boyutlu çözümleri için temeli RFB metoduna dayanan dengelenmiş sonlu elemanlar yöntemleri ele alınmaktadır. Ancak RFB metodunun aksine lokal türev denkleminin çözümünün hesaplanması, her eleman içinde uygun bir ağ tanıtılarak sayısal olarak yapılmaktadır. Böyle bir ağı inşa etmek için lokaldeki problemin kalıntısının $L_1$ normunu en aza indirgeyeceğiz. Ortaya çıkan yaklaşık çözüm, tam çözüme benzer bir davranış sergilemektedir. Sonuç olarak elde edilen bu ağ Brezzi ve çalışma arkadaşlarının ortaya koyduğu ağ ile örtüşmektedir.

Çalışmamıza difüzyon katsayısı küçük bir parametre olan KDR probleminin iki boyutlu çözümleri için dengelenmiş sonlu elemanlar yöntemleri ile devam edeceğiz. Bu çalışmada, üçgen elemanlar için lokal türev denkleminin çözümünü sayısal olarak hesaplayabilen bir ağ inşa eden bir yaklaşım ortaya koyacağız. Ortaya çıkan yaklaşık çözüm tam çözüme benzer bir davranış sergilemektedir.

Son olarak durağan olmayan KDR problemi için sayısal çözümler üreten algoritma geliştireceğiz. Zamandaki ayrıklaştırma için Crank- Nicolson operatörü kullanılırken uzaysal ayrıklaştırma için bir boyutta elde edilen sayısal yöntemi kullanacağız. Metodun etkinliği sayısal deneylerle gösterilmektedir.

# TABLE OF CONTENTS

# LIST OF FIGURES

# CHAPTER 1

# INTRODUCTION

Throughout this thesis, our main interest is in linear elliptic operators whose second-order derivatives are multiplied by some parameter $\epsilon$ that is allowed to be close to zero. These derivatives model diffusion while first-order derivatives are associated with transport processes. When $\epsilon$ is near zero and the elliptic differential operator has convective terms, it is called a convection-diffusion operator. Prof. Martin Stynes states that (Stynes, 2005): "Such operators, while still satisfying the definition of ellipticity, live dangerously by flirting with the non-elliptic world." Their convective terms have a major influence on both the theoretical and numerical solution of the problem and can not be ignored. We shall see that the true solution of convection-diffusion-reaction (CDR) problems may exhibit small subregions where the derivatives of the solution are very large, when some problem parameters are too big compared to others. We describe such behaviour by saying that the solution has a layer.

Mathematical models that involve a combination of convective and diffusive patterns are among the most widespread in all of science, engineering and other fields where mathematical modeling is important. Water quality problems, convective heat transfer problems, simulation of the semiconductor devices can be given as an example of these models. Also the linearization of the Navier-Stokes equation and drift-diffusion equation of semiconductor device modeling are important instances. In two dimensions, the CDR problem takes the following form,

$$
\begin{aligned}
\mathcal{L}u = -\epsilon\Delta u + \beta \cdot \nabla u + \sigma u &= f \quad \text{in } \Omega \subset \mathbb{R}^2\,, \\
u &= g \quad \text{on } \partial\Omega\,,
\end{aligned} \tag{1.1}
$$

where $0 < \epsilon \ll 1$, and the functions $\beta$, $\sigma$ and $f$ are assumed to be continuous on $\bar{\Omega}$. We also assume that $\sigma > 0$ on $\bar{\Omega}$. Here $\Omega$ is any bounded domain in $\mathbb{R}^2$. Moreover, the boundary $\partial\Omega$ can be divided into 3 parts:

$$
\text{inflow boundary } \partial^-\Omega = \{x \in \partial\Omega : \beta \cdot n < 0\}\,,
$$

$$\text{outflow boundary } \partial^+\Omega = \{\partial\Omega : \beta \cdot n > 0\}\,,$$
$$\text{characteristic flow boundary } \partial^0\Omega = \{\partial\Omega : \beta \cdot n = 0\}\,,$$

where $n$ is the outward-pointing unit normal to $\partial\Omega$.

Towards this end, we begin with a brief description of the framework of convection-diffusion problems (i.e., $\sigma = 0$): the structure of their solutions will be analyzed, with special emphasis on the main phenomena layers. See (Stynes, 2005) for a more leisurely and detailed exposition of this material. A typical solution $u$ will have boundary layers, narrow regions close to $\partial\Omega$ where $|\nabla u|$ is large.

**Example 1.1** *For instance, consider the following convection-diffusion equation as a model problem:*

$$-0.01\Delta u + u_x = 1 \quad \text{in } \Omega = (0,1) \times (0,1)\,,$$
$$u = 0 \quad \text{on } \partial\Omega\,. \tag{1.2}$$

*In Figure 1.1 we plot the solution $u(x,y)$ to the problem (1.1). On most of $\Omega$ one has*
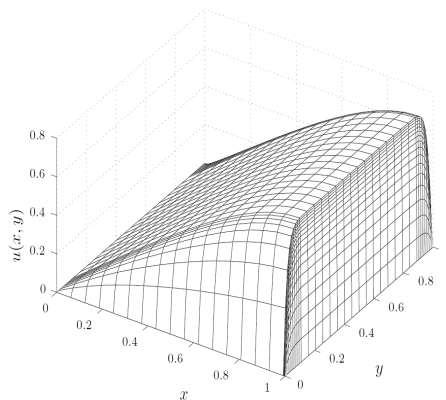


Figure 1.1. Exponential boundary layer with two characteristic boundary layers.

$u \approx u_0 = x$. The side $x = 1$ of $\bar{\Omega}$ is the outflow boundary $\partial^+\Omega$ and an exponential layer appears there. The tangential flow boundaries $y = 0$ and $y = 1$ have characteristic boundary layers that grow in strength as $x$ moves from 0 to 1 because of the increasing discrepancy between $u_0$ and the boundary condition.

Besides boundary layers, solutions of such problems can also have interior layers if there is a discontinuity in the boundary data on $\partial^-\Omega$.

**Example 1.2** *In Figure 1.2 we use the same differential operator as in Example 1.1, with $\epsilon = 10^{-6}$. A jump discontinuity has been introduced in the inflow boundary data:*

$$g(0,y) = \begin{cases} 1, & 0 \le y \le 1/2\,, \\ 0, & 1/2 < y \le 1\,. \end{cases}$$
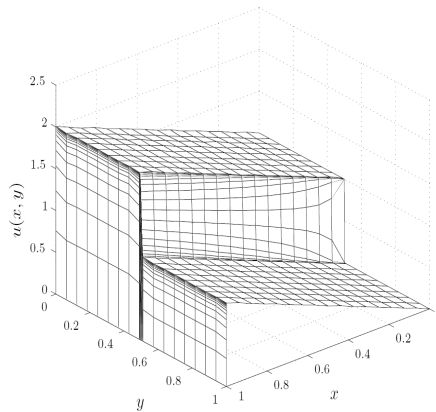


Figure 1.2. Solution of Example 1.2.

Typically in this model problem, but also in real fluid flow simulation, the numerical solution may produce unphysical oscillations that pollute the whole domain, while the exact solution shows boundary or internal layers. To overcome this difficulty, several approximate numerical analysis methods have evolved over the years; among them a commonly used method is the finite difference (Anderson et al., 1984) scheme. The finite difference model of a problem gives a pointwise approximation to the governing equations. This model is formed by writing difference equations for an array of grid points and improved as more nodes are used. With finite difference techniques we can treat some moderately difficult problems; but, for example, when we encounter irregular geometries or an unusual specification of boundary conditions, we find that finite difference techniques become inconvenient to use.

Unlike the finite difference method (FDM), finite element method (FEM) is a tool to obtain the approximate solution of the partial differential equations (PDE) by using a variational formulation; see, (Oden, 1991) for an introduction to FEM (see also (Ciarlet,

1978), (Hughes, 1987), (Johnson, 1987)). The variational formulation is an integral over the domain. By FEM, the domain is divided into elements, which are triangles in this study, and the PDE is approximated on each subdomain. As a consequence, if we employ a classical FEM with a discretization scale which is too big to resolve the layers, the solution that we obtain has in general large numerical oscillations spreading all over the domain, and can be completely unrelated to the exact solution (Stynes, 2005), (Morton, 1996). To properly resolve the layers, the discretization parameter must be of the same size of the ratio between diffusion and convection. In many problems, this choice would lead to a huge number of degrees of freedom making the finite element discretization computationally intractable. However, the advantages of geometrical flexibility, higher order accuracy and stability of the FEM encourage the use of FEMs in such problems. Moreover, finite element formulation is very flexible, i.e., it is easy to adapt changes in problem parameters, such as boundary conditions, physical properties, and geometric domains with little program modification.

A variety of finite elements approaches have been proposed to deal with the oscillation phenomenon in the numerical solution of the convection-diffusion equation; see, e.g., (Roos et al., 2008). Prof. K.W Morton states that (Morton, 1996): "Accurate modelling of the interaction between convective and diffusive processes is the most ubiquitous and challenging task in the numerical approximation of partial differential equations." One of the most successful classes of FEMs for treating convection-dominated problems is based on the stabilization formulations. The subject of stabilized FEMs has been studied for more than twenty years, and it is still attractive today; see, e.g., (Franca et al., 2006), (Franca and Frey, 1992) and many references cited therein. In the framework of finite element methods, a simple modification consists of adding a suitable amount of artificial diffusion. This idea was developed by Hughes and his co-workers (see; (Hughes et al., 1989), (Hughes and Mallet, 1986)). Their method, namely Streamline-Upwind/Petrov-Galerkin (SUPG) method, adds diffusion only in the streamline direction while preserving the consistency of the variational formulation. The SUPG method and its variants (such as the Galerkin Least-Squares method) have become the most popular numerical methods for this kind of problems among FEMs, which are based on augmenting the variational formulation by mesh-dependent terms in order to gain control over the derivatives of the solution (Brooks and Hughes, 1982), (Franca et al., 1992). The great advantage of this approach is not only its generality, but also its error analysis can be performed in many cases of interest. Despite the success of SUPG, the need for the proper choice of stabilizing parameter is considered as a major drawback of the method.

Motivated by this idea, intrinsically stable methods have been developed. Among them, in this thesis, we shall consider the Residual-Free Bubbles (RFB) method.

In particular, from an abstract standpoint, the RFB method is based on a local enrichment of the finite element space, so called bubble functions, without modification of the variational formulation; see (Brezzi et al., 1999), (Cangiani and Suli, 2005), (Franca et al., 2005a), (Franca et al., 2005b) and many references cited therein. The bubbles make the whole formulation intrinsically stable. The relationship between the use of bubble functions and stabilized methods was also studied in (Baiocchi et al., 1993), (Brezzi et al., 1992), (Brezzi and Russo, 1994). It turns out that, to find a more suitable value for the stabilizing parameter in the SUPG method, it is crucial to use special type of functions, so called the residual-free bubbles (RFB), defined by a local problem posed inside each element. The RFB method also allows one to prove error bounds (Brezzi et al., 2000), (Sangalli, 2000) and can be generalized to a much wider variety of problems (Brezzi et al., 1998b), (Franca and Russo, 1996). However it requires to solve a local differential equation which may not be easier than to solve the original one (Franca et al., 1998), (Franca and Tobiska, 2002). For the case described above, and if we employ continuous, piecewise-linear elements, this corresponds to solving, in each element $K$, the following boundary value problem (BVP):

$$
\begin{aligned}
-\epsilon \Delta b_K + \beta \cdot \nabla b_K &= 1 \quad \text{in } K, \\
b_K &= 0 \quad \text{on } \partial K.
\end{aligned}
\tag{1.3}
$$

For a practical implementation of the Residual-Free Bubble (RFB) formulation, the actual computation of the bubble functions is carried out numerically by introducing an appropriate subgrid. In recent times several authors tried to deal with problems of the type (1.3) by providing an approximate solution with the use of suitable subgrid node(s). This is the case of (Brezzi et al., 1998a), (Brezzi et al., 2005), where the subgrid is formed with a single internal node per triangle. From the incipit of RFB method, we know that this technique interprets the stabilization parameter as the mean value of the solution of a differential equation defined at the element level. Here we shall discuss the computation of the parameter by constructing convenient stabilizing subgrids on a triangular element in the framework of two-dimensional convection-diffusion problems.

The aim of the thesis is to develop stabilized FEMs for convection-diffusion-reaction (CDR) problems, particularly designed to treat the most interesting case of small

diffusion. We shall propose two new stabilized FEMs based on the RFB method. Unlike the RFB method, we shall present a smart cheaper way to compute approximately the solution of the bubble problem in each element in the hope to contribute a good stabilizing effect to the numerical method overall. The followings are description of the work presented in this thesis:

- In Chapter 2, we begin our journey by summarizing the general ideas behind the some of the most important stabilization methods to successfully solve the convection dominated convection-diffusion problems. In particular, we shall first review the Streamline Upwind/Petrov-Galerkin (SUPG). The underlying idea is to augment the variational formulation by mesh-dependent terms in order to gain control over the derivatives of the solution. Next, we give a review of residual-free bubble (RFB) method which is based on enriching the finite element spaces by suitable bubble functions. Finally, we describe two different approaches in which we discuss the construction of suitable sub-grid nodes to obtain efficient approximations to bubble functions.

- In Chapter 3, we devise a stabilization method for one-dimensional convection-diffusion reaction problems which is based on the RFB method and particularly designed to treat the most interesting case of small diffusion. We shall try to mimic the stabilizing effect of the RFBs by approximating them on a specially chosen sub-grid and using these approximations in place of RFB functions. This provides, as a consequence, a cheap, yet efficient approximations which retain the same qualitative behavior. The resulting numerical method has similar stability features to the RFB method for the whole range of problem parameters in which the transition from one regime to another is accurately captured by the algorithm. Numerical results confirming theoretical findings are presented for several examples. We shall show that the location of the sub-grid nodes suggested by the strategy herein coincides with the one in (Brezzi et al., 2003).

- In Chapter 4, we extend the ideas in Chapter 3 to obtain a stabilized FEM for two-dimensional convection-diffusion-reaction problems with a small diffusivity. We shall show that the location of the sub-grid nodes suggested by the strategy herein has a similar structure to the one suggested by Brezzi and his coworkers in a different framework for the convection-diffusion problems ( see (Brezzi et al., 2005)).

- Finally in Chapter 5, a stabilized finite element method for one-dimensional time-dependent convection- diffusion- reaction problems has been introduced and tested.

In particular, the discretization in time is performed by Crank-Nicolson finite difference scheme, for the space discretization we employ the method in Chapter 3.

# CHAPTER 2

# REVIEW OF STABILIZED FINITE ELEMENT METHODS

In this chapter, we begin by summarizing the general ideas behind the some of the most important stabilization methods, trying to underly their potentials and their limitations. In particular, we shall first review the Streamline-Upwind/Petrov-Galerkin (SUPG) and residual-free bubble (RFB) methods. Next, we discuss the explicit locations of sub-grid nodes on which we construct the pseudo bubble functions with two different approaches and describe the details of the numerical methods proposed. We will consider the following linear elliptic convection-diffusion problem in a polygonal domain $\Omega$:

$$
\begin{aligned}
-\epsilon \triangle u + \beta \cdot \nabla u &= f \quad \text{in } \Omega, \\
u &= 0 \quad \text{on } \partial\Omega.
\end{aligned}
\tag{2.1}
$$

Let $\mathcal{T}_h = \{K\}$ be an admissible (any two elements have disjoint closure, a vertex in common, or share a complete edge) triangulation of $\Omega$ into shape regular triangles $K$, and let $h_K = diam(K)$, $h = \max_K h_K$. We assume that the diffusion $\epsilon$ is a positive constant, and both the convection field $\beta$ and the right-hand side $f$ are piecewise constant with respect to the triangulation $\mathcal{T}_h$. Owing to these assumptions, the Lax-Milgram Lemma implies that problem (2.1) is well posed.

## 2.1. Classical Stabilized Methods

Classical stabilized finite element methods achieve the double goal of stability and accuracy in the solution of (2.1) in both the convection-dominated and diffusion-dominated regimes, by modifying the bilinear form. This section aims to give the reader a summary of current understanding of the SUPG and RFB methods, as applied to linear steady-state convection-diffusion problems.

### 2.1.1. SUPG Method for Boundary Value Problems

SUPG method, introduced by Brooks and Hughes in (Brooks and Hughes, 1982), (Hughes and Brooks, 1979) and (Hughes and Mallet, 1986) can be considered as the first successful stabilization technique to prevent oscillations in convection-dominated problems in the finite element method (FEM). The main steps are: introduce artificial diffusion in streamline direction only, interpret this as a modification of the test function of the advection terms and finally, enforce consistency, such that this modified test function is applied to all terms of the weak form. Then, the term artificial diffusion is not fully applicable any longer, because the stabilized weak form can not, in general, be manipulated such that only a diffusion term is extracted. The exact solution of the problem still satisfies the SUPG stabilized weak form.

Let us recall the classical variational formulation of problem (2.1): Find $u \in H_0^1(\Omega)$ such that

$$a(u, v) = (f, v) \quad \forall v \in H_0^1(\Omega), \tag{2.2}$$

where

$$a(u, v) = \epsilon \int_\Omega \nabla u \, \nabla v + \int_\Omega (\beta \cdot \nabla u) \, v, \tag{2.3}$$

is a continuous and coercive bilinear form on the Hilbert space $H_0^1(\Omega)$. A Galerkin approximation of problem (2.1) consists in taking a finite-dimensional subspace $V_h$ of $H_0^1(\Omega)$, and then solving the variational problem (2.2) in $V_h$. For the sake of simplicity, from now on we will restrict ourselves to the case of continuous, piecewise linear elements, i.e., we will consider the following finite element space:

$$V_L = \{v \in H_0^1(\Omega), \, v|_K \text{ linear } \forall K \in \mathcal{T}_h\}, \tag{2.4}$$

so that the approximation of (2.2) reads: Find $u_L \in V_L$ such that

$$a(u_L, v_L) = (f, v_L) \quad \forall v_L \in V_L. \tag{2.5}$$

If the problem is convection-dominated, then, unless the mesh size $h$ is of the same size of $\epsilon/|\beta|$, the solution of (2.5) will exhibit strong oscillations spreading all over the domain.

The idea behind the SUPG method is to add to the original bilinear form $a(\cdot, \cdot)$ a term which introduces a suitable amount of artificial diffusion in the direction of streamlines, but without upsetting consistency. In the case of problem (2.1), with linear elements, the SUPG method reads: Find $u_L \in V_L$ such that

$$a(u_L, v_L) + \sum_K \tau_K \int_K (\beta \cdot \nabla u_L - f)(\beta \cdot \nabla v_L) = (f, v_L) \quad \forall v_L \in V_L, \qquad (2.6)$$

where $\tau_K$ is a stabilization parameter depending on the local character of the discretization. Define $\tau_K$ element by element according to the size of $Pe_K = \frac{|\beta|_K h_K}{6\epsilon}$:

$$\tau_K = \begin{cases} \frac{h_K}{2|\beta|_K} & Pe_K \geq 1, \\ \\ \frac{h_K^2}{12\epsilon} & Pe_K < 1. \end{cases} \qquad (2.7)$$

Despite the success of the SUPG method, there are areas for improvement. For example, because the method is not monotone, it does not preserve the positivity of the solution, which is unphysical in some applications. Another disadvantage is that the amount of streamline diffusion has to be tuned depending on the problem at hand and in real world fluid flow simulations, tuning of the method can be difficult. This difficulty has motivated the introduction of intrinsically stable methods. In the following section, we shall review the Residual-Free Bubbles (RFB) method of Brezzi and Russo in (Brezzi and Russo, 1994).

## 2.1.2. RFB Method for Boundary Value Problems

The SUPG and RFB methods are closely related, as discussed in (Brezzi et al., 1997a). A detailed discussion of the advantages and disadvantages of these methods can be found in references (Brezzi et al., 1997b), (Brezzi and Marini, 2002a).

As we mentioned before, a possible drawback of the SUPG method is the sensitivity of the solution to the stabilization parameter $\tau_K$, whose value is not determined precisely by the available theory. A way to recover intrinsically the value of $\tau_K$ is to use

the RFB approach (see (Brezzi and Russo, 1994) , (Brezzi et al., 2003)) that will be recalled here. RFB method is based on a local enrichment of the finite-element space instead of a modification of the variational formulation. The idea is to add to the usual space of piecewise polynomials, the so called bubbles. Bubbles are functions whose support remains inside the elements of the triangulation. The numerical method turns out to be stable, at the price of having to solve local problems in order to approximate, and possibly eliminate, the infinite bubble degrees of freedom. In particular, the finite element space $V_L$ is enlarged in the following way. For each element $K$, we define the space of bubbles in $K$ as

$$B_K = H_0^1(K) \,, \tag{2.8}$$

and the enlarging space $V_B$ as

$$V_B = \bigoplus_K B_K \,. \tag{2.9}$$

We solve the problem (2.2) on

$$V_h = V_L \bigoplus V_B \,. \tag{2.10}$$

Then every $v_h \in V_h$ can uniquely be written in the form of $v_h = v_L + v_B$, where $v_L \in V_L$ and $v_B \in V_B$. The variational problem (2.2) is approximated as follows:
Find $u_h = u_L + u_B \in V_L + V_B$ such that,

$$
\begin{aligned}
a(u_L + u_B, v_L) &= (f, v_L) \quad \forall v_L \in V_L \,, & (2.11) \\
a(u_L + u_{B,K}, v_{B,K}) &= (f, v_{B,K}) \quad \forall v_{B,K} \in B_K \,. & (2.12)
\end{aligned}
$$

Solving (2.12) inside each $K$ for $u_{B,K}$ and substituting into the first equation, it can be shown that the effect of the bubbles can be identified with an additional term that has an identical structure with the mesh dependent term in the SUPG method. Consequently, the resulting scheme on $V_L$ reads: Find $u_L \in V_L$ such that,

$$a(u_L, v_L) + \sum_K \widehat{\tau}_K \int_K (\beta \cdot \nabla u_L - f)(\beta \cdot \nabla v_L) = (f, v_L) \quad \forall v_L \in V_L \,, \tag{2.13}$$

where

$$\widehat{\tau}_K = \frac{1}{|K|} \int_K b_K \,, \tag{2.14}$$

and $b_K$ solves the following boundary value problem on $K$:

$$\begin{cases} -\epsilon \triangle b_K + \beta \cdot \nabla b_K = 1 & \text{in } K \,, \\ b_K = 0 & \text{on } \partial K \,. \end{cases} \tag{2.15}$$

RFB method is a general methodology for solving partial differential equations. For a practical implementation of the RFB formulation, one has to approximate the infinitely many degrees of freedom of the bubble, in order to suitably approximate the local problems. In particular, the method (2.13) depends on the solution of the local problem (2.15) whose solution may be difficult as much as the original problem. Therefore several numerical methods were proposed to compute an approximate solution of the problem (2.15). The common point of these attempts was to construct a low-dimensional subspace $B_{h,K} \subset B_K$ in such a way that the solution of the discrete local problem, Find $b_{h,K} \in B_{h,K}$ such that,

$$a(b_{h,K}, b_h) = (1, b_h) \quad \forall b_h \in B_{h,K} \,, \tag{2.16}$$

could produce a solution $b_{h,K}$ such that,

$$\int_K b_{h,K} \simeq \int_K b_K \,. \tag{2.17}$$

We are left with the problem of evaluating, possibly in some approximate way, the integral of $b_K$ appearing in (2.14). If $\epsilon \ll |\beta|_K h_K$ the solution of (2.15) will be very close to a pyramid with one (or two) almost vertical faces on the outflow boundary of $K$ (the element boundary layer). The remaining faces of this pyramid have slope $1/|\beta|_K$ in the direction of $\beta|_K$. Hence, if we define $\widehat{h}_K$ as the length of longest segment parallel to $\beta|_K$ and contained in $K$, we have

$$\int_K b_K \approx \text{ volume of the pyramid} = \frac{|\mathrm{K}|}{3} \frac{\widehat{\mathrm{h}}_{\mathrm{K}}}{|\beta|_{\mathrm{K}}} \,, \tag{2.18}$$

so that

$$\widehat{\tau}_K = \frac{1}{|K|} \int_K b_K \approx \frac{\widehat{h}_K}{3|\beta|_K} \,. \tag{2.19}$$

However, this only work in the asymptotic case where $\epsilon \ll h$. The optimal value of the stabilization parameter $\widehat{\tau}$ in the other regimes still needs to be discussed carefully. This is the subject of the next section.

## 2.2.  Using Finite Element Method with Stabilizing Subgrid

In practice, it is important to devise numerical algorithms that provide cheap approximations to the bubble function, contributing a good stabilizing effect to the numerical method overall. Therefore, the actual computation of the bubble function could be carried out numerically by introducing a suitable subgrid. In this way a fully discrete procedure is obtained. The choice of the subgrid dictates how fine scales are incorporated into the coarse scale formulation. Now, we will introduce two different approaches to construct such subgrid for triangular elements.

### 2.2.1.  First Approach: Plain-Galerkin Scheme on the Augmented Grid

In this section, we will summarize a different point of view which was proposed in (Brezzi and Marini, 2002a), (Brezzi and Marini, 2002b) for two-dimensional convection-diffusion problems and further developed in (Brezzi et al., 2003) for one-dimensional convection-diffusion-reaction problems. The basic idea is to consider both the original grid and the subgrid at the same time as an augmented grid, and to solve with Standard Galerkin method on such Augmented Grid (SGAG). In practice, the internal nodes added with the subgrid can still be eliminated by static condensation, so that the method could still be regarded as a variant of the RFB approach. Once a convenient subgrid has been

constructed, it is easy to use the idea of a plain Galerkin code without any tricks. Essentially, the idea is to compute the coefficients of the subgrid stiffness matrix as functions of the distance between each internal node and its closest node at the boundary of the element. Then the distance is chosen in such a way that the coefficient corresponding to the adjacent boundary node becomes zero.

As announced, the key point is to construct a subgrid in each element $K$, and then solve the variational problem on the augmented space, essentially made of piecewise linear functions on the augmented grid, that is the union of the original grid and of the subgrid. We are going to take a subgrid that contains just one additional node $P = P_K$ in each element $K$. The node is then joined to the three vertices, thus splitting the triangle in three subgrid triangles. To further specify the strategy of choice, we prescribe that the location of $P_K$ should be chosen along one of the three medians of $K$. The choice of the median, and the position of $P_K$ on it will depend on the direction of $\beta$, and will be made precise in the sequel.

As a first step, suppose the $P_K$'s are given. In $V_h$ we choose a basis made of functions having value $1$ at one node and $0$ at the other nodes. The basis function attached to the each point $P_K$ will have support contained in $K$. The other three basis functions that are different from zero in $K$ will have value $1$ at one vertex, and $0$ at $P_K$ and at the other vertices.

As we are going to discuss each element separately, we drop most of the indices $K$, and take a local numbering for the vertices, that will be denoted by $V_i$ $(i = 1, 2, 3)$ using the counter-clockwise ordering. The basis functions that are different from $0$ on $K$ will then be denoted by $b_P, \varphi_1, \varphi_2, \varphi_3$, where

$$b_P(P) = 1, \ b_P(V_i) = 0, \ \ i = 1, 2, 3, \tag{2.20}$$

and

$$\varphi_i(V_i) = 1, \ \ \varphi_i(V_j) = 0 \ \ j \neq i, \ \ \varphi_i(P) = 0. \tag{2.21}$$

In order to specify the choice for $P$ we need some additional notation. As in Figure 2.1, we denote by $e_i$ ($i = 1,2,3$) the edges of $K$, with $e_i$ opposite to $V_i$; $|e_i|$ will denote the length of $e_i$, $n^i$ the outward unit normal to $e_i$, and $\nu_i = |e_i| n^i$. The actual numbering of the vertices will be chosen according to the direction of $\beta$. Then, as announced, $P$ will be a point on the median $m$ from $V_1$ to the midpoint $M$ of edge $e_1$, that is,

(a) two inflow edges.      (b) one inflow edge.
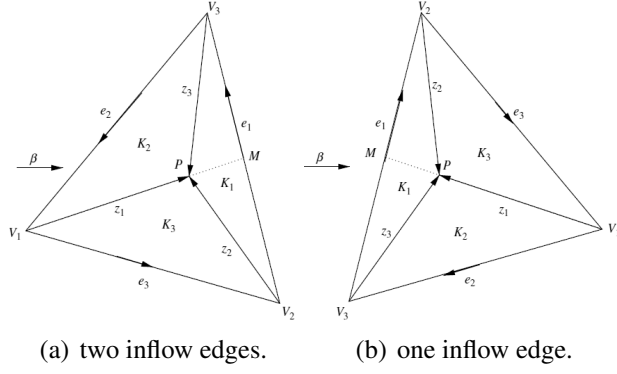
Figure 2.1. Notation and types of outflow boundary.

$m = (e_3 - e_2)/2$. Finally, we denote by $K_i$ the three sub-triangles obtained by connecting $P$ with the vertices $V_i$, by $|K_i|$ the area of $K_i$, and by $z_i$ the vectors from $V_i$ pointing to $P$.

**Case 1. The inflow boundary is made of two edges of K:** Referring to Figure 2.1, let $e_2, e_3$ be the two inflow edges. The position of $P$ along the median from $V_1$ will be determined by annihilating the sum of the contributions of $V_2$ and $V_3$ to $P$. More precisely, we look for $P = (1 - t)V_1 + tM \quad 0 < t < 1$, such that

$$a_K(\varphi_2, b_P) + a_K(\varphi_3, b_P) = 0. \tag{2.22}$$

Solving equation (2.22) for $t$ gives

$$t_1^* = 1 + \frac{\epsilon |e_1|^2}{\epsilon |e_2 - e_3|^2 - 2|K|(\beta, \nu_1)/3}. \tag{2.23}$$

However, we do not always take that given by (2.23) as actual value for $t$. Indeed, for $\epsilon$ not too small (that is, for diffusion-dominated problems) this type of stabilization would be unnecessary, and actually the value provided by (2.23) could be meaningless. Hence we take

$$\begin{cases} t = t_1^*, & \epsilon \leq \epsilon_1^* = \frac{2|K|(\beta, \nu_1)/3}{3|e_1|^2 + |e_2 - e_3|^2}, \\ t = 2/3, & \text{otherwise.} \end{cases} \tag{2.24}$$

Moreover for $0 < \epsilon < \epsilon_1^*$ we have $1 > t_1^* > 2/3$ so that, for every $\epsilon > 0$, we have

$$2/3 \leq t < 1 \, . \tag{2.25}$$

**Case 2.   The inflow boundary is made of one edge of K:** Referring to Figure 2.1, let $e_1$ be the inflow edge. In this case we determine the position of $P$ along the median from $V_1$ by annihilating the contribution of $V_1$ to $P$, that is, by imposing

$$a_K(\varphi_1, b_P) = 0 \, , \tag{2.26}$$

which gives

$$t_2^* = \frac{\epsilon(|e_2|^2 + |e_3|^2)}{\epsilon|e_2 - e_3|^2/2 - |K|(\beta, \nu_1)/3} \, . \tag{2.27}$$

As we did in Case 1, however, we do not take $t = t_2^*$ for every value of $\epsilon$, but only for convection-dominated problems. Hence we take

$$\begin{cases} t = t_2^*, & \epsilon \leq \epsilon_2^* = \frac{2|K|(-\beta, \nu_1)/3}{3(|e_2|^2 + |e_3|^2) - |e_2 - e_3|^2} \, , \\ t = 2/3, & \text{otherwise.} \end{cases} \tag{2.28}$$

Moreover for $0 < \epsilon < \epsilon_2^*$ we have $0 < t_2^* < 2/3$ so that, for every $\epsilon > 0$, we have

$$0 \leq t < 2/3 \, . \tag{2.29}$$

Here, for the model problem of linear convection-diffusion equations, we propose a simple criterion to choose a single internal node such that the corresponding plain-Galerkin scheme on the augmented grid provides the same a priori error estimates that are typically obtained with SUPG or RFB methods, see (Brezzi et al., 2005) and the references therein for further reading.

## 2.2.2. Second Approach: Optimal Bubbles

In this section, we will summarize another methodology, inspired by a different philosophy, proposed in (Brezzi et al., 1998a), for computing an approximate solution of (2.15), which is good enough to indicate suitable values for the parameter $\tau_K$. As we shall see in the following with more details, the approximate optimal bubble (that is also referred as pseudo residual-free bubble) is set to be piecewise linear, having the shape of a pyramid, with vertex in a point $P$ internal to $K$. The height of the pyramid will be determined by solving problem (2.15) in the Galerkin sense, while the location of $P$ is to be chosen in order to minimize the $L_1$-norm of the residual. The method that is generated in this way increases smoothly the amount of added streamline diffusion as $\epsilon$ decreases, and gives the same limit (for $\epsilon \to 0$) as the residual-free bubbles.

We want now to approximate the solution $b_K$, of (2.15) with functions of the type $\alpha b_P(x)$, where $\alpha$ and $P$ have to be suitably chosen. As a first step, we choose $\alpha$ as a function of $P$; we look for $\alpha = \alpha(P)$ such that

$$a(\alpha b_P, b_P)_K = \int_K b_P \,. \tag{2.30}$$

An easy computation gives,

$$\alpha(P) = \frac{\int_K b_P}{\epsilon \int_K |\nabla b_P|^2} \,. \tag{2.31}$$

We then set $B_P(x) = \alpha(P) b_P(x)$ and we see immediately that the corresponding stabilization parameter $\widetilde{\tau}_K$ is again given by (see (2.19))

$$\widetilde{\tau}_K = \frac{1}{|K|} \int_K B_P = \frac{1}{|K|} \frac{(\int_K b_P)^2}{\epsilon \int_K |\nabla b_P|^2} \,. \tag{2.32}$$

As a second step, we choose P. We require that

$$J(P) = \int_K | -\epsilon \triangle B_P + \beta|_K \cdot \nabla B_P - 1| \,, \tag{2.33}$$

is minimum. According to (Brezzi et al., 1998a) it is possible to write $J(P)$ as follows:

$$J(P) = \frac{2|K|}{3} + \int_K |\beta|_K \cdot \nabla B_P - 1| \,. \tag{2.34}$$

We note that the first term on the right-hand side of (2.34) is independent of $P$. For every fixed $P$, $g_i = g_i(P)$ are constants and defined as

$$g_i = (\beta \cdot \nabla B_P - 1)|_{K_i} \,. \tag{2.35}$$

We notice now that

$$\sum_{i=1}^{3} \int_{K_i} g_i = -|K| \,, \tag{2.36}$$

and finally, always for every $P \in K$, by triangle inequality we have

$$\sum_{i=1}^{3} \int_{K_i} |g_i| \geq |K| \,. \tag{2.37}$$

Using (2.36), we conclude that at least one of the $g_i$'s is negative. If, for some $\widetilde{P} \in K$, all the $g_i$'s are negative, then $\widetilde{P}$ minimizes (2.33), thanks to (2.36) and (2.37). Next, we state the following lemma:

**Lemma 2.2.2.1** *Let, for every $P \in K$, $g_i(P)$, $i = 1, 2, 3$, be defined as in (2.35). Then there exists a region $S \subset K$ such that $\forall P \in S$,*

$$g_i(P) \leq 0 \quad (i = 1, 2, 3). \tag{2.38}$$

**Proof** See (Brezzi et al., 1998a) for details. □

All the points belonging to the region $S$ obtained in Lemma 2.2.2.1 will produce the same value for (2.33). With this notation, standard computations give,

$$\widetilde{\tau}_K = \frac{1}{|K|} \int_K B_P = \frac{4|K|^2/9\epsilon}{\sum_j e_j^2/\lambda_j(P)}, \tag{2.39}$$

where $\lambda_i = \lambda_i(P) = |K_i|/|K|$ denotes the unknown barycentric coordinates of $P$. We see that the stabilizing parameter $\widetilde{\tau}_K$ will still depend on the choice of $P$ (in $S$) through the formula above and, in a sense, the stabilizing effect decreases when $P$ is approaching $\partial K$. If the problem is truly convection-dominated in $K$, the region $S$ will consist of a narrow strip around the outflow part of the boundary $\partial K$. Hence, in order to have the point $P$ quantatively, we can proceed as follows.



(a) single outflow edge.    (b) two outflow edges.
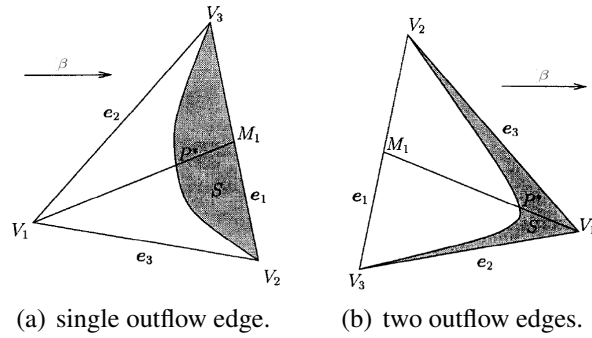
Figure 2.2. Types of outflow boundary.

**Case 1.   The outflow boundary is made of one edge of K:** Take the first case of Figure 2.2 (corresponding to a single outflow edge), and consider the line joining $V_1$, with the midpoint $M_1$ of $e_1$. Note that on this line only a small interval, close to $M_1$, will be in $S$. We then choose the point $P^*$ on the line $V_1 M_1$, closest to $V_1$, and belonging to $S$. This choice gives a simple procedure to compute $P^*$. Indeed, the problem now has only one parameter; the corresponding algorithm to spot $P^*$ is easier to implement than to explain, and consists simply in the translation of the previous description in a set of algebraic equations.

**Remark 2.1** *The criteria (2.33) was also used in (Neslitürk, 2006). Let us denote the set of points on the median $V_1 M_1$ as a function depending on a single parameter t: $P = (1 - t)V_1 + tM_1$. Following the same steps in (Neslitürk, 2006), the interval in*

*which the values of t minimize the integral (2.33) is obtained as*

$$1 > t \geq t_1^* = 1 + \frac{\epsilon|e_1|^2}{\epsilon|e_2 - e_3|^2 - \frac{4}{3}|K|(\beta, \nu_2)} \ . \tag{2.40}$$

**Case 2.  The outflow boundary is made of two edges of K:** If we take now the case of Figure 2.2 (corresponding to two outflow edges) we can consider, again, the line joining $V_1$, with the midpoint $M_1$ of $e_1$: on this line the region $S$ will now be found as a small interval whose closure contains $V_1$. As before, we choose $P^*$ as the point, in this interval, closest to $M_1$.

**Remark 2.2** *Similarly, following the same steps in  (Neslitürk, 2006), the interval in which the values of t minimize the integral (2.33) is obtained as*

$$0 < t \leq t_2^* = \frac{\epsilon(|e_2|^2 + |e_3|^2)}{\epsilon|e_2 - e_3|^2/2 - |K|(\beta, \nu_1)/3} \ . \tag{2.41}$$

Here, the concept of pseudo residual-free bubble, which is a certain approximation of the true residual-free bubble has been introduced. In some cases the method obtained by using the pseudo residual-free bubbles gives better results than the SUPG method, see (Brezzi et al., 1998a) and the references therein for further reading.

# CHAPTER 3

# A STABILIZED FINITE ELEMENT METHOD FOR CONVECTION-DIFFUSION-REACTION PROBLEMS IN 1-D

It is known that the enrichment of the polynomial finite element space of degree 1 by bubble functions results in a stabilized scheme of the SUPG-type for the convection-diffusion-reaction (CDR) problems. In particular, the residual-free bubbles (RFB) can assure stabilized methods, but they are usually difficult to compute, unless the configuration is simple. Therefore it is important to devise numerical algorithms that provide cheap approximations to the RFB functions, contributing a good stabilizing effect to the numerical method overall. Here we propose a stabilization technique based on the RFB method and particularly designed to treat the most interesting case of small diffusion. We replace the RFB functions by their cheap, yet efficient approximations which retain the same qualitative behavior. The approximate bubbles are computed on a suitable sub-grid, the choice of whose nodes are critical and determined by minimizing the residual of a local problem with respect to $L_1$ norm. The resulting numerical method has similar stability features with the RFB method for the whole range of problem parameters. This fact is also confirmed by numerical experiments. We also note that the location of the sub-grid nodes suggested by the strategy herein coincides with the one in (Brezzi et al., 2003).

## 3.1. Introduction

It is well known that the CDR problems may contain thin regions in which the solution varies abruptly. The plain Galerkin method may not work for such problems on reasonable discretizations, producing unphysical oscillations. The SUPG method, and its variants, are among the most popular approaches to overcome that difficulty, which are based on augmenting the variational formulation by mesh-dependent terms in order to gain control over the derivatives of the solution (Brooks and Hughes, 1982), (Franca et al., 1992), (Hughes et al., 1989). The great advantage of this approach is not only its generality, but also its error analysis can be performed in many cases of interest. Never-

theless, the need for the proper choice of stabilizing parameter is considered as a major drawback of the method.

Another approach consists of enriching the finite element spaces by bubble functions. The relationship between the use of bubble functions and stabilized methods was also studied in (Baiocchi et al., 1993), (Brezzi et al., 1992), (Brezzi and Russo, 1994). It turns out that, to find a more suitable value for the stabilizing parameter in the SUPG method, it is crucial to use special type of functions, so called the residual-free bubbles (RFB), defined by a local problem posed inside each element. The RFB method also allows one to prove error bounds (Brezzi et al., 2000), (Sangalli, 2000) and can be generalized to a much wider variety of problems (Brezzi et al., 1998b), (Franca and Russo, 1996). However it requires to solve a local differential equation which may not be easier than to solve the original one (Franca et al., 1998), (Franca and Tobiska, 2002).

Yet another way of stabilizing the Galerkin method is to stabilize by means of a suitable refinement around the layer so that, the stabilization is actually not needed anymore, like in the Shishkin meshes (Farrell et al., 2000). The drawback of these methodologies resides in that they require a priori knowledge of the layer locations.

Here, we will present a stabilization method for one-dimensional CDR problems, particularly designed to treat the most interesting case of small diffusion, but able to adapt from one regime to another continuously. It is based on the RFB method, in which, however, we replace the RFB functions by their cheap, yet efficient approximations, so called pseudo RFBs, which retain the same qualitative behavior as the RFBs. Similar approaches to obtain suitable approximations to the RFBs can be found in the literature (Brezzi et al., 1998a), (Brezzi and Marini, 2002a), (Brezzi et al., 2005), (Neslitürk, 2006), (Neslitürk, 2010). The pseudo bubbles are chosen to be piecewise linear on a suitable sub-grid that, the position of whose nodes are determined by minimizing the residual of local differential problems with respect to $L_1$ norm. The recipe for spotting sub-grid points is simple and their location coincides with the one in (Brezzi et al., 2003). The resulting numerical method has similar stability features to the RFB method for the whole range of problem parameters. This fact is confirmed by numerical experiments presented below.

## 3.2. Statement of the Problem

We will consider the following linear elliptic CDR problem in $I = (0,1)$:

$$\mathcal{L}u = -\epsilon u'' + \beta u' + \sigma u = f(x) \quad \text{with} \quad u(0) = u(1) = 0 \,. \tag{3.1}$$

Let $0 = x_0 < x_1 < x_2 < \ldots < x_{N-1} < x_N = 1$ and $\mathcal{T}_h = \{K\}$ be a decomposition of $I$ into subintervals $K = (x_{k-1}, x_k)$ where $k = 1, \ldots, N$. For the sake of simplicity, we shall assume that the decomposition is uniform, so that we can denote the length of the intervals in the subdivision by $h$. However, all our discussions will take place at the element level, and therefore, they will also be valid for quasi-uniform decompositions.

We assume that the diffusion coefficient $\epsilon$ is a positive constant, and that the convection field $\beta$ and the reaction field $\sigma$ are non-negative piecewise constants with respect to the decomposition $\mathcal{T}_h$. So, unless $\beta \equiv 0$ (pure reaction case), we can speak of inflow and outflow. When $\epsilon \ll |\beta|\, h + \sigma\, h^2$, the solution of the problem will have boundary layers for a generic $f$, that can be either only at the outflow, or at both ends of $I$, depending on the reciprocal values of $|\beta|\, h$ and $\sigma\, h^2$. In these cases, the pure Galerkin method will typically fail, showing strong oscillations near the boundary layers, and some stabilization is needed.

Here we will consider stabilizations based on the augmented space idea which includes the RFB strategy. We start by recalling the abstract variational formulation of problem (3.1): Find $u \in H_0^1(I)$ such that

$$a(u,v) = (f,v) \quad \forall v \in H_0^1(I) \,, \tag{3.2}$$

where

$$a(u,v) = \epsilon \int_I u' v' \, dx + \int_I (\beta\, u)' v \, dx + \int_I \sigma\, u\, v dx \,. \tag{3.3}$$

We now define $V_h$ as a finite-dimensional space, which is a subspace of $H_0^1(I)$. Then the standard Galerkin finite element method reads: Find $u_h \in V_h$ such that

$$a(u_h, v_h) = (f, v_h) \quad \forall v_h \in V_h \,. \tag{3.4}$$

Now, we decompose the space $V_h$ such that $V_h = V_L \bigoplus V_B$, where $V_L$ is the space of continuous piecewise linear polynomials and $V_B = \bigoplus_K B_K$ with $B_K = H_0^1(K)$ Then every $v_h \in V_h$ can be written in the form of $v_h = v_L + v_B$, where $v_L \in V_L$ and $v_B \in V_B$. We require the bubble component $u_B$ of $u_h$ to satisfy the original differential equations in $K$ strongly, i.e.

$$\mathcal{L}u_B = -\mathcal{L}u_L + f \quad \text{in } K, \tag{3.5}$$

subject to the boundary condition,

$$u_B = 0 \quad \text{on } \partial K. \tag{3.6}$$

By the classical static condensation procedure (Brezzi et al., 2003), the method used to compute an improved linear approximation due to the residual-free bubble effect reads: Find $u_h = u_L + u_B$ in $V_h$ such that

$$a(u_L, v_L) + a(u_B, v_L) = (f, v_L) \quad \forall v_L \in V_L. \tag{3.7}$$

The term $a(u_B, v_L)$ is responsible for the stabilization of the numerical method and the bubble component $u_B$ should be computed before we solve (3.7) for its linear part. Recall $u_B$ is identified by the linear part $u_L$ and the source function $f$ through (3.5)-(3.6), which may be as complicated as solving the original differential equation. Therefore, it is important to bring a simple recipe about to obtain a suitable approximation to the bubble component of the problem that provides a similar stabilizing effect into the numerical method. We discuss that approach in the following section. The discussion will take place in a typical element $K$, and therefore, we shall drop the index $K$ in the notation unless it is necessary.

## 3.3. Description of the Method

Let us define a sub-grid in a typical element $K = (x_{k-1}, x_k)$ by adding two points $z_1$ and $z_2$ with the property that

$$x_{k-1} < z_1 < z_2 < x_k, \tag{3.8}$$

on which, we approximate the bubble functions. The shape of approximations, which is essentially related with the location of sub-grid points, is crucial to get a good stabilization effect on the numerical method. Therefore the choice of points in the sub-grid must be fulfilled in a special manner. That will be accomplished by a minimization process with respect to $L_1$ norm in the presence of layers.

Let us assume that $f$ be a piecewise linear function with respect to the discretization. Then the residual in (3.5) becomes a linear function and it is reasonable to consider bubble functions $B_i$, $(i = 1, 2)$ defined by

$$\mathcal{L}B_i = -\mathcal{L}\psi_i \quad \text{in } K, \qquad B_i = 0 \quad \text{on } \partial K, \quad i = 1, 2, \tag{3.9}$$

where $\psi_1, \psi_2$ are the restrictions of the piecewise linear basis functions for $V_L$ to $K$ (Figure 3.1).
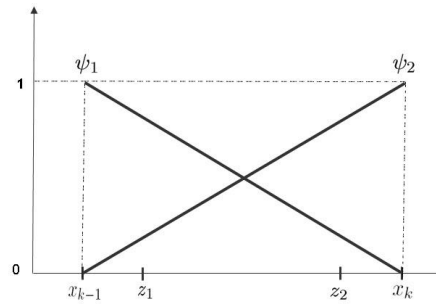


Figure 3.1. The restrictions of piecewise linear basis functions to a typical element $K$.

Further we define $B_f$,

$$\mathcal{L}B_f = f \quad \text{in } K, \qquad B_f = 0 \quad \text{on } \partial K. \tag{3.10}$$

Now if

$$u_L\big|_K = u_L(x_{k-1})\,\psi_1 + u_L(x_k)\,\psi_2\,, \tag{3.11}$$

then, we take

$$u_B\big|_K = u_L(x_{k-1})\,B_1 + u_L(x_k)\,B_2 + B_f\,. \tag{3.12}$$

Thus

$$
\begin{aligned}
\mathcal{L}u_B &= u_L(x_{k-1})\mathcal{L}B_1 + u_L(x_k)\mathcal{L}B_2 + \mathcal{L}B_f \\
&= u_L(x_{k-1})\left(-\mathcal{L}\psi_1\right) + u_L(x_k)\left(-\mathcal{L}\psi_2\right) + f \\
&= -\mathcal{L}(u_L(x_{k-1})\psi_1 + u_L(x_k)\psi_2) + f = -\mathcal{L}u_L + f \quad \text{in } K\,.
\end{aligned}
\tag{3.13}
$$

That is, the equation (3.5) is automatically satisfied with the present choice of bubble functions. The equation (3.9) is similar to the original problem (3.1) and may be difficult to solve. However, using the element geometry and the problem properties, it is possible to construct a cheap, yet efficient approximate bubbles, say $B_i^*$, over the sub-grid (3.8), having the same qualitative behavior with its continuous counterpart $B_i$ ($i = 1, 2$). The construction of such approximate bubble functions $B_i^*$ is given in the following.
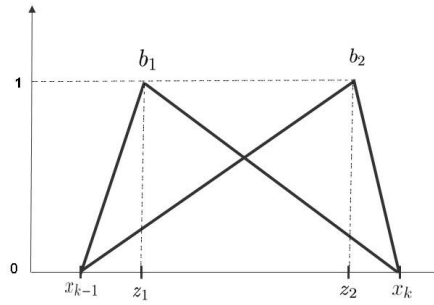


Figure 3.2. Basis functions employed in the approximation of bubble functions.

Let $B_i^*(x) = \alpha_i b_i(x)$ be the classical Galerkin approximation of $B_i$ through (3.9)

$$
a(B_i^*, b_i)_K = (-\mathcal{L}\psi_i, b_i)_K\,, \quad i = 1, 2\,,
\tag{3.14}
$$

where $b_i$ is a piecewise linear function with the following properties (Figure 3.2)

$$
b_i(x_{k-1}) = b_i(x_k) = 0\,, b_i(z_i) = 1, \qquad i = 1, 2\,.
\tag{3.15}
$$

Using integration by parts, the properties of bubble functions and the midpoint rule for

quadratic terms that appears, we get explicit expressions for $\alpha_1$ and $\alpha_2$, separately:

$$\alpha_1 = \frac{(-\mathcal{L}\psi_1, b_1)_K}{\epsilon||b_1'||^2_K + \sigma||b_1||^2_K} = \frac{3\beta + (\xi - 2h)\sigma}{2h(\frac{3\epsilon}{\xi(h-\xi)} + \sigma)}, \tag{3.16}$$

and

$$\alpha_2 = \frac{(-\mathcal{L}\psi_2, b_2)_K}{\epsilon||b_2'||^2_K + \sigma||b_2||^2_K} = -\frac{3\beta + (2h - \eta)\sigma}{2h(\frac{3\epsilon}{\eta(h-\eta)} + \sigma)}. \tag{3.17}$$

Note that $\alpha_2 < 0$. Now it remains to choose $z_i$, so that the stabilizing effect of bubble function $B_i$ is maintained in its discrete counterpart $B_i^*$ $(i = 1, 2)$. The main criteria that we use to determine the locations of the sub-grid nodes is to minimize $L_1$ norm of the residual coming out from the bubble equation (3.9) in the critical case where a layer structure exists. In other words, we choose $z_i$ such that

$$J_i = \int_K \left| \mathcal{L}B_i^* + \mathcal{L}\psi_i \right| dx, \quad i = 1, 2, \tag{3.18}$$

is minimum. That approach (3.18) were also used in (Brezzi et al., 1998a). Before we derive the explicit locations of sub-grid points that emerges from the criterion (3.18), let us make some general observations on their configuration. Set

$$\xi = z_1 - x_{k-1}, \quad \eta = x_k - z_2, \quad \delta = z_2 - z_1,$$
$$K_1 = [x_{k-1}, z_1], \quad K_2 = [z_1, z_2], \quad K_3 = [z_2, x_k]. \tag{3.19}$$

From (3.8) and (3.19), it is obvious that $\xi + \delta + \eta = h$. At the same time, we do not want $\delta$ to be too small, either, when compared with $\xi$ and $\eta$. Therefore we take,

$$\delta \geq \min\{\xi, \eta\}. \tag{3.20}$$

From the qualitative behavior of the problem (3.1), we always have $\eta \leq \xi$, which implies

$\delta \geq \eta$. Hence $\eta$ will always be the smallest of the three sub-lengths and, thus we have

$$\eta \leq h/3\,. \tag{3.21}$$

Now we are in a position to give the explicit description of sub-grid points for each type of problem regime.

### 3.3.1. Diffusion-Dominated Regime

In the present algorithm, the problem is assumed to be diffusion-dominated when $6\epsilon > \beta h + \sigma h^2/9$. In this regime, the stabilization is not needed, and a uniform sub-grid seems to be appropriate. Therefore we choose $\xi = \eta = \delta = h/3$.

### 3.3.2. Convection-Dominated Regime

In convection dominated case, we have a single exponential boundary layer at the outflow. Therefore it is enough to find an optimal location for $z_2$ only and place $z_1$ on an appropriate location with respect to the configuration of the problem. We assume that the problem is convection-dominated if $6\epsilon \leq \beta h + \sigma h^2/9$ with $3\beta \geq \sigma h$. The following lemma suggests an optimal position for $z_2$ by using (3.18).

**Lemma 3.3.2.1** *In convection-dominated case, the point* $\eta_e = \dfrac{-3\beta + \sqrt{9\beta^2 + 24\epsilon\sigma}}{2\sigma}$ *minimizes the integral (3.18) for* $i = 2$.

**Proof** Following the lines of (Brezzi et al., 1998a), it is possible to write the integral $J_2$ as follows:

$$J_2 = \int_K \left| -\epsilon {B_2^*}'' \right| dx + \int_K \left| \beta {B_2^*}' + \sigma B_2^* + \beta \psi_2' + \sigma \psi_2 \right| dx\,. \tag{3.22}$$

Let $g_2 = \beta {B_2^*}' + \sigma B_2^* + \beta \psi_2' + \sigma \psi_2$. Then, a direct calculation over $K$ gives,

$$\int_K g_2\, dx \;=\; \beta + \frac{\sigma h}{2}(\alpha_2 + 1) = \beta + \frac{\sigma h}{2}\left( 1 - \frac{3\beta + (2h - \eta)\sigma}{2h(\frac{3\epsilon}{\eta(h-\eta)} + \sigma)} \right)$$

$$= \frac{\beta[12\epsilon + \sigma\eta(h - \eta)] + \sigma[6\epsilon h + \sigma\eta^2(h - \eta)]}{12\epsilon + 4\sigma\eta(h - \eta)}, \tag{3.23}$$

which is always positive. Now we split the element $K$ by $z_2$ and investigate the sign of $g_2$ in each of these subdomains. In that direction, use $\sigma\eta \leq \sigma\frac{h}{3} < \beta$ and $\alpha_2 < 0$, to get

$$
\begin{aligned}
g_2\Big|_{K_3} &= -\alpha_2\frac{\beta}{\eta} - \alpha_2\frac{\sigma(x - x_k)}{\eta} + \frac{\beta}{h} + \frac{\sigma(x - x_{k-1})}{h} \\
&= -\frac{\alpha_2}{\eta}(\beta - \sigma(x_k - x)) + \frac{1}{h}(\beta + \sigma(x - x_{k-1})) \\
&\geq -\frac{\alpha_2}{\eta}(\beta - \sigma\eta) + \frac{1}{h}(\beta + \sigma(x - x_{k-1})) > 0.
\end{aligned}
\tag{3.24}
$$

Thus the second term on the right hand side of (3.22) attains its minimum if $g_2\big|_{K_1 \cup K_2}$ is non-negative, too. That is,

$$
\begin{aligned}
g_2\Big|_{K_1 \cup K_2} &= \alpha_2\frac{\beta}{h - \eta} + \alpha_2\frac{\sigma(x - x_{k-1})}{h - \eta} + \frac{\beta}{h} + \frac{\sigma(x - x_{k-1})}{h} \\
&= (\beta + \sigma(x - x_{k-1}))\left(\frac{\alpha_2}{h - \eta} + \frac{1}{h}\right) \\
&= (\beta + \sigma(x - x_{k-1}))\frac{-\sigma\eta^2 - 3\beta\eta + 6\epsilon}{2h(3\epsilon + \sigma\eta(h - \eta))},
\end{aligned}
\tag{3.25}
$$

is positive, only if

$$\eta \leq \frac{-3\beta + \sqrt{9\beta^2 + 24\epsilon\sigma}}{2\sigma}. \tag{3.26}$$

On the other hand, the first term on the right hand side of (3.22) is a locally decreasing function of $\eta$, since

$$\int_K \left| -\epsilon B_2^{*\,''} \right| dx = -\alpha_2\frac{\epsilon h}{\eta(h - \eta)}, \tag{3.27}$$

and

$$\frac{d}{d\eta}\left(-\alpha_2 \frac{\epsilon h}{\eta(h-\eta)}\right) = -\frac{\epsilon(h-2\eta)(3\beta + \sigma(2h-\eta))}{2\eta(h-\eta)(3\epsilon + \sigma\eta(h-\eta))} < 0. \tag{3.28}$$

This fact together with (3.26) determines an optimal value for $\eta$. $\qquad\square$

**Remark 3.1** *The value of $\eta_e$ coincides with the one suggested in (Brezzi et al., 2003).*

**Remark 3.2** *The value of $\alpha_2$ at $\eta_e$ is simply equal to $\dfrac{\eta_e}{h} - 1$.*

The choice of other lengths $\delta$ and $\xi$ should be consistent with the physics of the problem. Thus we take $\eta = \eta_e$, $\delta = \eta$ and $\xi$ is chosen accordingly ($\xi = h - 2\eta$).

### 3.3.3. Reaction-Dominated Regime

In reaction-dominated case, we have two parabolic boundary layers at both ends and the location of both sub-grid points $z_1$ and $z_2$ should be chosen in such a way that approximate bubble functions mimic the exact ones. Thus we spot the position of $z_2$ from Lemma 3.3.2.1 and it remains to find a proper location for $z_1$, which can be accomplished by minimizing the integral

$$J_1 = \int_K \left| \mathcal{L}B_1^* + \mathcal{L}\psi_1 \right| dx. \tag{3.29}$$

Before we find an optimal position for $z_1$, we need the following intermediate result. Note that the problem is assumed to be reaction-dominated if $6\epsilon \le \beta h + \sigma h^2/9$ and $3\beta < \sigma h$.

**Lemma 3.3.3.1** *Let $\alpha_1$ be as in (3.16). In reaction dominated regime we have*

$$\frac{\xi}{2h} - 1 < \alpha_1 < 0. \tag{3.30}$$

**Proof** The upper estimate can easily be obtained by using the fact that

$$3\beta + (\xi - 2h)\sigma < \sigma h + (\xi - 2h)\sigma = \sigma(\xi - h) < 0. \tag{3.31}$$

To show the lower bound, observe that,

$$\alpha_1 + 1 = \frac{3\beta + (\xi - 2h)\sigma}{2h(\frac{3\epsilon}{\xi(h-\xi)} + \sigma)} + 1 = \frac{6\epsilon h + \xi(h-\xi)(3\beta + \sigma\xi)}{6\epsilon h + 2\sigma\xi h(h-\xi)}$$

$$> \frac{6\epsilon h + \sigma\xi^2(h-\xi)}{6\epsilon h + 2\sigma\xi h(h-\xi)} > \frac{\xi(6\epsilon + \sigma\xi(h-\xi))}{2h(3\epsilon + \sigma\xi(h-\xi))} > \frac{\xi}{2h}. \qquad (3.32)$$

$\square$

The following lemma suggests an optimal position for $z_1$.

**Lemma 3.3.3.2** *In reaction-dominated case, the point* $\xi_e = \dfrac{3\beta + \sqrt{9\beta^2 + 24\epsilon\sigma}}{2\sigma}$ *mini-mizes the integral (3.29).*

**Proof** It is possible to write the integral $J_1$ as follows:

$$J_1 = \int_K \left| -\epsilon B_1^{*\,\prime\prime} \right| dx + \int_K \left| \beta B_1^{*\,\prime} + \sigma B_1^* + \beta\psi_1^\prime + \sigma\psi_1 \right| dx. \qquad (3.33)$$

Let $g_1 = \beta B_1^{*\,\prime} + \sigma B_1^* + \beta\psi_1^\prime + \sigma\psi_1$. Without loss of generality, assume $\xi > \dfrac{2\beta}{\sigma}$. Then we have,

$$\int_K g_1 \, dx = -\beta + \frac{\sigma h}{2}(\alpha_1 + 1) = -\beta + \frac{\sigma h}{2}\left( \frac{3\beta + (\xi - 2h)\sigma}{2h(\frac{3\epsilon}{\xi(h-\xi)} + \sigma)} + 1 \right)$$

$$= \frac{6\epsilon(\sigma h - 2\beta) + \sigma\xi(h-\xi)(\sigma\xi - \beta)}{4(3\epsilon + \sigma\xi(h-\xi))} > 0. \qquad (3.34)$$

Now split $K$ into two subregions by $z_1$ and calculate the integral of $g_1$ over each of these sub-domains:

$$\int_{K_1} g_1 \, dx = \alpha_1\beta + \alpha_1\frac{\sigma\xi}{2} - \beta\frac{\xi}{h} - \frac{\sigma\xi}{2h}(\xi - 2h) = \beta\left( \alpha_1 - \frac{\xi}{h} \right) + \sigma\xi\left( \frac{\alpha_1}{2} - \frac{\xi}{2h} + 1 \right)$$

$$\geq \beta\left( \alpha_1 - \frac{\xi}{h} \right) + 2\beta\left( \frac{\alpha_1}{2} - \frac{\xi}{2h} + 1 \right) = 2\beta\left( \alpha_1 - \frac{\xi}{h} + 1 \right), \qquad (3.35)$$

where we have used Lemma 3.3.3.1. Further we have

$$\int_{K_2 \cup K_3} g_1 \, dx = \left( \sigma \frac{h - \xi}{2} - \beta \right) \left( \alpha_1 - \frac{\xi}{h} + 1 \right). \qquad (3.36)$$

The common factor of the last terms in (3.35)-(3.36) can be rewritten as

$$\alpha_1 - \frac{\xi}{h} + 1 = \frac{3\beta + (\xi - 2h)\sigma}{2h(\frac{3\epsilon}{\xi(h-\xi)} + \sigma)} - \frac{\xi}{h} + 1 = \frac{(h - \xi)\left(-\sigma\xi^2 + 3\beta\xi + 6\epsilon\right)}{2h(3\epsilon + \sigma\xi(h - \xi))}.$$

Since $\int_K g_1 \, dx \geq 0$, the second term on the right hand side of (3.33) attains its minimum if both $\int_{K_1} g_1 \, dx$ and $\int_{K_2 \cup K_3} g_1 \, dx$ are positive. For sufficiently large $\sigma$, this is only possible if

$$\xi \leq \frac{3\beta + \sqrt{9\beta^2 + 24\epsilon\sigma}}{2\sigma}. \qquad (3.37)$$

On the other hand, we note that the first term on the right hand side of (3.33) is a locally decreasing function of $\xi$, since

$$\int_K | - \epsilon {B_1^*}'' | \, dx = -\alpha_1 \frac{\epsilon h}{\xi(h - \xi)}, \qquad (3.38)$$

and

$$\frac{d}{d\xi} \left( -\alpha_1 \frac{\epsilon h}{\xi(h - \xi)} \right) = -\frac{\epsilon(h - 2\xi)(-3\beta + 2\sigma h - \sigma\xi))}{2\xi(h - \xi)(3\epsilon + \sigma\xi(h - \xi))} < 0, \qquad (3.39)$$

for $\sigma h > 3\beta$. This fact together with (3.37) determines the optimal value $\xi_e$. $\qquad \square$

**Remark 3.3** *The value of $\xi_e$ coincides with the one suggested in (Brezzi et al., 2003).*

**Remark 3.4** *The value of $\alpha_1$ at $\xi_e$ is simply equal to $\dfrac{\xi_e}{h} - 1$.*

Hence we take $\eta = \eta_e$, $\xi = \min\{h - 2\eta, \xi_e\}$ and $\delta$ is chosen accordingly (i.e. $\delta = h - \eta - \xi$). We note that the points continuously get through from one regime to another in all cases.

Finally we recall that the pseudo bubble functions $B_i^*$ $(i = 1, 2)$ are approximations to $B_i$ on the sub-grid specified above, through (3.14) and they are used in place of $B_i$ to represent $u_B$ in (3.12). The approximate representation of $u_B$ by bubble functions $B_i^*$ $(i = 1, 2)$ is eventually used to solve (3.7) for its linear part. It still remains the computation of $B_f$, before we employ $u_B$ to solve the numerical method (3.7) for the linear part of $u_h$, which is accomplished in the following:

**Remark 3.5** *Since $f$ is assumed to be a linear function, it can be expressed in terms of $\psi_1, \psi_2$, that is*

$$f \mid_K = f(x_{k-1}) \, \psi_1 + f(x_k) \, \psi_2 \,. \tag{3.40}$$

*We want to express $B_f$ in terms of $B_i$, $i = 1, 2$. Set*

$$B_f = \lambda_1 B_1 + \lambda_2 B_2 \,. \tag{3.41}$$

*Using the definition of $B_f$ in (3.10) and the property that $\psi_1 + \psi_2 = 1$ (partition of unity) we have*

$$
\begin{aligned}
\mathcal{L}B_f &= \mathcal{L}(\lambda_1 B_1 + \lambda_2 B_2) = \lambda_1 \mathcal{L}B_1 + \lambda_2 \mathcal{L}B_2 = \lambda_1(-\mathcal{L}\psi_1) + \lambda_2(-\mathcal{L}\psi_2) \\
&= -\lambda_1(\beta\psi_1' + \sigma\psi_1) - \lambda_2(\beta\psi_2' + \sigma\psi_2) \\
&= (-\lambda_1\beta\psi_1' - \lambda_2\beta\psi_2')(\psi_1 + \psi_2) - \lambda_1\sigma\psi_1 - \lambda_2\sigma\psi_2 \\
&= -[\lambda_1(\beta\psi_1' + \sigma) + \lambda_2(\beta\psi_2')] \, \psi_1 - [\lambda_1(\beta\psi_1') + \lambda_2(\beta\psi_2' + \sigma)] \, \psi_2 \,,
\end{aligned}
$$

*which is only possible if*

$$
\begin{aligned}
-[\lambda_1(\beta\psi_1' + \sigma) + \lambda_2(\beta\psi_2')] &= f(x_{k-1}) \,, \\
-[\lambda_1(\beta\psi_1') + \lambda_2(\beta\psi_2' + \sigma)] &= f(x_k) \,.
\end{aligned}
\tag{3.42}
$$

*Once the constants $\lambda_1, \lambda_2$ are determined, we substitute them into equation (3.41) to get $B_f$.*

## 3.4. Numerical Results

In this section, we report some numerical experiments to illustrate the performance of the present algorithm in the interesting case of small diffusion which corresponds to the convection-dominated or reaction-dominated regimes depending on the ratio between the related problem parameters. We remark that the linear part of the numerical solution $u_L$ only are presented in all figures.

**Experiment 1:** We first consider the constant-coefficient case where $\beta = 1$ and $f(x) = 1$. We compute the approximate solution on both uniform and non-uniform meshes. The uniform mesh is generated by dividing the unit interval $[0, 1]$ into ten elements, i.e., the mesh size $h = 1/10$ and the grid points $x_i = i\,h$ where $i = 0, 1, .., 10$. The non-uniform mesh is randomly generated from the uniform mesh by adding a small fraction of $h$ to $x_i$, so that the grid point $x_i$ of the uniform mesh is replaced by a point between $x_i - \frac{h}{4}$ and $x_i + \frac{h}{4}$. We display the numerical results on the non-uniform mesh only because the results on the uniform mesh is similar, yet better. In Figure 3.3, we present the linear part of the numerical solution $u_L$ together with the exact solution $u$ for $\epsilon = 10^{-2}$ and various intensities of reaction ($\sigma = 0.1, 1, 10, 20, 50, 100$). The corresponding numerical results for $\epsilon = 10^{-5}$ are reported in Figure 3.4.

**Experiment 2:** We turn our attention to a variable-coefficient case for the same range of the problem parameters. We set $\beta = \frac{x+1}{6}$ and decompose the domain into a uniform discretization of 20 elements. Two different source functions are tested and the numerical results are displayed in Figure 3.5-3.8.

**Experiment 3:** We consider a more interesting variable-coefficient case which exhibits an internal layer. We set $\beta = -2(2x-1)$ and $f(x) = 4(2x-1)$ and we decompose the domain into a uniform discretization of 25 elements. We report the corresponding numerical results in Figure 3.9-3.10.

In this chapter, we wanted to stabilize the CDR problem by the RFB functions, starting from an augmented discretization of piecewise linear continuous finite elements. In particular we wanted to mimic the stabilizing effect of the RFBs by approximating them on a specially chosen sub-grid and using these approximations in place of RFB functions. The basic idea behind the present choice of the sub-grid is based on minimizing the residual of a local problem in a special manner. In all three experiments, we report that the numerical results are in good agreement with the physical configuration of the problem for a wide range of parameters, even when the mesh is coarse. The transition from one regime to another is accurately captured by the algorithm. The related results

are also comparable with the one in  (Brezzi et al., 2003).  Therefore we may conclude that the pseudo RFBs retain the stability features of RFBs and provide a robust, yet a cheap numerical method. We finally note that the choice of sub-grid proposed herein coincides with the one suggested in  (Brezzi et al., 2003).
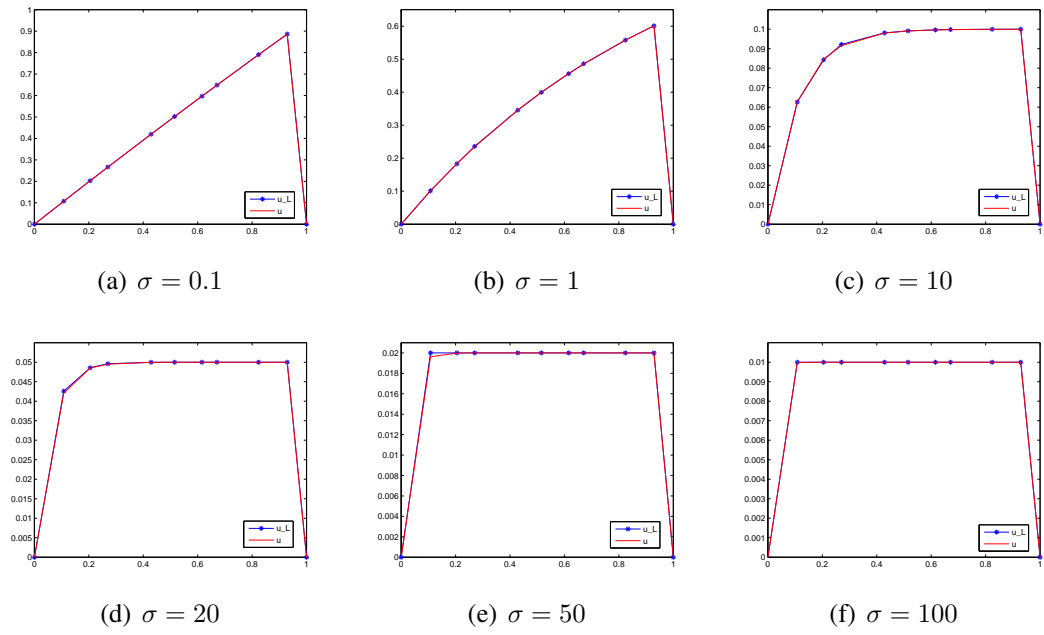
Figure 3.3. The linear part $u_L$ of the numerical solution and the exact solution $u$ for several values of $\sigma$ when $f(x) = 1$ and $\epsilon = 10^{-2}$.
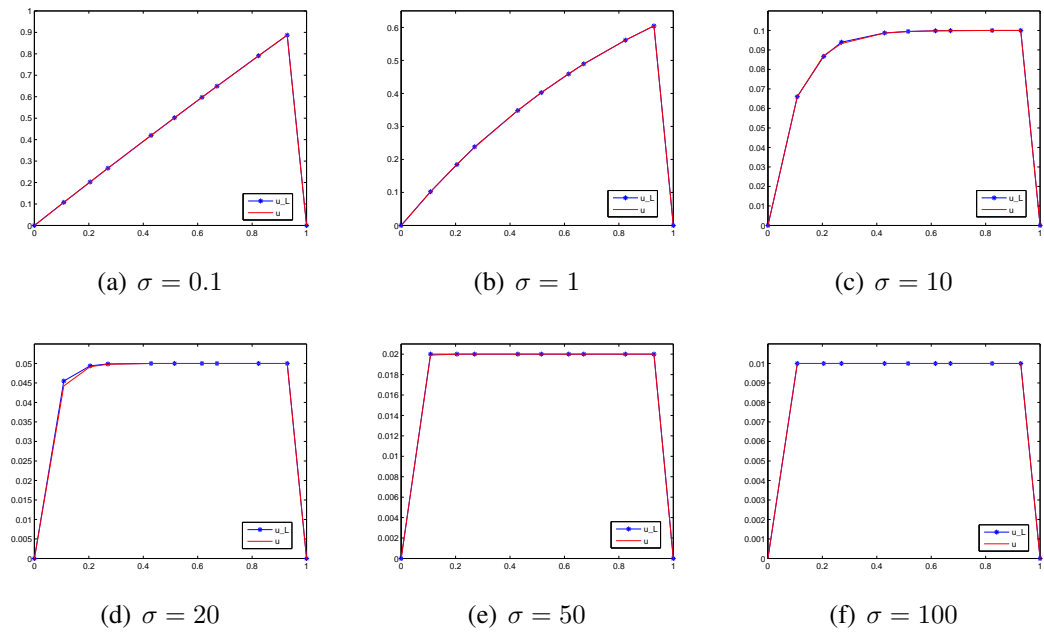


Figure 3.4. The linear part $u_L$ of the numerical solution and the exact solution $u$ for several values of $\sigma$ when $f(x) = 1$ and $\epsilon = 10^{-5}$.

Figure 3.5. The linear part $u_L$ of the numerical solution for several values of $\sigma$ when $f(x) = 1$, $\epsilon = 10^{-2}$ and $\beta = \frac{x+1}{6}$.



Figure 3.6. The linear part $u_L$ of the numerical solution for several values of $\sigma$ when $f(x) = 1$, $\epsilon = 10^{-5}$ and $\beta = \frac{x+1}{6}$.
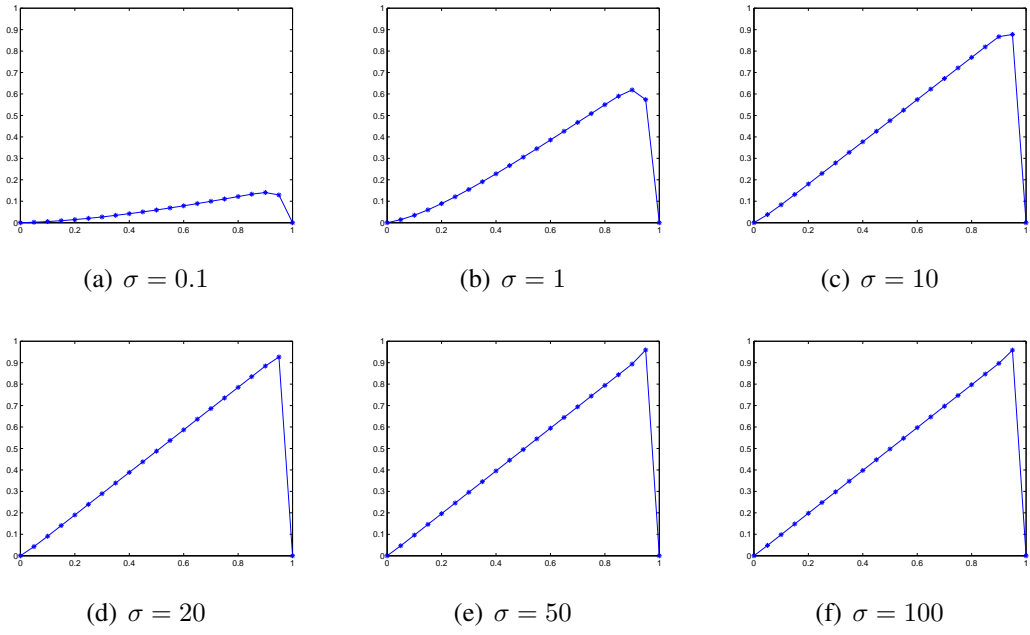
Figure 3.7. The linear part $u_L$ of the numerical solution for several values of $\sigma$ when $f(x) = \sigma x$, $\epsilon = 10^{-2}$ and $\beta = \frac{x+1}{6}$.
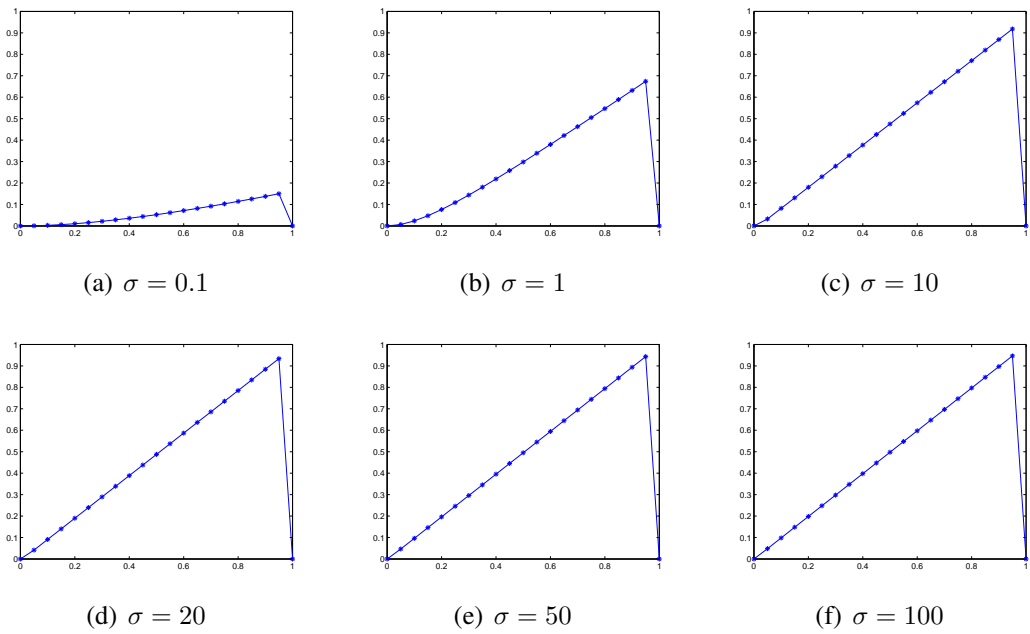


Figure 3.8. The linear part $u_L$ of the numerical solution for several values of $\sigma$ when $f(x) = \sigma x$, $\epsilon = 10^{-5}$ and $\beta = \frac{x+1}{6}$.

(a) $\sigma = 0.1$      (b) $\sigma = 1$      (c) $\sigma = 10$

(d) $\sigma = 20$      (e) $\sigma = 50$      (f) $\sigma = 100$
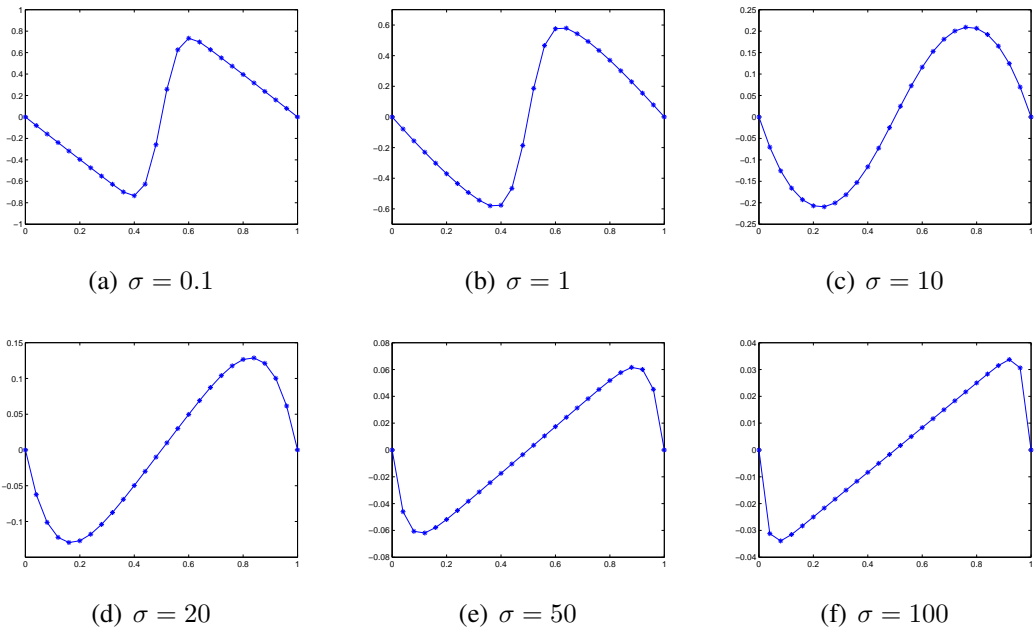
Figure 3.9. The linear part $u_L$ of the numerical solution for several values of $\sigma$ when $f(x) = 4(2x - 1)$, $\epsilon = 10^{-2}$ and $\beta = -2(2x - 1)$.



(a) $\sigma = 0.1$      (b) $\sigma = 1$      (c) $\sigma = 10$

(d) $\sigma = 20$      (e) $\sigma = 50$      (f) $\sigma = 100$
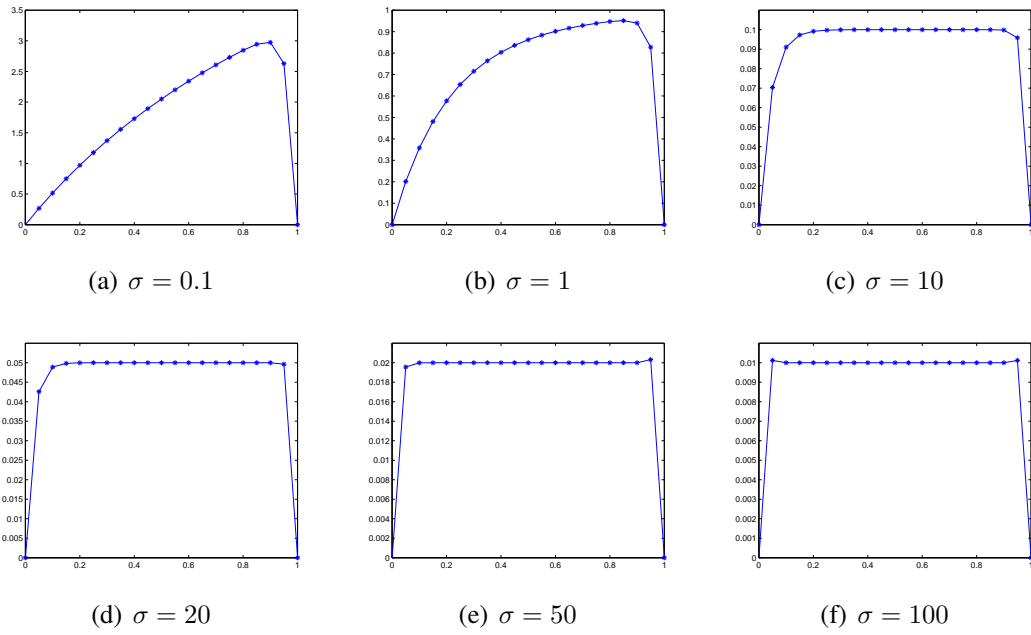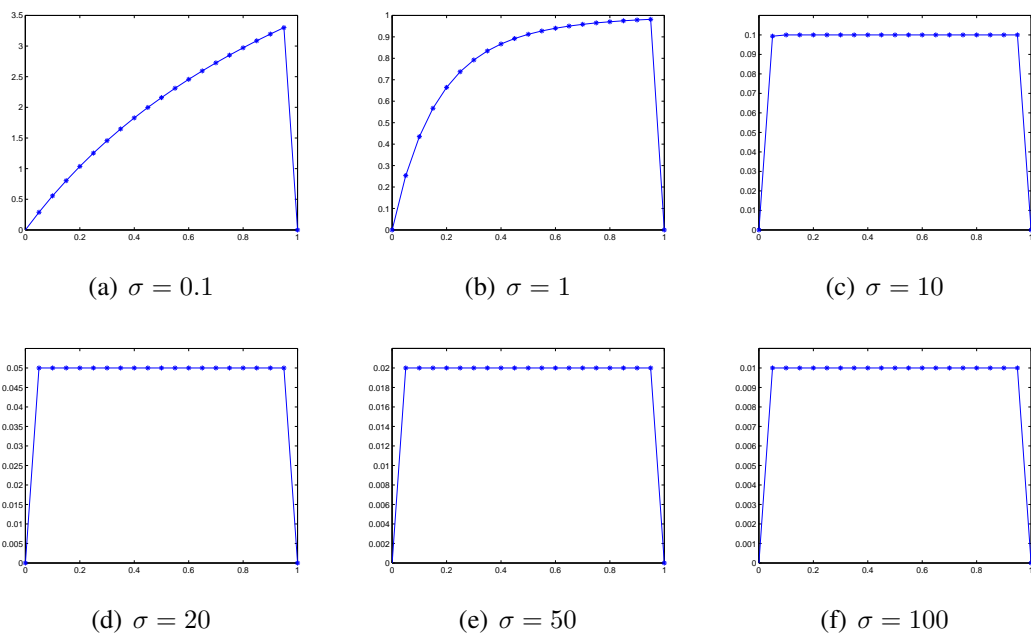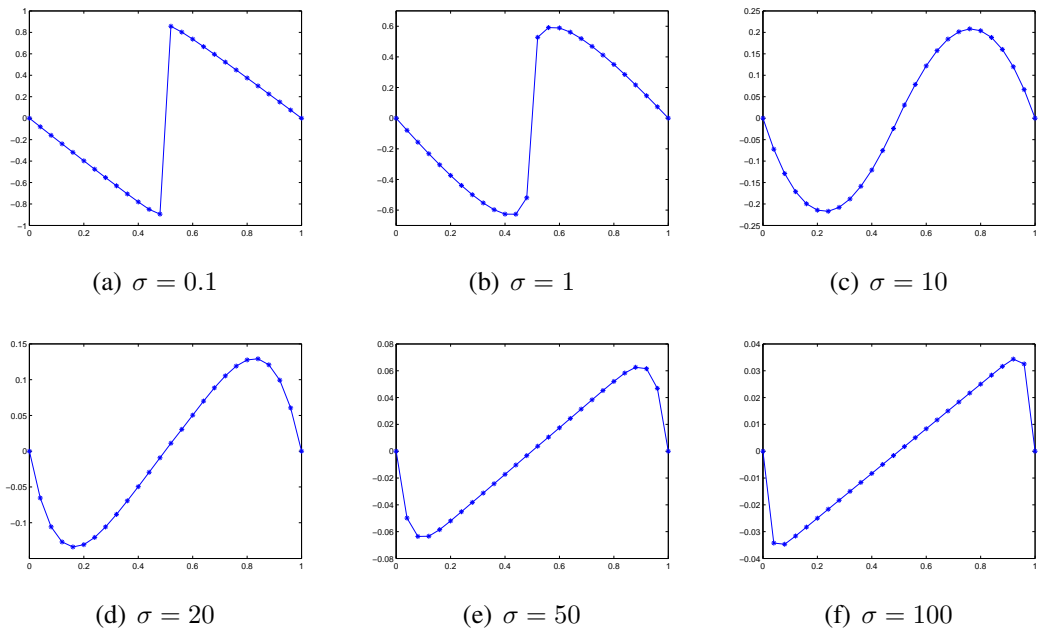
Figure 3.10. The linear part $u_L$ of the numerical solution for several values of $\sigma$ when $f(x) = 4(2x - 1)$, $\epsilon = 10^{-5}$ and $\beta = -2(2x - 1)$.

# CHAPTER 4

# A STABILIZED FINITE ELEMENT METHOD FOR CONVECTION-DIFFUSION-REACTION PROBLEMS IN 2-D

A stabilized finite element method is studied herein for two-dimensional convection diffusion-reaction problems. The method is based on the residual-free bubbles (RFB) method. However we replace the RFB functions by their cheap, yet efficient approximations, computed on a specially chosen subgrid, which retain the same qualitative behavior. Since the correct spot of subgrid points plays a crucial role in the approximation, it is important to determine their optimal location, which we do it through a minimization process with respect to $L_1$ norm. The resulting numerical method has similar stability features with the RFB method for the whole range of problem parameters and this fact is also confirmed by numerical experiments. We also note that the location of the sub-grid nodes suggested by the strategy herein has a similar structure to the one suggested by Brezzi and his coworkers in a different framework for the convection-diffusion problems ( see  (Brezzi et al., 2005)).

## 4.1.  Introduction

The exact solution of CDR problems may exhibit layer structures, small subregions where the derivatives of the solution are very large, when some problem parameters are too big compared to other(s). The plain Galerkin method may not work for such problems on reasonable discretizations, producing unphysical oscillations. A variety of finite elements approaches have been proposed to deal with such situations. The Streamline-Upwind Petrov-Galerkin (SUPG) method and its variants are among the most popular ones, which are based on augmenting the variational formulation by mesh-dependent terms in order to gain control over the derivatives of the solution (see  (Brooks and Hughes, 1982),  (Franca et al., 1992),  (Hughes et al., 1989)). The great advantage of this approach is not only its generality, but also its error analysis can be performed in many cases of interest. However, it is generally unclear to select the optimal parameter

value in the SUPG method.

Augmenting the polynomial finite element space by so called bubble functions gives rise to another class of numerical methods whose relation to the stabilized methods was studied in (Baiocchi et al., 1993), (Brezzi et al., 1992), (Brezzi and Russo, 1994). It turns out that to find a more suitable value for the stabilization parameter in the SUPG method, we can use special type of functions, so called the residual-free bubbles (RFB), defined by a local problem posed inside each element. The RFB method also allows one to prove error bounds (see (Brezzi et al., 2000), (Sangalli, 2000)) and can be generalized to a wider class of problems (see (Brezzi et al., 1998b), (Franca and Russo, 1996)). However, it requires to solve a local differential equation which may not be easier than to solve the original one (see (Franca et al., 1998), (Franca and Tobiska, 2002)).

A more recent approach employing pseudo-bubble functions, the bubble functions approximated on a specially chosen subgrid with a few nodes, has been proposed in the context of the RFB method. The grid is constructed so that small scale-effect of the exact solution could be accurately represented in the numerical approximation, through the use of the those bubble functions. This methodology has shown to be very effective, not only for the one-dimensional advection-dominated problems, but also for the reaction-dominated ones (see (Brezzi et al., 2003)). Alternatively, such grid points can also be constructed by minimizing the residual of local differential problems with respect to $L_1$ norm (see (Şendur and Neslitürk, 2012)).

Here, we will present a stabilization method for two-dimensional CDR problems, particularly designed to treat the interesting case of small diffusion, but able to adapt from one regime to anothers continuously. It is based on the RFB method in which the RFB functions are replaced by pseudo RFBs, which retain the same qualitative behavior as the RFBs. Similar approaches employing suitable approximations to the RFBs can be found in the literature (see (Brezzi et al., 1998a), (Brezzi and Marini, 2002a), (Brezzi et al., 2005), (Neslitürk, 2006), (Neslitürk, 2010)). The pseudo bubbles are computed by using a suitable subgrid inside each element. The choice of the subgrid nodes depends on minimizing the residual of a local differential problems with respect to $L_1$ norm. The resulting numerical method has similar stability features to the RFB method for the whole range of problem parameters. This fact is also confirmed by numerical experiments.

## 4.2. Statement of the Problem

We will consider the following linear elliptic CDR problem in a polygonal domain $\Omega$:

$$\begin{cases} \mathcal{L}u = -\epsilon \Delta u + \beta \cdot \nabla u + \sigma u = f & \text{in } \Omega\,, \\ u = 0 & \text{on } \partial\Omega\,. \end{cases} \tag{4.1}$$

Let $\mathcal{T}_h = \{K\}$ be a decomposition of $\Omega$ into triangles $K$, and let $h_K = \operatorname{diam}(K)$ with $h = \max_{K \in \mathcal{T}_h} h_K$. We assume that the diffusion coefficient $\epsilon$ is a positive constant and that the convection field $\beta$ and the reaction field $\sigma$ are non-negative piecewise constants with respect to the decomposition $\mathcal{T}_h$. Here we will consider stabilizations based on the augmented space idea with special emphasis on the RFB strategy. We start by recalling the abstract variational formulation of the problem (4.1): Find $u \in H_0^1(\Omega)$ such that

$$a(u, v) = (f, v) \quad \forall v \in H_0^1(\Omega)\,, \tag{4.2}$$

where

$$a(u, v) = \epsilon \int_\Omega \nabla u \cdot \nabla v + \int_\Omega (\beta \cdot \nabla u)\, v + \int_\Omega \sigma\, u\, v\,. \tag{4.3}$$

We now define $V_h$ as a finite-dimensional space, which is a subspace of $H_0^1(\Omega)$. Then the standard Galerkin finite element method reads: Find $u_h \in V_h$ such that

$$a(u_h, v_h) = (f, v_h) \quad \forall v_h \in V_h\,. \tag{4.4}$$

Now, we decompose the space $V_h$ such that $V_h = V_L \bigoplus V_B$, where $V_L$ is the space of continuous piecewise linear polynomials and $V_B = \bigoplus_K B_K$ with $B_K = H_0^1(K)$ Then every $v_h \in V_h$ can uniquely be written in the form of $v_h = v_L + v_B$, where $v_L \in V_L$ and $v_B \in V_B$. We require the bubble component $u_B$ of $u_h$ to satisfy the original differential equations in $K$ strongly, i.e.

$$\mathcal{L}u_B = -\mathcal{L}u_L + f \quad \text{in } K\,, \tag{4.5}$$

subject to the boundary condition,

$$u_B = 0 \quad \text{on } \partial K \,. \tag{4.6}$$

By the classical static condensation procedure (Brezzi et al., 2003), the method used to compute an improved linear approximation due to the residual-free bubble effect reads: Find $u_h = u_L + u_B$ in $V_h$ such that

$$a(u_L, v_L) + a(u_B, v_L) = (f, v_L) \quad \forall v_L \in V_L \,. \tag{4.7}$$

The term $a(u_B, v_L)$ is responsible for the stabilization of the numerical method and the bubble component $u_B$ should be computed before we solve (4.7) for its linear part. Recall $u_B$ is identified by the linear part $u_L$ and the source function $f$ through (4.5)-(4.6), which may be as complicated as solving the original differential equation. Therefore, it is important to bring a simple recipe about to obtain a suitable approximation to the bubble component of the problem that provides a similar stabilizing effect into the numerical method. Regarding the simplicity of element geometry, this approach can be turned into a workable method, whose details are given in the following section.

## 4.3. Description of the Method

Let $P_i$, $(i = 1, 2, 3)$ be an internal point of $K$, for which we will construct a pseudo bubble function. The quality of approximate bubble functions, which is crucial to get a good stabilization effect on the numerical method, is essentially related with the location of internal points. Therefore the choice of those points must be fulfilled in a special manner. That will be accomplished through a minimization process with respect to $L_1$ norm in the presence of layers.

To be more descriptive, we consider bubble functions $B_i$, $(i = 1, 2, 3)$ defined by

$$\mathcal{L} B_i = -\mathcal{L} \psi_i \quad \text{in } K, \qquad B_i = 0 \quad \text{on } \partial K, \quad i = 1, 2, 3 \,, \tag{4.8}$$

where $\psi_i$ are the restrictions of the piecewise linear basis functions for $V_L$ to $K$. Further

we define $B_f$,

$$\mathcal{L}B_f = f \quad \text{in } K, \qquad B_f = 0 \quad \text{on } \partial K. \tag{4.9}$$

Now if

$$u_L\big|_K = \sum_{i=1}^{3} c_i \psi_i \,, \tag{4.10}$$

then

$$u_B\big|_K = \sum_{i=1}^{3} c_i B_i + B_f \,, \tag{4.11}$$

with the same coefficient $c_i$. Thus

$$\mathcal{L}u_B = -\mathcal{L}u_L + f \quad \text{in } K \tag{4.12}$$

That is, the equation (4.5) is automatically satisfied with the present choice of bubble functions. The equation (4.8) is similar to the original problem (4.1) and may be difficult to solve. However, using the element geometry and the problem properties, it is possible to construct a cheap, yet efficient approximate bubbles, say $B_i^*$, having the same qualitative behavior with its continuous counterpart. The construction of such approximate bubble functions $B_i^*$ is introduced in the following.

Let $b_i$ be a piecewise linear function with

$$b_i(V_j) = 0 \text{ and } b_i(P_i) = 1 \quad \forall i, j = 1, 2, 3 \,, \tag{4.13}$$

where $V_i$ are the vertices of $K$. Further let $B_i^*(x) = \alpha_i b_i(x)$ be the classical Galerkin approximation of $B_i$ through (4.8)

$$a(B_i^*, b_i)_K = (-\mathcal{L}\psi_i, b_i)_K \quad , \quad i = 1, 2, 3 \,. \tag{4.14}$$

Using (4.14), an easy computation gives :

$$\alpha_i = \frac{(-\mathcal{L}\psi_i, b_i)}{\epsilon||\nabla b_i||_K^2 + \sigma||b_i||_K^2} \quad i = 1, 2, 3.$$ (4.15)

The main criteria that we use to determine the locations of the internal points is to minimize $L_1$ norm of the residual coming out from the bubble equation (4.8) in the critical case where a layer structure exists. In other words, we choose $P_i$ such that

$$J_i = \int_K \left| \mathcal{L}B_i^* + \mathcal{L}\psi_i \right|, \quad i = 1, 2, 3,$$ (4.16)

is minimum. Before we derive the explicit locations of the internal points, additional notation related to the element geometry should be introduced first. We denote the edges of $K$ by $e_i$ opposite to $V_i$, the length of $e_i$ by $|e_i|$, the midpoint of edge $e_i$ by $M_i$, the outward unit normal to $e_i$ by $n^i$, $\nu_i = |e_i|n^i$ and $\beta_{\nu_i} = (\beta \, , \, \nu_i)$ (see Figure 4.1). We further introduce the three sub-triangles $K_{i1}, K_{i2}, K_{i3}$ which are obtained by connecting the additional node $P_i$ with the vertices of $K$ and the area of $K_{ij}$ by $|K_{ij}|$. The actual numbering of the vertices will be chosen according to the direction of $\beta$. Now, let us choose the location of $P_i$ along the median from $V_i$, that is

$$P_i = (1 - t_i)M_i + t_i V_i \, , \quad 0 < t_i < 1 \, , \quad i = 1, 2, 3 \, .$$ (4.17)
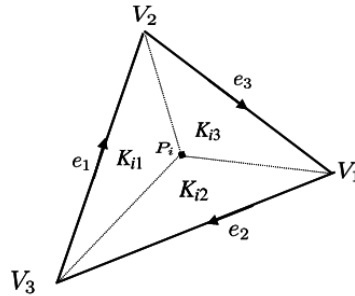


Figure 4.1. Splitting $K$ into three sub-regions by $P_i$.

In order to specify the problem regimes, we further introduce

$$\epsilon_i^* = \frac{2|K|(-3\beta_{\nu_i} + \sigma|K|)}{9(|e_1|^2 + |e_2|^2 + |e_3|^2)} \qquad i = 1, 2, 3. \tag{4.18}$$

To be able to determine the explicit locations of the internal points, we have to distinguish among the following cases:

### 4.3.1. Convection-Dominated Flows
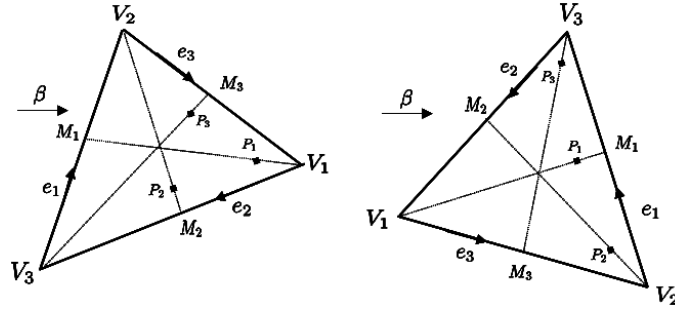


Figure 4.2. Configuration of internal nodes for convection-dominated regime: one inflow edge (left) and two inflow edges (right).

**Case 1. The inflow boundary is made of one edge of K:** Let the inflow boundary make up of one edge and let $e_1$ be the inflow one. Then, we have $\beta_{\nu_1} < 0$, $\beta_{\nu_2} > 0$ and $\beta_{\nu_3} > 0$. We assume that the problem is convection-dominated if

$$\epsilon \leq \epsilon_1^* \quad \text{with} \quad 2\sigma|\text{K}| < \min\{\beta_{\nu_2}, \beta_{\nu_3}\}. \tag{4.19}$$

Now, split K into three subregions by using the definition of $P_1$ in (4.17), so that (see Figure 4.1)

$$|K_{11}| = t_1|K| \qquad |K_{12}| = |K_{13}| = (1 - t_1)|K|/2. \tag{4.20}$$

Moreover, we have

$$\psi_1 = \frac{|K_{11}|}{|K|} b_1\big|_{K_{11}} \quad \text{and} \quad \nabla b_1\big|_{K_{1i}} = -|e_i|n_i/2|K_{1i}|, \quad i = 1, 2, 3. \tag{4.21}$$

Further, define

$$l_{ijk}(t) = |e_i|^2 + t|e_j - e_k|^2,$$
$$f_{ijk}(t) = 3\epsilon l_{ijk}(t) + 2(1-t)t\sigma|K|^2 \qquad i, j, k = 1, 2, 3. \tag{4.22}$$

Note that $l_{ijk}(t) \sim O(h^2)$ for any $t \in (0, 1)$. Now, using the properties of bubble function $b_1$ and the midpoint rule for quadratic terms that appears, we get an explicit expression for $\alpha_1$ depending on the parameter $t_1$:

$$\alpha_1 = \frac{(-\mathcal{L}\psi_1, b_1)}{\epsilon||\nabla b_1||_K^2 + \sigma||b_1||_K^2} = -\frac{(S_1 + t_1\sigma|K|)/12}{\frac{\epsilon\, l_{123}(t_1)}{4(1-t_1)t_1|K|} + \frac{\sigma|K|}{6}} = -\frac{t_1(1-t_1)|K|(S_1 + t_1\sigma|K|)}{f_{123}(t_1)}, \tag{4.23}$$

where $S_1 = -2\beta_{\nu_1} + \sigma|K|$. Note that $S_1 > 0$ and $\alpha_1 < 0$. Before finding an optimal position for $P_1$, we need the following intermediate result.

**Lemma 4.3.1.1** $\int_K |-\epsilon\Delta B_1^*|$ *is an increasing function of* $t_1$ *in the interval* $(1/2, 1)$.
**Proof** Following the lines of (Brezzi et al., 1998a), we have

$$\int_K |-\epsilon\Delta B_1^*| = -\epsilon\,\alpha_1 \sum_{i=1}^{3} \frac{|e_i|^2}{2|K_{1i}|}. \tag{4.24}$$

Define

$$\tilde{f}_{ijk}(t) = 2|K|\left( S_i\left((2t-1)|e_i|^2 + t^2|e_j - e_k|^2\right) + 2t^2\sigma|K|(|e_j|^2 + |e_k|^2) \right).$$

Now, use the expression for $\alpha_1$ in (4.23) to get

$$\frac{d}{dt_1} \int_K |-\epsilon \Delta B_1^*| = -\frac{d}{dt_1}\left( \epsilon\, \alpha_1 \left( \frac{|e_1|^2}{2\,|K_{11}|} + \frac{|e_2|^2}{2\,|K_{12}|} + \frac{|e_3|^2}{2\,|K_{13}|} \right)\right)$$

$$= \epsilon \frac{d}{dt_1}\left( \frac{t_1\,(1-t_1)\,(S_1 + t_1\sigma|K|)}{f_{123}(t_1)} \left( \frac{|e_1|^2}{2\,t_1} + \frac{|e_2|^2}{1-t_1} + \frac{|e_3|^2}{1-t_1} \right)\right)$$

$$= \epsilon \frac{d}{dt_1}\left( \frac{S_1 + t_1\sigma|K|}{f_{123}(t_1)} \left( \frac{1-t_1}{2}|e_1|^2 + t_1|e_2|^2 + t_1|e_3|^2 \right)\right)$$

$$= \epsilon\sigma|K|\frac{3\epsilon l_{123}^2(t_1) + \tilde{f}_{123}(t_1)}{2 f_{123}^2(t_1)}\,. \tag{4.25}$$

Hence the result immediately follows since $f_{123}(t_1) > 0$ and $\tilde{f}_{123}(t_1) > 0$ whenever $1/2 < t_1 < 1$. $\qquad\square$

The following lemma suggests an optimal position for $P_1$ along the median from $V_1$ by using (4.16).

**Lemma 4.3.1.2** *If the inflow boundary make up of one edge, then the point*

$$t_1^* = 1 - \frac{-\rho_1 + \sqrt{\rho_1^2 + \lambda_1}}{2\sigma\,|K|^2}\,,$$

*minimizes the integral $J_1$ in convection-dominated flows where*

$$\rho_1 = -2\beta_{\nu_1}|K| + 3\epsilon\,|e_2 - e_3|^2\,, \quad \lambda_1 = 24\epsilon\sigma|K|^2\left(|e_2|^2 + |e_3|^2\right).$$

**Proof**   It is possible to rewrite the integral $J_1$ in (4.16) as follows (see (Brezzi et al., 1998a)):

$$J_1 = \int_K |-\epsilon \Delta B_1^*| + \sum_{i=1}^{3} \int_{K_{1i}} |\beta \cdot \nabla B_1^* + \sigma B_1^* + \beta \cdot \nabla \psi_1 + \sigma \psi_1|\,. \tag{4.26}$$

Let $g_1 = \beta \cdot \nabla B_1^* + \sigma B_1^* + \beta \cdot \nabla \psi_1 + \sigma \psi_1$. When an appropriate form of $\alpha_1$ in (4.23) is substituted, we get

$$\int_K g_1 = -\frac{1}{2}\beta_{\nu_1} + \frac{\sigma|K|}{3}(\alpha_1 + 1)$$

$$= -\frac{1}{2}\beta_{\nu_1} + \frac{\sigma|K|}{3}\left(-\frac{t_1(1-t_1)|K|(S_1 + t_1\sigma|K|)}{f_{123}(t_1)} + 1\right)$$

$$= \frac{3\epsilon\, l_{123}(t_1)\left(-3\beta_{\nu_1} + 2\sigma|K|\right) + 2(1-t_1)t_1\,\sigma|K|^2\left(-\beta_{\nu_1} + (1-t_1)\sigma|K|\right)}{6f_{123}(t_1)},$$

where the last expression is obviously positive. Now, split $K$ into three subregions by $P_1$ (see Figure 4.1) and investigate the sign of $g_1$ over each of these sub-domains to get:

$$g_1\big|_{K_{12}} = (\beta \cdot \nabla B_1^* + \sigma B_1^* + \beta \cdot \nabla \psi_1 + \sigma \psi_1)\big|_{K_{12}}$$

$$= (\alpha_1 \beta \cdot \nabla b_1 + \alpha_1 \sigma b_1 + \beta \cdot \nabla \psi_1 + \sigma \psi_1)\big|_{K_{12}}$$

$$= \alpha_1 \beta \cdot \nabla b_1\big|_{K_{12}} + \alpha_1 \sigma b_1\big|_{K_{12}} + \frac{|K_{11}|}{|K|}\beta \cdot \nabla b_1\big|_{K_{11}} + \sigma \psi_1$$

$$= -\frac{\alpha_1}{2|K_{12}|}\beta_{\nu_2} + \alpha_1 \sigma b_1\big|_{K_{12}} - \frac{\beta_{\nu_1}}{2|K|} + \sigma \psi_1$$

$$> -\frac{\alpha_1}{2|K_{12}|}\beta_{\nu_2} + \alpha_1 \sigma - \frac{\beta_{\nu_1}}{2|K|} + \sigma \psi_1 \quad (\text{since } \max_{\mathbf{x}\in K_{12}} b_1(\mathbf{x}) < 1)$$

$$> -\frac{\alpha_1}{2|K_{12}|}\beta_{\nu_2} + \alpha_1 \sigma \quad (\text{since } \max_{\mathbf{x}\in K_{12}} \psi_1(\mathbf{x}) > 0 \text{ and } \beta_{\nu_1} < 0)$$

$$= -\alpha_1\left(\frac{\beta_{\nu_2}}{2|K_{12}|} - \sigma\right), \tag{4.27}$$

which is positive by condition (4.19). Similarly, we have

$$g_1\big|_{K_{13}} = (\beta \cdot \nabla B_1^* + \sigma B_1^* + \beta \cdot \nabla \psi_1 + \sigma \psi_1)\big|_{K_{13}}$$

$$= (\alpha_1 \beta \cdot \nabla b_1 + \alpha_1 \sigma b_1 + \beta \cdot \nabla \psi_1 + \sigma \psi_1)\big|_{K_{13}}$$

$$= \alpha_1 \beta \cdot \nabla b_1\big|_{K_{13}} + \alpha_1 \sigma b_1\big|_{K_{13}} + \frac{|K_{11}|}{|K|}\beta \cdot \nabla b_1\big|_{K_{11}} + \sigma \psi_1$$

$$= -\frac{\alpha_1}{2|K_{13}|}\beta_{\nu_3} + \alpha_1 \sigma b_1\big|_{K_{13}} - \frac{\beta_{\nu_1}}{2|K|} + \sigma \psi_1$$

$$> \quad -\frac{\alpha_1}{2|K_{13}|}\beta_{\nu_3} + \alpha_1\sigma - \frac{\beta_{\nu_1}}{2|K|} + \sigma\psi_1 \quad (\text{since } \max_{\mathbf{x}\in K_{13}} b_1(\mathbf{x}) < 1)$$

$$> \quad -\frac{\alpha_1}{2|K_{13}|}\beta_{\nu_3} + \alpha_1\sigma \quad (\text{since } \max_{\mathbf{x}\in K_{13}} \psi_1(\mathbf{x}) > 0 \text{ and } \beta_{\nu_1} < 0)$$

$$= \quad -\alpha_1\left(\frac{\beta_{\nu_3}}{2|K_{13}|} - \sigma\right), \tag{4.28}$$

which is also positive. Thus, the second term on the right hand side of (4.26) attains its minimum if $g_1|_{K_{11}}$ is positive, too:

$$g_1|_{K_{11}} \quad = \quad (\beta \cdot \nabla B_1^* + \sigma B_1^* + \beta \cdot \nabla \psi_1 + \sigma\psi_1)|_{K_{11}}$$

$$= \quad (\alpha_1\beta \cdot \nabla b_1 + \alpha_1\sigma b_1 + \beta \cdot \nabla \psi_1 + \sigma\psi_1)|_{K_{11}}$$

$$= \quad \alpha_1\beta \cdot \nabla b_1 |_{K_{11}} + \alpha_1\sigma b_1 |_{K_{11}} + \frac{|K_{11}|}{|K|}\beta \cdot \nabla b_1 |_{K_{11}} + \sigma\psi_1$$

$$= \quad -\frac{\alpha_1}{2|K_{11}|}\beta_{\nu_1} + \alpha_1\sigma b_1 |_{K_{11}} - \frac{\beta_{\nu_1}}{2|K|} + \sigma\psi_1$$

$$= \quad (\alpha_1 + t_1)\left(-\frac{\beta_{\nu_1}}{2|K|t_1} + \sigma b_1 |_{K_{11}}\right). \tag{4.29}$$

Now, use the expression for $\alpha_1$ in (4.23) to get

$$\alpha_1 + t_1 = -\frac{t_1\sigma|K|^2\left(t_1 - 1 + \frac{-\rho_1 + \sqrt{\rho_1^2 + \lambda_1}}{2\sigma|K|^2}\right)\left(t_1 - 1 - \frac{\rho_1 + \sqrt{\rho_1^2 + \lambda_1}}{2\sigma|K|^2}\right)}{3\epsilon l_{123}(t_1) + 2t_1(1 - t_1)\sigma|K|^2}.$$

The only root of the expression (4.29) in the interval $(0, 1)$ is $t_1^* = 1 - \frac{-\rho_1 + \sqrt{\rho_1^2 + \lambda_1}}{2\sigma|K|^2}$ and $g_1|_{K_{11}}$ is positive if $t_1 \geq t_1^*$. This fact together with Lemma 4.3.1.1 determines an optimal value for $t_1$. $\qquad\square$

**Remark 4.1** *As actual value for $t_1$, we do not always take that given by Lemma 4.3.1.2. Indeed, for $\epsilon$ not too small (that is, for diffusion dominated problems) this type of stabi-*

*lization would be unnecessary. Furthermore the value provided by Lemma 4.3.1.2 can even be meaningless. Therefore, we take*

$$
\begin{cases}
t_1 = t_1^*, & \text{if } \epsilon \leq \epsilon_1^*, \\
t_1 = 1/3, & \text{otherwise.}
\end{cases}
\tag{4.30}
$$

*which also gives a continuous dependence of $t_1$ upon $\epsilon$. Notice that for $\epsilon = \epsilon_1^*$ we have exactly $t_1 = t_1^* = 1/3$ and for $0 < \epsilon < \epsilon_1^*$ we have $1 > t_1^* > 1/3$. Therefore for every $\epsilon > 0$, we have*

$$
\frac{1}{3} \leq t_1 < 1.
\tag{4.31}
$$

*This can be seen as a weaker assumption than we made in Lemma 4.3.1.1, that was $t_1 \in (1/2, 1)$. However, it seems just as a technicality as we get reasonable results for values of $t_1$ in $(1/3, 1)$.*

**Remark 4.2** *We further observed that*

$$
\lim_{\sigma \to 0} t_1^* = 1 - \frac{6\epsilon \left(|e_2|^2 + |e_3|^2\right)}{-2\beta_{\nu_1}|K| + 3\epsilon |e_2 - e_3|^2},
\tag{4.32}
$$

*which has a similar structure to the one suggested by Brezzi and his coworkers in a different framework for the convection-diffusion problems ( see (Brezzi et al., 2005)).*

**Remark 4.3** *The value of $\alpha_1$ at $t_1^*$ is simply equal to $-t_1^*$.*

**Remark 4.4** *For convection dominated regime, the choice of other points, $P_2$ and $P_3$, should be consistent with the physics of the problem. Thus we take*

$$
t_1 = t_1^*, \, t_2 = t_3 = \min\{1/3, 1 - t_1^*\}.
\tag{4.33}
$$

**Remark 4.5** *In Figure 4.3, we display the behavior of approximate bubble functions in a typical element $K$ with one inflow edge, for various intensities of diffusion ($\epsilon = 10^{-2}, 10^{-3}, 10^{-4}$) when $\theta = 72^o$ and $\sigma = 0.001$. The first column of the figure presents the*
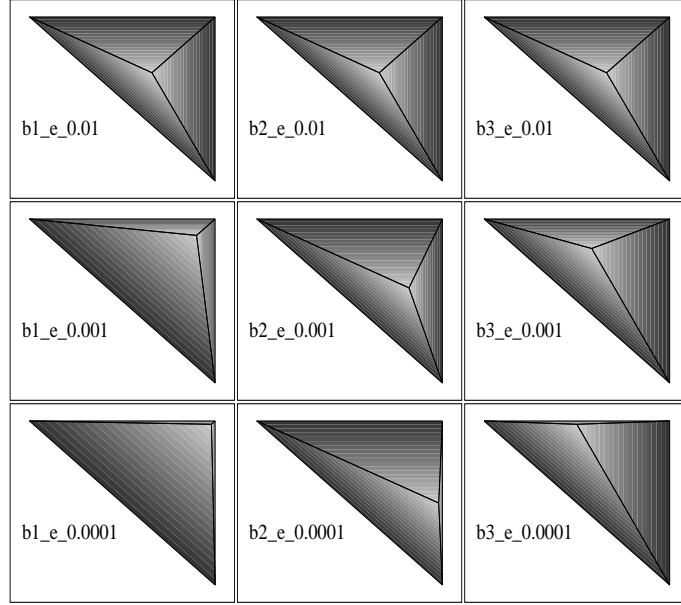
Figure 4.3. Bubble functions, $b_1$, $b_2$, $b_3$, in a typical element with one inflow edge, when $\theta = 72^o$, $N = 20$, $\epsilon = 10^{-2}, 10^{-3}, 10^{-4}$.

*bubble function, $b_1$, for decreasing values of diffusion ($\epsilon = 10^{-2}, 10^{-3}, 10^{-4}$). The corresponding numerical results for $b_2$ and $b_3$ are shown in column 2 and 3, respectively. Note that the self-adjustment of the subgrid node as the problem evolves from the diffusion-dominated regime into the convection-dominated one. The location is readjusted so that the pseudo bubbles contributes to the stability of the numerical method at its most.*

**Case 2.** **The inflow boundary is made of two edges of K:** Let the inflow boundary make up of two edges and let $e_2$ and $e_3$ be the inflow ones. Then, we have $\beta_{\nu_1} > 0$, $\beta_{\nu_2} < 0$ and $\beta_{\nu_3} < 0$. In this case, we assume that the problem is convection-dominated if

$$\epsilon \leq \min\{\epsilon_2^*, \epsilon_3^*\} \quad \text{with} \quad 2\sigma|K| < \beta_{\nu_1}. \tag{4.34}$$

Now, we will construct the explicit locations of $P_2$ and $P_3$, separately:

**Construction of $P_2$:** We determine the location of $P_2$ along the median from $V_2$ by condition (4.16) and split K into three subregions by using the definition of $P_2$ in (4.17), so that we have (see Figure 4.1)

$$|K_{21}| = |K_{23}| = (1 - t_2)|K|/2 \qquad |K_{22}| = t_2|K|. \tag{4.35}$$

Moreover, we have

$$\psi_2 = \frac{|K_{22}|}{|K|} b_2 \big|_{K_{22}} \quad \text{and} \quad \nabla b_2 \big|_{K_{2i}} = -|e_i| n_i / 2 |K_{2i}|, \ \ i = 1, 2, 3. \qquad (4.36)$$

Similarly, we get an explicit expression for $\alpha_2$:

$$\alpha_2 = \frac{(-\mathcal{L}\psi_2, b_2)}{\epsilon ||\nabla b_2||_K^2 + \sigma ||b_2||_K^2} = -\frac{(S_2 + t_2 \sigma |K|)/12}{\frac{\epsilon l_{213}(t_2)}{4(1-t_2)t_2|K|} + \frac{\sigma|K|}{6}} = -\frac{t_2(1-t_2)|K|(S_2 + t_2 \sigma |K|)}{f_{213}(t_2)} \qquad (4.37)$$

where $S_2 = -2\beta_{\nu_2} + \sigma|K|$. Note that $S_2 > 0$ and $\alpha_2 < 0$. Before finding an optimal position for $P_2$, we need the following intermediate result.

**Lemma 4.3.1.3** $\displaystyle\int_K |-\epsilon\Delta B_2^*|$ *is an increasing function of $t_2$ in the interval $(1/2, 1)$.*
**Proof** Following the lines of Lemma 4.3.1.1, we have

$$\int_K |-\epsilon\Delta B_2^*| = -\epsilon\,\alpha_2 \sum_{i=1}^{3} \frac{|e_i|^2}{2\,|K_{2i}|}. \qquad (4.38)$$

With the use of an explicit form of $\alpha_2$ in (4.37), we further have

$$\frac{d}{dt_2} \int_K |-\epsilon\Delta B_2^*| = \epsilon\sigma|K| \frac{3\epsilon l_{213}^2(t_2) + \tilde{f}_{213}(t_2)}{2 f_{213}^2(t_2)}. \qquad (4.39)$$

The result immediately follows since $f_{213}(t_2) > 0$ and $\tilde{f}_{213}(t_2) > 0$ whenever $1/2 < t_2 < 1$. $\qquad\qquad \square$

**Lemma 4.3.1.4** *Let $a, b, c$ be real numbers such that $a + b + c > 0$, $a > 0$ and $c < 0$. Then $|a| + |b| + |c|$ attains its minimum if $b > 0$.*
**Proof** Since $c < 0$ we have $a + b - c > a + b + c > 0$ and therefore $a - c > -b$. Then, the result immediately follows since

$$|a| + |b| + |c| = a + |b| - c > -b + |b| = \begin{cases} -2b & \text{if } b < 0, \\ 0 & \text{if } b > 0. \end{cases} \qquad (4.40)$$

$\square$

The following lemma suggests an optimal position for $P_2$ by using (4.16).

**Lemma 4.3.1.5** *If the inflow boundary make up of two edges then the point*

$$t_2^* = 1 - \frac{-\rho_2 + \sqrt{\rho_2^2 + \lambda_2}}{2\sigma |K|^2},$$

*minimizes the integral $J_2$ in convection-dominated flows where*

$$\rho_2 = -2\beta_{\nu_2}|K| + 3\epsilon \, |e_1 - e_3|^2, \quad \lambda_2 = 24\epsilon\sigma|K|^2\left(|e_1|^2 + |e_3|^2\right).$$

**Proof** The proof is similar to the one of Lemma 4.3.1.2. Let us recall that

$$J_2 = \int_K |-\epsilon\Delta B_2^*| + \sum_{i=1}^3 \int_{K_{2i}} |\beta \cdot \nabla B_2^* + \sigma B_2^* + \beta \cdot \nabla\psi_2 + \sigma\psi_2|. \qquad (4.41)$$

Let $g_2 = \beta \cdot \nabla B_2^* + \sigma B_2^* + \beta \cdot \nabla\psi_2 + \sigma\psi_2$. A direct calculation over $K$ gives,

$$
\begin{aligned}
\int_K g_2 &= -\frac{1}{2}\beta_{\nu_2} + \frac{\sigma|K|}{3}(\alpha_2 + 1) \\[2mm]
&= \frac{3\epsilon\, l_{213}(t_2)\left(-3\beta_{\nu_2} + 2\sigma|K|\right) + 2(1 - t_2)t_2\,\sigma|K|^2\left(-\beta_{\nu_2} + (1 - t_2)\sigma|K|\right)}{6 f_{213}(t_2)},
\end{aligned}
$$

where the last expression is obviously positive. Now split $K$ into three subregions by $P_2$ and investigate the sign of $g_2$ over each of these sub-domains:

$$
\begin{aligned}
g_2\big|_{K_{21}} &= (\beta \cdot \nabla B_2^* + \sigma B_2^* + \beta \cdot \nabla\psi_2 + \sigma\psi_2)\big|_{K_{21}} \\[2mm]
&= (\alpha_2\beta \cdot \nabla b_2 + \alpha_2\sigma b_2 + \beta \cdot \nabla\psi_2 + \sigma\psi_2)\big|_{K_{21}} \\[2mm]
&= \alpha_2\beta \cdot \nabla b_2\big|_{K_{21}} + \alpha_2\sigma b_2\big|_{K_{21}} + \frac{|K_{22}|}{|K|}\beta \cdot \nabla b_2\big|_{K_{22}} + \sigma\psi_2
\end{aligned}
$$

$$= -\frac{\alpha_2}{2|K_{21}|}\beta_{\nu_1} + \alpha_2\sigma b_2\big|_{K_{21}} - \frac{\beta_{\nu_2}}{2|K|} + \sigma\psi_2$$

$$> -\frac{\alpha_2}{2|K_{21}|}\beta_{\nu_1} + \alpha_2\sigma - \frac{\beta_{\nu_2}}{2|K|} + \sigma\psi_2 \quad (\text{since } \max_{\mathbf{x}\in K_{21}} b_2(\mathbf{x}) < 1)$$

$$> -\frac{\alpha_2}{2|K_{21}|}\beta_{\nu_1} + \alpha_2\sigma \quad (\text{since } \max_{\mathbf{x}\in K_{21}} \psi_2(\mathbf{x}) > 0 \text{ and } \beta_{\nu_2} < 0)$$

$$= -\alpha_2\left(\frac{\beta_{\nu_1}}{2|K_{21}|} - \sigma\right), \tag{4.42}$$

which is positive by condition (4.34). Similarly, we have

$$g_2\big|_{K_{23}} = (\beta\cdot\nabla B_2^* + \sigma B_2^* + \beta\cdot\nabla\psi_2 + \sigma\psi_2)\big|_{K_{23}}$$

$$= (\alpha_2\beta\cdot\nabla b_2 + \alpha_2\sigma b_2 + \beta\cdot\nabla\psi_2 + \sigma\psi_2)\big|_{K_{23}}$$

$$= \alpha_2\beta\cdot\nabla b_2\big|_{K_{23}} + \alpha_2\sigma b_2\big|_{K_{23}} + \frac{|K_{22}|}{|K|}\beta\cdot\nabla b_2\big|_{K_{22}} + \sigma\psi_2$$

$$= -\frac{\alpha_2}{2|K_{23}|}\beta_{\nu_3} + \alpha_2\sigma b_2\big|_{K_{23}} - \frac{\beta_{\nu_2}}{2|K|} + \sigma\frac{|K_{22}|}{|K|}b_2\big|_{K_{22}}$$

$$< -\frac{\alpha_2}{2|K_{23}|}\beta_{\nu_3} - \frac{\beta_{\nu_2}}{2|K|} + \sigma\frac{|K_{22}|}{|K|} \quad (\text{since } 0 < \max_{\mathbf{x}\in K_{22}\cup K_{23}} b_2(\mathbf{x}) < 1)$$

$$= \frac{l_{213}(t_2)(\,\epsilon - \tilde{\epsilon}_2(t_2)\,)(-\beta_{\nu_2} + 2t_2\sigma|K|)}{2|K|\,f_{213}(t_2)}, \tag{4.43}$$

where $\tilde{\epsilon}_2(t) = \dfrac{2|K|t\left(-\beta_{\nu_3}(-2\beta_{\nu_2}+(1+t)\sigma|K|)-(1-t)\sigma|K|(-\beta_{\nu_2}+2t\sigma|K|)\right)}{3\,l_{213}(t)\left(-\beta_{\nu_2}+2t\sigma|K|\right)}$. We note that the sign of the previous expression only depends on the sign of $\epsilon - \tilde{\epsilon}_2(t_2)$. According to the observation shown below,

$$\tilde{\epsilon}_2(t) = \frac{2|K|t\left(-\beta_{\nu_3}(-2\beta_{\nu_2}+(1+t)\sigma|K|)-(1-t)\sigma|K|(-\beta_{\nu_2}+2t\sigma|K|)\right)}{3\,l_{213}(t)\left(-\beta_{\nu_2}+2t\sigma|K|\right)}$$

$$> \frac{|K|\left(-\beta_{\nu_3}(-2\beta_{\nu_2} + \frac{3}{2}\sigma|K|) - \sigma|K|(-\frac{\beta_{\nu_2}}{2} + \sigma|K|)\right)}{6\left(|e_1|^2 + |e_3|^2\right)\left(-\beta_{\nu_2} + 2\sigma|K|\right)} \quad \text{(since } 1/2 < t < 1\text{)}$$

$$> \frac{|K|\left(-2\beta_{\nu_2} + \frac{3}{2}\sigma|K|\right)\left(-\beta_{\nu_3} - \frac{2}{3}\sigma|K|\right)}{6\left(|e_1|^2 + |e_3|^2\right)\left(-\beta_{\nu_2} + 2\sigma|K|\right)}$$

$$> \frac{|K|\left(-\beta_{\nu_3} - \frac{2}{3}\sigma|K|\right)}{8\left(|e_1|^2 + |e_3|^2\right)}\,, \tag{4.44}$$

where in the third and fourth line we used the following inequalities, respectively

$$-\frac{\beta_{\nu_2}}{2} + \sigma|K| \quad < \quad \frac{2}{3}(-2\beta_{\nu_2} + \frac{3}{2}\sigma|K|)\,,$$

$$-2\beta_{\nu_2} + \frac{3}{2}\sigma|K| \quad > \quad \frac{3}{4}(-\beta_{\nu_2} + 2\sigma|K|)\,, \tag{4.45}$$

$\epsilon - \tilde{\epsilon}_2(t_2)$ will apparently be negative in convection-dominated flows, which implies that $g_2\big|_{K_{23}} < 0$. Thus, the second term on the right hand side of (4.41) attains its minimum if $g_2\big|_{K_{22}}$ is non-negative, too (see Lemma 4.3.1.4):

$$g_2\big|_{K_{22}} \quad = \quad (\beta \cdot \nabla B_2^* + \sigma B_2^* + \beta \cdot \nabla \psi_2 + \sigma \psi_2)\big|_{K_{22}}$$

$$= \quad (\alpha_2 \beta \cdot \nabla b_2 + \alpha_2 \sigma b_2 + \beta \cdot \nabla \psi_2 + \sigma \psi_2)\big|_{K_{22}}$$

$$= \quad \alpha_2 \beta \cdot \nabla b_2\big|_{K_{22}} + \alpha_2 \sigma b_2\big|_{K_{22}} + \frac{|K_{22}|}{|K|}\beta \cdot \nabla b_2\big|_{K_{22}} + \sigma \psi_2$$

$$= \quad -\frac{\alpha_2}{2|K_{22}|}\beta_{\nu_2} + \alpha_2 \sigma b_2\big|_{K_{22}} - \frac{\beta_{\nu_2}}{2|K|} + \sigma \psi_2$$

$$= \quad (\alpha_2 + t_2)\left(-\frac{\beta_{\nu_2}}{2|K|t_2} + \sigma b_2\big|_{K_{22}}\right)\,. \tag{4.46}$$

56

Now, use the expression for $\alpha_2$ in (4.37) to get

$$\alpha_2 + t_2 = -\frac{t_2\sigma|K|^2\left(t_2 - 1 + \frac{-\rho_2 + \sqrt{\rho_2^2 + \lambda_2}}{2\sigma|K|^2}\right)\left(t_2 - 1 - \frac{\rho_2 + \sqrt{\rho_2^2 + \lambda_2}}{2\sigma|K|^2}\right)}{3\epsilon l_{213}(t_2) + 2t_2(1 - t_2)\sigma|K|^2} .$$

The only root of the expression (4.46) in the interval $(0, 1)$ is $t_2^* = 1 - \dfrac{-\rho_2 + \sqrt{\rho_2^2 + \lambda_2}}{2\sigma|K|^2}$ and $g_2|_{K_{22}}$ is positive if $t_2 \geq t_2^*$. This fact together with Lemma 4.3.1.3 determines an optimal value for $t_2$. $\qquad\square$

**Remark 4.6** *As actual value for $t_2$, we do not always take that given by Lemma 4.3.1.5. Indeed, for $\epsilon$ not too small (that is, for diffusion dominated problems) this type of stabilization would be unnecessary. Furthermore the value provided by Lemma 4.3.1.5 even can be meaningless.Therefore, we take*

$$\begin{cases} t_2 = t_2^*, & \text{if } \epsilon \leq \epsilon_2^*, \\ t_2 = 1/3, & \text{otherwise}, \end{cases} \tag{4.47}$$

*which also gives a continuous dependence of $t_2$ upon $\epsilon$. Notice that for $\epsilon = \epsilon_2^*$ we have exactly $t_2 = t_2^* = 1/3$ and for $0 < \epsilon < \epsilon_2^*$ we have $1 > t_2^* > 1/3$. Therefore for every $\epsilon > 0$, we have*

$$\frac{1}{3} \leq t_2 < 1. \tag{4.48}$$

*This can be seen a weaker assumption than we made in Lemma 4.3.1.3, that was $t_2 \in (1/2, 1)$. However, it seems just as a technicality as we get reasonable results for values of $t_2$ in $(1/3, 1)$.*

**Remark 4.7** *We further observed that*

$$\lim_{\sigma \to 0} t_2^* = 1 - \frac{6\epsilon\left(|e_1|^2 + |e_3|^2\right)}{-2\beta_{\nu_2}|K| + 3\epsilon|e_1 - e_3|^2} , \tag{4.49}$$

*which has a similar structure to the one suggested by Brezzi and his coworkers in a different framework for the convection-diffusion problems ( see (Brezzi et al., 2005)).*

**Remark 4.8** *The value of $\alpha_2$ at $t_2^*$ is simply equal to $-t_2^*$.*

**Construction of $P_3$:** The position of $P_3$ along the median from $V_3$ will be determined by minimizing the value of the integral $J_3$. Let us start with splitting K into three subregions by $P_3$ (see Figure 4.1)

$$|K_{31}| = |K_{32}| = (1 - t_3)|K|/2 \qquad |K_{33}| = t_3|K|. \tag{4.50}$$

Similarly we have several explicit expressions for $\alpha_3$:

$$\alpha_3 = \frac{(-\mathcal{L}\psi_3, b_3)}{\epsilon||\nabla b_3||_K^2 + \sigma||b_3||_K^2} = -\frac{(S_3 + t_3\sigma|K|)/12}{\frac{\epsilon\,l_{312}(t_3)}{4(1-t_3)t_3|K|} + \frac{\sigma|K|}{6}} = -\frac{t_3\,(1 - t_3)\,|K|(S_3 + t_3\sigma|K|)}{f_{312}(t_3)}, \tag{4.51}$$

where $S_3 = -2\beta_{\nu_3} + \sigma|K|$. Note that $S_3 > 0$ and $\alpha_3 < 0$. Before finding an optimal position for $P_3$, we need the following intermediate result.

**Lemma 4.3.1.6** $\displaystyle\int_K |-\epsilon\Delta B_3^*|$ *is an increasing function of $t_3$ in the interval $(1/2, 1)$.*

**Proof** The proof is similar to the one in Lemma 4.3.1.3, so we skip it. $\qquad\square$

The following lemma suggests an optimal position for $P_3$ by using (4.16).

**Lemma 4.3.1.7** *If the inflow boundary make up of two edges then the value*

$$t_3^* = 1 - \frac{-\rho_3 + \sqrt{\rho_3^2 + \lambda_3}}{2\sigma\,|K|^2},$$

*minimizes the integral $J_3$ in convection-dominated flows where*

$$\rho_3 = -2\beta_{\nu_3}|K| + 3\epsilon\,|e_1 - e_2|^2, \quad \lambda_3 = 24\epsilon\sigma|K|^2\left(|e_1|^2 + |e_2|^2\right).$$

**Proof** The proof can be recovered from Lemma 4.3.1.5, by just changing the roles of $e_2$ and $e_3$, and replacing subindex 2 by subindex 3, accordingly. $\qquad\square$

**Remark 4.9** *As actual value for $t_3$, we do not always take that given by Lemma 4.3.1.7. Indeed, for $\epsilon$ not too small (that is, for diffusion dominated problems) this type of stabi-*

*lization would be unnecessary. Furthermore the value provided by Lemma 4.3.1.7 even can be meaningless. Therefore, we take*

$$
\begin{cases}
t_3 = t_3^*, & \text{if } \epsilon \leq \epsilon_3^*, \\
t_3 = 1/3, & \text{otherwise.}
\end{cases}
\tag{4.52}
$$

*which also gives a continuous dependence of $t_3$ upon $\epsilon$. Notice that for $\epsilon = \epsilon_3^*$ we have exactly $t_3 = t_3^* = 1/3$ and for $0 < \epsilon < \epsilon_3^*$ we have $1 > t_3^* > 1/3$. Therefore for every $\epsilon > 0$, we have*

$$
\frac{1}{3} \leq t_3 < 1.
\tag{4.53}
$$

*This can be seen a weaker assumption than we made in Lemma 4.3.1.6, that was $t_3 \in (1/2, 1)$. However, it seems just as a technicality as we get reasonable results for values of $t_3$ in $(1/3, 1)$.*

**Remark 4.10** *We further observed that*

$$
\lim_{\sigma \to 0} t_3^* = 1 - \frac{6\epsilon \left( |e_1|^2 + |e_2|^2 \right)}{-2\beta_{\nu_3}|K| + 3\epsilon \, |e_2 - e_1|^2} \,,
\tag{4.54}
$$

*which has a similar structure to the one suggested by Brezzi and his coworkers in a different framework for the convection-diffusion problems (Brezzi et al., 2005).*

**Remark 4.11** *The value of $\alpha_3$ at $t_3^*$ is simply equal to $-t_3^*$.*

**Remark 4.12** *For convection dominated regime, the choice of other point, $P_1$, should be consistent with the physics of the problem. Thus we take*

$$
t_2 = t_2^*, \; t_3 = t_3^*, \; t_1 = \min\{1/3, 1 - t_2^*, \, 1 - t_3^*\}.
$$

**Remark 4.13** *In Figure 4.4, we display the behavior of approximate bubble functions in a typical element $K$ with two inflow edges, for various intensities of diffusion ($\epsilon = 10^{-2}, 10^{-3}, 10^{-4}$) when $\theta = 72^o$ and $\sigma = 0.001$. The first column of the figure presents the bubble function, $b_1$, for decreasing values of diffusion ($\epsilon = 10^{-2}, 10^{-3}, 10^{-4}$). The*

*corresponding numerical results for $b_2$ and $b_3$ are shown in column 2 and 3, respectively.*
*As in the case of $P_1$, the location is automatically adjusted between problem regimes as*
*the problem evolves so that the pseudo bubbles contributes to the stability of the overall*
*numerical method at its most.*
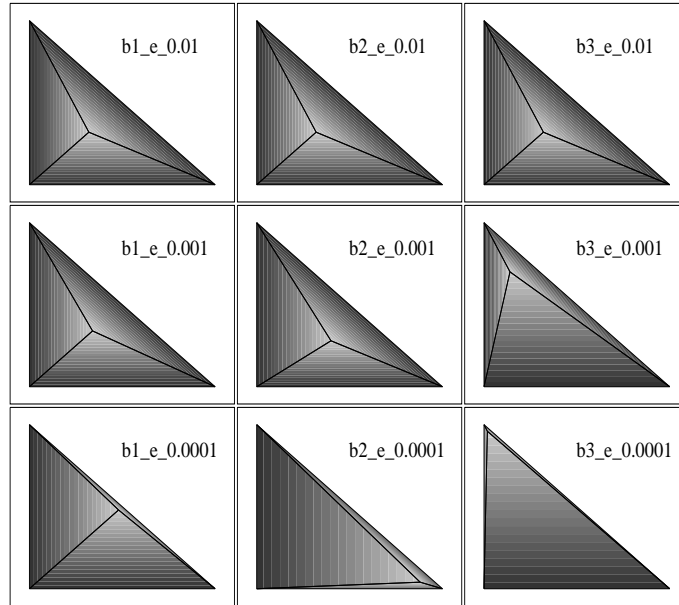


Figure 4.4. Bubble functions, $b_1$, $b_2$, $b_3$, in a typical element having two inflow edges, when $\theta = 72^o$, $N = 20$, $\epsilon = 10^{-2}, 10^{-3}, 10^{-4}$.

### 4.3.1.1. Numerical Results

In this section, we report some numerical experiments to illustrate the performance of the present algorithm in convection-dominated flows. We also report the $L_2$ and $H_1$ errors.

**Experiment 1: An example with analytical solution:** In the convergence analysis, we consider a simple problem on a unit square that can be solved analytically. Consider the

Figure 4.5. Configuration of Experiment 1.

unit square subject to the following boundary conditions (see Figure 4.5):

$$u = \begin{cases} 0, & \text{if } y = 0\,, \ 0 \leq x \leq 1, \\ 0, & \text{if } x = 1\,, \ 0 \leq y \leq 1, \\ 0, & \text{if } y = 1\,, \ 0 \leq x \leq 1, \\ \sin(\pi y), & \text{if } x = 0\,, \ 0 \leq y \leq 1. \end{cases} \tag{4.55}$$
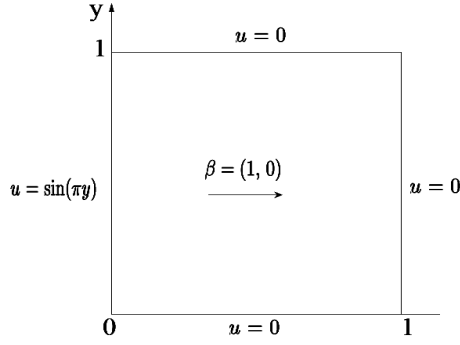
We set $\beta = (1,\,0)$ and $f(x) = 0$ in $\Omega$. Using separation of variables, the exact solution is given by:

$$u(x, y) = \frac{e^{x/2\epsilon} \ \sinh\left(-m(1 - x)\right) \ \sin(\pi y)}{\sinh\left(-m\right)} \quad \text{where} \quad \text{m} = \sqrt{1 + 4\epsilon(\epsilon \pi^2 + \sigma)}/2\epsilon\,.$$

Next, we take a set of uniform triangular meshes which are made up of $N = 10, 20, 40$ elements, respectively, in $x$ and $y$ directions. In Figure 4.6, we present the log-log plots of errors in $L_2$ and $H_1$ norms for different values of the mesh parameter N when $\sigma = 0.001$ and $\epsilon = 1, 10^{-2}, 10^{-3}, 10^{-4}$. Although the improvement is apparent as the mesh is refined, a slight degradation in the approximation can be observed for parameter values in mid-regime.
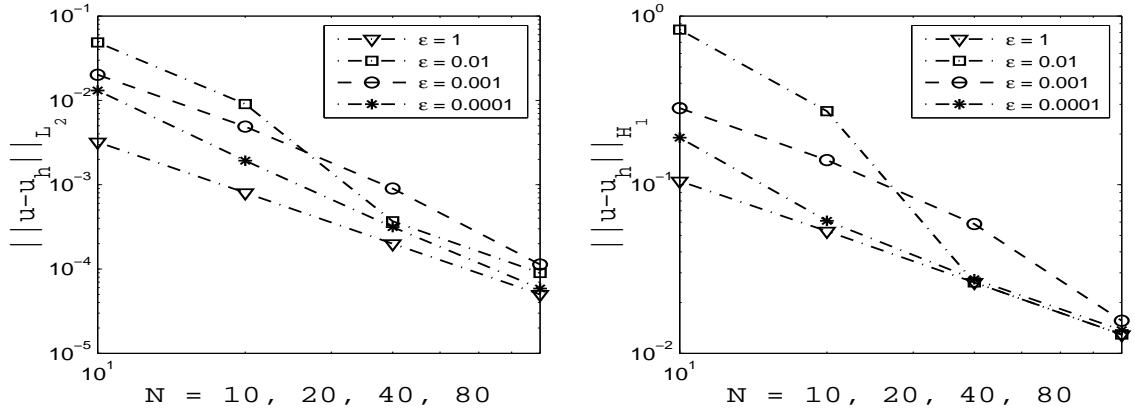
Figure 4.6. The error in $L_2$ (left) and $H_1$ (right) norm for several values of $\epsilon$ when $\beta = (1, 0)$, $\sigma = 0.001$ and $f(x) = 0$.
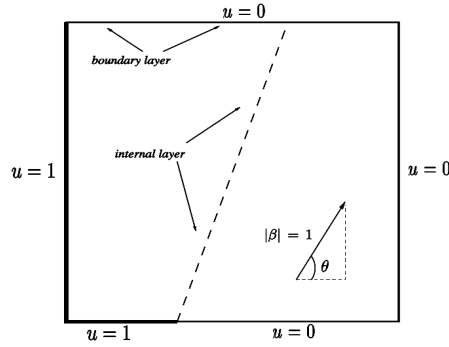


Figure 4.7. Configuration of Experiment 2.

**Experiment 2: Propagation of discontinuity:** We consider a test case shown in Fig. 4.7 for which the exact solution exhibits an internal and a boundary layer. We first take a set of uniform triangular meshes which are made up of $N = 10, 20, 40$ elements, respectively, in $x$ and $y$ directions (Figure 4.8). In Figures 4.9 - 4.11, we plot the solutions obtained with the present method for several values of $\theta$ when the convection dominates the flow, that is, $\epsilon = 0.001$, $\sigma = 0.001$ and $f = 0$. Although the present method could not eliminate the oscillations at all, it captures the layers well like in the RFB method, even on coarse meshes.

Figure 4.8. Triangular elements used in discretization of the problem domain.
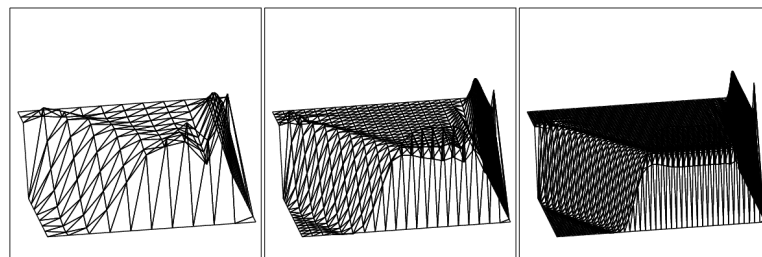


(a) The numerical solution for $\epsilon = 0.001$, $\sigma = 0.001$ and $f = 0$. (Left to right $N = 10, 20, 40$).
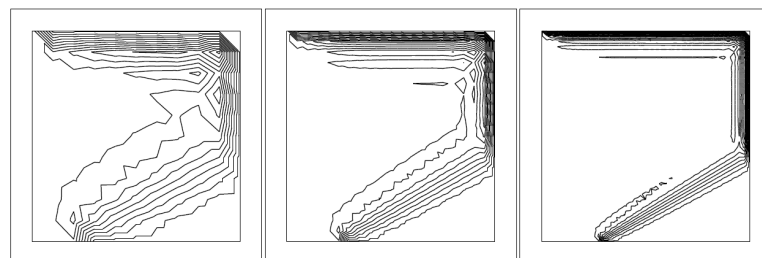


(b) Corresponding contour plots.

Figure 4.9. Numerical solution for $\theta = 30^o$.
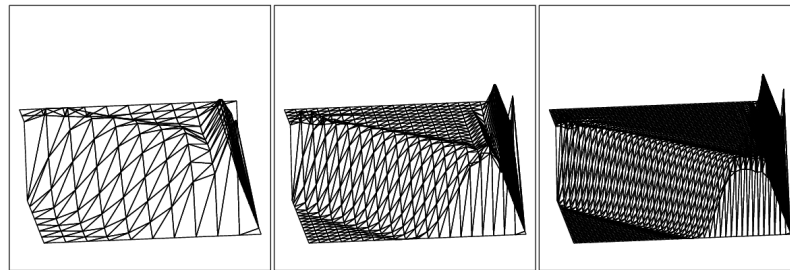
(a) The numerical solution for $\epsilon = 0.001$, $\sigma = 0.001$ and $f = 0$. (Left to right $N = 10, 20, 40$).
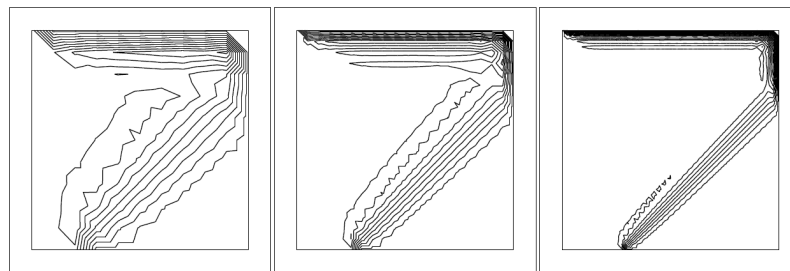


(b) Corresponding contour plots.

Figure 4.10. Numerical solution for $\theta = 45^o$.
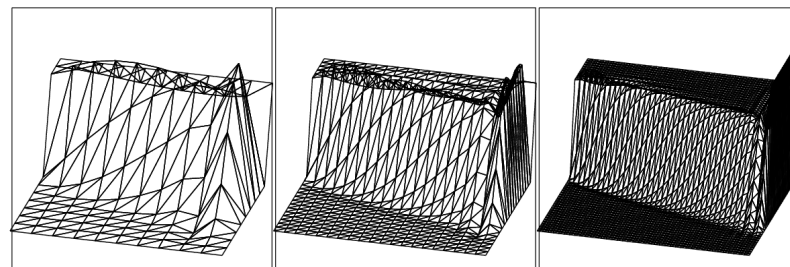


(a) The numerical solution for $\epsilon = 0.001$, $\sigma = 0.001$ and $f = 0$. (Left to right $N = 10, 20, 40$).
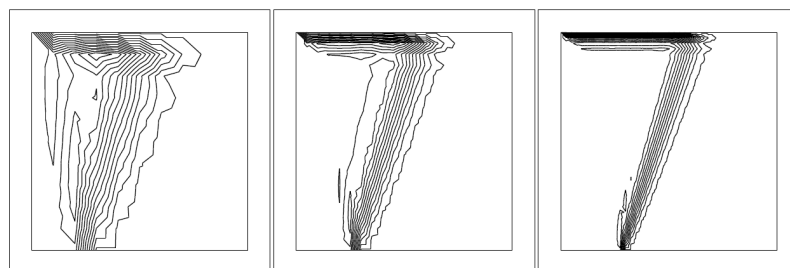


(b) Corresponding contour plots.

Figure 4.11. Numerical solution for $\theta = 72^o$.

Next, we consider the same problem on a non-uniform triangular mesh which are made up of $N = 10, 20, 40$ elements, respectively, in $x$ and $y$ directions (see Fig 4.12(a)). In Figure 4.12(b), we display the numerical results obtained with the present method for $\theta = 72^o$, $\epsilon = 0.01$, $\sigma = 0.001$ and $f = 0$. The numerical solutions show that the method is robust as the results are consistent with the physical configuration of the problem.



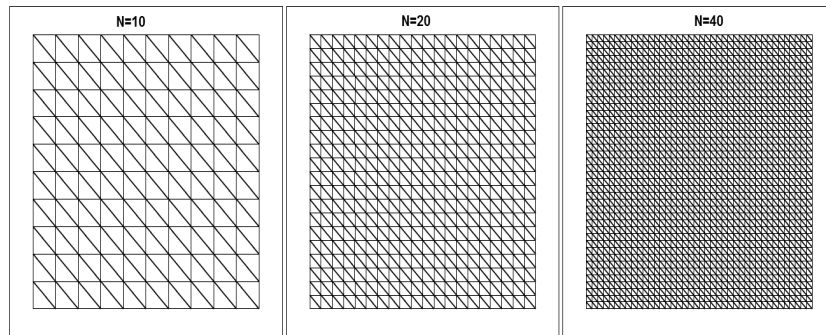(a) Triangular elements used in discretization of the problem domain.



(b) Numerical solution for $\epsilon = 0.01$, $\theta = 72^o$ (Left to right $N = 10, 20, 40$).
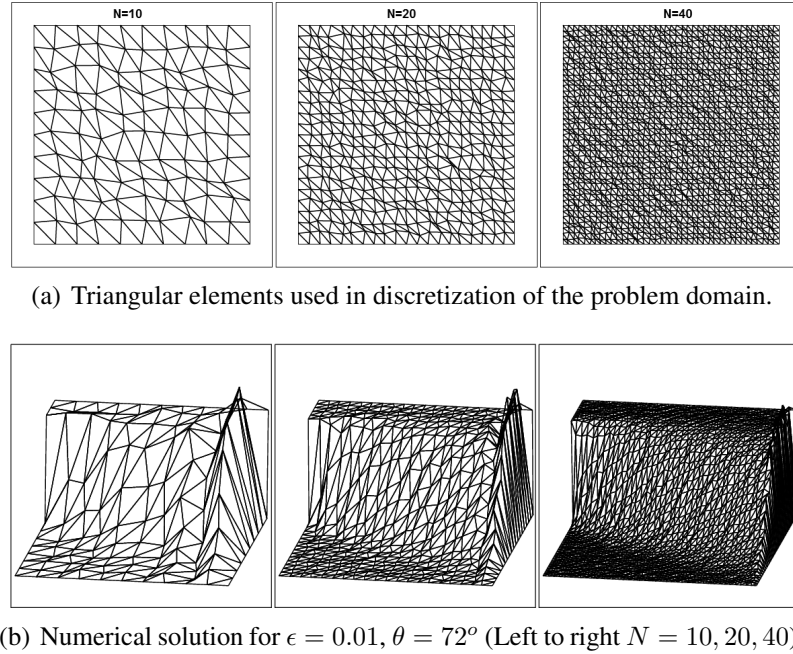
Figure 4.12. Numerical solution for $\theta = 72^o$ on a non-uniform triangular mesh.

## 4.3.2. Reaction-Dominated Flows

Now, we are in a position to determine the explicit locations of the internal nodes in reaction-dominated regimes. Let us proceed as follows:

**Case 1.** **The inflow boundary is made of two edges of K:** Let the inflow boundary make up of two edges and let $e_2$ and $e_3$ be the inflow ones. We assume that the problem is reaction-dominated if

$$\epsilon \leq \min\{\epsilon_2^*, \epsilon_3^*\} \quad \text{with} \quad \sigma|K| > 3\beta_{\nu_1} . \tag{4.56}$$
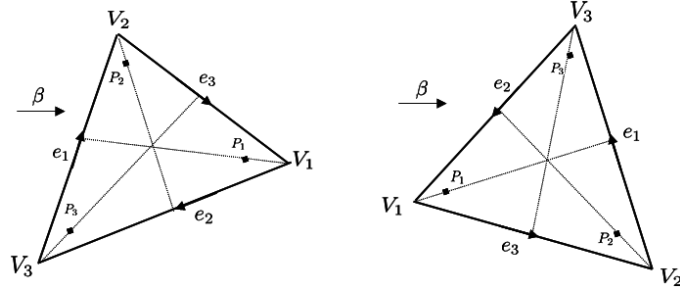
Figure 4.13. Configuration of internal nodes for reaction-dominated regime: one inflow edge (left) and two inflow edges (right).

Here, we take the positions of $P_2$ and $P_3$ as suggested in Section 4.3.1 for convection-dominated regime . Therefore, it remains us to find a proper location for $P_1$. The position of $P_1$ along the median from $V_1$ is determined by condition (4.16).

To be able to apply the condition (4.16), split $K$ into three subregions by using the definition of $P_1$ in (4.17) so that we have (see Figure 4.1)

$$|K_{11}| = t_1|K| \qquad |K_{12}| = |K_{13}| = (1 - t_1)|K|/2 \,. \tag{4.57}$$

Similarly, we have the following expression for $\alpha_1$:

$$\alpha_1 \;=\; -\frac{t_1 \,(1 - t_1)\,|K|(S_1 + t_1\sigma|K|)}{f_{123}(t_1)} \,, \tag{4.58}$$

where $S_1 = \sigma|K| - 2\beta_{\nu_1}$. Note that $S_1 > 0$ and $\alpha_1 < 0$ by condition (4.56). Before finding an optimal position for $P_1$, we need the following two intermediate results.

**Lemma 4.3.2.1** *Let $\alpha_1$ be as in (4.58). Then, we have*

$$1 + \alpha_1 > \frac{1 - t_1}{2} \,.$$

**Proof**  With the use of an explicit form of $\alpha_1$ in (4.58), we have

$$1 + \alpha_1 - \frac{1 - t_1}{2} \;=\; 1 - \frac{t_1 \,(1 - t_1)\,|K|(S_1 + t_1\sigma|K|)}{f_{123}(t_1)} - \frac{1 - t_1}{2}$$

66

$$= \frac{4t(1 - t_1)\beta_{\nu_1}|K| + 3\epsilon(1 + t_1)l_{123}(t_1)}{2f_{123}(t_1)} \,. \tag{4.59}$$

Hence the result immediately follows since $f_{123}(t_1) > 0$ whenever $0 < t_1 < 1$. $\qquad\square$

Next, we prove the following lemma:

**Lemma 4.3.2.2** *In reaction-dominated flows,* $\int_K | - \epsilon \Delta B_1^*|$ *is an increasing function of* $t_1$ *in the interval* $(1/2, 1)$.

**Proof**    Following the lines of  (Brezzi et al., 1998a), we get

$$\int_K | - \epsilon \Delta B_1^*| = -\epsilon\, \alpha_1 \sum_{i=1}^{3} \frac{|e_i|^2}{2\,|K_i|} \,. \tag{4.60}$$

Now, use the expression for $\alpha_1$ in (4.58) to get

$$
\begin{aligned}
\frac{d}{dt_1} \int_K | - \epsilon \Delta B_1^*| &= -\frac{d}{dt_1}\left( \epsilon\, \alpha_1 \left( \frac{|e_1|^2}{2\,|K_{11}|} + \frac{|e_2|^2}{2\,|K_{12}|} + \frac{|e_3|^2}{2\,|K_{13}|} \right) \right) \\
&= \epsilon \frac{d}{dt_1}\left( \frac{t_1\,(1 - t_1)\,(S_1 + t_1\sigma|K|)}{f_{123}(t_1)} \left( \frac{|e_1|^2}{2\,t_1} + \frac{|e_2|^2}{1 - t_1} + \frac{|e_3|^2}{1 - t_1} \right) \right) \\
&= \epsilon \frac{d}{dt_1}\left( \frac{S_1 + t_1\sigma|K|}{f_{123}(t_1)} \left( \frac{1 - t_1}{2}|e_1|^2 + t_1|e_2|^2 + t_1|e_3|^2 \right) \right) \\
&= \epsilon\sigma|K| \frac{3\epsilon l_{123}^2(t_1) + \tilde{f}_{123}(t_1)}{2f_{123}^2(t_1)} \,. \tag{4.61}
\end{aligned}
$$

The result immediately follows since $f_{123}(t_1) > 0$ and $\tilde{f}_{123}(t_1) > 0$ under the condition (4.56) with $1/2 < t_1 < 1$. $\qquad\square$

The following lemma suggests an optimal position for $P_1$ by using (4.16).

**Lemma 4.3.2.3** *If the inflow boundary make up of two edges, then the point*

$$t_1^{**} = 1 - \frac{\rho_1 + \sqrt{\rho_1^2 + \lambda_1}}{2\sigma\,|K|^2} \,,$$

*minimizes the integral $J_1$ in reaction-dominated flows where*

$$\rho_1 = 2\beta_{\nu_1}|K| - 3\epsilon\,|e_2 - e_3|^2\,, \quad \lambda_1 = 24\epsilon\sigma|K|^2\Big(|e_2|^2 + |e_3|^2\Big)\,.$$

**Proof**   According to (Brezzi et al., 1998a), it is possible to rewrite the integral $J_1$ in (4.16) as follows:

$$J_1 = \int_K |-\epsilon\Delta B_1^*| + \sum_{i=1}^{3}\int_{K_{1i}} |\beta\cdot\nabla B_1^* + \sigma B_1^* + \beta\cdot\nabla\psi_1 + \sigma\psi_1|\,. \qquad (4.62)$$

Let $g_1 = \beta\cdot\nabla B_1^* + \sigma B_1^* + \beta\cdot\nabla\psi_1 + \sigma\psi_1$. Without loss of generality, we assume that $\beta_{\nu_2} > \beta_{\nu_3}$ and

$$2\sigma|K_{13}| > -4\beta_{\nu_3} + \beta_{\nu_1}\,. \qquad (4.63)$$

A direct calculation over $K$ gives,

$$
\begin{aligned}
\int_K g_1 &= -\frac{1}{2}\beta_{\nu_1} + \frac{\sigma|K|}{3}(\alpha_1 + 1)\\[2mm]
&= -\frac{1}{2}\beta_{\nu_1} + \frac{\sigma|K|}{3}\left(-\frac{t_1\,(1-t_1)\,|K|(S_1 + t_1\sigma|K|)}{f_{123}(t_1)} + 1\right)\\[2mm]
&= \frac{3\epsilon l_{123}(t_1)\Big(2\sigma|K| - 3\beta_{\nu_1}\Big) + 2t_1(1-t_1)\sigma|K|^2\Big(-\beta_{\nu_1} + (1-t_1)\sigma|K|\Big)}{6f_{123}(t_1)}\,,
\end{aligned}
$$

which is positive since $f_{123}(t_1) > 0$ and

$$
\begin{aligned}
2\sigma|K| - 3\beta_{\nu_1} &> 6\beta_{\nu_1} - 3\beta_{\nu_1} = 3\beta_{\nu_1} > 0 \quad \text{(by condition (4.56) )}\,,\\
-\beta_{\nu_1} + (1-t_1)\sigma|K| &= -\beta_{\nu_1} + 2\sigma|K_{13}| > 0 \quad \text{(by assumption (4.63) )}\,.
\end{aligned}
$$

If we split $K$ into three subregions by $P_1$ (see Figure 4.1) and calculate the integral of $g_1$ over each of these sub-domains, we get:

$$\int_{K_{12}} g_1 = -\frac{\alpha_1}{2}\beta_{\nu_2} - \frac{|K_{12}|}{2|K|}\beta_{\nu_1} + \frac{\sigma|K_{12}|}{3}\left(\alpha_1 + \frac{|K_{11}|}{|K|} + 1\right)$$

$$> -\frac{\alpha_1}{2}\beta_{\nu_3} - \frac{|K_{12}|}{2|K|}\beta_{\nu_1} + \frac{\sigma|K_{12}|}{3}\left(\alpha_1 + \frac{|K_{11}|}{|K|} + 1\right) \quad (\text{since } \beta_{\nu_2} > \beta_{\nu_3})$$

$$= -\frac{\alpha_1}{2}\beta_{\nu_3} - \frac{|K_{13}|}{2|K|}\beta_{\nu_1} + \frac{\sigma|K_{13}|}{3}\left(\alpha_1 + \frac{|K_{11}|}{|K|} + 1\right) = \int_{K_{13}} g_1.$$

Moreover, we have

$$\int_{K_{13}} g_1 = -\frac{\alpha_1}{2}\beta_{\nu_3} - \frac{|K_{13}|}{2|K|}\beta_{\nu_1} + \frac{\sigma|K_{13}|}{3}\left(\alpha_1 + \frac{|K_{11}|}{|K|} + 1\right)$$

$$= -\frac{\alpha_1}{2}\beta_{\nu_3} + \frac{\sigma|K_{13}|}{3}(\alpha_1 + 1) + \frac{|K_{13}|}{|K|}\left(\frac{-\beta_{\nu_1}}{2} + \frac{\sigma|K_{11}|}{3}\right)$$

$$> \frac{\beta_{\nu_3}}{2}\frac{1+t_1}{2} + \frac{\sigma|K_{13}|}{3}\frac{1-t_1}{2} + \frac{|K_{13}|}{|K|}\left(\frac{-\beta_{\nu_1}}{2} + \frac{\sigma|K_{11}|}{3}\right) \text{(Lemma 4.3.2.1)}$$

$$= \frac{\beta_{\nu_3}}{2}\frac{1+t_1}{2} + \sigma|K|\frac{(1-t_1)^2}{12} + \frac{1-t_1}{2}\left(\frac{-\beta_{\nu_1}}{2} + \frac{t_1}{3}\sigma|K|\right)$$

$$= \frac{\beta_{\nu_3}}{2}\frac{1+t_1}{2} - \frac{\beta_{\nu_1}}{2}\frac{1-t_1}{2} + \sigma|K|\frac{(1-t_1)(1+t_1)}{12}$$

$$> \frac{\beta_{\nu_3}}{2} - \frac{\beta_{\nu_1}}{8} + \sigma|K|\frac{1-t_1}{8} \quad (\text{since } 1/2 < t_1 < 1)$$

$$= \frac{1}{8}\left(4\beta_{\nu_3} - \beta_{\nu_1} + 2\sigma|K_{13}|\right),$$

which is positive by assumption (4.63). Thus the second term on the right hand side of (4.62) attains its minimum if $\int_{K_{11}} g_1$ is non-negative, too:

$$\int_{K_{11}} g_1 = -\frac{\alpha_1}{2}\beta_{\nu_1} - \frac{|K_{11}|}{2|K|}\beta_{\nu_1} + \frac{\sigma|K_{11}|}{3}\left(\alpha_1 + \frac{|K_{11}|}{|K|}\right)$$

$$= \left(\alpha_1 + \frac{|K_{11}|}{|K|}\right)\left(-\frac{\beta_{\nu_1}}{2} + \frac{\sigma|K_{11}|}{3}\right)$$

$$= \left(\alpha_1 + t_1\right)\left(-\frac{\beta_{\nu_1}}{2} + \frac{t_1\sigma|K|}{3}\right)$$

$$= -\frac{t_1\left(-3\beta_{\nu_1} + 2\sigma|K|t_1\right)\left(t_1^2\sigma|K|^2 - t_1(-\rho_1 + 2\sigma|K|^2) + |K|S_1 - 3\epsilon|e_1|^2\right)}{6f_{123}(t_1)}$$

$$= -\frac{2t_1\sigma^2|K|^3\left(t_1 - \frac{3\beta_{\nu_1}}{2\sigma|K|}\right)\left(t_1 - 1 + \frac{\rho_1 + \sqrt{\rho_1^2 + \lambda_1}}{2\sigma|K|^2}\right)\left(t_1 - 1 - \frac{-\rho_1 + \sqrt{\rho_1^2 + \lambda_1}}{2\sigma|K|^2}\right)}{6f_{123}(t_1)} \,.$$

The only root of the last expression in $(0, 1)$ is $t_1^{**} = 1 - \dfrac{\rho_1 + \sqrt{\rho_1^2 + \lambda_1}}{2\sigma|K|^2}$ and thus, $\displaystyle\int_{K_{11}} g_1$ is positive if $t_1 \geq t_1^{**}$. This fact together with Lemma 4.3.2.2 determines an optimal value for $t_1^{**}$. $\qquad\square$

**Remark 4.14** *Note that*

$$\lim_{\sigma \to 0} t_1^{**} = -\infty \quad \text{and} \quad \lim_{\sigma \to \infty} t_1^{**} = 1 \,.$$

*Thus the points $P_1$, $P_2$ and $P_3$ should be chosen as*

$$t_2 = t_2^*, \ t_3 = t_3^*, \ t_1 = \max\{\min\{1/3, 1 - t_2, 1 - t_3\}, t_1^{**}\} \,,$$

*where we borrow $t_2^*$ and $t_3^*$ from Section 4.3.1.*

**Remark 4.15** *In Figure 4.14, we present the behavior of approximate bubble functions in a typical element $K$ with two inflow edges for various values of reaction ($\sigma = 10, 100, 500$) when $\epsilon = 10^{-3}$ and $\theta = 72^o$. The first column of the figure presents the bubble function, $b_1$, for increasing values of reaction ($\sigma = 10, 100, 500$). The corresponding numerical results for $b_2$ and $b_3$ are shown in column 2 and 3, respectively. Note that the self-adjustment of the subgrid node as the problem evolves from the convection-dominated regime into the reaction-dominated one. The location is readjusted so that the pseudo bubbles contributes to the stability of the numerical method at its most.*

**Case 2.   The inflow boundary is made of one edge of K:** Let the inflow boundary make up of one edge and $e_1$ is the inflow one. In this case, we assume that the problem is
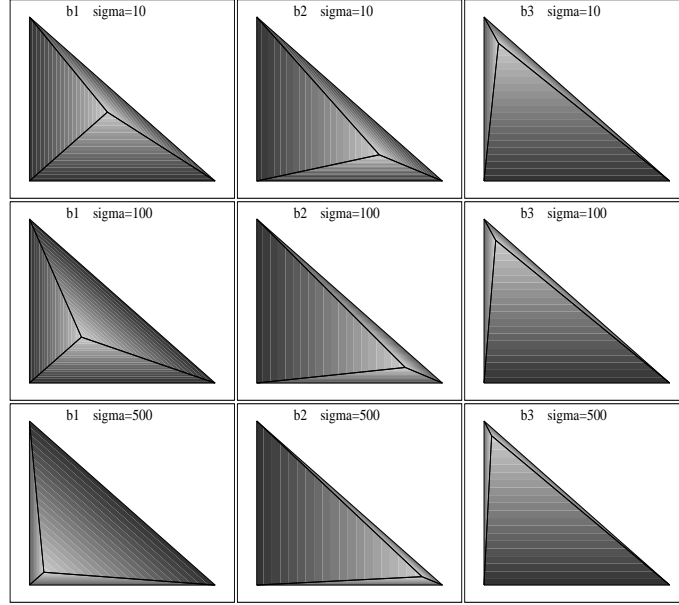
Figure 4.14. Bubble functions, $b_1$, $b_2$, $b_3$, in a typical element with two inflow edges, when $\epsilon = 10^{-3}$, $\theta = 72^o$, $N = 10$ and $\sigma = 10, 100, 500$.

reaction-dominated under the following conditions

$$\epsilon \leq \epsilon_1^* \quad \text{with} \quad \sigma|\mathrm{K}| > 3 \max\{\beta_{\nu_2}, \beta_{\nu_3}\}. \tag{4.64}$$

Here, we take the position of $P_1$ from convection-dominated regime which is suggested in Section 4.3.1 and it remains to find proper locations for $P_2$ and $P_3$. Now, we will construct the explicit locations of $P_2$ and $P_3$, separately:

**Construction of** $P_2$**:** The location of $P_2$ along the median from $V_2$ is determined by condition (4.16). If K is split into three subregions by using the definition of $P_2$ in (4.17), we have (see Figure 4.1)

$$|K_{21}| = |K_{23}| = (1 - t_2)|K|/2 \qquad |K_{22}| = t_2|K|. \tag{4.65}$$

Similarly, we get an explicit expression for $\alpha_2$:

$$\alpha_2 = -\frac{t_2 (1 - t_2) |K|(S_2 + t_2\sigma|K|)}{f_{213}(t_2)}, \tag{4.66}$$

where $S_2 = \sigma|K| - 2\beta_{\nu_2}$. Note that $S_2 > 0$ and $\alpha_2 < 0$ by condition (4.64). Before finding an optimal position for $P_2$, we need the following two intermediate results.

**Lemma 4.3.2.4** *Let $\alpha_2$ be as in (4.66). Then, we have:*

$$1 + \alpha_2 > \frac{1 - t_2}{2}.$$

**Proof** With the use of an explicit form of $\alpha_2$ in (4.66), we have

$$\begin{aligned}
1 + \alpha_2 - \frac{1 - t_2}{2} &= 1 - \frac{t_2 (1 - t_2) |K|(S_2 + t_2\sigma|K|)}{f_{213}(t_2)} - \frac{1 - t_2}{2} \\
&= \frac{4\beta_{\nu_2} t_2(1 - t_2) + 3\epsilon(1 + t_2) l_{213}(t_2)}{2 f_{213}(t_2)}. \tag{4.67}
\end{aligned}$$

The result immediately follows since $f_{213}(t_2) > 0$ whenever $0 < t_2 < 1$. $\square$

Next, we prove the following lemma:

**Lemma 4.3.2.5** *In reaction-dominated flows, $\int_K |-\epsilon\Delta B_2^*|$ is an increasing function of $t_2$ in the interval $(1/2, 1)$.*

**Proof** Following the lines of Lemma 4.3.2.2, we have

$$\int_K |-\epsilon\Delta B_2^*| = -\epsilon\,\alpha_2 \sum_{i=1}^{3} \frac{|e_i|^2}{2\,|K_{2i}|}. \tag{4.68}$$

Use the explicit expression for $\alpha_2$ in (4.66) to get

$$\frac{d}{dt_2}\int_K |-\epsilon\Delta B_2^*| = \epsilon\sigma|K|\frac{3\epsilon l_{213}^2(t_2) + \tilde{f}_{213}(t_2)}{2 f_{213}^2(t_2)}. \tag{4.69}$$

The result immediately follows since $f_{213}(t_2) > 0$, $\tilde{f}_{213}(t_2) > 0$ under the condition (4.64) with $1/2 < t_2 < 1$. $\square$

The following lemma suggests an optimal position for $P_2$ by using (4.16).

**Lemma 4.3.2.6** *If the inflow boundary make up of one edge, then the point*

$$t_2^{**} = 1 - \frac{\rho_2 + \sqrt{\rho_2^2 + \lambda_2}}{2\sigma |K|^2},$$

*minimizes the integral $J_2$ in reaction-dominated flows where*

$$\rho_2 = 2\beta_{\nu_2}|K| - 3\epsilon |e_1 - e_3|^2 \qquad \lambda_2 = 24\epsilon\sigma|K|^2\left(|e_1|^2 + |e_3|^2\right).$$

**Proof**    The proof is similar to the one in Lemma 4.3.2.3. Let us recall that

$$J_2 = \int_K |-\epsilon\Delta B_2^*| + \sum_{i=1}^3 \int_{K_{2i}} |\beta \cdot \nabla B_2^* + \sigma B_2^* + \beta \cdot \nabla\psi_2 + \sigma\psi_2| \qquad (4.70)$$

Let $g_2 = \beta \cdot \nabla B_2^* + \sigma B_2^* + \beta \cdot \nabla\psi_2 + \sigma\psi_2$. Without loss of generality, we may assume that

$$2\sigma|K_{21}| > -4\beta_{\nu_1} + \beta_{\nu_2}, \qquad (4.71)$$

in reaction-dominated flows. Then a direct calculation over $K$ gives,

$$
\begin{aligned}
\int_K g_2 &= -\frac{1}{2}\beta_{\nu_2} + \frac{\sigma|K|}{3}(\alpha_2 + 1) \\[2mm]
&= \frac{3\epsilon l_{213}(t_2)\left(2\sigma|K| - 3\beta_{\nu_2}\right) + 2t_2(1 - t_2)\sigma|K|^2\left(-\beta_{\nu_2} + (1 - t_2)\sigma|K|\right)}{6f_{213}(t_2)},
\end{aligned}
$$

which is positive since $f_{213}(t_2) > 0$ and

$$
\begin{aligned}
2\sigma|K| - 3\beta_{\nu_2} &> 6\beta_{\nu_2} - 3\beta_{\nu_2} = 3\beta_{\nu_2} > 0 \quad \text{(by condition (4.64))}, \\
-\beta_{\nu_2} + (1 - t_2)\sigma|K| &= -\beta_{\nu_2} + 2\sigma|K_{21}| > 0 \quad \text{(by assumption (4.71))}.
\end{aligned}
$$

Now split $K$ into three subregions by $P_2$ (see Figure 4.1) and calculate the integral of $g_2$ over each of these sub-domains so that we get

$$
\begin{aligned}
\int_{K_{21}} g_2 &= -\frac{\alpha_2}{2}\beta_{\nu_1} - \frac{|K_{21}|}{2|K|}\beta_{\nu_2} + \frac{\sigma|K_{21}|}{3}\left(\alpha_2 + \frac{|K_{22}|}{|K|} + 1\right) \\
&= -\frac{\alpha_2}{2}\beta_{\nu_1} + \frac{\sigma|K_{21}|}{3}(\alpha_2 + 1) + \frac{|K_{21}|}{|K|}\left(\frac{-\beta_{\nu_2}}{2} + \frac{\sigma|K_{22}|}{3}\right) \\
&> \frac{\beta_{\nu_1}}{2}\frac{1+t}{2} + \frac{\sigma|K_{21}|}{3}\frac{1-t_2}{2} + \frac{|K_{21}|}{|K|}\left(\frac{-\beta_{\nu_2}}{2} + \frac{\sigma|K_{22}|}{3}\right) \quad (\text{Lemma } 4.3.2.4) \\
&= \frac{\beta_{\nu_1}}{2}\frac{1+t_2}{2} + \frac{(1-t_2)^2}{12}\sigma|K| + \frac{1-t_2}{2}\left(\frac{-\beta_{\nu_2}}{2} + \frac{t_2}{3}\sigma|K|\right) \\
&= \frac{\beta_{\nu_1}}{2}\frac{1+t_2}{2} - \frac{\beta_{\nu_2}}{2}\frac{1-t_2}{2} + \sigma|K|\frac{(1-t_2)(1+t_2)}{12} \\
&= \frac{\beta_{\nu_1}}{2}\frac{1+t_2}{2} - \frac{\beta_{\nu_2}}{2}\frac{1-t_2}{2} + \sigma|K_{21}|\frac{1+t_2}{6} \\
&> \frac{\beta_{\nu_1}}{2} - \frac{\beta_{\nu_2}}{8} + \frac{\sigma|K_{21}|}{4} \quad (\text{since } 1/2 < t_2 < 1). \\
&= \frac{1}{8}\left(4\beta_{\nu_1} - \beta_{\nu_2} + 2\sigma|K_{21}|\right),
\end{aligned}
$$

which is positive by assumption (4.71). Further, we have

$$
\begin{aligned}
\int_{K_{23}} g_2 &= -\frac{\alpha_2}{2}\beta_{\nu_3} - \frac{|K_{23}|}{2|K|}\beta_{\nu_2} + \frac{\sigma|K_{23}|}{3}\left(\alpha_2 + \frac{|K_{22}|}{|K|} + 1\right) \\
&= -\frac{\alpha_2}{2}\beta_{\nu_3} + \frac{\sigma|K_{23}|}{3}(\alpha_2 + 1) + \frac{|K_{23}|}{|K|}\left(-\frac{\beta_{\nu_2}}{2} + \frac{\sigma|K_{22}|}{3}\right) \\
&> -\frac{\alpha_2}{2}\beta_{\nu_3} + \frac{\sigma|K_{23}|}{3}\frac{1-t_2}{2} + \frac{|K_{23}|}{|K|}\left(-\frac{\beta_{\nu_2}}{2} + \frac{\sigma|K_{22}|}{3}\right) \quad (\text{Lemma } 4.3.2.4)
\end{aligned}
$$

which is also positive by condition (4.64). Thus the second term on the right hand side of (4.70) attains its minimum if $\displaystyle\int_{K_{22}} g_2$ is non-negative, too:

$$\int_{K_{22}} g_2 = -\frac{\alpha_2}{2}\beta_{\nu_2} - \frac{|K_{22}|}{2|K|}\beta_{\nu_2} + \frac{\sigma|K_{22}|}{3}\left(\alpha_2 + \frac{|K_{22}|}{|K|}\right)$$

$$= \left(\alpha_2 + \frac{|K_{22}|}{|K|}\right)\left(-\frac{\beta_{\nu_2}}{2} + \frac{\sigma|K_{22}|}{3}\right)$$

$$= \left(\alpha_2 + t_2\right)\left(-\frac{\beta_{\nu_2}}{2} + \frac{t_2\sigma|K|}{3}\right)$$

$$= -\frac{t_2\left(-3\beta_{\nu_2} + 2t_2\sigma|K|\right)\left(t_2^2\sigma|K|^2 - t_2(-\rho_2 + 2\sigma|K|^2) + S_2|K| - 3\epsilon|e_2|^2\right)}{6f_{213}(t_2)}$$

$$= -\frac{2t_2\,\sigma^2\,|K|^3\left(t_2 - \frac{3\beta_{\nu_2}}{2\sigma|K|}\right)\left(t_2 - 1 + \frac{\rho_2 + \sqrt{\rho_2^2 + \lambda_2}}{2\sigma|K|^2}\right)\left(t_2 - 1 - \frac{-\rho_2 + \sqrt{\rho_2^2 + \lambda_2}}{2\sigma|K|^2}\right)}{6f_{213}(t_2)}\ .$$

The only root of the last expression in $(0,\,1)$ is $t_2^{**} = 1 - \dfrac{\rho_2 + \sqrt{\rho_2^2 + \lambda_2}}{2\sigma\,|K|^2}$ and thus, $\int_{K_{22}} g_2$ is positive if $t_2 \geq t_2^{**}$. This fact together with Lemma 4.3.2.5 determines an optimal value for $t_2^{**}$. $\qquad\square$

**Construction of $P_3$:** Now we determine the location of $P_3$ in the same manner. Let us start with splitting K into three subregions by $P_3$ (see Figure 4.1)

$$|K_{31}| = |K_{32}| = (1 - t_3)|K|/2 \qquad |K_{33}| = t_3|K|\,. \tag{4.72}$$

Similarly, we have several explicit expressions for $\alpha_3$:

$$\alpha_3 = -\frac{t_3\,(1 - t_3)\,|K|(S_3 + t_3\sigma|K|)}{f_{312}(t_3)}\,, \tag{4.73}$$

where $S_3 = \sigma|K| - 2\beta_{\nu_3}$. Note that $S_3 > 0$ and $\alpha_3 < 0$ by condition (4.64). Before finding an optimal position for $P_3$, we need the following two intermediate results.

**Lemma 4.3.2.7** *Let $\alpha_3$ be as in (4.73). Then, we have:*

$$1 + \alpha_3 > \frac{1 - t_3}{2}\,.$$

**Proof**   The proof is similar to the one in Lemma 4.3.2.4, so we skip it.   □

Next, we prove the following lemma:

**Lemma 4.3.2.8**  *In reaction-dominated flows, $\int_K |-\epsilon\Delta B_3^*|$  is an increasing function of $t_3$ in the interval $(1/2, 1)$.*

**Proof**   We skip it as its proof is similar to the one in Lemma 4.3.2.5.   □

The following lemma suggests an optimal position for $P_3$ by using (4.16).

**Lemma 4.3.2.9**  *If the inflow boundary make up of one edge then the point*

$$t_3^{**} = 1 - \frac{\rho_3 + \sqrt{\rho_3^2 + \lambda_3}}{2\sigma\,|K|^2}\,,$$

*minimizes the integral $J_3$ in reaction-dominated flows where*

$$\rho_3 = 2\beta_{\nu_3}|K| - 3\epsilon\,|e_1 - e_2|^2\,, \quad \lambda_3 = 24\epsilon\sigma|K|^2\Big(|e_1|^2 + |e_2|^2\Big)\,.$$

**Proof**   The proof can be recovered from Lemma 4.3.2.6 by just changing the roles of $e_2$ and $e_3$ and replacing subindex 2 by subindex 3 accordingly.   □

**Remark 4.16**  *Note that*

$$\lim_{\sigma\to 0} t_2^{**} = -\infty \quad \text{and} \quad \lim_{\sigma\to\infty} t_2^{**} = 1\,,$$
$$\lim_{\sigma\to 0} t_3^{**} = -\infty \quad \text{and} \quad \lim_{\sigma\to\infty} t_3^{**} = 1\,.$$

*Thus the points $P_1$, $P_2$ and $P_3$ should be taken as*

$$t_1 = t_1^*,\ t_2 = \max\{1 - t_1, t_2^{**}\},\ t_3 = \max\{1 - t_1, t_3^{**}\}\,,$$

*where we borrow $t_1^*$ from Section 4.3.1.*

**Remark 4.17**  *In Figure 4.15, we present the behavior of approximate bubble functions in a typical element $K$ with one inflow edge for various values of reaction ($\sigma = 10, 100, 500$) when $\epsilon = 10^{-3}$ and $\theta = 72^o$. The first column of the figure presents the bubble function, $b_1$,*
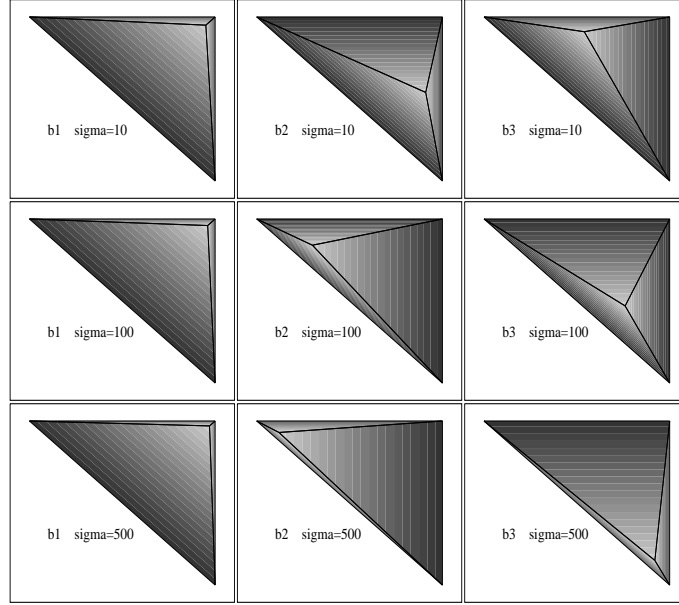
Figure 4.15. Bubble functions, $b_1$, $b_2$, $b_3$, in a typical element with one inflow edge, when $\epsilon = 10^{-3}$, $\theta = 72^o$, $N = 10$ and $\sigma = 10, 100, 500$.

*for increasing values of reaction ($\sigma = 10, 100, 500$). The corresponding numerical results for $b_2$ and $b_3$ are shown in column 2 and 3, respectively. Note that the self-adjustment of the subgrid node as the problem evolves from convection-dominated regime into the reaction-dominated one. The location is readjusted so that the pseudo bubbles contributes to the stability of the numerical method at its most.*

**Remark 4.18** *It still remains the computation of $B_f$, before we employ $u_B$ to solve the numerical method (4.7) for the linear part of $u_h$, which is accomplished in the following: Since $f$ is assumed to be a linear function, it can be expressed in terms of $\psi_1, \psi_2, \psi_3$, that is*

$$f \mid_K = f(V_1)\,\psi_1 + f(V_2)\,\psi_2 + f(V_3)\,\psi_3\,. \tag{4.74}$$

*We want to express $B_f$ in terms of $B_i$, $i = 1, 2, 3$. Set*

$$B_f = \lambda_1 B_1 + \lambda_2 B_2 + \lambda_3 B_3\,. \tag{4.75}$$

*Using the definition of $B_f$ in (4.9) and the property that $\psi_1 + \psi_2 + \psi_3 = 1$ (partition of unity) we have*

$$
\begin{aligned}
\mathcal{L}B_f &= \mathcal{L}(\lambda_1 B_1 + \lambda_2 B_2 + \lambda_3 B_3) = \lambda_1 \mathcal{L}B_1 + \lambda_2 \mathcal{L}B_2 + \lambda_3 \mathcal{L}B_3 \\
&= \lambda_1(-\mathcal{L}\psi_1) + \lambda_2(-\mathcal{L}\psi_2) + \lambda_3(-\mathcal{L}\psi_3) \\
&= -\lambda_1(\beta \cdot \nabla\psi_1 + \sigma\psi_1) - \lambda_2(\beta \cdot \nabla\psi_2 + \sigma\psi_2) - \lambda_3(\beta \cdot \nabla\psi_3 + \sigma\psi_3) \\
&= (-\lambda_1\beta \cdot \nabla\psi_1 - \lambda_2\beta \cdot \nabla\psi_2 - \lambda_3\beta \cdot \nabla\psi_3)(\psi_1 + \psi_2 + \psi_3) \\
&\quad - \lambda_1\sigma\psi_1 - \lambda_2\sigma\psi_2 - \lambda_3\sigma\psi_3 \\
&= -[\lambda_1(\beta \cdot \nabla\psi_1 + \sigma) + \lambda_2\beta \cdot \nabla\psi_2 + \lambda_3\beta \cdot \nabla\psi_3]\,\psi_1 \\
&\quad - [\lambda_1\beta \cdot \nabla\psi_1 + \lambda_2(\beta \cdot \nabla\psi_2 + \sigma) + \lambda_3\beta \cdot \nabla\psi_3]\,\psi_2 \\
&\quad - [\lambda_1\beta \cdot \nabla\psi_1 + \sigma + \lambda_2\beta \cdot \nabla\psi_2 + \lambda_3(\beta \cdot \nabla\psi_3 + \sigma)]\,\psi_3\,,
\end{aligned}
\tag{4.76}
$$

*which is only possible if*

$$
\begin{aligned}
-[\lambda_1(\beta \cdot \nabla\psi_1 + \sigma) + \lambda_2\beta \cdot \nabla\psi_2 + \lambda_3\beta \cdot \nabla\psi_3] &= f(V_1)\,, \\
-[\lambda_1\beta \cdot \nabla\psi_1 + \lambda_2(\beta \cdot \nabla\psi_2 + \sigma) + \lambda_3\beta \cdot \nabla\psi_3] &= f(V_2)\,, \\
-[\lambda_1\beta \cdot \nabla\psi_1 + \sigma + \lambda_2\beta \cdot \nabla\psi_2 + \lambda_3(\beta \cdot \nabla\psi_3 + \sigma)] &= f(V_3)\,.
\end{aligned}
$$

*Once the constants $\lambda_1, \lambda_2, \lambda_3$ are determined, we substitute them into equation (4.75) to get $B_f$.*

### 4.3.2.1. Numerical Results

In this section, we report some numerical experiments exhibiting boundary/internal layers to illustrate the efficiency and performance of the proposed finite element method in the interesting case of small diffusion which corresponds to the convection-dominated or reaction-dominated regimes. Although the present method could not eliminate the oscillations at all, it captures the layers well like in the RFB method. In all cases, it is clear that our numerical solution is able to show the dominant characteristics of the exact solution for a wide range of problem parameters at all levels of the mesh employed. We also note that, the transition from one regime to another is accurately captured by the

algorithm.
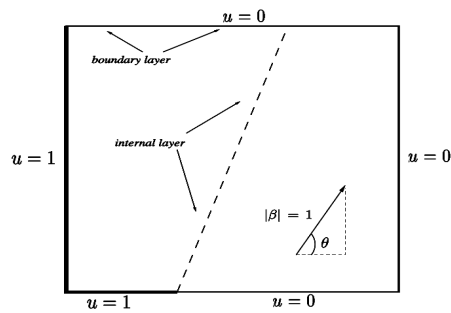


Figure 4.16. Configuration of test problem by Asensio et. al (2004).

**Experiment 1: Test problem by Asensio et. al (2004):** We start the numerical experiments with a test problem which is taken from (Asensio et al., 2004). Boundary conditions are displayed in Figure 4.16. We take a pair of uniform triangular meshes which are made up of $N = 10, 20$ elements, respectively, in $x$ and $y$ directions (Figure 4.17).

In Figures 4.18-4.19, we take $\epsilon = 10^{-4}$ and $\theta = 72^0$ and plot the solutions obtained with the present method for various values of reaction ($\sigma = f = 0.001, 1, 10, 20, 50, 1000$).
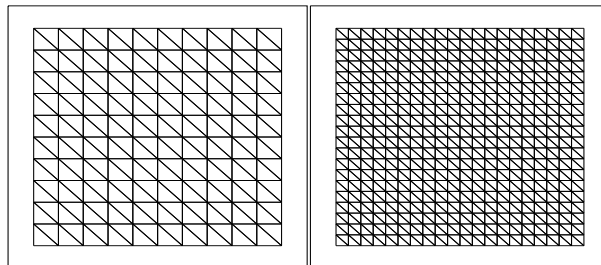


Figure 4.17. Triangular elements used in discretization of the problem domain: N=10 (left) and N=20 (right).

solution, min = -0.0470793, max = 1.44208

solution, min = 0, max = 1.29023

solution, min = 0, max = 1.47618

solution, min = 0, max = 1.39467

solution, min = 0, max = 1.33169

solution, min = 0, max = 1.32981

Figure 4.18. Numerical solution for $\theta = 72^o$, $\epsilon = 10^{-4}$ and $N = 10$.



solution, min = -0.0348126, max = 1.46803

solution, min = 0, max = 1.28822

solution, min = 0, max = 1.55278

solution, min = 0, max = 1.48521

solution, min = 0, max = 1.38203

solution, min = 0, max = 1.32698
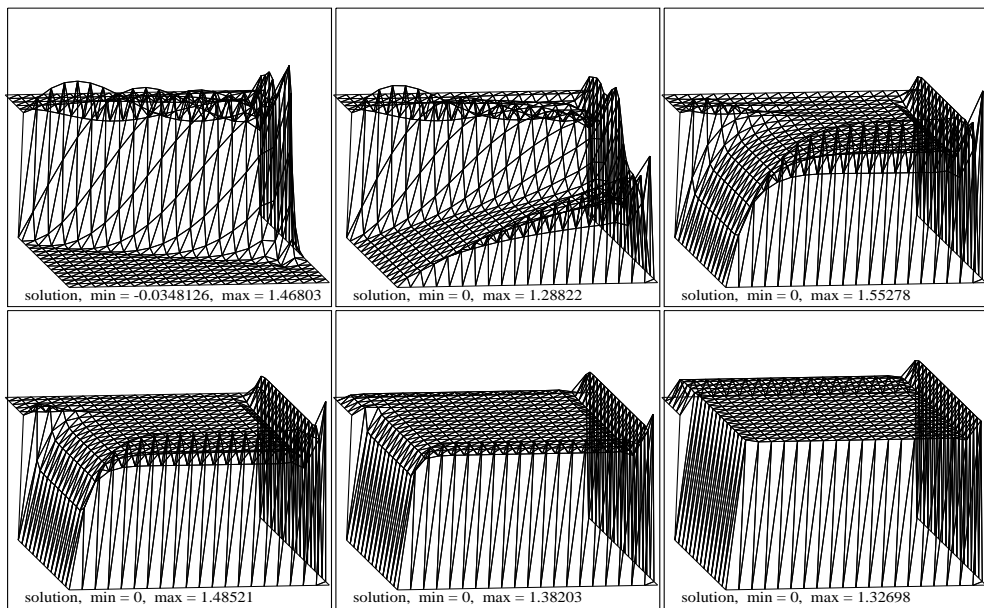
Figure 4.19. Numerical solution for $\theta = 72^o$, $\epsilon = 10^{-4}$ and $N = 20$.

**Experiment 2:** Next, we consider a different problem on the same mesh used in Experiment 1. Take $\epsilon = 10^{-3}$, $\theta = 72^0$ and $f = 1$ with $u = 0$ on $\partial\Omega$. In Figures 4.20-4.21, we plot the solutions obtained with the present method for various values of reaction ($\sigma = 0.001, 1, 10, 20, 50, 100$). It is significant that the method works well in both problems, in both convection and reaction-dominated cases, even on coarse meshes. In addition, the smooth transition between regimes can be observed.



(a) $\sigma = 0.001$       (b) $\sigma = 1$       (c) $\sigma = 10$

(d) $\sigma = 20$       (e) $\sigma = 50$       (f) $\sigma = 100$
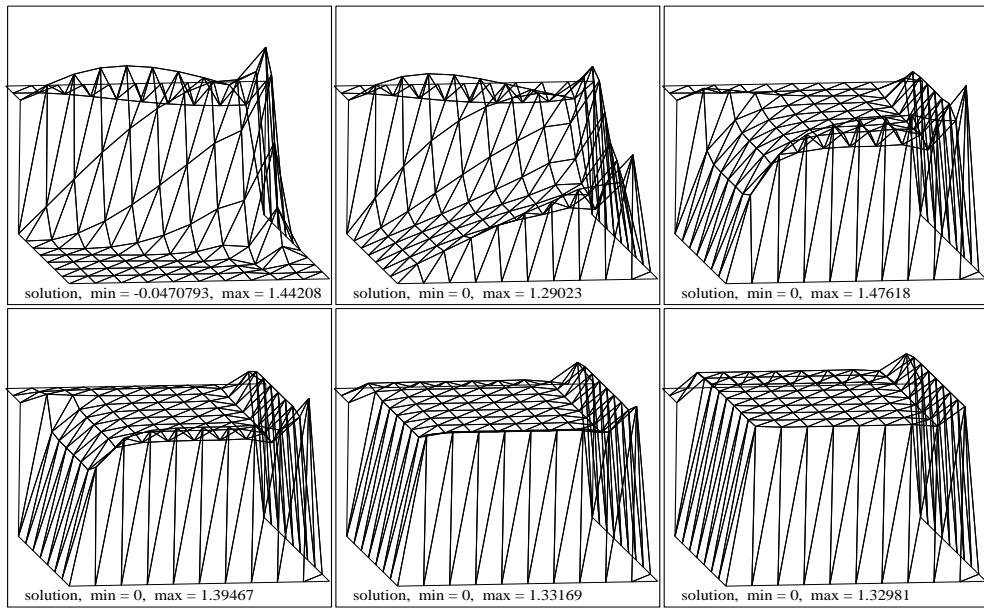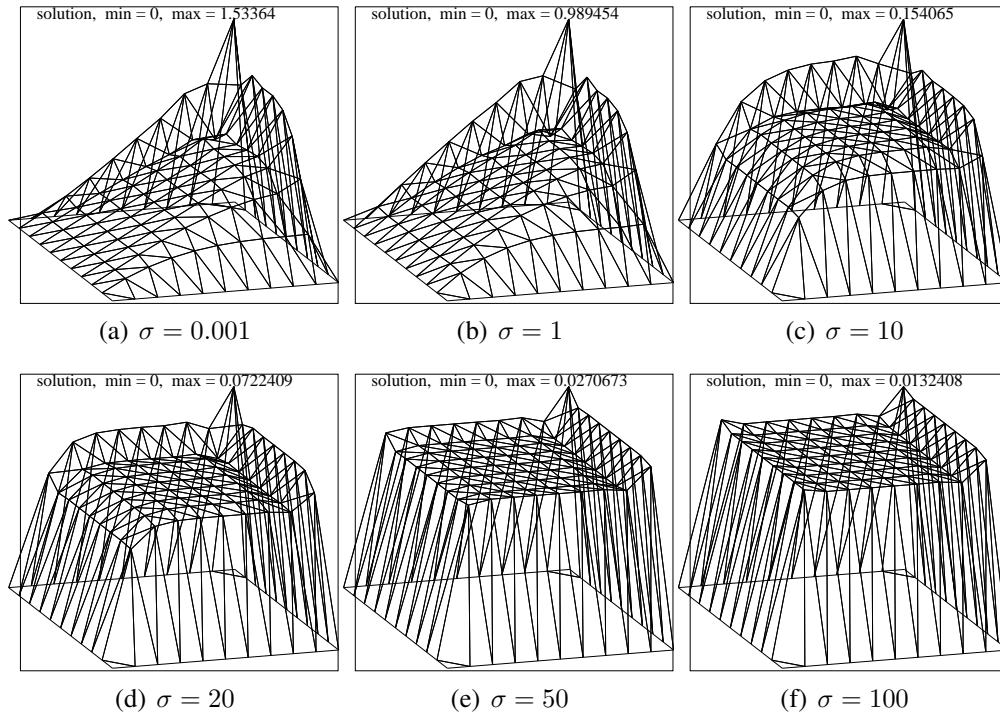
Figure 4.20. Numerical solution for $\theta = 72^o$, $\epsilon = 10^{-3}$ and $N = 10$.
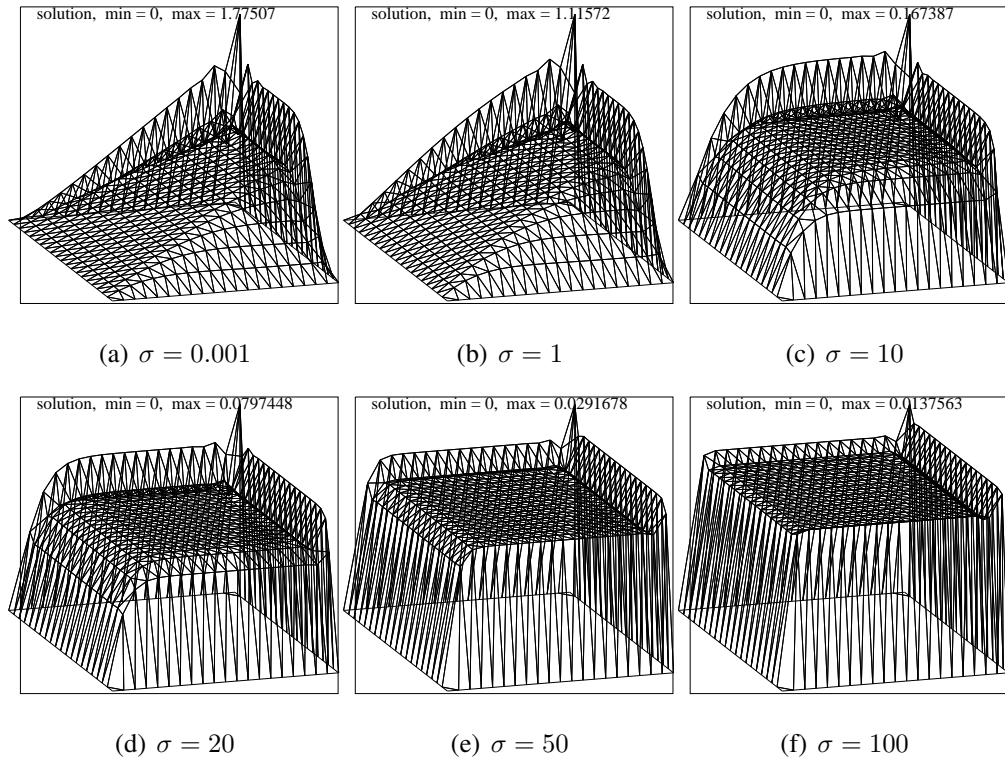
Figure 4.21. Numerical solution for $\theta = 72^o$, $\epsilon = 10^{-3}$ and $N = 20$.
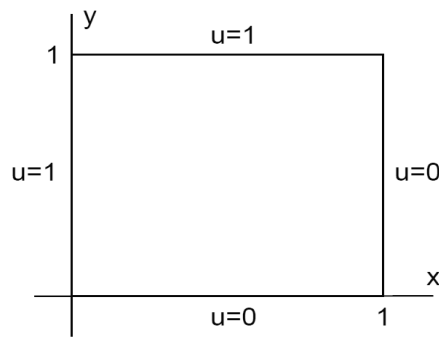


Figure 4.22. Configuration of test problem by Franca and Valentin (2000).

**Experiment 3: Test problem by Franca and Valentin (2000):** Next, we consider a test problem which is taken from (Franca and Valentin, 2000). We display the boundary conditions in Figure 4.22. In Figure 4.23, we plot the solutions obtained with the present method when $\epsilon = 10^{-4}$, $\beta = (0.15, 0.1)$ and $f = 0$ for various values of reaction ($\sigma = 10, 10^2, 10^3$).

solution, min = -0.13034, max = 1    solution, min = -0.116681, max = 1    solution, min = -0.113474, max = 1

solution, min = -0.184943, max = 1    solution, min = -0.122981, max = 1    solution, min = -0.115838, max = 1
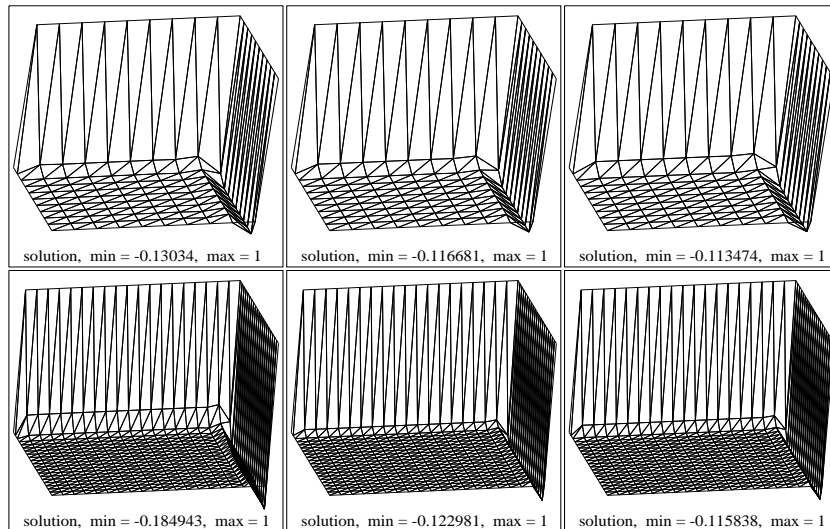
Figure 4.23. Numerical solution for $\epsilon = 10^{-4}$ (Top to bottom: $N = 10, 20$  Left to right: $\sigma = 10, 10^2, 10^3$).

Now, the same problem is tested on a pair of non-uniform triangular meshes, made up of $N = 10, 20$ elements, respectively, in $x$ and $y$ directions (see Figure 4.24). The numerical solutions in Figure 4.25 show that the method is robust as the results are consistent with the physical configuration of the problem on both uniform and non-uniform meshes.
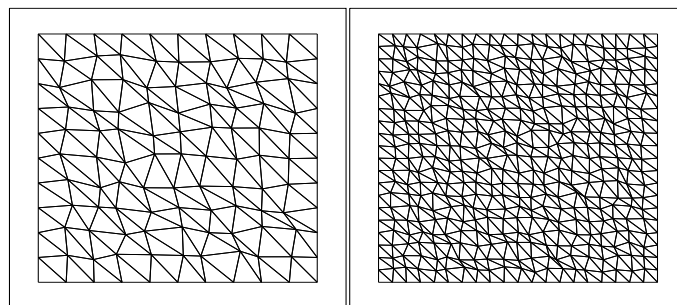


Figure 4.24. Triangular elements used in discretization of the problem domain: N=10 (left) and N=20 (right).
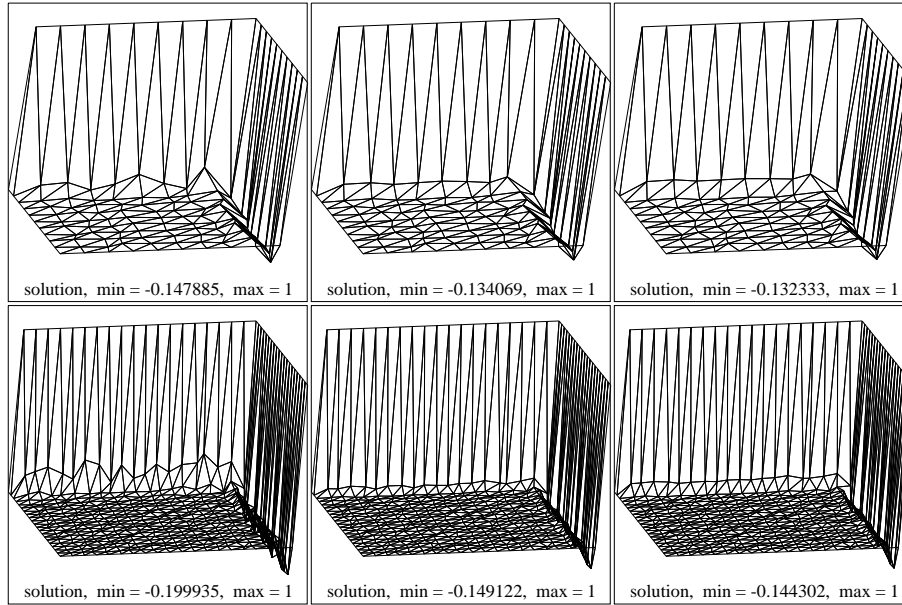
Figure 4.25. Top to bottom: $N = 10, 20$   Left to right: $\sigma = 10, 10^2, 10^3$.

**Experiment 4: Test problem by Codina (1998):** Finally, we consider a test problem which is taken from (Codina, 1998). The mesh is displayed in Figure 4.17. The source term has been taken as $f = 1$, constant and the diffusion coefficient has been set to $\epsilon = 10^{-4}$. The velocity vector has been taken as $\beta = |\beta|(\cos \pi/3, \, \sin \pi/3)$ so that it is not aligned with the finite element mesh. Three different cases have been considered, corresponding to dominant convection, dominant reaction and combination of convection and reaction effects. These cases are:

**Case A:** $|\beta| = 1$   and   $\sigma = 0.0001$

**Case B:** $|\beta| = 0.0001$   and   $\sigma = 1$

**Case C:** $|\beta| = 0.5$   and   $\sigma = 1$
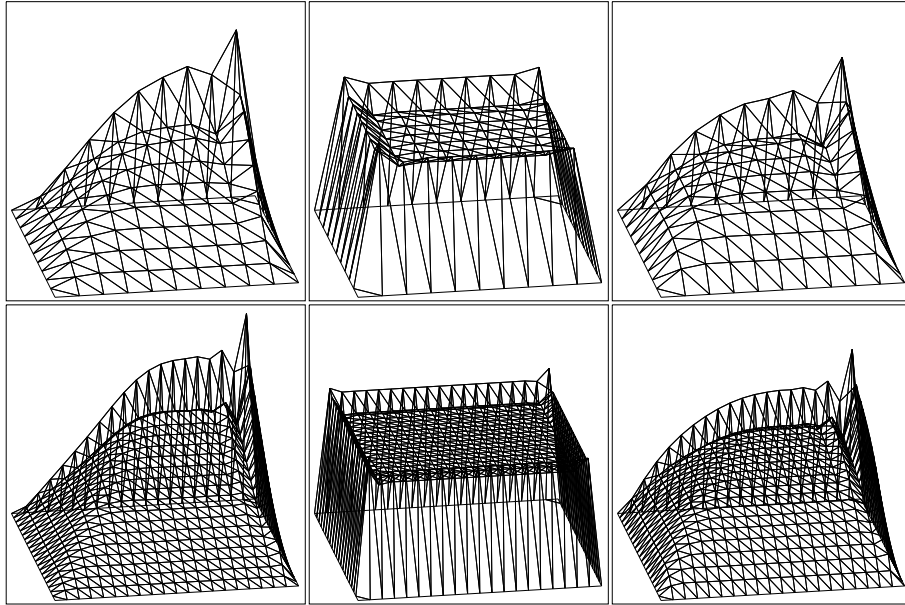
Figure 4.26. Top to bottom: $N = 10, 20$     Left to right: Case A, Case B, Case C.

In this chapter, a stabilized finite element method for 2-dimensional convection-diffusion-reaction problems has been introduced and tested. In particular, we wanted to mimic the stabilizing effect of the RFBs by approximating them on a specially chosen internal nodes and using these approximations in place of RFB functions. The basic idea behind the present choice of the internal nodes are based on minimizing the residual of local problems in a special manner. Based on the numerical experiences gained with this work, the pseudo RFBs retain the stability features of RFBs for the whole range of problem parameters.

# CHAPTER 5

# A STABILIZED FINITE ELEMENT METHOD FOR TIME-DEPENDENT CONVECTION-DIFFUSION-REACTION PROBLEMS IN 1-D

A numerical scheme for one-dimensional time-dependent convection-diffusion-reaction problem is studied herein. While discretization in time is performed by Crank − Nicolson finite difference scheme, for the space discretization we employ a recent method suggested by Neslitürk et al.

## 5.1. Introduction

The exact solution of convection-diffusion-reaction (C-D-R) problems may exhibit layer structures, small subregions where the derivatives of the solution are very large, when some problem parameters are too big compared to other(s). In order to retain an acceptable degree of accuracy for wide ranges of values of the physical coefficients, Galerkin method requires fine meshes . A variety of finite elements approaches have been proposed to deal with such situations. The streamline- diffusion discontinuous Galerkin method is the most popular one which is proposed in (Johnson et al., 1984).

There are several approach for discretizing time-dependent convection-diffusion-reaction problems by finite element methods. Firstly, space-time elements could be used. This results in very large systems of linear equations since all unknowns in the space-time cylinder are coupled. Secondly, semidiscretization as intermediate steps can be used. Here, we distinguish between the horizontal and vertical methods of lines. The vertical method of lines discretizes first in space and then in time while the horizontal method of lines (or Rothe's method) applies first a time discretisation which is followed by a discretisation in space.

It is worth mentioning that in (Ahmed et al., 2011), a time-dependent convection-diffusion-reactions problem is discretized in space by a continuous finite element method with local projection stabilization and in time by a discontinuous Galerkin method. Fur-

thermore the error estimates for the semidiscrete problem after discretizing in space only and for the fully discrete problem are also presented.

Numerical studies of stabilized finite element methods for time-dependent convection diffusion-reaction problems can also be found in (Codina, 1998). This study emphasizes the similarities and differences between them, including SUPG method. In (John and Schmeyer, 2008), the SUPG method was compared in comprehensive studies with other stabilized finite element methods. The key idea in these studies was to employ the Rothe's method on each discrete time level.

The Rothe's method was also studied for time-dependent convection–diffusion–reaction problems in (Harari and Hauke, 2007). An analogy to steady convection–diffusion–reaction problems with a modified reaction coefficient by the Rothe's method of discretizing in time prior to spatial discretization provides an upper bound on the time step for the onset of spatial instability. In this work authors show that the spatial stabilization removes this pathology, leading to stabilized implicit time-integration schemes that are free of spurious oscillations at small time steps. Furthermore, the stabilization of convection and reaction-dominated phenomena, as well as implementational issues such as the form of the time-marching schemes and initialization is addressed as a future work.

The coupling of other stabilization techniques in the one dimensional case with the finite difference time integration in particular the vertical and horizontal method of lines has been discussed in (Asensio et al., 2007). In this study, some numerical schemes for the one dimensional unsteady linear convection–diffusion–reaction problem are presented. Discretization in time is performed by using the Crank−Nicolson finite difference scheme, while for the space discretization some classical stabilized finite element schemes including Link-Cutting Bubble (LCB) strategy are proposed. The best results have been obtained discretizing first the time variable and then using the LCB formulation (referred as FDt-FEs). Authors show that the robustness and accuracy of the LCB strategy for convection and reaction-dominated regimes in the steady problem allow for the combination FDt-FEs with LCB to be a good method for the unsteady problem. Furthermore, the extension of the present results to the multidimensional case is addressed as the subject of future research. A recent use of Rothe's method is due to (Henao et al., 2010), that approximates transient transport equations by using Rothe's method combined with the Unusual Stabilized Finite Element Method (USFEM).

Here, we will discuss how to combine stabilized finite element methods designed for the stationary problem with simple time discretization schemes, for the numerical solution of the time-dependent problem. The key idea is to employ the Rothe method or

the horizontal method of lines. For the temporal discretization, we will employ fractional-step $\theta$-schemes. This approach reveals that the time-discrete equation that is solved in each time level of implicit schemes is, in fact, a Galerkin approximation of a steady equation with modified coefficients. In order to stabilize the Galerkin discretization , the stabilization method suggested in (Şendur and Neslitürk, 2012) is implemented.

## 5.2. Statement of the Problem

The partial differential equation that we want to solve numerically is, in the scalar transient case,

$$u_t + \mathcal{L}u = u_t(x,t) - \epsilon\, u_{xx}(x,t) + \beta\, u_x(x,t) + \sigma\, u(x,t) = f(x,t) \quad \text{in } \Omega \times (0,T)\,(5.1)$$

where $u$ is the unknown, $\beta$ is the convection velocity, $\epsilon > 0$ is the diffusion coefficient, $\sigma \geq 0$ is the reaction coefficient and $f$ is the source term. The 1-dimensional domain where the problem is to be solved has been denoted by $\Omega$ and the time interval by $(0,T)$. For simplicity, equation (5.1) will be equipped with the homogeneous Dirichlet boundary condition

$$u = 0 \quad \text{on } \partial\Omega, \quad t \in (0,T)\,, \tag{5.2}$$

and an initial condition of the form

$$u = u^0 \quad \text{in } \Omega, \quad t = 0\,. \tag{5.3}$$

We assume that $\beta$, $\sigma$ are non-negative piecewise constants, $f \in L^2(\,(0,T); L^2(\Omega)\,)$ and $u^0 \in L^2(\Omega)$. Under these assumptions, existence and uniqueness of the solution are guaranteed (see (Raviart and Thomas, 1992)).

We start by recalling the abstract variational formulation of the problem (5.1):

$$
\begin{cases}
\text{Find u} \in \mathrm{L}^2(\,(0,\mathrm{T});\mathrm{V}\,) \cap \mathcal{C}^0(\,[0,\mathrm{T}];\mathrm{L}^2(\Omega)\,) \text{ such that}, \\
(\dot{u}(t), v) + a(u(t), v) = (f(t), v), \quad \forall v \in H_0^1(\Omega), \\
u(0) = u^0,
\end{cases}
\tag{5.4}
$$

where $V = H_0^1(\Omega)$. Here, the superposed dot denotes the time differentiation and $(\cdot, \cdot)$ is the $L^2(\Omega)$ inner product and the bilinear operator is

$$
a(u, v) = \epsilon \int_\Omega u_x\, v_x + \int_\Omega \beta\, u_x\, v + \int_\Omega \sigma\, u\, v.
\tag{5.5}
$$

As announced, we aim to provide a simple recipe for the adaptation of stabilization methods particularly designed for stationary problems to the time-dependent case. In order to obtain reasonable approximations to problem (5.1), we consider the following strategies:

- FDt-FEs: first discretize the time derivative and then discretize in space

- FEs-FDt: first discretize the space derivative and then discretize in time

## 5.3. Discrete Formulations

We next describe in detail the methods that result from the two approaches mentioned in the previous part.

### 5.3.1. First Strategy: Horizontal Method of Lines

In this part, we will apply horizontal method of lines (or Rothe's method) to discretize time-dependent convection-diffusion-reaction problem in which time discretisation is followed by a discretisation in space.

Let $0 = x_0 < x_1 < x_2 < \ldots < x_{N-1} < x_N = 1$ and $\mathcal{T}_h = \{K\}$ be a decomposition of $\Omega$ into subintervals $K = (x_{k-1}, x_k)$ where $k = 1, \ldots, N$ and $\{0 = t_0 < t_1 \ldots < t_M = T\}$ be a uniform partition of time interval with $\Delta t_m = t_{m+1} - t_m$. We will consider fractional-step $\theta$-schemes as temporal discretization of (5.4). These schemes applied to

(5.4) lead at the discrete time $t_n$ to an equation of the form

$$\frac{u^{n+1}}{\Delta t_n} + \theta_1 \mathcal{L} u^{n+1} = \frac{u^n}{\Delta t_n} - \theta_2 \mathcal{L} u^n + \theta_3 f^n + \theta_4 f^{n+1} \quad \text{with} \quad u^0 = u(0),$$
$$n = 0, 1, ..., N - 1. \quad (5.6)$$

The backward Euler scheme is obtained for $\theta_1 = \theta_4 = 1$, $\theta_2 = \theta_3 = 0$ and the Crank–Nicolson scheme for $\theta_1 = \theta_2 = \theta_3 = \theta_4 = 1/2$. Equation (5.6) can be considered as a steady convection–diffusion–reaction equation at $t_n$. It is possible to write (5.6) in a more compact form: Define the associated stationary differential operator

$$\widetilde{\mathcal{L}}(\cdot) = \theta_1 \mathcal{L}(\cdot) + \frac{1}{\Delta t_n} I(\cdot) = -\theta_1 \epsilon \, \partial_{xx}(\cdot) + \theta_1 \beta \, \partial_x(\cdot) + (\theta_1 \sigma + \frac{1}{\Delta t_n})(\cdot), \quad (5.7)$$

where $I$ denotes the identity operator and corresponding bilinear form is defined as

$$\widetilde{a}(u, v) = \theta_1 a(u, v) + \frac{1}{\Delta t_n}(u, v) \qquad \forall u, v \in H_0^1(\Omega). \quad (5.8)$$

Then, (5.6) can be rewritten as,

$$\widetilde{\mathcal{L}} u^{n+1} = \frac{u^n}{\Delta t_n} - \theta_2 \mathcal{L} u^n + \theta_3 f^n + \theta_4 f^{n+1}, \quad n = 0, 1, ..., N - 1, \quad (5.9)$$

and correspondingly,

$$\widetilde{a}(u^{n+1}, v) = \frac{1}{\Delta t_n}(u^n, v) - \theta_2 a(u^n, v) + (\theta_3 f^n + \theta_4 f^{n+1}, v) \qquad \forall v \in H_0^1(\Omega).$$
$$(5.10)$$

Note that, equation (5.9) can be considered as steady convection–diffusion–reaction equation at $t_n$ with the following coefficients:

$$\widetilde{\epsilon} = \theta_1 \epsilon, \quad \widetilde{\beta} = \theta_1 \beta, \quad \widetilde{\sigma} = \theta_1 \sigma + \frac{1}{\Delta t_n}.$$

Since the time discretization is carried out, we discretize problem (5.10) in space by means of the stabilized methods designed for the steady problem which is considered in the following sub-sections:

**The SUPG Method:**

The original SUPG formulation was designed for the steady version of equation (5.1) as a method to avoid the numerical oscillations found using the classical Galerkin approach when $\epsilon$ is very small. The extension to the transient problem that we consider here is based on a previous discretization in time of the continuous equation (5.1). The SUPG method reads: Find $u_h^{n+1} \in V_h$ such that $\forall v_h \in V_h$

$$
\widetilde{a}(u_h^{n+1}, v_h) + \sum_{K \in \mathcal{T}_h} \widetilde{\tau}_K \left( \widetilde{\mathcal{L}} u_h^{n+1}, \, \widetilde{\beta} \, v_h' \right)_K = \frac{1}{\Delta t_n}(u_h^n, v_h) - \theta_2 a(u_h^n, v_h) +
$$

$$
+(\theta_3 f^n + \theta_4 f^{n+1}, v_h) + \sum_{K \in \mathcal{T}_h} \widetilde{\tau}_K \left( \frac{u_h^n}{\Delta t_n} - \theta_2 \mathcal{L} u_h^n + \theta_3 f^n + \theta_4 f^{n+1}, \, \widetilde{\beta} \, v_h' \right)_K
$$
$$
n = 0, 1, ..., N - 1 \,, \tag{5.11}
$$

where $(\cdot, \cdot)_K$ denotes the integration over $K$. It remains to define the algorithmic parameter $\tau_K$, which is also called *intrinsic time*. Appropriate parameters for the SUPG method have been studied in (Asensio et al., 2007). It was proposed to set the parameter on the mesh cell K to be

$$
\widetilde{\tau}_K = \left( \frac{12\widetilde{\epsilon}}{h_K^2} + \frac{2|\widetilde{\beta}|}{h_K} + 2\widetilde{\sigma} \right)^{-1}, \tag{5.12}
$$

which emanates from the maximum principle. Here $h_K$ is an appropriate measure for the size of the mesh cell $K$.

**Stabilized FEM in (Şendur and Neslitürk, 2012):**

In this part, in order to discretize equation (5.10) in space, the stabilization method, which is designed for the stationary problem is implemented (Şendur and Neslitürk, 2012). From now on, we refer to it as P-RFB. This method is based on the RFB method in which the RFB functions are replaced by pseudo RFBs, which retain the same qualitative be-

havior as the RFBs. The pseudo bubbles are computed by using a suitable subgrid inside each element. The choice of the subgrid nodes depends on minimizing the residual of a local differential problems with respect to $L_1$ norm. The resulting numerical method has similar stability features to the RFB method for the whole range of problem parameters. Before we apply this method, we need to re-construct the pseudo bubbles. The key idea is to re-calculate the approximate bubbles according to the operator $\widetilde{\mathcal{L}}$, rather than $\mathcal{L}$. The details of this method can be found in (Şendur and Neslitürk, 2012). Once the subgrid points are re-calculated, pseudo bubbles can be re-computed to re-construct the bubble space $\widetilde{V}_B$. Then we can apply the P-RFB strategy to (5.9), that is

$$
\begin{cases}
\text{Given } u_h^0 \in V_h = V_L \bigoplus \widetilde{V}_B \text{ find } u_h^{n+1} \in V_h \text{ such that } \forall v_h \in V_h\,, \\
\widetilde{a}(u_h^{n+1}, v_h) = \frac{1}{\Delta t_n}(u_h^n, v_h) - \theta_2 a(u_h^n, v_h) + (\theta_3 f^n + \theta_4 f^{n+1}, v_h)\,, \\
n = 0, 1, ... N - 1\,,
\end{cases}
\tag{5.13}
$$

where $V_L$ is the space of continuous piecewise linear polynomials.

**Remark 5.1** *With the present approach, we obtained a family of stationary convection-diffusion-reaction problems with modified coefficients. In view of equation (5.7), these problems could be reaction-dominated for small time steps even if the reaction term $\sigma$ is not dominant. Therefore, applying schemes that stabilizes the steady convection–diffusion–reaction equation have potential advantages to cope with the destabilizing effects at small time steps.*

## 5.3.2. Second Strategy: Vertical Method of Lines

In this part, we will apply vertical method of lines to discretize time-dependent convection-diffusion-reaction problem in which space discretisation is followed by a discretisation in time. If we discretize (5.1) first in space by SUPG method, we are lead to the scheme

$$
\frac{d}{dt}(u_h(t), v_h) + a(u_h(t), v_h) + \sum_{K \in \mathcal{T}_h} \tau_K \left( \frac{du_h(t)}{dt} + \mathcal{L}u_h\,, \widetilde{\beta}\, v_h' \right)_K
$$

$$
= (f(t), v_h) + \sum_{K \in \mathcal{T}_h} \tau_K \left( f(t)\,, \widetilde{\beta}\, v_h' \right)_K.
\tag{5.14}
$$

Now, we have a system of ODEs which can be integrated in time by using fractional-step $\theta$-scheme to get the following fully discretized problem: Find $u_h^{n+1} \in V_h$ such that $\forall v_h \in V_h$

$$\widetilde{a}(u_h^{n+1}, v_h) + \sum_{K \in \mathcal{T}_h} \tau_K \left( \widetilde{\mathcal{L}} u_h^{n+1} , \widetilde{\beta} \, v_h^{'} \right)_K = \frac{1}{\Delta t_n}(u_h^n, v_h) - \theta_2 a(u_h^n, v_h) +$$

$$+(\theta_3 f^n + \theta_4 f^{n+1}, v_h) + \sum_{K \in \mathcal{T}_h} \tau_K \left( \frac{u_h^n}{\Delta t_n} - \theta_2 \mathcal{L} u_h^n + \theta_3 f^n + \theta_4 f^{n+1} , \widetilde{\beta} \, v_h^{'} \right)_K$$

$$n = 0, 1, ..., N - 1, \tag{5.15}$$

where the algorithmic parameter $\tau_K$ is set to be

$$\tau_K = \left( \frac{12\epsilon}{h_K^2} + \frac{2|\beta|}{h_K} + 2\sigma \right)^{-1}. \tag{5.16}$$

If we discretize (5.1) first in space by the P-RFB method, we get the following scheme:

$$\begin{cases} \text{Find } u_h \in V_h = V_L \oplus V_B \text{ such that}, \\ \frac{d}{dt}(u_h(t), v_h) + a(u_h(t), v_h) = (f, v_h) \quad \forall v_h \in V_h. \end{cases} \tag{5.17}$$

Note that, construction of internal nodes and therefore the bubble space $V_B$ is done according to the steady operator $\mathcal{L}$. Now, we have a system of ODEs which can be integrated in time by using fractional-step $\theta$-scheme to get the following fully discretized problem:

$$\begin{cases} \text{Find } u_h^{n+1} \in V_h \text{ such that } \forall v_h \in V_h, \\ \frac{(u_h^{n+1}, v_h) - (u_h^n, v_h)}{\Delta t_n} + \theta_1 a(u_h^{n+1}, v_h) + \theta_2 a(u_h^n, v_h) = (\theta_3 f^n + \theta_4 f^{n+1}, v_h), \\ n = 0, 1, ...N - 1. \end{cases} \tag{5.18}$$

## 5.4. Numerical Results

In this section, we report some numerical experiments to illustrate the performance of several stabilization methods; SUPG method, P-RFB -strategy and LCB-strategy, some

of which are described in the previous section. In our calculations, we take a uniform partition of $\Omega$ into subintervals of length $h = |\Omega|/N$. For the time integration, we use Crank$-$Nicolson scheme with fixed time step $\Delta t$. For each value of $h$, every experiment is carried out with different values of the time step $\Delta t$. In particular we have run the computations for

$$CFL = \frac{\Delta t |\beta|}{h} = 0.1, 0.5, 1 \,. \qquad (5.19)$$

Moreover, in order to make the comparison of each strategy easy, all figures consist of a pair of graphics: FDt-FEs approach is always on the left figure while FEs-FDt is the one on the right figure.

**Experiment 1:** First, we consider the following problem:

$$
\begin{aligned}
&u_t - \epsilon u_{xx} + \beta u_x + \sigma u = \sigma \,, \\
&u(0, t) = 0 \quad u(1, t) = 0, \; t \in [0, \; 1] \,, \qquad (5.20) \\
&u(x, 0) = 0 \,.
\end{aligned}
$$

We set $\epsilon = 10^{-6}$, $\beta = 1$, $N = 40$. We plot the solutions obtained with the present method for various values of Courant number ($CFL = 0.1, 0.5, 1$) and reaction ($\sigma = 1, \, 10, \, 50$). We display the results in Figure 5.1- Figure 5.3. We do not display the numerical results when discretization in time is followed by space discretization with SUPG method because the results are too oscillatory. It seems that when using the FDt-FEs approach, the diffusion added in the streamline direction by the classical stabilization techniques, is not enough to eliminate the spurious oscillations. With the FEs-FDt approach, the oscillations are not completely removed from the approximation. When we apply the LCB method or P-RFB strategy, it can be clearly observed that spurious oscillations have almost disappeared.
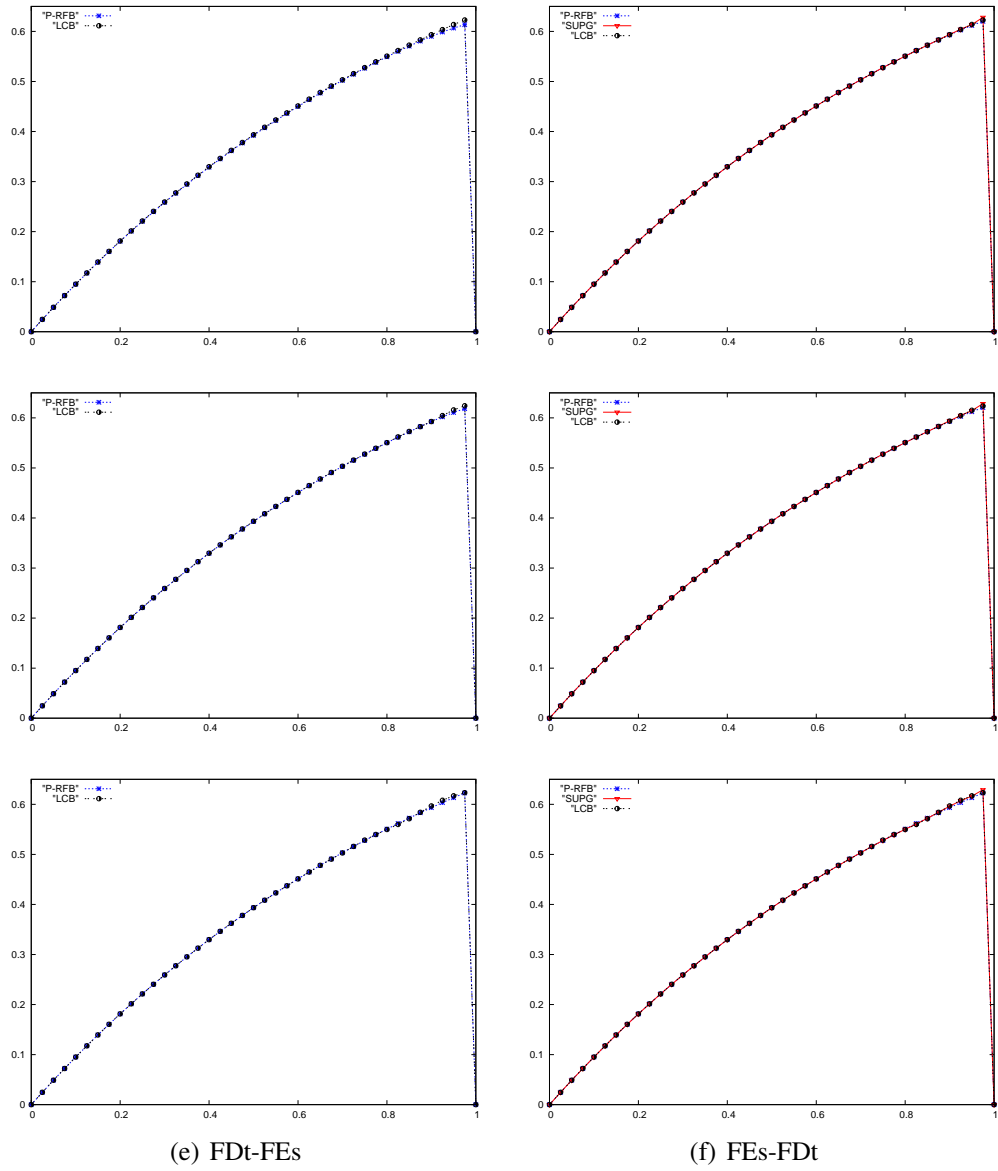
(e) FDt-FEs        (f) FEs-FDt

Figure 5.1. The numerical solution when $\epsilon = 10^{-6}$, $\beta = 1$, $\sigma = 1$. Top to bottom: $CFL = 0.1,\ 0.5,\ 1$.

(e) FDt-FEs    (f) FEs-FDt

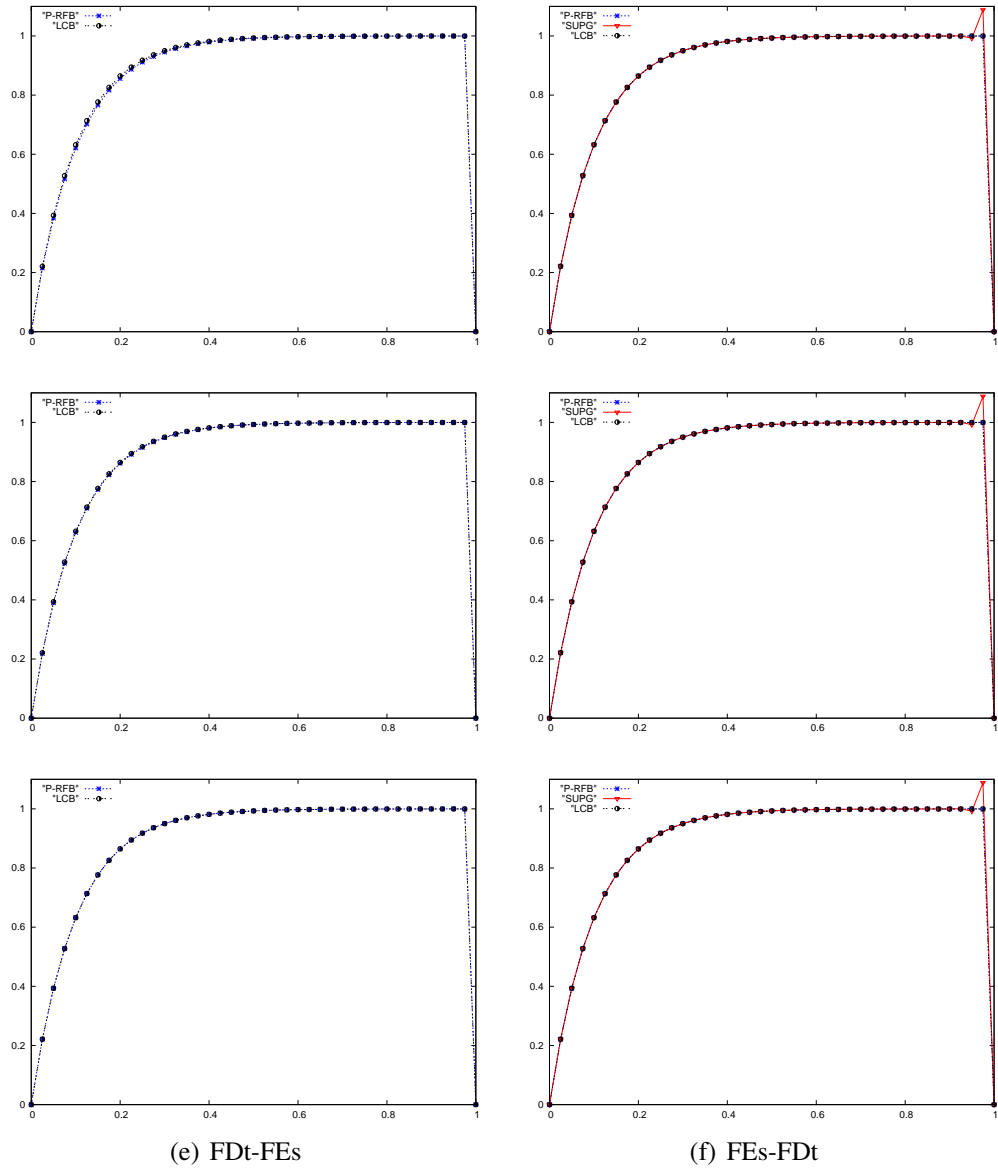Figure 5.2. The numerical solution when $\epsilon = 10^{-6}$, $\beta = 1$, $\sigma = 10$. Top to bottom: $CFL = 0.1,\ 0.5,\ 1$.

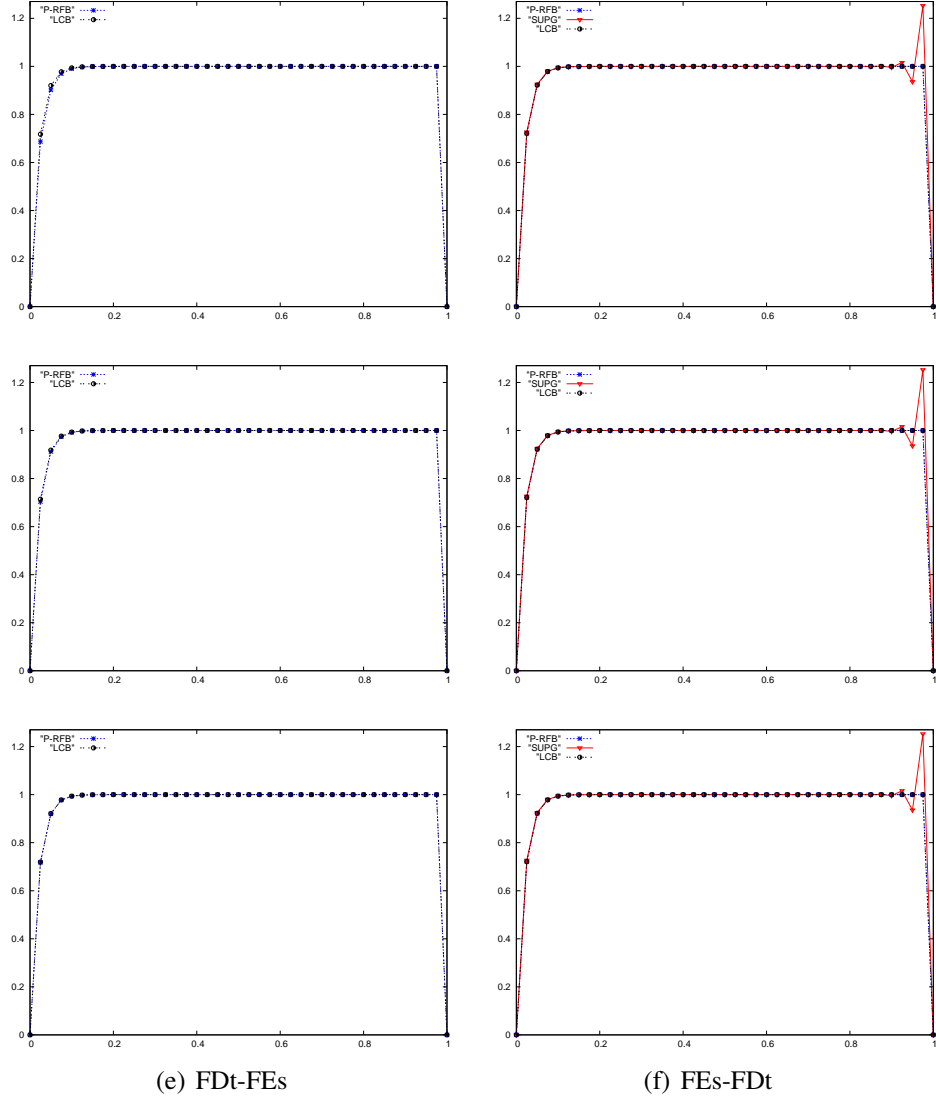(e) FDt-FEs                                    (f) FEs-FDt

Figure 5.3. The numerical solution when $\epsilon = 10^{-6}$, $\beta = 1$, $\sigma = 50$. Top to bottom:
$CFL = 0.1, 0.5, 1$.

**Experiment 2:** Next, we consider a test case shown in problem (5.21) for which the exact solution exhibits a boundary layer. We set $\epsilon = 10^{-6}$, $\beta = 1$, decompose the domain into a uniform discretization of 40 elements, i.e. $N = 40$ and plot the solutions obtained with the present method for various values of Courant number ($CFL = 0.1, 0.5, 1$) and reaction ($\sigma = 10^{-4}, 10$). We display the results in Figure 5.4- Figure5.5.

$$
\begin{cases}
u_t - \epsilon u_{xx} + \beta u_x + \sigma u = 0, \\
u(0,t) = 0 \quad u(1,t) = 1, \qquad t \in [0, \ 0.5], \\
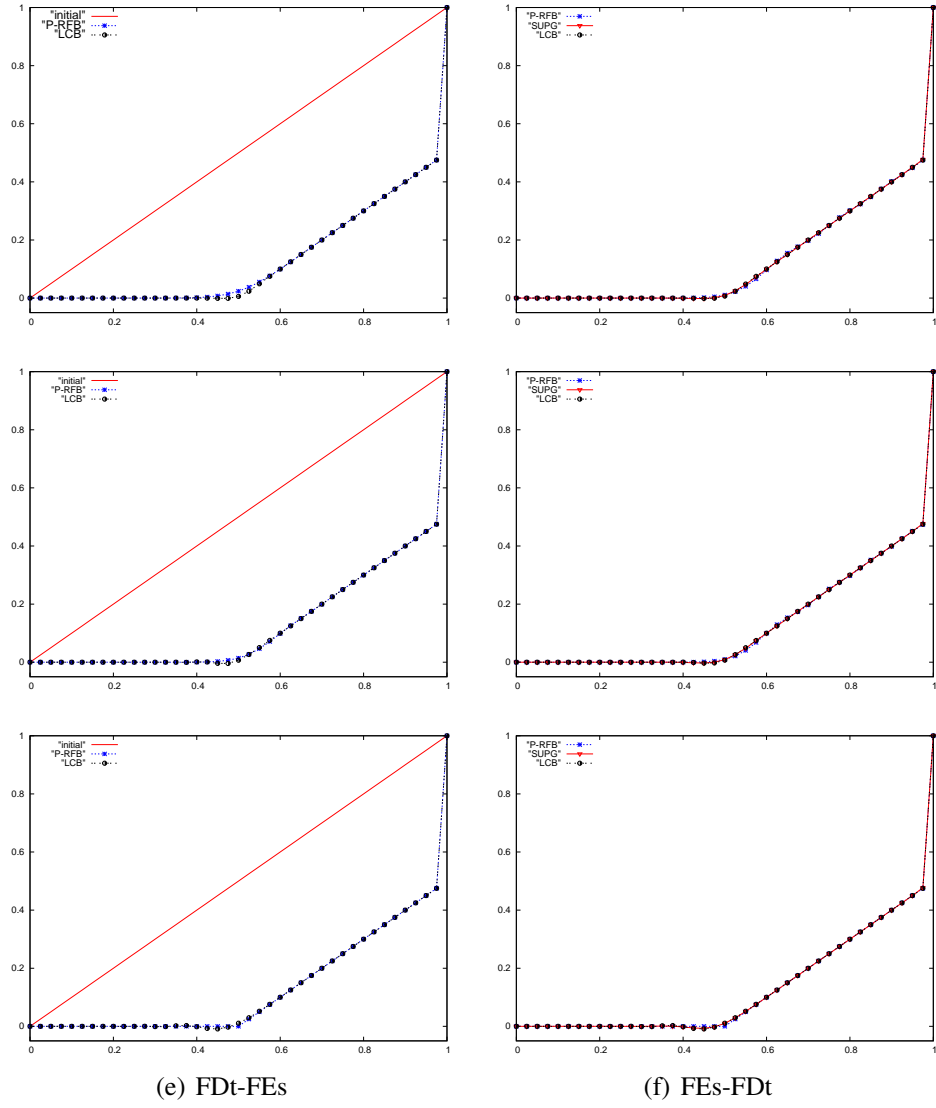u(x,0) = x\,.
\end{cases} \tag{5.21}
$$

97

(e) FDt-FEs

(f) FEs-FDt

Figure 5.4. The numerical solution when $\epsilon = 10^{-6}$, $\beta = 1$, $\sigma = 10^{-4}$. Top to bottom: $CFL = 0.1,\ 0.5,\ 1$.

(e) FDt-FEs                    (f) FEs-FDt
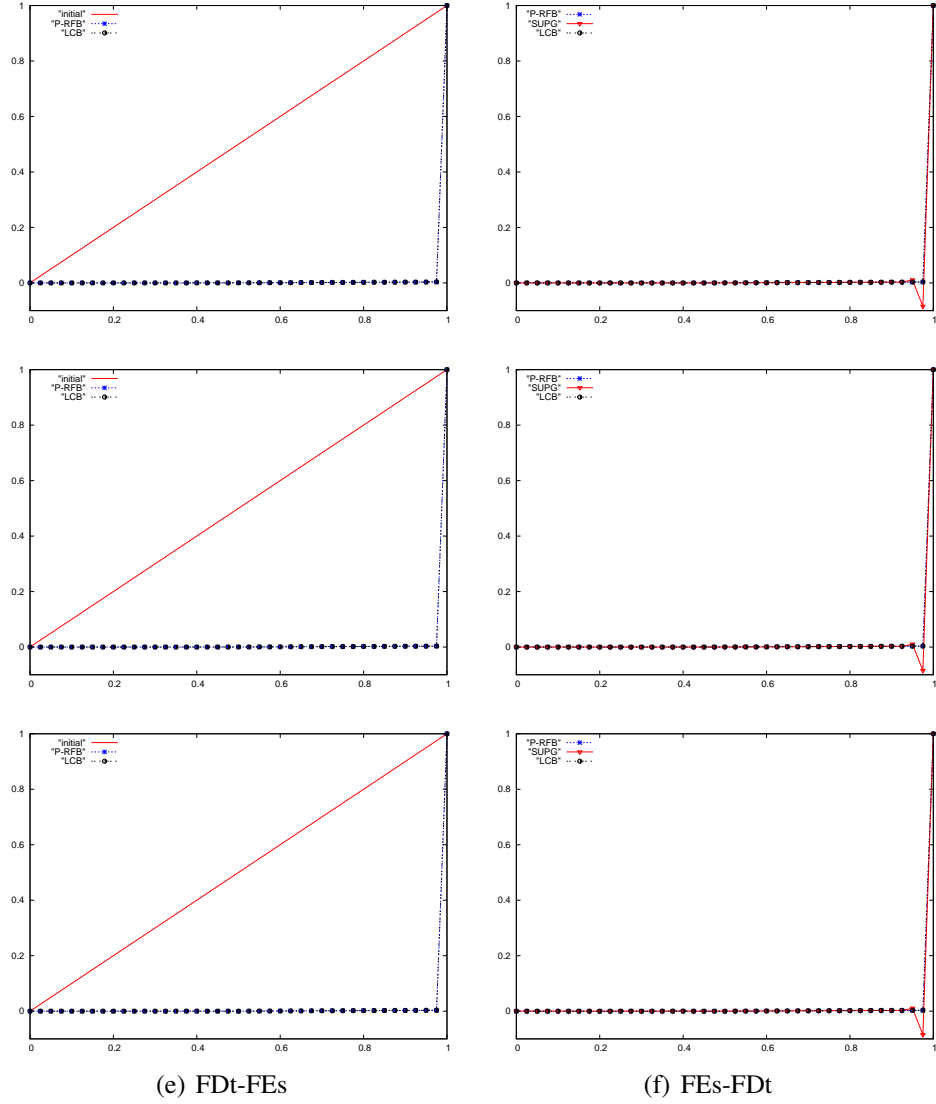
Figure 5.5. The numerical solution when $\epsilon = 10^{-6}$, $\beta = 1$, $\sigma = 10$. Top to bottom: $CFL = 0.1$, $0.5$, $1$.

We have presented several numerical approaches with different stabilization techniques to solve problem (5.1) for a wide range of parameters. We have compared the classical stabilization methods adapted to transport problems, with the new ones based on (Şendur and Neslitürk, 2012). We have tested all the schemes presented with different problems. Numerical results confirm the good performance of the method in (Şendur and Neslitürk, 2012) as compared with SUPG method and LCB strategy.

# CHAPTER 6

# CONCLUSIONS

It is well known that the exact solution of CDR problems may contain layers when some problem parameters are too big compared to others. The classical numerical methods, such as the central finite difference method or the plain Galerkin method, are inadequate when the diffusive term is small compared to the convective term. To deal with such situation, we first study a stabilized finite element method for one-dimensional convection-diffusion-reaction problems. We first discretize the domain into uniform elements. Then, we define a sub-grid in a typical element $K$ by adding two internal points on which, we approximate the bubble functions. The shape of approximations, which is essentially related with the location of sub-grid points, is crucial to get a good stabilization effect on the numerical method. Therefore the choice of points in the sub-grid is accomplished by a minimization process with respect to $L_1$ norm in the presence of layers. The stabilized FEM is shown to avoid the oscillations in the solution produced by the standard FEM.

Next, the idea in one-dimensional case is extended to obtain a stabilized finite element method for two-dimensional convection-diffusion-reaction problems. The method is based on a finite element formulation on an enriched space, in which the standard piecewise polynomial functions are enriched by means of bubbles, i.e., functions whose support remains inside the elements. The RFB functions are replaced by their cheap, yet efficient approximations which retain the same qualitative behavior. The approximate bubbles are computed on a suitable sub-grid, the choice of whose nodes are critical and determined by minimizing the residual of a local problem with respect to $L_1$ norm. The resulting numerical method is shown to have similar stability features with the RFB method for the whole range of problem parameters and this fact is also confirmed by numerical experiments.

Finally, some numerical schemes for the unsteady CDR problem, in one space dimension is presented. This is a combined stabilized method in which the discretization in time is performed by Crank−Nicolson method and for the spatial disceretization several stabilized finite element schemes including more recent strategy proposed in (Şendur and Neslitürk, 2012) is used.

# REFERENCES

Ahmed, N., Matthies, G., Tobiska, L. and Xiec, H., 2011: Discontinuous Galerkin time stepping with local projection stabilization for transient convection-diffusion-reaction problems. *Comput. Methods Appl. Mech. Engrg.*, **200**, 1747–1756.

Anderson, D. A., Tannehill, J. C. and Pletcher, R. H., 1984: *Computational Fluid Mechanics and Heat Transfer.* Hemisphere, Washington, DC.

Asensio, M. I., Russo, A. and Sangalli, G., 2004: The Residual-Free Bubble numerical method with quadratic elements. *Math. Models Methods Appl. Sci.* **14**, 641–661.

Asensio, M. I., Ayuso, B. and Sangalli, G., 2007: Coupling stabilized finite element methods with finite difference time integration for advection–diffusion–reaction problems. *Comput. Methods Appl. Mech. Engrg.*, **196**, 3475–3491.

Baiocchi, C., Brezzi, F. and Franca, L. P., 1993: Virtual bubbles and the GaLS. *Comput. Methods Appl. Mech. Engrg.* **105**, 125–141.

Brezzi, F., Bristeau, M. O., Franca, L. P., Mallet, M. and Roge, G., 1992: A relationship between stabilized finite element methods and the Galerkin method with bubble functions. *Comput. Methods Appl. Mech. Engrg.* **96**, 117–129.

Brezzi, F., Marini, D. and Russo, A., 1998a: Applications of pseudo residual-free bubbles to the stabilization of convection-diffusion problems. *Comput. Methods Appl. Mech. Engrg.*, **166**, 51–63.

Brezzi, F., Franca, L. P. and Hughes, T. J. R., 1997a: $b = \int g$. *Comput. Methods Appl. Mech. Engrg.* **145**, 329–339.

Brezzi, F., Franca, L. P., Hughes, T. J. R. and Russo, A., 1997b: Stabilization techniques and subgrid scales capturing. The State of the Art in Numerical Analysis, *I.S. Duff and G.A. Watson Eds.1*, 1996, *IMA Conference Series 63*, Oxford University Press, 1997, 391–406.

Brezzi, F., Franca, L. P. and Russo, A., 1998b: Further considerations on residual-free bubbles for advective-diffusive equations. *Comput. Methods Appl. Mech. Engrg.* **166**, 25–33.

Brezzi, F., Hauke, G. , Marini, D. and Sangalli, G., 2003: Link-cutting bubbles for the stabilization of convection-diffusion-reaction problems. *Math. Models Meth. Appl. Sci.*, **13**, 445–461.

Brezzi, F., Hughes, T. J. R., Marini, D., Russo, A. and Suli, E., 1999: A priori error

analysis of residual-free bubbles for advection-diffusion problems. *SIAM J. Numer. Anal.* **36**, 1933–1948.

Brezzi, F. and Marini, D., 2002a: Augmented spaces, two-level methods and stabilizing sub-grids. *Inter. J. Numer. Methods Fluids*, **40**, 31–46.

Brezzi, F. and Marini, D., 2002b: Subgrid phenomena and numerical schemes. *Mathematical Modeling and Numerical Simulation in Continuum Mechanics, I. Babuska, P.G. Ciarlet, T. Miyoshi, Eds., Springer Lect. Notes in Comput. Sci. Eng.*, **19**, 73–90.

Brezzi, F., Marini, D. and Russo, A., 2005: On the choice of a stabilizing sub-grid for convection-diffusion problems. *Comput. Methods Appl. Mech. Engrg.*, **194**, 127–148.

Brezzi, F., Marini, D. and Suli, E., 2000: Residual-free bubbles for advection-diffusion problems: The general error analysis. *Numer. Math.* **85**, 31–47.

Brezzi, F. and Russo, A., 1994: Choosing bubbles for advection-diffusion problems. *Math. Models Meth. Appl. Sci.*, **4**, 571–587.

Brooks, A. N. and Hughes, T. J. R., 1982: Streamline upwind/Petrov-Galerkin formulations for convection-dominated flows with particular emphasis on the incompressible Navier-Stokes equations. *Comput. Methods Appl. Mech. Engrg.*, **32**, 199–259.

Cangiani, A. and Suli, E., 2005: Enhanced RFB method. *Numer. Math.* **101**, 273–308.

Ciarlet, P. G., 1978: *The finite element methods for elliptic problems.* North-Holland, Amsterdam.

Codina R., 1998: Comparison of some finite element methods for solving the diffusion-convection-reaction equation. *Comput. Methods Appl. Mech. Engrg.*, **156**, 185–210.

Codina R., 2000: On stabilized finite element methods for linear systems of convection-diffusion-reaction equations. *Comput. Methods Appl. Mech. Engrg.*, **188**, 61–82.

Farrell, P. A., Hegarty, A. F., Miller, J. J. H., O'Riordan, E. and Shishkin, G. I., 2000: *Robust computational techniques for boundary layers.* Chapman and Hall/CRC Press, Boca Raton, FL.

Franca, L. P. and Frey, S. L., 1992: Stabilized finite element methods: II. The incompressible Navier- Stokes equations. *Comput. Methods Appl. Mech. Engrg.*, **99**, 209–233.

Franca, L. P., Frey, S. L. and Hughes, T. J. R., 1992: Stabilized finite element methods: I. Application to the advective-diffusive model. *Comput. Methods Appl. Mech. Engrg.*, **95**, 253–276.

Franca, L. P., Hauke, G. and Masud, A., 2006: Revisiting stabilized finite element methods for the advective-diffusive equation. *Comput. Methods Appl. Mech. Engrg.*, **195**, 1560–1572.

Franca, L. P., Madureira, A. L. and Valentin, F., 2005a: Towards multiscale functions: Enriching finite element spaces with local but not bubble-like functions. *Comput. Methods Appl. Mech. Engrg.*, **194**, 3006–3021.

Franca, L. P., Neslituurk, A. and Stynes, M., 1998: On the stability of residual-free bubbles for convection-diffusion problems and their approximation by a two-level finite element method. *Comput. Methods Appl. Mech. Engrg.*, **166**, 35–49.

Franca, L. P., Ramalho, J. V. A. and Valentin, F., 2005b: Multiscale and residual-free bubble functions for reaction-advection-diffusion problems. *Internat. J. Multiscale Comput. Engrg.*, **3**, 297–312.

Franca, L. P., Russo, A., 1996: Deriving upwinding, mass lumping and selective reduced integration by RFB. *Appl. Math. Lett.* **9**, 83–88.

Franca, L. P., Tobiska, L., 2002: Stability of the residual free bubble method for bilinear finite elements on rectangular grids. *IMA J. Numer. Anal.* **22**, 73–87.

Franca, L. P. and Valentin, F., 2000: On an improved unusual stabilized finite element method for the advective-reactive-diffusive equation. *Comput. Methods Appl.Mech. Engrg.*, **190**, 1785–1800.

Harari, I. and Hauke, G., 2007: Semidiscrete formulations for transient transport at small time steps. *Int. J. Numer. Meth. Fluids*, **54**, 731–743.

Henao, C. A. A., Coutinho, A. L. G. A. and Franca, L. P., 2010: A stabilized method for transient transport equations. *Comput Mech*, **46**, 199–204.

Hughes, T. J. R., 1987: *The finite element method: Linear static and dynamic finite element analysis.* Prentice-Hall, Englewood Cliffs, New Jersey.

Hughes, T. J. R. and Brooks, A. N., 1979: A multidimensional upwind scheme with no crosswind diffusion. *ASME Monograph AMD-34.*, **34**, ASME, New York.

Hughes, T. J. R., Franca, L. P. and Hulbert, G., 1989: A new finite element formulation for computational fluid dynamics: VIII. The Galerkin/least-squares method for advective-diffusive equations. *Comput. Methods Appl. Mech. Engrg.* **73**, 173–189.

Hughes, T. J. R. and Mallet, M., 1986: A new finite-element formulation for computational fluid dynamics. III. The generalized streamline operator for multidimensional advective-diffusive systems. *Comput. Methods Appl. Mech. Engng.*, **58**, 305–328.

John, V. and Schmeyer, E., 2008: Finite element methods for time-dependent convection–diffusion–reaction equations with small diffusion. *Comput. Methods Appl. Mech. Engrg.*, **198**, 475–494.

Johnson, C., Navert, U. and Pitkaranta J., 1984: Finite element methods for linear hyperbolic problems. *Comput. Methods Appl. Mech. Engrg.*, **45**, 285–312.

Johnson, C., 1987: *Numerical solution of partial differential equations by the finite element method.* Cambridge University Press, Cambridge.

Morton, K. W., 1996: *Numerical Solution of Convection-Diffusion Problems.* Chapman and Hall, London, UK.

Neslitürk, A. I., 2006: A stabilizing sub-grid for convection-diffusion problem. *M3AS*, **16**, 211–231.

Neslitürk, A. I., 2010: On the choice of stabilizing subgrid for convection-diffusion problem on rectangular grids. *Comput. Math. Appl.* **59**, 3687–3699.

Oden, J. T., 1991: Finite elements: An introduction, Handbook of Numerical Analysis *Elsevier* **2**, 3–15.

Raviart, P.A. and Thomas, J.M., 1992: Introduction al analyse numerique des equations aux derivees partielles, Masson, Paris.

Roos, H. G., Stynes, M. and Tobiska, L., 2008: Robust Numerical Methods For Singularly Perturbed Differential Equations. Verlag, Berlin: Springer.

Sangalli, G., 2000: Global and local error analysis for the residual-free bubbles method applied to advection-dominated problems. *SIAM J.Numer.Anal.*, **38**, 1496-1522.

Stynes, M., 2005: Steady-state convection-diffusion problems. *Acta Numerica*, Cambridge University Press, Cambridge, 445–508.

Şendur, A. and Neslıturk, A. I., 2012: Applications of the pseudo residual-free bubbles to the stabilization of convection-diffusion-reaction problems. *Calcolo*, **49**, 1–19.

# VITA

**Date of Birth and Place:** 21.08.1984, Aydın - Turkey

## EDUCATION

**2007 - 2012 Doctor of Philosophy in Mathematics**

Graduate School of Engineering and Sciences, İzmir Institute of Technology, İzmir -Turkey.

Thesis Title: ENRICHED FINITE ELEMENTS METHOD FOR
            CONVECTION-DIFFUSION-REACTION PROBLEMS

Supervisor: Prof. Dr. Oktay PASHAEV

Co-Supervisor: Prof. Dr. Ali İ. NESLİTÜRK

**2002 - 2007 Bachelor of Mathematics**

Department of Mathematics, Dokuz Eylül University, İzmir - Turkey.

## PROFESSIONAL EXPERIENCE

**2007 - 2012 Research and Teaching Assistant**

Department of Mathematics, İzmir Institute of Technology, İzmir - Turkey.

## PUBLICATIONS

Şendur, A. and Neslitürk, A. I., Applications of the pseudo residual-free bubbles to the stabilization of convection-diffusion-reaction problems, Calcolo $49\ (2012),\ 1-19.$

Filiz, A., Neslitürk, A. I. and Şendur, A., A fully discrete $\epsilon$-uniform method for singular perturbation problems on equidistant meshes, International Journal of Computer Mathematics. $89\ (2012),\ 190-199.$