INSTITUTO NACIONAL DE MATEMÁTICA PURA E APLICADA

DOMAIN DECOMPOSITION ANALYSIS FOR HETEROGENEOUS DARCY'S FLOW

ANÁLISE DE DECOMPOSIÇÃO DE DOMÍNIO
PARA FLUXOS DE DARCY HETEROGÊNEOS

Author: Juan Carlos Galvis Arrieta

Adviser: Prof. Dr. Marcus Sarkis

March 2008

Contents

List of Tables

List of Figures

To my parents Mari and Rafael.
To my brothers Pedro and Manuel.
To my nephew Juan Manuel.

Abstract

This thesis focuses on finite element and domain decomposition applications to three mathematical models related to fluid flow in porous media. The models have application in a variety of fields including areas such as petroleum engineering, environmental sciences, hydrology and biology, among others. The first model considered is the Darcy-Stokes coupling. We study the well posedness of the continuous model. We introduce a discretization, obtain the well posedness of the discrete model and derive a priori error estimates. We also design and analyze two domain decomposition preconditioners. The second model is the pressure equation with discontinuous coefficients. Here we design and analyze several domain decomposition preconditioners for the resulting linear system of a Discontinuous Galerkin type discretization. The third subject is the study of the stochastic pressure equation (without replacing the ordinary product by the Wick product). We use the white noise measure constructed from a Hilbert space and an operator to define and characterize adequate spaces for its solution. The approximation consists of a truncated Chaos expansion. We verify the well posedness of the discrete model and provide a priori error estimates. In all cases numerical experiments verify the theoretical results.

Keywords: finite element, porous media flow, multiphysics, multidomain, inf-sup condition, error estimates, mortar, non-matching grids, Stokes-Darcy coupling, domain decomposition preconditioners, discontinuous coefficients, interior penalty discretization, Discontinuous Galerkin, white noise analysis, Wiener Chaos expansion, stochastic simulation, ordinary product stochastic pressure equation

Resumo

Esta tese concentra-se em aplicações da análise de elementos finitos e decomposição de domínios a três modelos relacionados com o fluxo de fluidos em meios porosos. Estes modelos tem aplicações em varias áreas como engenharia de petróleo, ciências ambientais, hidrologia e biologia, entre outras. O primeiro modelo considerado é o acoplamento Darcy-Stokes. Estudamos a boa colocação do modelo contínuo. Introduzimos uma discretização, obtemos a boa colocação do modelo discreto e estimativas de erro *a priori*. Também desenvolvemos e analisamos precondicionadores de decomposição de domínios. O segundo modelo é a equação da pressão com coeficientes descontínuos. Desenvolvemos e analisamos vários preconditionadores de decomposição de domínio para o sistema linear que resulta de uma discretização do tipo Galerkin Descontínuo. O terceiro tópico é o estudo da equação da pressão estocástica (com produto ordinário ao invés do produto Wick). Usamos a medida de ruido branco construída de um espaço de Hilbert e um operador para definir e caracterizar espaços adequados para sua solução. A aproximação consiste de uma expansão em Caos truncada. Verificamos a boa colocação do modelo discreto e apresentamos estimativas de erro *a priori*. Nos três casos, experimentos numéricos confirmam os resultados teóricos.

Palavras-chaves: elementos finitos, fluxo em meio poroso, multi-física, multi-domínio, condição inf-sup, estimativas de erro, mortar, malhas não alinhadas, acoplamento Stokes-Darcy, precondicionadores de decomposição de domínio, coeficientes descontínuos, Galerkin descontínuo, análise de ruído branco, expansão em caos de Wiener, simulação estocástica, equação da pressão estocástica com produto ordinário.

Acknowledgments

I thank the financial support of 'Programa Estudantes-Convênio de Pós-Graduação -**PEC-PG** (CAPES)' and CNPq.

I am thankful to my adviser Professor Marcus Sarkis who always keep surprising me with his wide range of knowledge. He was always available for scientific conversations. I thank to Professor Maksymilian Dryja for accepting my participation in the research project concerning domain decomposition preconditioners for discontinuous Galerkin discretizations. I thank to Professors Jorge Zubelli, Dan Marchesin, Tarek Mathew, Abimael Dourado Loula, Carlos Tomei and Henrique Versieux for taking part in the examination committee.

I further express my thanks to the Fluid Lab people (Wanderson, Duilio, ET, Pablo, Ana, Grigori, Julio, Prof. André Nachbin, Sergio). I also want to thank to 'INSTITUTO NACIONAL DE MATEMÁTICA PURA E APLICADA- IMPA' for providing an excellent atmosphere of work; special thanks to 'ensino' (Andrea, Jusenildo) and room 210 (Rogerinho).

I thank to LG, Milton, Johel and other IMPA students for the scientific conversations.

I thank to all my friends in Rio de Janeiro. Special thanks to my friends from Colombia, Chile, Argentina, Brazil, Ecuador, Bolivia, Peru, Iran, Paraguay, Uruguay, Venezuela, Portugal, France and Spain. Further special thanks to the *Bar Bina Club*: {damián,pancho}@impa.br, {david,julian}@cbpf.bf and its many temporary members. I also thank to the people of the apartment 604 at *Marques de Olinda* 61 (Javier, Freddy, Airon).

I am grateful to my family and my girlfriend in Colombia for their encouragement and support.

Part I

Preliminaries and Summaries

Chapter 1

Introduction

Modeling fluid flow in porous media appears in a wide range of applications including areas such as petroleum engineering, environmental sciences, biology, hydrology and geology, among others. In this chapter we present some general ideas concerning the modeling of fluid flow in porous media and we introduce the topics studied in this thesis.

1.1 A snapshot of modeling of fluid flow in porous media

Modeling of fluid flow in porous media involves several important theoretical and numerical questions. The development of new mathematical and computational framework is fundamental for dealing efficiently with several important difficulties inherent to the porous medium modeling such as heterogeneities and lack of measurements. Any modeling project involving porous media flow, like in the numerical simulation of (multiphase) reservoir problems, has to deal with these difficulties. The numerical methods used to overcome these challenges should be efficient and take advantages of the increasing understanding of the mathematical models and of the growing computational capacity of modern computers .

The Darcy's law is one of the main tools in the (continuous) modeling of porous media flow. For a single fluid in the porous domain $D \subset \mathbb{R}^d$, this law can be written as

$$\boldsymbol{u} = -\frac{K}{\mu}\nabla p + \boldsymbol{F}, \tag{1.1}$$

where $\boldsymbol{u} : D \to \mathbb{R}^d$ is the seepage velocity of the fluid, the function $p : D \to \mathbb{R}$ is the fluid pressure, the scalar μ is the fluid viscosity and the function $\boldsymbol{F} : D \to \mathbb{R}^d$ represents an external force. The parameter K is the absolute permeability tensor that summarizes the capacity of the medium of letting the fluid flow. In general, K is a symmetric positive definite matrix and, as should be expected, is a very complicated object that offers several challenges and difficulties to overcome in its numerical treatment.

In the numerical modeling of fluid flows in porous media we have to solve equation (1.1). For instance, when dealing with the mathematical modeling of transport of pollutants or oil recovery processes, we have to face a system of,

possibly stochastic, partial differential equations which models the two-phase flow in a porous medium. The system is composed of two equations, a transport equation for the saturation (the relative volume of one of the two fluids) coupled with an equation for the velocity field, which is given by the Darcy's law and the incompressibility condition of the flow. With no sources or sinks, and neglecting the effects of gravity and capillarity, these equations are of the form (Ewing [1983], Bedrikovetsky [1993]):

$$\begin{aligned} \mathbf{u} &= \lambda(s)K\nabla p \\ \nabla \cdot \mathbf{u} &= 0 \\ \frac{\partial s}{\partial t} + \nabla \cdot (F(s)\mathbf{u}) &= 0. \end{aligned} \tag{1.2}$$

Here, $\mathbf{u}$ is the total seepage velocity, s is the water saturation, K is the permeability and p is the pressure. The constitutive functions $\lambda(s)$ and $F(s)$ represent the total mobility and the fractional flow of the water, respectively. The numerical approximation of (1.2) needs to deal with an equation of the form

$$-\nabla.\,(\kappa\nabla p) = f \text{ in } D, \tag{1.3}$$

where $\kappa = \lambda(s)K$. Instead of equation (1.3), its mixed formulation can appear, i.e., an equation of the form

$$\begin{cases} \kappa^{-1}\boldsymbol{u} + \nabla p &= 0, \text{ in } D \\ \nabla \cdot \boldsymbol{u} &= f, \text{ on } D. \end{cases} \tag{1.4}$$

Approximating the solution of (1.3) or (1.4) requires the computation of the solution of a large, sparse, ill conditioned and (maybe) indefinite linear system. The solution of this kind of linear system requires large CPU time and a lot of memory resources. In general, solving (1.3) or (1.4) is the computational bottleneck of the overall numerical computation. The situation is similar in other examples such as three phase flow models and other problems; see Ewing [1983], Bedrikovetsky [1993].

1.2 Summary of the thesis

We study three important subjects related with the numerical approximation of the solutions of equations (1.3) and (1.4).

- **Fluid flow across the well/reservoir interface:** It is in the surroundings of the well/reservoir interface where occurs the most interesting dynamics. We cannot expect that the Darcy's law by itself can offer a good modeling choice because the governing equations of the flow are very different on each side of the interface. Then, it is desirable to have a reliable theoretical and numerical model that mimics this complex coupling situation as accurate as possible.
- **Heterogeneity of the permeability field:** In reservoir modeling applications, the permeability of the rock can vary in order of magnitude. The variation in the permeability leads to difficulties in the approximation of some

quantities such as pressure and seepage velocity since it is expected that these quantities behave according to the fluctuations of the permeability. Numerical methods designed for homogeneous permeability do not work properly in the presence of heterogeneities. Then it is desirable to use adequate numerical schemes designed for heterogeneous permeability fields. These schemes should be reliable and efficient in the computation of the quantities of interest (seepage velocity, pressure, etc).

- **Uncertainty of the permeability field:** The heterogeneities of the porous medium occur at different scales, from the porous to the reservoir scale. This multiplicity of scales makes impossible the accurate knowledge of the properties of the rock. These properties determine the flow and/or transport of fluids in the porous medium. This uncertainty is incorporated to the model by letting some coefficients of the governing equations to be random or stochastic. Then, it is desirable to have numerical schemes that capture the behavior of the solution that depends on this uncertainty. We want this numerical methods to be reliable and efficient.

This thesis is the compilation of five (published or submitted for publication) works addressing mathematical and numerical analysis issues of the three main topics just described.

- **Fluid flow across the well/reservoir interface:** We consider this topic from the well-posedness of the linear stationary model to the construction and analysis of finite element discretizations and domain decomposition preconditioners.
- **Heterogeneity of the permeability field:** We focus on the efficient computation of the numerical solution of the pressure equation with discontinuous coefficients. We consider a Discontinuous Galerkin discretization for properly handling of the discontinuous coefficients and concentrate our effort in the construction of domain decomposition preconditioners.
- **Uncertainty of the permeability field:** We consider the stochastic pressure equation. We study the existence of solutions of this stochastic partial differential equation and propose a finite element approximation.

In all the cases numerical experiments confirm the theoretical results.

1.3 Structure of the thesis

The thesis is divided in two parts. Part I contains introduction, summary of the results, general conclusions and possible future research for each one of the three topics. Part II is divided in five chapters where each one, of these chapters contains a published or submitted for publication work (with minor notational modifications).

In Chapter 2 we summarize Chapters 5 and 6 on the coupling between fluid flow and porous media flow.

In Chapter 3 we present some preliminaries on Discontinuous Galerkin discretization of elliptic problems with discontinuous coefficients, and we summarize Chapters 7 and 8 on domain decomposition preconditioners for this discretization.

In Chapter 4 we present some preliminaries on white noise theory and summarize Chapter 9 on finite element analysis for the (ordinary product) stochastic pressure equations.

1.4 List of papers and contribution

- Chapter 5: Galvis and Sarkis [2007b].
- Chapter 6: Galvis and Sarkis [2007a].
- Chapter 7: Dryja, Galvis and Sarkis [2007a].
- Chapter 8: Dryja, Galvis and Sarkis [2007b].
- Chapter 9: Galvis and Sarkis [2008a].

In the papers Galvis and Sarkis [2007a,b, 2008a], I develop both the theoretical and computational experiments in collaboration with Professor Marcus Sarkis. The work Galvis and Sarkis [2007b] is based on the preliminary work Galvis [2004] while Galvis and Sarkis [2007a] extends and improves the preliminary work Galvis and Sarkis [2006].

In the paper Dryja, Galvis and Sarkis [2007b], the theoretical and computational tools have been achieved in collaboration with Professors Maksymilian Dryja and Marcus Sarkis. The papers Dryja, Galvis and Sarkis [2007a,b] are based on the preliminary work Dryja and Sarkis [2006]. My main contribution to Dryja, Galvis and Sarkis [2007a] consists of developing the numerical experiments.

Bibliography

Bedrikovetsky, P. (1993). *Mathematical Theory of Oil and Gas Recovery*. Kluwer Academic, London.

Dryja, M., Galvis, J., and Sarkis, M. (2007a). Balancing domain decomposition methods for discontinuous Galerkin discretization. In at al., U. L., editor, *Domain Decomposition Methods in Science and Engineering XVII*, Lecture Notes in Computational Science and Engineering, pages 271–278. Springer-Verlag.

Dryja, M., Galvis, J., and Sarkis, M. (2007b). BDDC methods for discontinuous Galerkin discretization of elliptic problems. *J. Complexity*, 23:715–739.

Dryja, M. and Sarkis, M. (2006). A Neumann-Neumann method for DG discretization of elliptic problems. Technical Report Serie A 456, Intituto de Mathemática Pura e Aplicada. http://www.preprint.impa.br/Shadows/SERIE_A/2006/456.html.

Ewing, R. E. (1983). *The mathematics of reservoir simulation*, volume 1 of *Frontiers in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

Galvis, J. (2004). Finite elements for well-reservoir coupling. Master's thesis, Instituto Nacional de Matemática Pura e Aplicada. TR-B011 / 2005, http://www.preprint.impa.br/cgi-bin/MMMsearch.cgi.

Galvis, J. and Sarkis, M. (2006). Balancing domain decomposition methods for mortar coupling Stokes-Darcy systems. In Keyes, D. and Widlund, O. B., editors, *Domain Decomposition Methods in Science and Engineering XVI*, volume 55 of *Lecture Notes in Computational Science and Engineering*, pages 373–380. Springer.

Galvis, J. and Sarkis, M. (2007a). BDD and FETI methods for mortar coupling of Stokes-Darcy systems. Submitted.

Galvis, J. and Sarkis, M. (2007b). Non-matching mortar discretization analysis for the coupling Stokes-Darcy equations. *Electron. Trans. Numer. Anal.*, 26:350–384.

Galvis, J. and Sarkis, M. (2008a). A priori error estimates for wiener-chaos finite element approximations for the darcy's equation in random porous media. Submitted.

Chapter 2

Darcy-Stokes Coupling

In this chapter we introduce the basic ideas on the coupling of free fluid flow with porous media flow. We present some preliminaries on Darcy-Stokes coupling and describe the results of Chapters 5 and 6.

2.1 The continuous model

An incompressible fluid in a region $D^f \subset \mathbb{R}^n$ can flow both ways across an interface Γ into a saturated porous medium domain $D^p \subset \mathbb{R}^n$. The model consists of *Stokes' equations* in the fluid region and *Darcy's law* for the filtration velocity in the porous medium region; see Figure 2.1. Define the interface $\Gamma := \overline{D^f} \cap \overline{D^p}$ and the domain $D = \text{int}(\overline{D^f} \cup \overline{D^p})$. The equations are:

$$\text{Stokes' equations} \left\{ \begin{array}{rcll} -\nabla \cdot T(\boldsymbol{u}^f, p^f) & = & \boldsymbol{f}^f & \text{in } D^f \\ \nabla \cdot \boldsymbol{u}^f & = & g^f & \text{in } D^f \\ \boldsymbol{u}^f & = & \boldsymbol{h}^f & \text{on } \Gamma^f := \partial D^f \setminus \Gamma \end{array} \right. \tag{2.1}$$

and

$$\text{Darcy's equations} \left\{ \begin{array}{rcll} \boldsymbol{u}^p & = & -\frac{\kappa}{\nu} \nabla p^p & \text{in } D^p \\ \nabla \cdot \boldsymbol{u}^p & = & g^p & \text{in } D^p \\ \boldsymbol{u}^p \cdot \boldsymbol{\eta}^p & = & h^p & \text{on } \Gamma^p := \partial D^p \setminus \Gamma. \end{array} \right. \tag{2.2}$$

Here $T(\boldsymbol{v}, p) := -pI + 2\nu \boldsymbol{D}\boldsymbol{v}$, where ν is the fluid viscosity, $\boldsymbol{D}\boldsymbol{v} := \frac{1}{2}(\nabla \boldsymbol{v} + \nabla \boldsymbol{v}^T)$ is the linearized strain tensor and κ denotes the rock permeability. For simplicity on the analysis, we assume that κ is a real positive constant.

We impose the following conditions:

1. **Interface matching conditions across** Γ:
   (a) *Conservation of mass across* Γ*:* $\boldsymbol{u}^f \cdot \boldsymbol{\eta}^f + \boldsymbol{u}^p \cdot \boldsymbol{\eta}^p = 0$ on Γ where $\boldsymbol{\eta}^i$ is the unit outward normal to D^i, $i = f, p$.
   (b) *Balance of normal forces across* Γ*:* $p^f - 2\nu \boldsymbol{\eta}^{fT} \boldsymbol{D}(\boldsymbol{u}^f)\boldsymbol{\eta}^f = p^p$ on Γ.
   (c) *Beavers-Joseph-Saffman condition:* This condition gives an expression for the component of the Cauchy stress tensor in the tangential direction

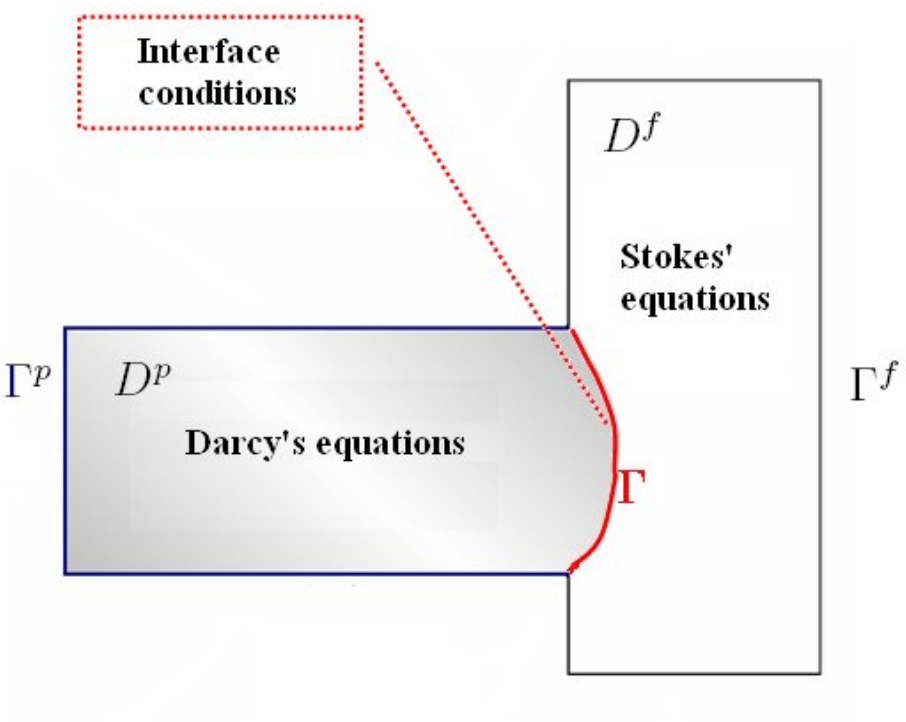

Figure 2.1: Darcy-Stokes coupling configuration. In the porous subdomain D^p we use Darcy's equations and in the fluid subodmain D^f we use Stokes' equations. We impose interface matching conditions on the interface Γ, normal seepage velocity on Γ^p and fluid velocity on Γ^f.

of the interface Γ; see Beavers and Joseph [1967], Jäger and Mikelić [2000] and references therein. It is expressed by:

$$\boldsymbol{u}^f \cdot \boldsymbol{\tau}_l = -\frac{\sqrt{\kappa}}{\alpha^f} 2\boldsymbol{\eta}^{fT} \boldsymbol{D}(\boldsymbol{u}^f)\boldsymbol{\tau}_l \quad l = 1, n-1; \text{ on } \Gamma,$$

where $\boldsymbol{\tau}_l$, $\ell = 1, n-1$, are the unit tangential vectors on Γ. The parameter α^f must be experimentally determined and depends on the porous geometry close to Γ among other things.

2. **Compatibility condition:** The divergence and boundary data satisfy,

$$\langle g^f, 1\rangle_{D^f} + \langle g^p, 1\rangle_{D^p} - \langle \boldsymbol{h}^f \cdot \boldsymbol{\eta}^f, 1\rangle_{\Gamma^f} - \langle h^p, 1\rangle_{\Gamma^p} = 0.$$

In Chapter 5 we consider this model, i.e., equations (2.1) and (2.2) together with conditions 1.(a),(b),(c) and 2. above. In Section 5.4 we derive a weak formulation for the model presented above and verify the inf-sup (also called LBB) condition. The complete inf-sup analysis presented in Section 5.4.2 implies existence, uniqueness and continuous dependence on the data of the weak solution of this coupled system of equations; see Girault and Raviart [1986] and Brezzi and Fortin [1991]. The inf-sup analysis of the continuous model uses tools developed in Section 5.2 and in Layton et al. [2002].

2.2 The discrete model

In Section 5.5.1 we choose particular finite element spaces for each subdomain and couple them using discrete Lagrange multipliers. We allow the two subdomain discretizations to be nonmatching across the interface. Nonmatching meshes are important for applications due to the different nature of the equations in each subdomain, therefore, it is desirable to allow the possibility of handling different triangulations and independent subdomain mesh refinements.

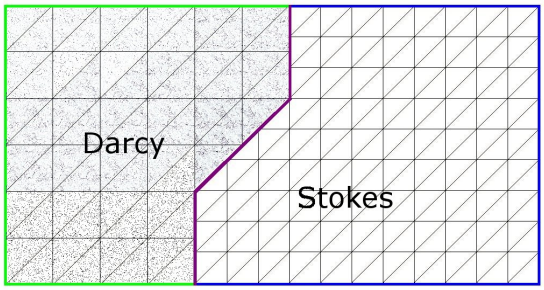

Figure 2.2: Nonmatching triangulations for Darcy-Stokes coupling.

On the fluid side we use $P_2 \backslash P_1$ triangular Taylor-Hood finite elements. For the porous region we use the lowest order Raviart-Thomas finite elements. The Taylor-Hood and Raviart-Thomas finite elements are coupled using a discrete Lagrange multiplier. We use piecewise constant functions on the interface as discrete Lagrange multipliers. In Figure 2.3 we show the degree of freedom configuration of the discrete coupling.

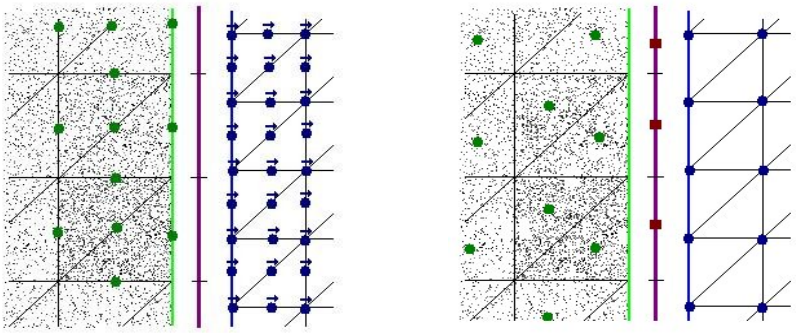

Figure 2.3: On the left picture: degrees of freedom for Raviart-Thomas velocity (left subdomain •) and Taylor-Hood velocity (right subdomain $\vec{\bullet}$). On the right picture: degrees of freedom for for Raviart-Thomas pressure (left subdomain •) and Taylor-Hood pressure (right subdomain •). P_0 elements on the interface for Lagrange multipliers (■).

In Section 5.5.2 we develop the corresponding (inf-sup) analysis of the chosen finite element approximation. This analysis implies the well-posedness of the discrete model.

2.2.1 A priori error estimates

Section 5.6 is devoted to a priori error estimates. Let $\|\cdot\|_a$ denote the sum of properly scaled $H^1(D_f)$ and $L^2(D_p)$ norms and $h := \max\{h^f, h^p\}$ be the maximum of the mesh sizes in each subdomain. In Proposition 5.27 we derive an a priori error estimate of the form

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_a \leq Ch\left(\sqrt{\nu}|\boldsymbol{u}^f|_{H^2(D^f)^2} + \sqrt{\frac{\nu}{\kappa}}|\boldsymbol{u}^p|_{H^1(D^p)^2}\right) + Ch^p\frac{1}{\sqrt{\nu}}|p^f|_{H^1(D^f)}. \quad (2.3)$$

Here, the velocity $\boldsymbol{u} = (\boldsymbol{u}^f, \boldsymbol{u}^p)$ is the exact solution of the continuous saddle point problem and $\boldsymbol{u}_h = (\boldsymbol{u}_h^f, \boldsymbol{u}_h^p)$ is its finite element approximation. The estimate (2.3) is obtained in such a way that the constant C is independent of the parameters ν, κ and the mesh ratio h^f/h^p. The estimate (2.3) and the inf-sup stability imply a priori error estimates for the pressure and the Lagrange multipliers; see Proposition 5.29. An improvement of the error estimates (2.3) is derived for the case where, on the interface, the mesh of the fluid side is a refinement of the porous side mesh. The improvement is in the sense that, in the presence of this refinement condition, the third term in the right hand side above does not appear. Section 5.7 includes numerical experiments to confirm the theory.

2.3 Preconditioning for Darcy-Stokes coupling

With the discretization chosen in Section 2.2 we obtain a sparse symmetric indefinite saddle point linear system $Au = f$ where the matrix A is of the form

$$A = \left[\begin{array}{c|c|c} K^f & 0 & M^{fT} \\ \hline 0 & K^p & -M^{pT} \\ \hline M^f & M^p & 0 \end{array}\right] = \left[\begin{array}{cc|cc|c} A^f & B^{fT} & 0 & 0 & C^{fT} \\ B^f & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & A^p & B^{pT} & -C^{pT} \\ 0 & 0 & B^p & 0 & 0 \\ \hline C^f & 0 & -C^p & 0 & 0 \end{array}\right]. \quad (2.4)$$

Here, matrix A^f corresponds to ν times the discrete version of the linearized stress tensor on D^f. The matrix A^p corresponds to ν/κ times a discrete L^2-norm on D^p. Matrix $-B^i$ is the discrete divergence in D^i, $i = f, p$, and matrices C^f and C^p correspond to the matrix form of the discrete conservation of mass on the interface Γ. In the block matrix (2.4), matrix K^f corresponds to the matrix form of the Stokes' equation, K^p corresponds to the Darcy's equation and the remaining nonzero blocks impose the discrete conservation of mass on the interface.

When dealing with discrete Stokes or Darcy problems, it is mandatory the use of preconditioners since these large sparse matrices are indefinite and ill conditioned. Note that in the linear system associated to (2.4) there are three saddle point (or three restrictions imposed) at once. Therefore, we can expect also the necessity of a preconditioner for the Darcy-Stokes system. It is worth mentioning that the condition of the matrix (2.4) depends heavily on the permeability parameter κ. The smaller the value of κ the worse is the condition number and this relation is linear. We recall that κ represents the permeability of the rock and, in

practice, its typical value is very small.

There are several possible choices for solving and preconditioning Stokes and Darcy problems. Recall that both of them, as well as the coupled system, are particular cases of saddle point problems. In the literature there are many choices for computing the solution and constructing preconditioners of saddle point problems; see Toselli and Widlund [2005] and references therein. The matrix (2.4) does not have a classical saddle point structure but a saddle point structure with two different restrictions. The design and analysis of preconditioners for (2.4) is more complex than in the classical saddle point problem.

In Chapter 6 we design and analyze two domain decomposition preconditioners based on Schur complement formulations associated with matrix (2.4). For both preconditioners, we derive condition number estimates of order $C_1(1 + \frac{1}{\kappa})$; see Theorems 6.1 and 6.3. The parameter κ represents de permeability of the porous medium. In case where the fluid discretization is finer than the porous side discretization, we derive a better condition number estimate of order $C_2(\frac{\kappa+1}{\kappa+(h^p)^2})$; see Theorem 6.6. Here h^p is the mesh size of the porous side triangulation. The constants C_1 and C_2 are independent of the permeability κ, the fluid viscosity ν, and the mesh ratio across the interface. For references on Domain decomposition based iterative methods and preconditioners for elliptic, Stokes, Darcy and many others (systems of) partial differential equations are considered extensively; see Toselli and Widlund [2005], Mathew [2008], Quarteroni and Valli [1999], Smith et al. [1996] and references there in. Numerical experiments confirm the sharpness of the theoretical estimates for the condition of the preconditioned operators; see Section 6.7.

2.4 Final comments

We make some comments on possible extensions of the analysis presented in Chapters 5 and 6.

- The Darcy-Stokes coupling provides a linear stationary model for the simulation of a free fluid that can filtrate through a porous media. It is possible to consider more complicated models such us Darcy-Navier Stokes coupling. We recall that the Stokes system is a basic part of any approximation of the Navier-Stokes system. More complicated models for the porous media flow can also be studied: the Brinkman equations (Darcy's terms plus a viscous term that accounts for flow through medium where the grains of the media are porous themselves) and Forchheimer equations (nonlinear version of Darcy's equations for high velocities).
- The matching conditions on the interface and the conservation of mass compatibility conditions play a major role in the theoretical and numerical analysis of any heterogeneous domain decomposition model. In particular when more general models are considered, an appropriate condition on the interface must be derived. In general, this is a very difficult and fundamental part of the modeling process.

- For the discrete part, we note that the finite element analysis considered here can be extended to a wide range choices of finite elements.
- The preconditioners designed and analyzed in Chapter 6 can be extended to the three dimensional case and to other discretizations, e.g., the $P2/P0$ coupled with Raviart-Thomas; see Discacciati et al. [2002], Layton et al. [2002] and Rivière and Yotov [2005].
- The domain decomposition preconditioner of Chapter 6 are based in the two subdomain decomposition of $D = D^f \cup D^p$. A possible future research is the design, analysis and implementation of (more sophisticated) domain decomposition methods based on a general decomposition of the domain D.

Bibliography

Beavers, G. S. and Joseph, D. D. (1967). Boundary conditions at a naturally permeable wall. *J. Fluid Mech.*, 30:197–207.

Brezzi, F. and Fortin, M. (1991). *Mixed and hybrid finite element methods*, volume 15 of *Springer Series in Computational Mathematics*. Springer-Verlag, New York.

Discacciati, M., Miglio, E., and Quarteroni, A. (2002). Mathematical and numerical models for coupling surface and groundwater flows. *Appl. Numer. Math.*, 43(1-2):57–74. 19th Dundee Biennial Conference on Numerical Analysis (2001).

Girault, V. and Raviart, P.-A. (1986). *Finite element methods for Navier-Stokes equations*, volume 5 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin. Theory and algorithms.

Jäger, W. and Mikelić, A. (2000). On the interface boundary condition of Beavers, Joseph, and Saffman. *SIAM J. Appl. Math.*, 60(4):1111–1127.

Layton, W. J., Schieweck, F., and Yotov, I. (2002). Coupling fluid flow with porous media flow. *SIAM J. Numer. Anal.*, 40(6):2195–2218 (2003).

Mathew, T. (2008). *Domain Decomposition Methods for the Numerical Solution of Partial Differential Equations*. Lecture Notes in Computational Science and Engineering. Springer.

Quarteroni, A. and Valli, A. (1999). *Domain decomposition methods for partial differential equations*. Numerical Mathematics and Scientific Computation. The Clarendon Press Oxford University Press, New York. , Oxford Science Publications.

Rivière, B. and Yotov, I. (2005). Locally conservative coupling of Stokes and Darcy flows. *SIAM J. Numer. Anal.*, 42(5):1959–1977.

Smith, B. F., Bjørstad, P. E., and Gropp, W. D. (1996). *Domain decomposition*. Cambridge University Press, Cambridge. Parallel multilevel methods for elliptic partial differential equations.

Toselli, A. and Widlund, O. (2005). *Domain decomposition methods—algorithms and theory*, volume 34 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin.

Chapter 3

Discontinuous Galerkin Discretization of Elliptic Problems

Discontinuous Galerkin (DG) methods are becoming more and more popular for the numerical approximation of partial differential equations since they are well suited for dealing with complex geometries, discontinuous coefficients and local or patch refinements. In this chapter we describe the relevant ideas of Chapters 7 and 8.

3.1 The continuous model

Let $\{D_i\}_{i=1}^N$ be a polygonal partition of $D \subset \mathbb{R}^2$. Assume that the subdomain D_i is of diameter $O(H_i)$. We consider the equation

$$\begin{cases} -\nabla \cdot (\rho \nabla u) &= f, \text{ in } D \\ u &= 0, \text{ on } \partial D, \end{cases} \tag{3.1}$$

where $\rho(x) = \rho_i$, $x \in D_i$, $i = 1, \ldots, N$, i.e., ρ is piecewise constant in $\{D_i\}_{i=1}^N$.

3.2 DG discretization

In this section we describe the discrete model and present its main futures and advantages. We also use a numerical example for motivating the necessity of preconditioning. For the general theory of DG discretizations we refer to Arnold [1982], Arnold et al. [2002, 2000] and also to Dutra do Carmo and Vinicius [2000], Dutra do Carmo et al. [2000a,b] and references therein.

3.2.1 Discrete problem

Let $\mathcal{T}_i$ be a triangulation of D_i, $i = 1, \ldots, N$. The resulting triangulation on D is in general nonmatching across ∂D_i. Let $X_i(D_i)$ be the regular finite element space of piecewise linear continuous functions in D_i. Denote by $X_h(D)$ the global space associated to all nodal values in the disjoint union of the subdomains $\{D_i\}_{i=1}^N$; see

Figure 3.1.

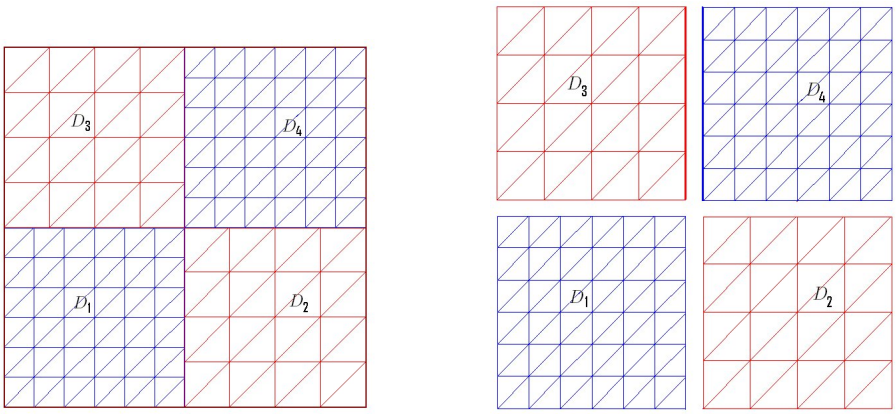

Figure 3.1: On the left picture we display a shape regular triangulation in each subdomain. On the right picture we show finite element meshes in each subdomain.

The discrete problems obtained from Discontinuous Galerkin methods involve the solution of large sparse linear systems where the degrees of freedom are associated to all the nodal values of $X_h(D)$; see Sections 7.2 and 8.2. These discretizations allow jumps of the solution across the interfaces ∂D_i, $i = 1, \ldots, N$. To control the jumps we penalize the $L^2(\partial D_i)$ jump of the solution. In Dryja [2003] it is established that the linear system resulting from the DG discretization (of Chapters 7 and 8) is positive definite when the penalty parameter is large enough (order one). Dryja also obtains a priori error estimates. The condition number of the resulting large sparse linear system is $O\left(\frac{\rho_{\max}}{\rho_{\min}} \frac{1}{h^2}\right)$ where $h = \max_i h_i$ and h_i is the discretization parameter of D_i, $i = 1, \ldots, N$.

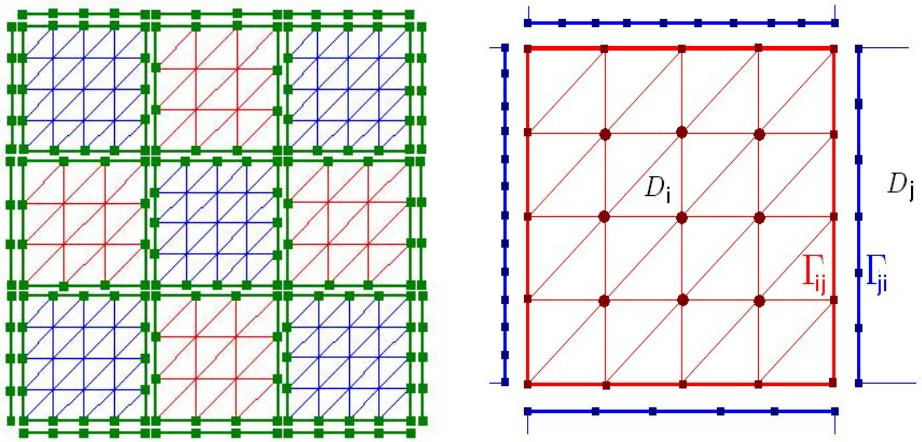

Figure 3.2: On the left: Degrees of freedom associated to Γ, the disjoint union of ∂D_i, $i = 1, 2, \ldots, N$, (■). In this case the square subdomain D is the union of 3×3 square subdomains D_i, $i = 1, 2, \ldots, 9$. On the right: Local degrees of freedom associated to D_i. They are classified in interior (●) and boundary (■) degrees of freedom.

The paper Dryja and Sarkis [2006] shows how to derive a Schur system, that

is, how to obtain a reduced system involving only *boundary* degrees of freedom for each subdomain D_i. We note that this is not a standard Schur complement reduction as in classical substructuring methods; see Chapter 4 in Toselli and Widlund [2005]. The reduced formulation involves only degrees of freedom associated to Γ, that is, the disjoint union of all boundary inherited discretizations; see Figure 3.2.

The Schur complement system is also symmetric and positive definite when the penalty parameter is large enough. Then, we can solve the original linear system with the CG algorithm where each iteration involves only interface degrees of freedom plus local solvers. The condition of the resulting Schur complement linear system is $O\left(\frac{\rho_{\max}}{\rho_{\min}}\frac{1}{Hh}\right)$ where $H = \max_i H_i$ with H_i being the diameter of D_i, $i = 1, 2, \ldots, N$. We mention also that this Schur complement is a subassembling of *local* Schur complements. The i-th Schur complement involves only the boundary degrees of freedom of subdomain D_i; see Figure 3.2. The problem is suitable for parallelization.

3.2.2 Example

Consider the following example. Let $D = [-1, 1] \times [-1, 1] \subset \mathbb{R}^2$, and consider the exact solution given by

$$u(x_1, x_2) = \begin{cases} u_1(x_1)u_2(x_2), & (x_1, x_2) \in D_1 := [-1, 0] \times [-1, 0] \\ u_1(k_1x_1)u_2(x_2), & (x_1, x_2) \in D_2 := [-1, 0] \times [0, 1] \\ u_1(x_1)u_2(k_2x_2), & (x_1, x_2) \in D_3 := [0, 1] \times [-1, 0] \\ u_1(k_1x_1)u_2(k_2x_2), & (x_1, x_2) \in D_4 := [0, 1] \times [0, 1], \end{cases} \tag{3.2}$$

which is solution of (3.1) with $\rho = k_1k_2 1_{D_1} + k_2 1_{D_2} + k_1 1_{D_3} + 1_{D_4}$.

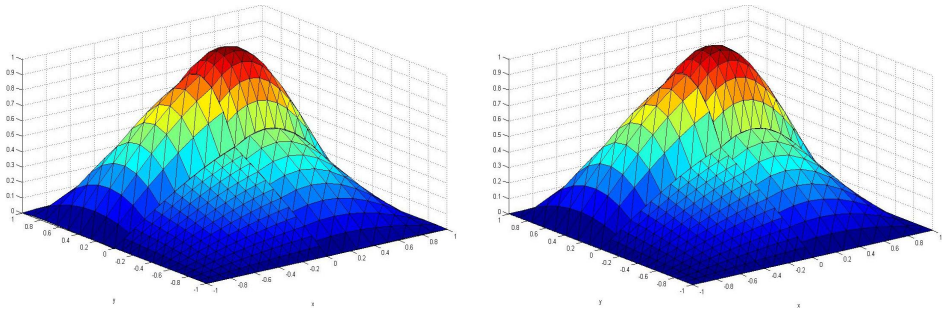

Figure 3.3: Computed (left) and exact solution (3.2) (right) for $k_1 = 2$, $k_2 = 5$, $u_1(x_1) = \sin(\frac{\pi*(x+1)}{k_1+1})$ and $u_2(x_2) = 4(x + 1)(k_2 - x)/(k_2 + 1)^2$.

We start with mesh sizes $h_1 = h_4 = \frac{1}{3}$ and $h_2 = h_3 = \frac{1}{2}$ where h_i is the mesh size of D_i, $i = 1, 2, 3, 4$. In Figure 3.3 we show the computed and exact solution for the second level of refinement. In Table 3.1 we show the $L^2(D)$, $H^1(D)$-broken and $L^2(\Gamma)$ approximation errors and the for different levels of refinements. We also show the condition number estimates for the Schur complement problem. We see a decay factor of 4 in the L^2-error, factor of 2 in the $H^1(D)$-broken norm error and factor of 3 for the interface error. We also observe a linear growth of the logarithm of the condition number and the number of iterations until convergence (i.e., until the initial residual is reduced by 10^{-6}).

$h\downarrow$	$\|u-u^h\|_{L^2(D)}$	$\sum_{i=1}^4 \|u-u^h\|_{H^1(D_i)}$	$\sum_{i,j} \|u-u^h\|_{L^2(\Gamma_{ij})}$	Cond	Iter
$\times 2^1$	0.0334560	0.478163	0.1842924	67.68	40
$\times 2^2$	0.0096112	0.241894	0.0667458	135.51	55
$\times 2^3$	0.0025424	0.120784	0.0236767	271.04	79
$\times 2^4$	0.0006518	0.060236	0.0083684	541.60	113
$\times 2^5$	0.0001649	0.030065	0.0029569	1082.29	161
$\times 2^5$	4.1469e-05	0.015017	0.0010450	2163.36	229

Table 3.1: Error of the DG approximation of the exact solution (3.2) with $\delta = 4$ (columns 1-3). Condition number and iteration count of the CG algorithm (columns 4 and 5).

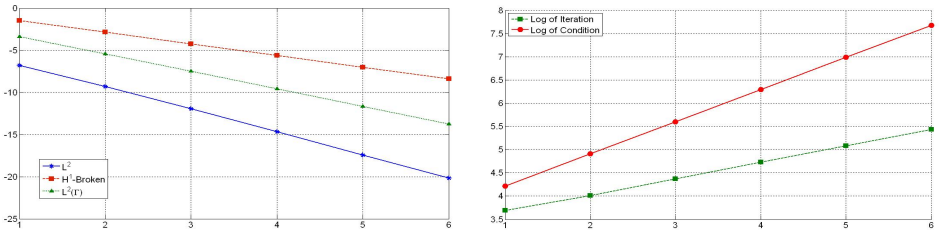

Figure 3.4: Logarithm of the $L^2(D)$, H^1-broken, and $L^2(\Gamma)$ errors (left). Logarithm of the iteration and condition number (right).

3.2.3 Main issues of DG discretization

The main advantages of the DG discretization for elliptic problem with discontinuous coefficients are:

- proper handling of complicated geometries
- proper handling of discontinuities along the interface of a decomposition of the domain
- patch refinements
- Schur complement problem

These features make the DG discretization an excellent choice for computing approximations of the solution of the elliptic problem with discontinuous coefficients.

The Schur complement problem involves only degrees of freedom associated with the interfaces between subdomain and can be solved using the CG algorithm where in each iteration we have to solve one local Dirichlet problems per subdomain. The condition number is of order $O(\frac{\rho_{\max}}{\rho_{\min}}\frac{1}{Hh})$ and therefore the number of CG iteration grows like $O(h^{-\frac{1}{2}})$; see also the fourth and fifth columns in Table 3.1 and Figure 3.4. This convergence depends also on the number of subdomains, the local meshes and the jump of the coefficient. Then, *preconditioning is mandatory* for an efficient computation of the solution of the Schur complement problem. We want the performance of the preconditioners to be independent of the number of subdomains and the local triangulations sizes, as well as the jump of the coefficients. The objective in Chapters 7 and 8 is the design and analysis of preconditioners with performance independent of the jumps of the coefficients and the local mesh sizes.

3.3 Domain decomposition preconditioners

In Dryja and Sarkis [2006], the authors design and analyze a Neumann-Neumann type method for solving the Schur complement problem associated to the DG discretization of an elliptic problem with discontinuous coefficient. For a complete analysis of these preconditioners and their hybrid versions, we refer also the document in preparation Dryja, Galvis and Sarkis [2008b] which includes also numerical experiments.
Dryja et al. [2008].

In Chapter 7 we design and analyze a Balancing Domain Decomposition (BDD) method for solving the resulting Schur complement problem. For BDD methods we refer to Mandel [1993] and Toselli and Widlund [2005].

In Chapter 8 we design and analyze Balancing Domain Decomposition with Constraints (BDDC) methods for solving the Schur complement problem. For the general theory of BDDC methods for conforming elements, we refer to Dohrmann [2003], Mandel et al. [2005] and Li and Widlund [2006].

Under the interface condition introduced in Section 7.4.1 and Assumption 8.4, and using the abstract Schwarz theory, we derive condition number estimates for the BDD and BDDC preconditioner introduced above; see Sections 7.4 and 8.7, respectively. The condition number estimate is of the form $C(1+\max_i \log \frac{H_i}{h_i})^2$ in all the cases. The constant C is independent of h_i, H_i and the jumps of the coefficients. The methods are well suited for parallel computations and can be straightforwardly extended to three dimensional problems. Results of numerical tests confirm the theoretical results and the necessity of the imposed assumptions.

3.4 Final comments

We make some comments on possible extensions of the results presented in Chapters 7 and 8.

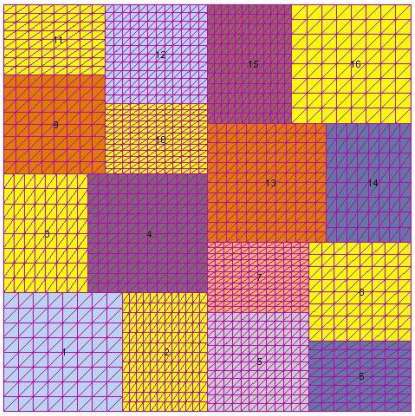

Figure 3.5: Geometrically nonconforming partition of D with a shape regular triangulation in each subdomain.

- In Chapters 7 and 8 we consider a geometrically conforming partition of the domain D. The objective of the work in preparation Dryja, Galvis and Sarkis [2008a] is the extension of the preconditioners designed and analyzed in Chapters 7 and 8, as well as the additive preconditioner of Dryja and Sarkis [2006], to the case of geometrically nonconforming partitions of the domain D; see Figure 3.5. Dryja et al. [2008a]
- A possible future research is the extension of these methods for the case of discontinuous coefficients varying inside each subdomain.

Bibliography

Arnold, D. N. (1982). An interior penalty finite element method with discontinuous elements. *SIAM J. Numer. Anal.*, 19(4):742–760.

Arnold, D. N., Brezzi, F., Cockburn, B., and Marini, D. (2000). Discontinuous Galerkin methods for elliptic problems. In *Discontinuous Galerkin methods (Newport, RI, 1999)*, volume 11 of *Lect. Notes Comput. Sci. Eng.*, pages 89–101. Springer, Berlin.

Arnold, D. N., Brezzi, F., Cockburn, B., and Martin, D. (2002). Unified analysis of discontinuous Galerkin method for elliptic problems. *SIAM J. Numer. Anal.*, 39(5):1749–1779.

Dohrmann, C. R. (2003). A preconditioner for substructuring based on constrained energy minimization. *SIAM J. Sci. Comput.*, 25(1):246–258 (electronic).

Dryja, M. (2003). On discontinuous Galerkin methods for elliptic problems with discontinuous coefficients. *Comput. Methods Appl. Math.*, 3(1):76–85.

Dryja, M., Galvis, J., and Sarkis, M. (2008a). A Neumann-Neumann method for dg discretization of elliptic problems on geometrically non-conforming substructures. In preparation.

Dryja, M., Galvis, J., and Sarkis, M. (2008b). Neumann-Neumann methods for DG discretization of elliptic problems. In preparation.

Dryja, M. and Sarkis, M. (2006). A Neumann-Neumann method for DG discretization of elliptic problems. Technical Report Serie A 456, Intituto de Mathemática Pura e Aplicada. http://www.preprint.impa.br/Shadows/SERIE_A/2006/456.html.

Dutra do Carmo, E., Celani Duarte, A., and Alves Rochinha, F. (2000a). Consistent discontinuous finite elements in elastodynamics. *Comput. Methods Appl. Mech. Engrg.*, 190(193–223).

Dutra do Carmo, E., Celani Duarte, A., and Alves Rochinha, F. (2000b). Discontinuous finite element formulations applied to cracked elastic domains. *Comput. Methods Appl. Mech. Engrg.*, 185:21–36.

Dutra do Carmo, E. and Vinicius, C. D. A. (2000). A discontinuous finite element-based domain decomposition method. *Comput. Methods Appl. Mech. Engrg.*, 190:825–843.

Li, J. and Widlund, O. (2006). FETI-DP, BDDC, and block Cholesky methods. *Internat. J. Numer. Methods Engrg.*, 66(2):250–271.

Mandel, J. (1993). Balancing domain decomposition. *Comm. Numer. Methods Engrg.*, 9(3):233–241.

Mandel, J., Dohrmann, C. R., and Tezaur, R. (2005). An algebraic theory for primal and dual substructuring methods by constraints. *Appl. Numer. Math.*, 54(2):167–193.

Toselli, A. and Widlund, O. (2005). *Domain decomposition methods—algorithms and theory*, volume 34 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin.

Chapter 4

Elliptic Partial Differential Equations with Random Coefficients

Partial differential equations with random coefficients are a great source of very interesting theoretical and numerical questions. Models involving partial differential equations with incorporated uncertainties appear in a large variety of fields such as fluid dynamics, petroleum engineering, economics, environmental sciences and many others. In this chapter we describe the main ideas of Chapter 9. The Section 4.3 provides some preliminaries from white noise theory.

4.1 The continuous and discrete model

In Chapter 9 we deal with the ordinary product random (or stochastic) pressure equation with log-normal coefficient,

$$\mu(\omega)\text{-almost sure} \begin{cases} -\nabla_x \cdot (\kappa(x,w)\nabla_x u(x,w)) &= f(x,\omega), \text{ for all } x \in D \\ u(x,\omega) &= 0, \text{ for all } x \in \partial D, \end{cases} \tag{4.1}$$

where μ is a probability measure in a suitable probability space. The coefficient κ is the exponential of Gaussian process in D. The corresponding Wick product equations, where the main term in (4.1) is replaced by $\kappa(x,w) \diamond \nabla_x u(x,w)$ with $\diamond$ denoting the Wick product, is easier to analyze; see Holden et al. [1996] and Benth and Theting [2002].

We use the white noise theory to construct and characterize adequate spaces for the solution of the ordinary product stochastic pressure equation (4.1). Motivated from Roman and Sarkis [2006] and Benth and Gjerde [1998] we consider the white noise probability space associated with the Hilbert space H and the operator A; see Section 4.3 and references therein for examples of H and A. The probability is defined in the sigma-field of Borel subsets of $\mathcal{S}'$, the dual of the nuclear countable Hilbert space $\mathcal{S}$. Here, the space $\mathcal{S}$ is the countably Hilbert space constructed from H and A. The probability measure μ is given by the Bochner-Minlos theorem and

it is characterized by

$$E_\mu e^{i\langle\cdot,\xi\rangle} := \int_{\mathcal{S}'} e^{i\langle\omega,\xi\rangle} d\mu(\omega) = e^{-\frac{1}{2}\|\xi\|_H^2}, \text{ for all } \xi \in \mathcal{S}. \tag{4.2}$$

In Section 9.3, we pose the problem (4.1) in the weak sense in the $\mathcal{U}_s^1$ spaces, $s \in \mathbb{R}$. Here $\mathcal{U}_s^1$ is the tensor product

$$\mathcal{U}_s^1 = H^1(D) \otimes (L^2)_s \text{ where we denote } (L^2)_s := L^2(\mathcal{S}', e^{s\|\omega\|_{-\theta}^2} d\mu(\omega)),$$

with the tensor product norm. The norm $\|\cdot\|_{-\theta}^2$ is defined in Section 4.3.2. Recall that $H^1(D)$ is the standard Sobolev space of order one. We provide conditions on the right-hand side f for the existence and uniqueness of solutions in the space $\mathcal{U}_s^1$. When writing the weak form of the equation we choose different spaces for the solution and test functions. We prove the corresponding inf-sup condition; see Theorem 9.6.

In Section 9.4 we prove that every $u \in \mathcal{U}_s^1$ can be expressed as a series in term of the special stochastic polynomials $H_{\sigma(s)^2,\boldsymbol{\alpha}}(\omega)$ with coefficients in $H^1(D)$,

$$u(x,\omega) = \sum_{\boldsymbol{\alpha}\in\boldsymbol{\mathcal{J}}} u_{\boldsymbol{\alpha},s}(x) H_{\sigma(s)^2,\boldsymbol{\alpha}}(\omega) \text{ with } u_{\boldsymbol{\alpha},s} \in H^1(D) \text{ for all } \boldsymbol{\alpha} \in \boldsymbol{\mathcal{J}}. \tag{4.3}$$

Here $\boldsymbol{\mathcal{J}}$ is the set of all infinite dimensional compact support multi-indices; see Remark 4.7.

A generalization (to the $\mathcal{U}_s$ space) of the approximation proposed in Benth and Theting [2002] and Roman and Sarkis [2006] is considered. This approximation was introduced originally in the space $H^1(D) \otimes (L^2)_0$. Our approximation is a truncated expansion (4.3) with the coefficients restricted to $X^h(D)$, a (spatial) finite element approximation of $H^1(D)$. The approximation of the solution of the weak version of (4.1) is of the form

$$u^{N,K,h}(x,\omega) = \sum_{\boldsymbol{\alpha}\in\boldsymbol{\mathcal{J}}^{N,K}} u^h_{\boldsymbol{\alpha},s}(x) H_{\sigma(s)^2,\boldsymbol{\alpha}}(\omega) \text{ with } u_{\boldsymbol{\alpha},s} \in X^h(D) \text{ for all } \boldsymbol{\alpha} \in \boldsymbol{\mathcal{J}}^{N,K}, \tag{4.4}$$

where the set $\boldsymbol{\mathcal{J}}^{N,K}$ includes only multi-indices $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots)$ with total degree $\alpha_1 + \alpha_2 + \cdots \leq N$ and such that $\alpha_\ell = 0$ for all $\ell > K$, and each coefficient $u^h_{\boldsymbol{\alpha},s}$ is a usual (spatial) finite element function; see Section 9.4. Note that after computing the coefficients of the approximation $u^{N,K,h}$ in (4.4), we can simulate realizations by simply simulating the stochastic polynomials $H_{\sigma(s)^2,\boldsymbol{\alpha}}$. These polynomials are product of Hermite polynomials of standard normal random variables and they are easy to simulate; see Remark 4.7.

Our discrete formulation seeks an approximation of solution in the form (4.4) that satisfies the weak form of (4.1) posed in a finite dimensional space. In the discrete formulation we also use different spaces for the solution and test functions. We prove the inf-sup stability of this approximation (Lemma 9.21) and provide a priori error estimates (Theorem 9.24) for a wide class of norms that depends on

the choice of a sequence of weights; see Section 4.3.

The chosen approximation leads to the solution of a positive definite symmetric linear system with the complexity of a coupled system of $\binom{N+K}{K}$ elliptic equations in $D \subset \mathbb{R}^d$. Here K and N are the number of variables and the maximum degree of the stochastic polynomials expansion, respectively; see Section 9.6.

We generalize and improve the results of Cao [2006] and Benth and Gjerde [1998] on approximation of a (generalized) process by its truncated Chaos expansion; see Corollary 9.15, Example 9.14 and Remark 9.16.

The general set up in Chapter 9 allows us several choices of the Hilbert space H and the operator A. This implies different modeling choices and different a priori error estimates. Three possible convenient choices of H and A are discussed in Section 9.7. In Section 9.8 we choose a particular case and present numerical experiments in order to confirm the theory.

4.2 Finals comments

We make some comments on possible future work concerning Chapter 9.

- The objective in Galvis and Sarkis [2008] is the stochastic regularity results for the ordinary product stochastic pressure equation using the weighted norms introduced in Chapter 9 to deduce the a priori error estimates.
- Other immediate research matter is the domain decomposition analysis of the resulting linear system. We recall that the resulting matrix has a block structure where each block is the discretization of an elliptic operator. This matrix is not block-sparse as in the case of Chapter 3.
- A possible future research is to study the extension of our results to more general coefficients rather than log-normal. It is even possible to consider other measures instead of the Gaussian measure, e.g., the Poisson measure.
- Other very interesting research concern is the approximation of the corresponding parabolic problem. Additional temporal noise can be added to the model.

Bibliography

Becnel, J. J. (2006). Equivalence of topologies and Borel fields for countably-Hilbert spaces. *Proc. Amer. Math. Soc.*, 134(2):581–590 (electronic).

Benth, F. E. and Gjerde, J. (1998). Convergence rates for finite element approximations of stochastic partial differential equations. *Stochastics Stochastics Rep.*, 63(3-4):313–326.

Benth, F. E. and Theting, T. G. (2002). Some regularity results for the stochastic pressure equation of Wick-type. *Stochastic Anal. Appl.*, 20(6):1191–1223.

Berezanskiĭ, Y. M. (1986). *Selfadjoint operators in spaces of functions of infinitely many variables*, volume 63 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI. Translated from the Russian by H. H. McFaden, Translation edited by Ben Silver.

Bogachev, V. I. (1998). *Gaussian measures*, volume 62 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI.

Cao, Y. (2006). On convergence rate of Wiener-Ito expansion for generalized random variables. *Stochastics*, 78(3):179–187.

Cochran, W. G., Kuo, H.-H., and Sengupta, A. (1998). A new class of white noise generalized functions. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.*, 1(1):43–67.

Da Prato, G. (2006). *An introduction to infinite-dimensional analysis*. Universitext. Springer-Verlag, Berlin. Revised and extended from the 2001 original by Da Prato.

Da Prato, G. and Zabczyk, J. (1992). *Stochastic equations in infinite dimensions*, volume 44 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge.

Galvis, J. and Sarkis, M. (2008b). Regularity results for the ordinary product stochastic pressure equation. In preparation.

Hida, T. (1980). *Brownian motion*, volume 11 of *Applications of Mathematics*. Springer-Verlag, New York. Translated from the Japanese by the author and T. P. Speed.

Hida, T., Kuo, H.-H., Potthoff, J., and Streit, L. (1993). *White noise*, volume 253 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht. An infinite-dimensional calculus.

Holden, H., Øksendal, B., Ubøe, J., and Zhang, T. (1996). *Stochastic partial differential equations*. Probability and its Applications. Birkhäuser Boston Inc., Boston, MA. A modeling, white noise functional approach.

Kuo, H.-H. (1996). *White noise distribution theory*. Probability and Stochastics Series. CRC Press, Boca Raton, FL.

Kuo, H.-H. (2002). A quarter century of white noise theory. In *Quantum information, IV (Nagoya, 2001)*, pages 1–37. World Sci. Publ., River Edge, NJ.

Obata, N. (1994). *White noise calculus and Fock space*, volume 1577 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin.

Roman, L. J. and Sarkis, M. (2006). Stochastic Galerkin method for elliptic SPDEs: a white noise approach. *Discrete Contin. Dyn. Syst. Ser. B*, 6(4):941–955 (electronic).

Shigekawa, I. (2004). *Stochastic analysis*, volume 224 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI. Translated from the 1998 Japanese original by the author, Iwanami Series in Modern Mathematics.

4.3 Appendix A: Preliminaries from white noise analysis

We present a very short compendium of results from the Withe Noise theory and infinite dimensional analysis. We refer to Kuo [1996, 2002], Hida [1980], Hida et al. [1993], Holden et al. [1996], Obata [1994], Da Prato [2006], Shigekawa [2004], Berezanskiĭ [1986], Bogachev [1998] and references there in.

4.3.1 The Bochner-Minlos theorem

The Bochner-Minlos theorem is an important tool for the introduction of probability measures in infinite dimensional spaces. We refer to Hida [1980], Holden et al. [1996], Hida et al. [1993] and also to Kuo [1996], Shigekawa [2004] and Berezanskiĭ [1986] for similar results. For the special case of measures we refer to Da Prato [2006], Bogachev [1998] and references therein. The presentation of this section follows Da Prato and Zabczyk [1992], Hida [1980] and Hida et al. [1993].

Let X, Y be (separable) Hilbert spaces. $T \in L(X,Y)$ is said to be *nuclear operator* (denoted by $T \in \mathcal{L}_1(X,Y)$) if there exit two sequences $\{b_k\} \subset Y$ and $\{a_k\} \subset X$ such that

$$Tx = \sum_1^\infty (x, a_k)_X b_k \text{ and } \sum_1^\infty \|a_k\|_X \|b_k\|_Y < \infty.$$

We write $T = \sum_1^\infty b_k \oplus a_k$ and define

$$\|T\|_{\mathcal{L}_1} = \inf_{T=\sum_1^\infty b_k \oplus a_k} \sum_1^\infty \|a_k\|_X \|b_k\|_Y.$$

A vector space $\mathcal{S}$ is said to be a *countably Hilbert* space if it is topologized by a countably many compatible Hilbert norms $\|\cdot\|_n$, with respect to which $\mathcal{S}$ is complete. The Hilbert norms $\{\|\cdot\|_n\}$ being compatible means that if we have $\lim_k \|h_k\|_m = 0$ and $\|h_k\|$ is Cauchy in $\|\cdot\|_n$, then $\lim_k \|h_k\|_n = 0$.

Let $\mathcal{S}_n$ be the completion of H with respect to the norm $\|\cdot\|_n$. Then by definition, $\mathcal{S} = \cap_n \mathcal{S}_n$ with the projective limit topology. Suppose $\{\|\cdot\|_n\}$ are arranged in the increasing order, then we have the inclusions $H \supset S_1 \supset \ldots \supset S_n \supset \ldots$.

Take $\|\cdot\|_0 = \|\cdot\|_H$ as base norm to construct the dual spaces $\mathcal{S}'_n$, $n = 1, \ldots$. We have $H = \mathcal{S}_0 = \mathcal{S}'_0 \subset \mathcal{S}'_1 \subset \ldots \subset \mathcal{S}'_n \subset \ldots$.

Let $\mathcal{S}$ be a countably Hilbert. If for any m there exists $n > m$ such that the injection mapping $I_m^n : \mathcal{S}_n \to \mathcal{S}_m$ is nuclear, then $\mathcal{S}$ is called a *countably Hilbert nuclear space* or simply a *nuclear space.*

Remark 4.1 *Sometimes in the definition of countably nuclear space, instead of nuclear inclusion, it is required the existence of a Hilbert-Schmidt inclusion map.*

Remark 4.2 *When we identify $H \equiv H^*$ the triple $\mathcal{S} \subset H \subset \mathcal{S}^*$ is called Gelfand triple. If H is a space of functions, the space $\mathcal{S}$ is called the set of regular or test functions and $\mathcal{S}'$ the set of generalized functions or distributions.*

Consider the triple $\mathcal{S} \subset H \subset \mathcal{S}'$, where H is the base Hilbert space and $\mathcal{S}$ is the nuclear space. The σ-field $\mathcal{B}(\mathcal{S}')$ of $\mathcal{S}'$ is the weak* Borel σ-field; see Becnel [2006] and Obata [1994].

Lemma 4.3 (Bochner-Minlos theorem) *Let $C : \mathcal{S} \to \mathbb{R}$ be a characteristic functional, that is, the functional C is:*

1. *continuous on $\mathcal{S}$,*
2. *positive definite,*
3. *$C(0) = 1$.*

Then, there exist a unique probability measure μ on $(\mathcal{S}', \mathcal{B}(\mathcal{S}'))$ such that

$$C(\xi) = \int_{\mathcal{S}'} e^{i\langle w,\xi\rangle} d\mu(\omega), \text{ for all } \xi \in \mathcal{S}.$$

Moreover, if C is continuous with respect to $\|\cdot\|_p$ and if $n(> p)$ is such that the injection $I_p^n : \mathcal{S}_n \mapsto \mathcal{S}_p$ is of Hilbert-Schmidt type, then $\mu_C(\mathcal{S}_{-n}) = 1$.

4.3.2 Construction of nuclear spaces from a Hilbert space and an operator

We now present a construction of nuclear spaces based on a Hilbert space and an operator; see Obata [1994] and Kuo [1996].

Let $A : D(A) \subset H \to H$ be a densely defined operator such that there exists an orthonormal basis $\{\eta_j\}$ for H satisfying:

1. $A\eta_j = \lambda_j \eta_j$, $j = 1, 2, \ldots$.
2. $1 < \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_j \leq \ldots$.
3. $\sum_{j=1}^{\infty} \lambda_j^{-2\theta} < \infty$ for some $\theta > 0$.

We have that

1. From Item 2. the operator A^{-1} is bounded with norm given by $\|A^{-1}\| = \lambda_1^{-1} < 1$.
2. From Item 3. we have that $A^{-\theta}$ is a Hilbert-Schmidt operator ($A^{-\theta} \in \mathcal{L}_2$) with Hilbert-Schmidt norm $\|A^{-\theta}\|_{\mathcal{L}_2}^2 = \sum_{j=1}^{\infty} \lambda_j^{-2\theta}$.
3. From Item 3. we have that $A^{-2\theta}$ is nuclear ($A^{-2\theta} \in \mathcal{L}_1$) and and $\|A^{-2\theta}\|_{\mathcal{L}_1} = \sum_{j=1}^{\infty} \lambda_j^{-2\theta}$.

For $p > 0$ define $\mathcal{S}_p := \{\xi \in H; \ \ \|\xi\|_p < \infty\}$ where

$$\|\xi\|_p^2 := \|A^p \xi\|_H^2 = \sum_{j=0}^{\infty} \lambda_j^{2p} (\xi, \eta_j)_H^2.$$

For $p < 0$ define $\mathcal{S}_p$ as the dual space of $\mathcal{S}_{-p}$. It is easy to see that $\|\cdot\|_{-p} = \|A^{-p}\cdot\|_H$ and that the duality pairing between $\mathcal{S}_p$ and $\mathcal{S}_{-p}$ is an extension of the H inner product. We also define

$$\mathcal{S} = \cap_{p\geq 0}\mathcal{S}_p \text{ (with the projective limit topology)}$$

and $\mathcal{S}'$ as the dual space of $\mathcal{S}$. We say that $\mathcal{S}$ is the nuclear Spaces associated with H and A and we have

$$\mathcal{S} \subset \ldots \subset \mathcal{S}_{p+1} \subset \mathcal{S}_p \subset \ldots \subset H \subset \ldots \subset \mathcal{S}_{-p} \subset \mathcal{S}_{-(p+1)} \subset \ldots \subset \mathcal{S}'.$$

This countable Hilbert space is the standard countable Hilbert space constructed from (H, A). For a wider discussion on the topology (or topologies) that can be defined on nuclear spaces we refer to Becnel [2006] and Obata [1994].

We have that $p \geq q$ implies $\mathcal{S}_p \subset \mathcal{S}_q$. The inclusion $\mathcal{S}_{p+\theta} \subset \mathcal{S}_p$ is Hilbert-Schmidt.

Example 4.4 *Consider the densely defined differential operator*

$$A_1 = -\frac{d^2}{dx^2} + x^2 + 1. \tag{4.5}$$

We have an $L^2(\mathbb{R})$ orthonormal system of eigenfunctions of A_1 which are the Hermite functions $e_n(x) := \frac{1}{\sqrt{\sqrt{\pi}(n-1)!}} e^{-\frac{1}{2}x^2} h_{n-1}(\sqrt{2}x)$, $\quad n = 1, 2, \ldots$, where h_n is the n degree Hermite polynomial. We have $A_1 e_n = (2n)e_n$, $n = 1, 2, \ldots$. Then we can construct a nuclear space from $H = L^2(\mathbb{R})$ and the operator (4.5). It is easy to see that the resulting nuclear space is $\mathcal{S}(\mathbb{R})$ the Schwartz space of C^∞ rapidly decreasing functions. This construction can be extended to $\mathbb{R}^d$. Define

$$\eta_j := e_{\boldsymbol{n}^{(j)}} = e_{n_1^{(j)}} \otimes \ldots \otimes e_{n_d^{(j)}},\ j = 1, 2, \ldots. \tag{4.6}$$

We have $A_1^{\otimes d}\eta_j = \lambda_j \eta_j$ where $\lambda_j := \prod_{k=1}^d (2n_k^{(j)})$, $j = 1, 2, \ldots$. Note that $\boldsymbol{n}^{(1)} = (1, \ldots, 1) \in \mathbb{R}^d$, $\lambda_1 = 2^d$ and that $1 < \lambda_1 \leq \lambda_2 \leq \ldots$.

4.3.3 The space (L^2) and the Chaos expansions in terms of Fourier-Hermite polynomials

Denote $(L^2) := L^2(\mathcal{S}', \mu)$. We need to consider multi-index of arbitrary length. To simplify the notation, we regard multi-indices as elements of the space $(\mathbb{N}_0^{\mathbb{N}})_c$ of all sequences $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots)$ with elements $\alpha_j \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$ and with compact support, i.e., with only finitely many $\alpha_j \neq 0$. We write $\boldsymbol{\mathcal{J}} = (\mathbb{N}_0^{\mathbb{N}})_c$. Given $\boldsymbol{\alpha} \in \boldsymbol{\mathcal{J}}$, define the order and length of $\boldsymbol{\alpha}$, denoted by $d(\boldsymbol{\alpha})$ and $|\boldsymbol{\alpha}|$ respectively, by

$$d(\boldsymbol{\alpha}) := \max\{j : \alpha_j \neq 0\} \qquad \text{and} \qquad |\boldsymbol{\alpha}| := \alpha_1 + \alpha_2 + \ldots + \alpha_{d(\boldsymbol{\alpha})}.$$

Take a complete orthonormal system $\{\eta_j\} \subset H$. For $\boldsymbol{\alpha} \in \mathcal{J}$, the $\boldsymbol{\alpha}$-th *Fourier-Hermite* polynomial base on $\{\eta_j\}$ is the polynomial

$$H_{\boldsymbol{\alpha}}(\omega) := \prod_{i=1}^{d(\boldsymbol{\alpha})} h_{\alpha_i}(\langle \omega, \eta_i \rangle) \text{ for all } \omega \in \mathcal{S}'(\mathbb{R}^d),$$

where $\{h_k\}_{k=1}^\infty$ is the family of Hermite polynomials.

Lemma 4.5 *(Holden et al. [1996], Theorem 2.2.3) The family of Fourier-Hermite polynomials* $\{H_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha}\in\mathcal{J}}$ *constitutes an orthogonal basis for* (L^2)*. Moreover,* $\|H_{\boldsymbol{\alpha}}\|^2_{(L^2)} = \boldsymbol{\alpha}!$ *for all* $\boldsymbol{\alpha}\in\boldsymbol{\mathcal{J}}$*.*

Lemma 4.6 *(Wiener-Itô chaos expansion theorem, Holden et al. [1996], Theorem 2.2.4) Every* $u\in(L^2)$ *has a unique representation*

$$u(\omega)=\sum_{\boldsymbol{\alpha}\in\mathcal{J}} u_{\boldsymbol{\alpha}}H_{\boldsymbol{\alpha}}(\omega), \tag{4.7}$$

where $u_{\boldsymbol{\alpha}}\in\mathbb{R}$ *for all* $\boldsymbol{\alpha}\in\boldsymbol{\mathcal{J}}$*. Moreover, we have* $\|u\|^2_{(L^2)}=\sum_{\boldsymbol{\alpha}\in\mathcal{J}}\boldsymbol{\alpha}!u^2_{\boldsymbol{\alpha}}$*.*

Remark 4.7 *In Section 9.3 we introduce the family of spaces* $(L^2)_s$*,* $s\in\mathbb{R}$ *defined by*

$$(L^2)_s := L^2(\mathcal{S}', e^{s\|\omega\|^2_{-\theta}}d\mu(\omega)), \tag{4.8}$$

with norm $\|v\|^2_{(L^2)_s} := \int_{\mathcal{S}'}|v|^2e^{s\|\omega\|^2_{-\theta}}d\mu$ *for* $v: S'\to\mathbb{R}$*. Using the Fernique's theorem, there is a constant* $c_F>0$*, such that for all* $s<c_F$*, Lemmas 4.5 and 4.6 can be extended to the space* $(L^2)_s$*. When* $\mathcal{S}'$ *is constructed from the Hilbert space* H *and the operator* A *as in Section 4.3.2, then* $c_F=\frac{\lambda_1^2}{2}$*. In this case, for* $u\in(L^2)_s$ *and* $s<\frac{\lambda_1^2}{2}$*, we have*

$$u(\omega)=\sum_{\boldsymbol{\alpha}\in\mathcal{J}} u_{\boldsymbol{\alpha},s}H_{\sigma(s),\boldsymbol{\alpha}}(\omega), \tag{4.9}$$

where the σ*-Fourier-Hermite polynomials are defined by*

$$H_{\sigma^2(s),\boldsymbol{\alpha}}(\omega) := \frac{1}{\sqrt{\sigma_*(s)}}\prod_{j=1}^{d(\boldsymbol{\alpha})} h_{\sigma_j^2(s),\alpha_j}(\langle\omega,\eta_j\rangle),\quad \omega\in\mathcal{S}',$$

where $\sigma_*=\sigma_*(s):=\int_{\mathcal{S}'}e^{s\|\omega\|^2_{-\theta}}d\mu(\omega)$ *and* $\sigma_j(s):=\left(1-\frac{2s}{\lambda_j^{2\theta}}\right)^{-\frac{1}{2}}$*. Here* $h_{\sigma,n}$ *is the* σ*-Hermite polynomial.*

Remark 4.8 *For the Chaos expansion in terms of* multiple Wiener-Itô integrals *and* Wick ordered polynomials *we refer to Hida [1980], Hida et al. [1993], Holden et al. [1996], Kuo [1996] and Obata [1994].*

Part II

Papers

Chapter 5

Non-matching Mortar Discretization Analysis for the Coupling Stokes/Darcy Equations

We consider the coupling across an interface of fluid and porous media flows with Beavers-Joseph-Saffman transmission conditions. Under an adequate choice of Lagrange multipliers on the interface we analyze inf-sup conditions and optimal a priori error estimates associated with the continuous and discrete formulations of this Stokes-Darcy system. We allow the meshes of the two regions to be non-matching across the interface. Using mortar finite element analysis and appropriate scaled norms we show that the constants that appear on the a priori error bounds do not depend on the viscosity, permeability and ratio of mesh parameters. Numerical experiments are presented.

5.1 Introduction

We analyze the coupling across an interface of fluid and porous media flows. This problem appears in several applications such as well-reservoir coupling in petroleum engineering, transport of substances across groundwater and surface water, and (bio)fluid-organ interactions. More precisely, we consider the following situation: an incompressible fluid in a region D_f can flow both ways across an interface Γ into a saturated porous medium domain D_p. The model studied here consists of *Stokes' equations* in the fluid region D_f and *Darcy's law* for the filtration velocity in the porous medium region D_p. The transmission conditions we consider on the interface Γ are the Beavers-Joseph-Saffman conditions which are widely accepted by the scientific community; see Beavers and Joseph [1967], Jäger and Mikelić [2000] and Saffman [1971]. In this paper we study inf-sup conditions and a priori error estimates associated with the continuous and discrete formulations of this Stokes-Darcy system. There are previous works addressing such issues, Layton et al. [2002], Rivière and Yotov [2005], Discacciati and Quarteroni [2004], Burman and Hansbo [2007], as well as related problems such as Stokes-Laplacian systems, Discacciati et al. [2002], Quarteroni et al. [2002], Discacciati [2004], Stokes-Navier Stokes, Girault et al. [2005], Quarteroni and Valli [1999], and

preconditioned iterative methods, Discacciati and Quarteroni [2003], Discacciati [2004], Discacciati and Quarteroni [2004], Galvis and Sarkis [2006], among others Mardal et al. [2002], Arbogast and Lehr [2006].

This paper is organized as follows: in Section 5.2 we discuss norms and seminorms of dual spaces on subsets. The differential systems are introduced in Section 5.3, where velocity and normal flux are considered as the boundary data for the Stokes part $\Gamma_f = \partial D_f \backslash \Gamma$ and the Darcy part $\Gamma_p = \partial D_p \backslash \Gamma$, respectively; for other formulations and boundary data see Discacciati et al. [2002] and Discacciati and Quarteroni [2003]. The transmission conditions on the interface Γ, known as Beavers-Joseph-Saffman conditions, are then introduced. In Section 5.4 we analyze weak formulations of the continuous model and we discuss the choice $H^{1/2}(\Gamma)$ as the space for Lagrange multipliers in order to couple these two systems of partial differential equations. In Layton et al. [2002], Layton, Schieweck, and Yotov developed existence and uniqueness of the weak solution for this problem. They were able to show the inf-sup condition on the smaller space $H_{00}^{1/2}(\Gamma)$. Recall that $H_{00}^{1/2}(\Gamma)$ is the subspace of functions in $H^{1/2}(\partial D_p)$ that vanish on $\partial D_p \setminus \Gamma$. In this paper, we use tools developed in Section 5.2 and in Layton et al. [2002] to present a complete analysis for the inf-sup condition with Lagrange multipliers on the space $H^{1/2}(\Gamma)$. We note that from the physical point of view the space $H^{1/2}(\Gamma)$ is the correct choice since the Lagrange multipliers are related to the Darcy pressure on the interface Γ and the value of the Darcy pressure at $\Gamma \cap \partial(D_f \cup D_p)$ is not prescribed when flux boundary condition is imposed on the porous side exterior boundary Γ_p. We note however that in the case where the pressure is imposed as the boundary condition on the Darcy exterior boundary Γ_p, the space $H_{00}^{1/2}(\Gamma)$ would be the correct choice; see Discacciati et al. [2002]. In Section 5.5 we derive the discrete inf-sup conditions and in Section 5.6 the a priori error estimates. We consider the triangular $P2\backslash P1$ Taylor Hood elements space for the free flow region D_f and the lowest order Raviart-Thomas for the Darcy region D_p. In Layton et al. [2002], Layton, Schieweck and Yotov developed a priori error estimates for the matching case, while in Rivière and Yotov [2005], and also in Burman and Hansbo [2007], the authors considered the non-matching case using Discontinuous Galerkin finite element discretizations. In this paper we consider the coupling via Lagrange multipliers and we develop an analysis based on mortar finite elements techniques (see Bernardi et al. [1994] and Wohlmuth [2000]) and scaled norms in order to obtain constants independent of the permeability, viscosity and ratio of mesh parameters. We also pay special attention to the constants appearing in the a priori error estimates. In Appendix B (Section 5.10) we provide the construction of the Fortin interpolation for P2\P1 Taylor Hood elements. In Section 5.7 we test numerically the algorithms and in Section 5.8 we make some conclusions.

5.2 Preliminaries and notations

Let D be a bounded Lipschitz continuous domain and let $\Gamma \subset \partial D$ and $\Gamma^c := \partial D \backslash \Gamma$ be of non-vanishing $(n-1)$-dimensional measure with respect to ∂D. Here $n = 2$ or 3.

To avoid the proliferation of constants, we will use the notation $A \preceq B$ to represent the inequality $A \leq (constant) \cdot B$.

Lemma 5.1 *Given $\mu \in H^{1/2}(\Gamma)$, define $E_\Gamma^{1/2}\mu := \gamma_0\varphi$ where γ_0 is the trace on ∂D and φ is the weak solution of*

$$\begin{cases} -\Delta\varphi = 0 & \text{in } D \\ \varphi = \mu & \text{on } \Gamma \\ \partial_{\boldsymbol{\eta}}\varphi = 0 & \text{on } \Gamma^c. \end{cases}$$

Then $E_\Gamma^{1/2}\mu \in H^{1/2}(\partial D)$ and $\|E_\Gamma^{1/2}\mu\|_{H^{1/2}(\partial D)} \preceq \|\mu\|_{H^{1/2}(\Gamma)}$.

For $\mu \in H^{1/2}(\Gamma)$ let $E_{00,\Gamma}^{1/2}\mu$ denote the extension by zero on Γ^c. Remember that $E_{00,\Gamma}^{1/2}\mu \in H^{1/2}(\partial D)$ if and only if $\mu \in H_{00}^{1/2}(\Gamma)$. We have the following result:

Lemma 5.2 *For all $\mu \in H^{1/2}(\partial D)$ there exist $\mu_\Gamma \in H^{1/2}(\Gamma)$ and $\mu_{\Gamma^c} \in H_{00}^{1/2}(\Gamma^c)$ such that $\mu = E_\Gamma^{1/2}\mu_\Gamma + E_{00,\Gamma^c}^{1/2}\mu_{\Gamma^c}$. This decomposition is unique.*

Proof. Let $\mu \in H^{1/2}(\partial D)$. Take $\mu_\Gamma = \mu|_\Gamma$ and $\mu_{\Gamma^c} = \varphi|_{\Gamma^c}$ where $\varphi = \mu - E_\Gamma^{1/2}\mu_\Gamma$. Observe that $\mu_\Gamma \in H^{1/2}(\Gamma)$ and

$$\|E_\Gamma^{1/2}\mu_\Gamma\|_{H^{1/2}(\partial D)} \preceq \|\mu_\Gamma\|_{H^{1/2}(\Gamma)} \leq \|\mu\|_{H^{1/2}(\partial D)},$$

therefore, $\varphi \in H^{1/2}(\partial D)$. Observe also that $E_{00,\Gamma^c}^{1/2}\mu_{\Gamma^c} = \varphi$ because μ and $E_\Gamma^{1/2}\mu_\Gamma$ coincide on Γ. For the uniqueness, if $0 = E_\Gamma^{1/2}\mu_\Gamma + E_{00,\Gamma^c}^{1/2}\mu_{\Gamma^c}$ then $E_\Gamma^{1/2}\mu_\Gamma$ is the trace of the weak solution of the problem: $-\Delta\varphi = 0$ in D, $\varphi = 0$ on Γ, $\partial_{\boldsymbol{\eta}}\varphi = 0$ on Γ^c. Then $\mu_\Gamma = 0$. ∎

We have two dual spaces associated with Γ, the space $H_{00}^{-1/2}(\Gamma)$ (the dual of $H_{00}^{1/2}(\Gamma)$) and $H^{-1/2}(\Gamma)$ (the dual space of $H^{1/2}(\Gamma)$). The first space is larger than the second one.

Definition 5.3 *If $f \in H^{-1/2}(\partial D)$, then $f|_{\Gamma^c} = 0$ means by definition that:*

$$\langle f, E_{00,\Gamma^c}^{1/2}\mu\rangle_{\partial D} = 0 \quad \text{for all } \mu \in H_{00}^{1/2}(\Gamma^c).$$

A useful result related with this definition is the following:

Lemma 5.4 *If $f \in H^{-1/2}(\partial D)$, there are $f_\Gamma \in H^{-1/2}(\Gamma)$ and $f_{\Gamma^c} \in H_{00}^{-1/2}(\Gamma^c)$ such that, for all $\mu \in H^{1/2}(\partial D)$, let $\mu = E_\Gamma^{1/2}\mu_\Gamma + E_{00,\Gamma^c}^{1/2}\mu_{\Gamma^c}$ as defined in Lemma 5.2, we have:*

$$\langle f, \mu\rangle_{\partial D} = \langle f_\Gamma, \mu_\Gamma\rangle_\Gamma + \langle f_{\Gamma^c}, \mu_{\Gamma^c}\rangle_{\Gamma^c}. \tag{5.1}$$

Proof. For $\mu_\Gamma \in H^{1/2}(\Gamma)$ and $\mu_{\Gamma^c} \in H_{00}^{1/2}(\Gamma^c)$ define:

$$\langle f_\Gamma, \mu_\Gamma\rangle_\Gamma := \langle f, E_\Gamma^{1/2}\mu_\Gamma\rangle_{\partial D} \qquad \langle f_{\Gamma^c}, \mu_{\Gamma^c}\rangle_{\Gamma^c} := \langle f, E_{00,\Gamma^c}^{1/2}\mu_{\Gamma^c}\rangle_{\partial D}.$$

We obtain

$$\langle f_\Gamma, \mu_\Gamma\rangle_\Gamma \leq \|f\|_{H^{-1/2}(\partial D)}\|E_\Gamma^{1/2}\mu_\Gamma\|_{H^{1/2}(\partial D)} \preceq \|f\|_{H^{-1/2}(\partial D)}\|\mu_\Gamma\|_{H^{1/2}(\Gamma)},$$

and so $f_\Gamma \in H^{-1/2}(\Gamma)$. Analogously $f_{\Gamma^c} \in H_{00}^{-1/2}(\Gamma^c)$. Moreover:

$$\langle f_\Gamma, \mu_\Gamma\rangle_\Gamma + \langle f_{\Gamma^c}, \mu_{\Gamma^c}\rangle_{\Gamma^c} = \langle f, E_\Gamma^{1/2}\mu_\Gamma + E_{00,\Gamma^c}^{1/2}\mu_{\Gamma^c}\rangle_{\partial D} = \langle f, \mu\rangle_{\partial D}.$$

■

Remark 5.5 *In particular, if $f \in H^{-1/2}(\partial D)$ and $f|_{\Gamma^c} = 0$, see Definition 5.3 above, we have from (5.1) that:*

$$\langle f, \mu\rangle_{\partial D} = \langle f_\Gamma, \mu_\Gamma\rangle_\Gamma.$$

Hence, functionals in $H^{-1/2}(\partial D)$ which are zero when restricted to $\partial D \setminus \Gamma$ can be identified with functionals in $H^{-1/2}(\Gamma)$.

Remark 5.6 *Given $f_\Gamma \in H^{-1/2}(\Gamma)$ we can define $f \in H^{-1/2}(\partial D)$ by $\langle f, \mu\rangle_{\partial D} := \langle f_\Gamma, \mu_\Gamma\rangle_\Gamma$, where $\mu = E_\Gamma^{1/2}\mu_\Gamma + E_{00,\Gamma^c}^{1/2}\mu_{\Gamma^c}$ as defined in Lemma 5.2. We have a similar result for $f_\Gamma \in H_{00}^{-1/2}(\Gamma^c)$.*

Define the space $\boldsymbol{H}(\mathrm{div}, D)$ by

$$\boldsymbol{H}(\mathrm{div}, D) := \left\{\boldsymbol{v} \in \boldsymbol{L}^2(D) :\ \nabla\cdot\boldsymbol{v} \in L^2(D)\right\},$$

with the norm

$$\|\boldsymbol{v}\|^2_{\boldsymbol{H}(\mathrm{div},\, D)} := \|\boldsymbol{v}\|^2_{\boldsymbol{L}^2(D)} + \|\nabla\cdot\boldsymbol{v}\|^2_{L^2(D)}. \tag{5.2}$$

Recall that if $\boldsymbol{v} \in \boldsymbol{H}(\mathrm{div}, D)$ then $\boldsymbol{v}\cdot\boldsymbol{\eta} \in H^{-1/2}(\partial D)$. For the next result see Wohlmuth et al. [2000].

Lemma 5.7 *For each $\boldsymbol{u} \in \boldsymbol{H}(\mathrm{div}, D)$ with $\int_{\partial D} \boldsymbol{u}\cdot\boldsymbol{\eta} = \langle \boldsymbol{u}\cdot\boldsymbol{\eta}, 1\rangle_{\partial D} = 0$ we have:*

$$\sup_{\substack{\phi\in H^{1/2}(\partial D)\\ \phi\neq \text{ constant}}} \frac{\langle \boldsymbol{u}\cdot\boldsymbol{\eta}, \phi\rangle_{\partial D}}{|\phi|_{H^{1/2}(\partial D)}} \preceq \|\boldsymbol{u}\cdot\boldsymbol{\eta}\|_{H^{-1/2}(\partial D)} \leq \sup_{\substack{\phi\in H^{1/2}(\partial D)\\ \phi\neq \text{ constant}}} \frac{\langle \boldsymbol{u}\cdot\boldsymbol{\eta}, \phi\rangle_{\partial D}}{|\phi|_{H^{1/2}(\partial D)}}.$$

with a constant which depends only on D.

Using an argument similar to the one given in Wohlmuth et al. [2000] we have:

Lemma 5.8 *For each $f \in H^{-1/2}(\Gamma)$ with $\int_\Gamma f = \langle f, 1\rangle_\Gamma = 0$, we have:*

$$\sup_{\substack{\phi\in H^{1/2}(\Gamma)\\ \phi\neq \text{ constant}}} \frac{\langle f, \phi\rangle_\Gamma}{|\phi|_{H^{1/2}(\Gamma)}} \preceq \|f\|_{H^{-1/2}(\Gamma)} \leq \sup_{\substack{\phi\in H^{1/2}(\Gamma)\\ \phi\neq \text{ constant}}} \frac{\langle f, \phi\rangle_\Gamma}{|\phi|_{H^{1/2}(\Gamma)}},$$

with a constant which depends only on Γ.

Proof. Observe that if α is a constant then $\langle f, \alpha\rangle_\Gamma = \alpha\langle f, 1\rangle_\Gamma = 0$ and for $\phi \in H^{1/2}(\Gamma)$ non-constant we have

$$\frac{\langle f, \phi\rangle_\Gamma}{\|\phi\|_{H^{1/2}(\Gamma)}} \leq \frac{\langle f, \phi\rangle_\Gamma}{|\phi|_{H^{1/2}(\Gamma)}},$$

then

$$\|f\|_{H^{-1/2}(\Gamma)} = \sup_{\substack{\phi\in H^{1/2}(\Gamma)\\ \phi\neq\text{constant}}} \frac{\langle f, \phi\rangle_\Gamma}{\|\phi\|_{H^{1/2}(\Gamma)}} \leq \sup_{\substack{\phi\in H^{1/2}(\Gamma)\\ \phi\neq\text{constant}}} \frac{\langle f, \phi\rangle_\Gamma}{|\phi|_{H^{1/2}(\Gamma)}}$$

which gives the right inequality. Using a Poincaré inequality, there exists a positive constant which depends only on Γ, such that

$$\|\psi\|^2_{H^{1/2}(\Gamma)} \preceq |\psi|^2_{H^{1/2}(\Gamma)}$$

holds for all $\psi \in H^{1/2}(\Gamma)$ with $\int_\Gamma \psi = 0$. For $\phi \in H^{1/2}(\Gamma)$ non-constant we have:

$$\psi := \phi - \int_\Gamma \phi \neq 0,$$

and

$$\frac{\langle f, \psi\rangle_\Gamma}{\|\psi\|_{H^{1/2}(\Gamma)}} = \frac{\langle f, \phi\rangle_\Gamma}{\|\psi\|_{H^{1/2}(\partial D)}} \succeq \frac{\langle f, \phi\rangle}{|\psi|_{H^{1/2}(\Gamma)}} = \frac{\langle f, \phi\rangle}{|\phi|_{H^{1/2}(\Gamma)}}.$$

■

This gives an equivalent norm in the subspace of $H^{-1/2}(\Gamma)$ of zero average functionals.

Definition 5.9 *For $f \in H^{-1/2}(\Gamma)$, f with zero average ($\langle f, 1\rangle_\Gamma = 0$), define:*

$$|f|_{H^{-1/2}(\Gamma)} := \sup_{\substack{\phi\in H^{1/2}(\Gamma)\\ \phi\neq \text{ constant}}} \frac{\langle f, \phi\rangle_\Gamma}{|\phi|_{H^{1/2}(\Gamma)}} = \sup_{\substack{\phi\in H^{1/2}(\Gamma)\\ \int_\Gamma \phi=0, \phi\neq 0,}} \frac{\langle f, \phi\rangle_\Gamma}{|\phi|_{H^{1/2}(\Gamma)}}.$$

We have the following result:

Lemma 5.10 *For $\mu \in H^{1/2}(\Gamma)$ with $\int_\Gamma \mu = 0$ we have:*

$$|\mu|_{H^{1/2}(\Gamma)} = \sup_{\substack{f\in H^{-1/2}(\Gamma)\\ \langle f,1\rangle=0}} \frac{\langle f, \mu\rangle_\Gamma}{|f|_{H^{-1/2}(\Gamma)}}.$$

Proof. Consider $(H^{1/2}(\Gamma) \cap L^2_0(\Gamma))^*$, the dual space of $H^{1/2}(\Gamma) \cap L^2_0(\Gamma)$, and observe that a functional $f_0 \in (H^{1/2}(\Gamma) \cap L^2_0(\Gamma))^*$ can be extended to one in $H^{1/2}(\Gamma)^*$, say f, by the following formula: $\langle f, \phi\rangle := \langle f_0, \phi_0\rangle$ where $\phi \in H^{1/2}(\Gamma)$ and $\phi_0 := \phi - \int_\Gamma \phi$. ■

5.3 P.D.E model

In general, $D_f,\ D_p \subset \mathbb{R}^n$, $\Gamma = \overline{D}_f \cap \overline{D}_p$, $D = \text{int}(\overline{D}_f \cup \overline{D}_p)$, D_f and D_p are Lipschitz, so it is possible to define outward unit normal vectors, denoted by $\boldsymbol{\eta}_j$, $j = f, p$. The tangent vectors on Γ are denoted by $\boldsymbol{\tau}_1$ ($n = 2$), or $\boldsymbol{\tau}_l$, $l = 1, 2$ ($n = 3$). In order to avoid a setting that is too general, when $n = 2$ we consider $D_f = (1, 2) \times (0, 1)$ and $D_p = (0, 1) \times (0, 1)$ or a regular Lipschitz perturbation of this configuration. Analogous conditions are consider for the case $n = 3$.

Define $\Gamma_j := \partial D_j \setminus \Gamma$, $j = f, p$. Velocities are denoted by $\boldsymbol{u}_j : D_j \to \mathbb{R}^n$, $j = f, p$. Pressures are $p_j : D_j \to \mathbb{R}$, $j = f, p$.

As it was mentioned previously, Stokes' equations are the the model for the fluid region. The model basically consists of conservation of mass and conservation of momentum, and we have:

$$\begin{cases} -\nabla \cdot T(\boldsymbol{u}_f, p_f) &= \boldsymbol{f}_f \quad \text{in } D_f \\ \nabla \cdot \boldsymbol{u}_f &= g_f \quad \text{in } D_f \\ \boldsymbol{u}_f &= \boldsymbol{h}_f \quad \text{on } \Gamma_f. \end{cases} \tag{5.3}$$

Here $T(\boldsymbol{v}, p) := -pI + 2\nu \boldsymbol{D}\boldsymbol{v}$ where ν is the fluid viscosity and $\boldsymbol{D}\boldsymbol{v} := \frac{1}{2}(\nabla \boldsymbol{v} + \nabla^T \boldsymbol{v})$ is the linearized strain tensor.

For the porous domain D_p, Darcy's law is used, i.e., $(\boldsymbol{u}_p, p_p)$ satisfies on D_p:

$$\begin{cases} \boldsymbol{u}_p &= -\frac{\kappa}{\nu} \nabla p_p + \boldsymbol{f}_p \quad \text{in } D_p \quad \text{(Darcy's law)} \\ \nabla \cdot \boldsymbol{u}_p &= g_p \quad \text{in } D_p \\ \boldsymbol{u}_p \cdot \boldsymbol{\eta}_p &= h_p \quad \text{on } \Gamma_p. \end{cases} \tag{5.4}$$

In general κ is a symmetric and a uniformly positive definite tensor that represents the rock permeability. For simplicity on the analysis we assume that κ is a real positive constant. Recall that ν is the fluid viscosity.

We also impose the compatibility condition

$$\int_{D_f} g_f + \int_{D_p} g_p - \int_{\Gamma_f} \boldsymbol{h}_f \cdot \boldsymbol{\eta}_f - \int_{\Gamma_p} h_p = 0. \tag{5.5}$$

The systems presented above must be coupled across the interface Γ. The following conditions are imposed (see Layton et al. [2002], Discacciati and Quarteroni [2003], Discacciati et al. [2002], Discacciati and Quarteroni [2004] and references therein):

Conservation of mass across Γ*:* It is expressed by:

$$\boldsymbol{u}_f \cdot \boldsymbol{\eta}_f + \boldsymbol{u}_p \cdot \boldsymbol{\eta}_p = 0 \text{ on } \Gamma. \tag{5.6}$$

This means that the fluid that is leaving a region enters in the other one.

Balance of normal forces across Γ*:* From Cauchy formula we see that

$$\Sigma(\boldsymbol{u}_f, p_f) := T(\boldsymbol{u}_f, p_f)\boldsymbol{\eta}_f$$

is the force on ∂D_f acting on the fluid volume inside D_f, i.e., Σ is the Cauchy stress (or traction) vector. The force on Γ from D_f side is then $\Sigma(\boldsymbol{u}_f, p_f)$. The only force acting on the interface from D_p side is the one given by p_p in the direction of $\boldsymbol{\eta}_p$ and must be equal to the component of Σ in this direction. We get

$$p_f - 2\nu \boldsymbol{\eta}_f^T \boldsymbol{D}(\boldsymbol{u}_f)\boldsymbol{\eta}_f = p_p \quad \text{on } \Gamma. \tag{5.7}$$

The other components of Σ are more delicate and treated below.

Beavers-Joseph-Saffman condition: This condition is a kind of empirical law that gives an expression for the tangential component of Σ. It is expressed by:

$$\boldsymbol{u}_f \cdot \boldsymbol{\tau}_l = -\frac{\sqrt{\kappa}}{\alpha_f} 2\boldsymbol{\eta}_f^T \boldsymbol{D}(\boldsymbol{u}_f)\boldsymbol{\tau}_l \quad \text{on } \Gamma,\ l = 1, n-1. \tag{5.8}$$

In the general case, κ is a symmetric and uniformly positive definite tensor, and κ in (5.8) is replaced by $\tau_l \cdot \kappa \cdot \tau_l$.

A related condition is

$$(\boldsymbol{u}_f - \boldsymbol{u}_p) \cdot \boldsymbol{\tau}_l = -\frac{\sqrt{\kappa}}{\alpha_f} 2\boldsymbol{\eta}_f^T \boldsymbol{D}(\boldsymbol{u}_f)\boldsymbol{\tau}_l \quad \text{on } \Gamma,\ l = 1, n-1,$$

which is known as the Beavers-Joseph condition. But it turns out in practice that the component of $\boldsymbol{u}_p$ in $\boldsymbol{\tau}_l$ direction is small compared with that of $\boldsymbol{u}_f$. When more general cases are considered, suitable interface conditions have to be imposed. An analytical way to find the right interface conditions is via homogenization (see Hornung [1997]).

5.4 Weak formulations and inf-sup analysis

In this section we derive and analyze several weak formulations associated with the Stokes-Darcy system presented in Section 5.3.

5.4.1 Weak formulations

According to Appendix 5.9, it is enough to consider the case $g_f = 0$ and $\boldsymbol{h}_f = \boldsymbol{0}$ in (5.3) and $g_p = 0$ and $h_p = 0$ in (5.4).

For D_f define

$$\boldsymbol{X}_f := H_0^1(D_f, \Gamma_f)^n \text{ and } M_f := L^2(D_f). \tag{5.9}$$

where $H_0^1(D_f, \Gamma_f)^n$ means by definition the subspace of functions $\boldsymbol{v}_f$ such that each component of $\boldsymbol{v}_f$ belongs to $H^1(D_f)$ and vanishes on Γ_f.

For D_p we introduce the following spaces:

$$\boldsymbol{X}_p := \boldsymbol{H}_0(\mathrm{div}, D_p, \Gamma_p) \text{ and } M_p := L^2(D_p), \tag{5.10}$$

where $\boldsymbol{H}_0(\mathrm{div}, D_p, \Gamma_p)$ is defined as the subspace of $\boldsymbol{H}(\mathrm{div}, D_p)$ of functions with vanishing normal component on Γ_p in the sense of Definition 5.3. Recall that if

$\boldsymbol{u}_p \in \boldsymbol{H}(\mathrm{div}, D_p)$ then $\boldsymbol{u}_p \cdot \boldsymbol{\eta}_p \in H^{-1/2}(\partial D_p)$; see (5.2).

Define $\boldsymbol{X} := \boldsymbol{X}_f \times \boldsymbol{X}_p$ with the usual norm, i.e., given $\boldsymbol{v} = (\boldsymbol{v}_f, \boldsymbol{v}_p) \in \boldsymbol{X}$,

$$\|\boldsymbol{v}\|^2_{\boldsymbol{X}} := |\boldsymbol{v}_f|^2_{H^1(D_f)^n} + \|\boldsymbol{v}_p\|^2_{\boldsymbol{H}(\mathrm{div},\, D_p)}. \tag{5.11}$$

We also set $M := M_f \times M_p$ with the norm $\|q\|^2_M := \|q_f\|^2_{L^2(D_f)} + \|q_f\|^2_{L^2(D_p)}$.

In order to derive a weak formulation we first proceed formally and then we introduce the adequate rigorous framework.

We start with the Stokes' equation (5.3). For all $\boldsymbol{v}_f \in \boldsymbol{X}_f$ we have:

$$(-2\nu\nabla\cdot\boldsymbol{D}\boldsymbol{u}_f, \boldsymbol{v}_f)_{D_f} + (\nabla p_f, \boldsymbol{v}_f)_{D_f} = (\boldsymbol{f}_f, \boldsymbol{v}_f)_{D_f}. \tag{5.12}$$

From the Green formula we have

$$\begin{aligned} -(\Delta\boldsymbol{u}_f, \boldsymbol{v}_f)_{D_f} &= (\nabla\boldsymbol{u}_f, \nabla\boldsymbol{v}_f)_{D_f} - (\nabla\boldsymbol{u}_f\boldsymbol{\eta}_f, \boldsymbol{v}_f)_\Gamma \\ &= (\nabla\boldsymbol{u}_f, \nabla\boldsymbol{v}_f)_{D_f} - \langle\boldsymbol{\eta}_f^T\nabla\boldsymbol{u}_f\boldsymbol{\eta}_f, \boldsymbol{v}_f\cdot\boldsymbol{\eta}_f\rangle_\Gamma \\ &\quad - \sum_{l=1}^{n-1}\langle\boldsymbol{\tau}_l^T\nabla\boldsymbol{u}_f\boldsymbol{\eta}_f, \boldsymbol{v}_f\cdot\boldsymbol{\tau}_l\rangle_\Gamma, \end{aligned}$$

and

$$\begin{aligned} -(\nabla\cdot\nabla\boldsymbol{u}_f^T, \boldsymbol{v}_f)_{D_f} &= (\nabla\boldsymbol{u}_f^T, \nabla\boldsymbol{v}_f)_{D_f} - \langle\nabla\boldsymbol{u}_f^T\boldsymbol{\eta}_f, \boldsymbol{v}_f\rangle_\Gamma \\ &= (\nabla\boldsymbol{u}_f^T, \nabla\boldsymbol{v}_f)_{D_f} - \langle\boldsymbol{\eta}_f^T\nabla\boldsymbol{u}_f^T\boldsymbol{\eta}_f, \boldsymbol{v}_f\cdot\boldsymbol{\eta}_f\rangle_\Gamma \\ &\quad - \sum_{l=1}^{n-1}\langle\boldsymbol{\tau}_l^T\nabla\boldsymbol{u}_f^T\boldsymbol{\eta}_f, \boldsymbol{v}_f\cdot\boldsymbol{\tau}_l\rangle_\Gamma, \end{aligned}$$

then

$$\begin{aligned} -(2\nabla\cdot\boldsymbol{D}\boldsymbol{u}_f, \boldsymbol{v}_f)_{D_f} &= 2(\boldsymbol{D}\boldsymbol{u}_f, \boldsymbol{D}\boldsymbol{v}_f)_{D_f} - 2\langle\boldsymbol{\eta}_f^T\boldsymbol{D}\boldsymbol{u}_f\boldsymbol{\eta}_f, \boldsymbol{v}_f\cdot\boldsymbol{\eta}_f\rangle_\Gamma \\ &\quad - 2\sum_{l=1}^{n-1}\langle\boldsymbol{\tau}_l^T\boldsymbol{D}\boldsymbol{u}_f\boldsymbol{\eta}_f, \boldsymbol{v}_f\cdot\boldsymbol{\tau}_l\rangle_\Gamma. \end{aligned}$$

For the second term on (5.12) we have:

$$(\nabla p_f, \boldsymbol{v}_f)_{D_f} = \langle p_f, \boldsymbol{v}_f\cdot\boldsymbol{\eta}_f\rangle_\Gamma - (p_f, \nabla\cdot\boldsymbol{v}_f)_{D_f}. \tag{5.13}$$

For $\boldsymbol{u}_f, \boldsymbol{v}_f \in \boldsymbol{X}_f$ and $q_f \in M_f$ define:

$$a_f(\boldsymbol{u}_f, \boldsymbol{v}_f) := 2\nu(\boldsymbol{D}\boldsymbol{u}_f, \boldsymbol{D}\boldsymbol{v}_f)_{D_f} + \sum_{l=1}^{n-1}\frac{\nu\alpha_f}{\sqrt{\kappa}}\langle\boldsymbol{u}_f\cdot\boldsymbol{\tau}_l, \boldsymbol{v}_f\cdot\boldsymbol{\tau}_l\rangle_\Gamma, \tag{5.14}$$

$$b_f(\boldsymbol{v}_f, q_f) := -(q_f, \nabla\cdot\boldsymbol{v}_f)_{D_f}. \tag{5.15}$$

By replacing (5.13) in (5.12), and using the condition (5.8), we obtain for all $\boldsymbol{v}_f \in \boldsymbol{X}_f$ and $q_f \in M_f$

$$\begin{cases} a_f(\boldsymbol{u}_f, \boldsymbol{v}_f) + b_f(\boldsymbol{v}_f, p_f) + \langle p_f - 2\nu\boldsymbol{\eta}_f^T\boldsymbol{D}(\boldsymbol{u}_f)\boldsymbol{\eta}_f, \boldsymbol{v}_f\cdot\boldsymbol{\eta}_f\rangle_\Gamma & = (\boldsymbol{f}_f, \boldsymbol{v}_f)_{D_f} \\ b_f(\boldsymbol{u}_f, q_f) & = 0. \end{cases} \tag{5.16}$$

Analogously, defining

$$a_p(\boldsymbol{u}_p, \boldsymbol{v}_p) := (\frac{\nu}{\kappa}\boldsymbol{u}_p, \boldsymbol{v}_p)_{D_p} \quad \text{for all } \boldsymbol{u}_p, \boldsymbol{v}_p \in \boldsymbol{X}_p, \tag{5.17}$$

$$b_p(\boldsymbol{v}_p, q_p) := -(q_p, \nabla\cdot\boldsymbol{v}_p)_{D_p} \text{ for all } \boldsymbol{v}_p \in \boldsymbol{X}_p \text{ and } q_p \in M_p,$$

we have for all $\boldsymbol{v}_p \in \boldsymbol{X}_p$ and $q_p \in M_p$

$$\begin{cases} a_p(\boldsymbol{u}_p, \boldsymbol{v}_p) + b_p(\boldsymbol{v}_p, p_p) + \langle p_p, \boldsymbol{v}_p\cdot\boldsymbol{\eta}_p\rangle_\Gamma & = (\boldsymbol{f}_p, \boldsymbol{v}_p)_{D_p} \\ b_p(\boldsymbol{u}_p, q_p) & = 0. \end{cases} \tag{5.18}$$

To couple the two subproblems (5.16) and (5.18) we use balance of normal forces (5.7) and a Lagrange multiplier which also approximate the Darcy pressure on the interface Γ. Introduce the Lagrange multiplier λ :

$$\lambda = p_p = p_f - 2\nu\boldsymbol{\eta}_f^T\boldsymbol{D}(\boldsymbol{u}_f)\boldsymbol{\eta}_f = p_f - 2\nu\boldsymbol{\eta}_f^T\nabla\boldsymbol{u}\boldsymbol{\eta}_f. \tag{5.19}$$

Then we get :

$$\begin{cases} a_f(\boldsymbol{u}_f, \boldsymbol{v}_f) + b_f(\boldsymbol{v}_f, p_f) + \langle \boldsymbol{v}_f\cdot\boldsymbol{\eta}_f, \lambda\rangle_\Gamma & = (\boldsymbol{f}_f, \boldsymbol{v}_f)_{D_f} & \text{for all } \boldsymbol{v}_f \in \boldsymbol{X}_f \\ a_p(\boldsymbol{u}_p, \boldsymbol{v}_p) + b_p(\boldsymbol{v}_p, p_p) + \langle \boldsymbol{v}_p\cdot\boldsymbol{\eta}_p, \lambda\rangle_\Gamma & = (\boldsymbol{f}_p, \boldsymbol{v}_p)_{D_p} & \text{for all } \boldsymbol{v}_p \in \boldsymbol{X}_p \\ b_f(\boldsymbol{u}_f, q_f) & = 0 & \text{for all } q_f \in M_f \\ b_p(\boldsymbol{u}_p, q_p) & = 0 & \text{for all } q_f \in M_p \\ \langle \boldsymbol{u}_f\cdot\boldsymbol{\eta}_f + \boldsymbol{u}_p\cdot\boldsymbol{\eta}_p, \mu\rangle_\Gamma & = 0 & \text{for all } \mu \in \Lambda \end{cases} \tag{5.20}$$

where the space Λ is defined below.

Define $a : \boldsymbol{X} \times \boldsymbol{X} \to \mathbb{R}$ and $b : \boldsymbol{X} \times M \to \mathbb{R}$ by:

$$a(\boldsymbol{u}, \boldsymbol{v}) := a_f(\boldsymbol{u}_f, \boldsymbol{v}_f) + a_p(\boldsymbol{u}_p, \boldsymbol{v}_p), \tag{5.21}$$

$$b(\boldsymbol{v}, q) := b_f(\boldsymbol{v}_f, q_f) + b_p(\boldsymbol{v}_p, q_p), \tag{5.22}$$

Using (5.6) we obtain,

$$\begin{cases} a(\boldsymbol{u}, \boldsymbol{v}) + b(\boldsymbol{v}, p) + \langle \boldsymbol{v}_f\cdot\boldsymbol{\eta}_f + \boldsymbol{v}_p\cdot\boldsymbol{\eta}_p, \lambda\rangle_\Gamma & = (\boldsymbol{f}_f, \boldsymbol{v}_f)_{D_f} + (\boldsymbol{f}_p, \boldsymbol{v}_p)_{D_p} \\ b(\boldsymbol{u}, q) & = 0 \\ \langle \boldsymbol{u}_f\cdot\boldsymbol{\eta}_f + \boldsymbol{u}_p\cdot\boldsymbol{\eta}_p, \mu\rangle_\Gamma & = 0. \end{cases} \tag{5.23}$$

Note that if p is a solution of (5.23), then p plus any constant is also a solution of (5.23); this follows directly from applying the divergence theorem on the first equation of (5.23) and using (5.19). In addition, using the the divergence theorem on the second equation of (5.23) and the compatibility condition (5.5) we have that the equation (5.23) is automatically satisfied for constant test functions $q \in M$. Therefore, we can replace the space M in (5.20) by the following subspace of M

$$M^\circ := \left\{ q = (q_f, q_p) \in M \ : \ \int_{D_f} q_f + \int_{D_p} q_p = 0 \right\}. \tag{5.24}$$

We have to choose a suitable function space Λ for λ. Observe that on the porous exterior boundary Γ_p we consider zero flux as boundary condition, i.e., $\boldsymbol{v}_p\cdot\boldsymbol{\eta}_p = 0$ on Γ_p. Recalling Definition 5.3, this means that

$$\langle \boldsymbol{v}_p\cdot\boldsymbol{\eta}_p, E^{1/2}_{00,\Gamma_p}\phi\rangle_{\partial D_p} = 0 \quad \text{for all } \phi \in H^{1/2}_{00}(\Gamma_p),$$

where $E^{1/2}_{00,\Gamma_p}$ denotes the extension by zero on $\Gamma^c_p = \Gamma$. Then, according to Lemma 5.4 and Remark 5.5 we can think of $\boldsymbol{v}_p \cdot \boldsymbol{\eta}_p$ as a distribution in $H^{-1/2}(\Gamma)$, more precisely, we can define $\boldsymbol{v}_p \cdot \boldsymbol{\eta}_p|_\Gamma \in H^{-1/2}(\Gamma)$ as

$$\langle \boldsymbol{v}_p \cdot \boldsymbol{\eta}_p|_\Gamma, \phi \rangle_\Gamma := \langle \boldsymbol{v}_p \cdot \boldsymbol{\eta}_p, E^{1/2}_\Gamma \phi \rangle_{\partial D_p}, \quad \phi \in H^{1/2}(\Gamma), \tag{5.25}$$

where $E^{1/2}_\Gamma$ is the extension operator defined in Lemma 5.1. This is the main mathematical motivation for choosing Λ as $H^{1/2}(\Gamma)$ rather than $H^{1/2}_{00}(\Gamma)$. On the fluid exterior boundary Γ_f we are using Dirichlet boundary condition, i.e., $\boldsymbol{v}_f = \boldsymbol{0}$ on Γ_f. Then $\boldsymbol{v}_f \cdot \boldsymbol{\eta}_f|_\Gamma \in H^{1/2}_{00}(\Gamma)$ relatively to ∂D_f. Then $\boldsymbol{v}_f \cdot \boldsymbol{\eta}_f|_\Gamma \in H^{1/2}_{00}(\Gamma)$ relatively to ∂D_p. Here we use the fact that $H^{1/2}_{00}(\Gamma)$, which is the trace of $H^1_0(D_f, \Gamma_f)$, is equivalent to the trace of $H^1_0(D_p, \Gamma_p)$ if the shape and measure of D_f are of the similar size of those of D_p; see Grisvard [1985] and Nečas [1967]. Since $H^{1/2}_{00}(\Gamma) \subset H^{-1/2}(\Gamma)$ we conclude that $\boldsymbol{v}_f \cdot \boldsymbol{\eta}_f|_\Gamma \in H^{-1/2}(\Gamma)$. In what follows we denote $\boldsymbol{v}_p \cdot \boldsymbol{\eta}_p|_\Gamma$ simply by $\boldsymbol{v}_p \cdot \boldsymbol{\eta}_p$ and $\boldsymbol{v}_f \cdot \boldsymbol{\eta}_f|_\Gamma$ by $\boldsymbol{v}_f \cdot \boldsymbol{\eta}_f$.

From the previous discussion we conclude that $\boldsymbol{v}_f \cdot \boldsymbol{\eta}_f + \boldsymbol{v}_p \cdot \boldsymbol{\eta}_p \in H^{-1/2}(\Gamma)$ and so we choose for λ the space

$$\Lambda := H^{1/2}(\Gamma) \quad \text{with} \quad \|\cdot\|^2_\Lambda := \|\cdot\|^2_{H^{1/2}(\Gamma)} = \|\cdot\|^2_{L^2(\Gamma)} + |\cdot|^2_{H^{1/2}(\Gamma)} \tag{5.26}$$

and define $b_\Gamma : \boldsymbol{X} \times \Lambda \to \mathbb{R}$ by:

$$b_\Gamma(\boldsymbol{v}, \mu) := \langle \boldsymbol{v}_f \cdot \boldsymbol{\eta}_f, \mu \rangle_\Gamma + \langle \boldsymbol{v}_p \cdot \boldsymbol{\eta}_p, \mu \rangle_\Gamma, \quad \boldsymbol{v} = (\boldsymbol{v}_f, \boldsymbol{v}_p) \in \boldsymbol{X}, \quad \mu \in \Lambda, \tag{5.27}$$

with the second duality pairing as in (5.25).

From Lemma 5.4 we have the following result.

Lemma 5.11 *$b_\Gamma : \boldsymbol{X} \times \Lambda \to \mathbb{R}$ defined in (5.27) and (5.25) is continuous.*

Another reason for choosing $H^{1/2}(\Gamma)$ instead of $H^{1/2}_{00}(\Gamma)$ is because the Lagrange multiplier represents the porous pressure on Γ, see (5.19), and hence there is no physical reason for the pressure p_p to vanish on $\Gamma \cap \partial D$ when flux boundary conditions are imposed on the porous side exterior boundary Γ_p. The space we choose for Λ is richer than $H^{1/2}_{00}(\Gamma)$, therefore the equation

$$b_\Gamma(\boldsymbol{u}, \mu) = 0 \qquad \text{for all } \mu \in \Lambda = H^{1/2}(\Gamma)$$

applied to $\boldsymbol{u}$ is a stronger condition than considering μ on the space $H^{1/2}_{00}(\Gamma)$. As a result, better mass conservation near $\Gamma \cap \partial D$ is achieved. On the other hand, choosing $H^{1/2}_{00}(\Gamma)$ as the spaces of Lagrange multipliers associated to the porous pressure would be more appropriate if zero pressure was imposed on ∂D; see Discacciati et al. [2002].

First weak formulation

We finally arrive to the weak formulation of the problem: Find $(\boldsymbol{u}^1, p^1, \lambda^1) \in \boldsymbol{X} \times M^\circ \times \Lambda$ such that:

$$\begin{cases} a(\boldsymbol{u}^1, \boldsymbol{v}) + b(\boldsymbol{v}, p^1) + b_\Gamma(\boldsymbol{v}, \lambda^1) & = \ell(\boldsymbol{v}) \quad \text{for all } \boldsymbol{v} \in \boldsymbol{X} \\ b(\boldsymbol{u}^1, q) & = 0 \quad \text{for all } q \in M^\circ \\ b_\Gamma(\boldsymbol{u}^1, \mu) & = 0 \quad \text{for all } \mu \in \Lambda, \end{cases} \tag{5.28}$$

where

$$\ell(\boldsymbol{v}) := (\boldsymbol{f}_f, \boldsymbol{v}_f)_{D_f} + (\boldsymbol{f}_p, \boldsymbol{v}_p)_{D_p} \text{ for all } \boldsymbol{v} \in \boldsymbol{X}. \tag{5.29}$$

and the bilinear forms a, b and b_Γ are defined in (5.21), (5.22) and, (5.27) and (5.25), respectively.

Next we introduce two other weak formulations and we refer to them as the second and the third weak formulations; see (5.32) and (5.35). The second weak formulation is an intermediate step for deriving the third weak formulation. The third formulation is the most fundamental one among the three formulations and it is where most of the analysis is carried on. Once the inf-sup condition is established for the third weak formulation, the inf-sup for the other two formulations follow straightforwardly; see Remark 5.18. The analysis of the third weak formulation is based on seminorms and on the theoretical tools developed in Section 5.2. The three weak formulations are all *equivalent* in the following sense (see Remarks 5.12 and 5.13):

1. If we know a solution $(\hat{\boldsymbol{u}}, \hat{p}, \hat{\lambda})$ for one weak formulation, then we can construct a solution for the other two weak formulations. This construction is done by removing or by recovering the mean value of the fluid and porous pressure solutions and the mean value of the Lagrange multiplier solution.

2. All three weak formulations have the same velocity solutions.

The Proposition 5.25 establishes the inf-sup condition for the third weak formulation, therefore, the existence and uniqueness of the solution follow; see Subsection 5.4.1. Hence, existence of a solution for the first and second weak formulations follows from Remarks 5.12 and 5.13. Finally, the Remark 5.18 establishes the inf-sup conditions for the first and second weak formulations and therefore, the uniqueness of their solution.

Second weak formulation

Now we introduce an equivalent weak formulation for (5.28) by eliminating the velocities with non-zero mean normal jump across Γ and also the Lagrange multipliers that are constants; see Remark 5.12 below. Define

$$\boldsymbol{X}^\circ = \left\{ \boldsymbol{v} = (\boldsymbol{v}_f, \boldsymbol{v}_p) \in \boldsymbol{X} \ : \ b_\Gamma(\boldsymbol{v}, 1) = \int_\Gamma \boldsymbol{v}_f \cdot \boldsymbol{\eta}_f + \boldsymbol{v}_p \cdot \boldsymbol{\eta}_p = 0 \right\} \tag{5.30}$$

and

$$\Lambda^\circ := H^{1/2}(\Gamma) \cap L^2_0(\Gamma) \quad \text{with norm } |\cdot|_{\Lambda^\circ} := |\cdot|_{H^{1/2}(\Gamma)}. \tag{5.31}$$

The second weak formulation: Find $(\boldsymbol{u}^2, p^2, \lambda^2) \in \boldsymbol{X}^\circ \times M^\circ \times \Lambda^\circ$ such that:

$$\left\{ \begin{array}{lll} a(\boldsymbol{u}^2, \boldsymbol{v}) + b(\boldsymbol{v}, p^2) + b_\Gamma(\boldsymbol{v}, \lambda^2) & = \ell(\boldsymbol{v}) & \text{for all } \boldsymbol{v} \in \boldsymbol{X}^\circ \\ b(\boldsymbol{u}^2, q) & = 0 & \text{for all } q \in M^\circ \\ b_\Gamma(\boldsymbol{u}^2, \mu) & = 0 & \text{for all } \mu \in \Lambda^\circ. \end{array} \right. \tag{5.32}$$

Remark 5.12 *It is easy to see that if $(\boldsymbol{u}^1, p^1, \lambda^1) \in \boldsymbol{X} \times M^\circ \times \Lambda$ solves the weak formulation (5.28) then $\boldsymbol{u}^1 \in \boldsymbol{X}^\circ$ and $(\boldsymbol{u}^1, p^1, \lambda^2)$ solves (5.32) with $\lambda^2 = \lambda^1 - \frac{1}{|\Gamma|}\int_\Gamma \lambda^1$. To see the converse, let $(\boldsymbol{u}^2, p^2, \lambda^2) \in \boldsymbol{X}^\circ \times M^\circ \times \Lambda^\circ$ be a solution of (5.32). Construct $\boldsymbol{w} = (0, \boldsymbol{w}_p) \in \boldsymbol{X}$ such that*

$$\boldsymbol{w}_p \cdot \boldsymbol{\eta}_p = \frac{1}{|\Gamma|} \text{ on } \Gamma \text{ and } \boldsymbol{w}_p \cdot \boldsymbol{\eta}_p = 0 \text{ on } \Gamma_p,$$

and define

$$\bar{\lambda} := \ell(\boldsymbol{w}) - a(\boldsymbol{u}^2, \boldsymbol{w}) - b(\boldsymbol{w}, p^2)$$

and set $\lambda^1 := \lambda^2 + \bar{\lambda}$. Then $(\boldsymbol{u}^2, p^2, \lambda^1)$ solves (5.28). Indeed, observe that $b_\Gamma(\boldsymbol{w}, \lambda^1) = \bar{\lambda}$ and that for $\boldsymbol{v} = (\boldsymbol{v}_f, \boldsymbol{v}_p) \in \boldsymbol{X}$ we can find α such that $\boldsymbol{v}^2 := \boldsymbol{v} + \alpha \boldsymbol{w} \in \boldsymbol{X}^\circ$. Hence, we obtain

$$\begin{array}{rcl} a(\boldsymbol{u}^2, \boldsymbol{v}) + b(\boldsymbol{v}, p^2) + b_\Gamma(\boldsymbol{v}, \lambda^1) & = & \{a(\boldsymbol{u}^2, \boldsymbol{v}^2) + b(\boldsymbol{v}^2, p^2) + b_\Gamma(\boldsymbol{v}^2, \lambda^2)\} \\ & & -\alpha\{a(\boldsymbol{u}^2, \boldsymbol{w}) + b(\boldsymbol{w}, p^2) + b_\Gamma(\boldsymbol{w}, \lambda^1)\} \\ & = & \ell(\boldsymbol{v}^2) - \alpha\{a(\boldsymbol{u}^2, \boldsymbol{w}) + b(\boldsymbol{w}, p^2) + \bar{\lambda}\} \\ & = & \ell(\boldsymbol{v}^2) - \alpha\ell(\boldsymbol{w}) = \ell(\boldsymbol{v}). \end{array}$$

The second and third equations of (5.28) are also easily verified.

Third weak formulation

We can continue with the elimination of piecewise constant pressures on each subdomain together with velocities with non-zero mean normal component on Γ. Define

$$\boldsymbol{X}^{\circ\circ} = \left\{ \boldsymbol{v} = (\boldsymbol{v}_f, \boldsymbol{v}_p) \in \boldsymbol{X}^\circ \ : \ \int_\Gamma \boldsymbol{v}_f \cdot \boldsymbol{\eta}_f = 0 \text{ and } \int_\Gamma \boldsymbol{v}_p \cdot \boldsymbol{\eta}_p = 0 \right\} \tag{5.33}$$

and

$$M^{\circ\circ} := \left\{ q = (q_f, q_p) \in M_f \times M_p \ : \ \int_{D_f} q_f = 0 \text{ and } \int_{D_p} q_p = 0 \right\}, \tag{5.34}$$

and consider the following formulation: Find $(\boldsymbol{u}^3, p^3, \lambda^3) \in \boldsymbol{X}^{\circ\circ} \times M^{\circ\circ} \times \Lambda^\circ$ such that:

$$\left\{ \begin{array}{lll} a(\boldsymbol{u}^3, \boldsymbol{v}) + b(\boldsymbol{v}, p^3) + b_\Gamma(\boldsymbol{v}, \lambda^3) & = \ell(\boldsymbol{v}) & \text{for all } \boldsymbol{v} \in \boldsymbol{X}^{\circ\circ} \\ b(\boldsymbol{u}^3, q) & = 0 & \text{for all } q \in M^{\circ\circ} \\ b_\Gamma(\boldsymbol{u}^3, \mu) & = 0 & \text{for all } \mu \in \Lambda^\circ. \end{array} \right. \tag{5.35}$$

Remark 5.13 *Let $(\boldsymbol{u}^2, p^2, \lambda^2) \in \boldsymbol{X}^\circ \times M^\circ \times \Lambda^\circ$ be a solution of (5.32). We next show that $\boldsymbol{u}^2 \in \boldsymbol{X}^{\circ\circ}$. Consider the following piecewise constant pressure $p^c = (1, -\frac{|D_f|}{|D_p|}) \in M^\circ$. From the second equation in (5.32) we have*

$$0 = \int_{D_f} \nabla\cdot\boldsymbol{u}_f^2 - \frac{|D_f|}{|D_p|}\int_{D_p} \nabla\cdot\boldsymbol{u}_p^2 = \int_\Gamma \boldsymbol{u}_f^2\cdot\boldsymbol{\eta}_f - \frac{|D_f|}{|D_p|}\int_\Gamma \boldsymbol{u}_p^2\cdot\boldsymbol{\eta}_p,$$

and since $\boldsymbol{u}^2 \in \boldsymbol{X}^\circ$, i.e.,

$$\int_\Gamma \boldsymbol{u}_f^2\cdot\boldsymbol{\eta}_f + \int_\Gamma \boldsymbol{u}_p^2\cdot\boldsymbol{\eta}_p = 0,$$

we obtain $\int_\Gamma \boldsymbol{u}_f^2\cdot\boldsymbol{\eta}_f = \int_\Gamma \boldsymbol{u}_p^2\cdot\boldsymbol{\eta}_p = 0$, therefore, $\boldsymbol{u}^2 \in \boldsymbol{X}^{\circ\circ}$. Now set

$$p^3 := \left(p_f^2 - \frac{1}{|D_f|}\int_{D_f} p_f^2, p_p^2 - \frac{1}{|D_p|}\int_{D_p} p_p^2\right) \in M^{\circ\circ}.$$

Then $b(\boldsymbol{v}, p^3) = b(\boldsymbol{v}, p^2)$ for all $\boldsymbol{v} \in \boldsymbol{X}^{\circ\circ}$ and we conclude that $(\boldsymbol{u}^2, p^3, \lambda^2)$ solves (5.35).

Now the converse. Suppose $(\boldsymbol{u}^3, p^3, \lambda^3) \in \boldsymbol{X}^{\circ\circ} \times M^{\circ\circ} \times \Lambda^\circ$ solves (5.35). Let $\boldsymbol{z} = (\boldsymbol{z}_f, \boldsymbol{z}_p) \in \boldsymbol{X}^\circ$ be any function such that $\int_\Gamma \boldsymbol{z}_f\cdot\boldsymbol{\eta}_f = -\int_\Gamma \boldsymbol{z}_p\cdot\boldsymbol{\eta}_p = \frac{|D_p|}{|D_f|+|D_p|}$. Then

$$\begin{aligned} b(\boldsymbol{z}, p^c) &= \int_{D_f} \nabla\cdot\boldsymbol{z}_f - \frac{|D_f|}{|D_p|}\int_{D_p} \nabla\cdot\boldsymbol{z}_p \\ &= \int_\Gamma \boldsymbol{z}_f\cdot\boldsymbol{\eta}_f - \frac{|D_f|}{|D_p|}\int_\Gamma \boldsymbol{z}_p\cdot\boldsymbol{\eta}_p = 1. \end{aligned}$$

Define

$$\gamma := \ell(\boldsymbol{z}) - a(\boldsymbol{u}^3, \boldsymbol{z}) - b(\boldsymbol{z}, p^3) - b_\Gamma(\boldsymbol{z}, \lambda^3)$$

and $p^2 := p^3 + \gamma p^c$ where, as before, $p^c = (1, -\frac{|D_f|}{|D_p|})$. Next we show that $(\boldsymbol{u}^3, p^2, \lambda^3)$ solves (5.32). Indeed, if $(\boldsymbol{v}, q, \mu) \in \boldsymbol{X}^\circ \times M^\circ \times \Lambda^\circ$ we can find ϵ such that $\boldsymbol{v}^3 := \boldsymbol{v} + \epsilon\boldsymbol{z} \in \boldsymbol{X}^{\circ\circ}$. Then we have

$$\begin{aligned} & a(\boldsymbol{u}^3, \boldsymbol{v}) + b(\boldsymbol{v}, p^2) + b_\Gamma(\boldsymbol{v}, \lambda^3) \\ &= \{a(\boldsymbol{u}^3, \boldsymbol{v}^3) + b(\boldsymbol{v}^3, p^3) + b_\Gamma(\boldsymbol{v}^3, \lambda^3)\} + \gamma b(\boldsymbol{v}^3, p^c) \\ &\quad -\epsilon\{a(\boldsymbol{u}^3, \boldsymbol{z}) + b(\boldsymbol{z}, p^3) + b_\Gamma(\boldsymbol{z}, \lambda^3) + \gamma b(\boldsymbol{z}, p^c)\} \\ &= \ell(\boldsymbol{v}^3) - \epsilon\ell(\boldsymbol{z}) = \ell(\boldsymbol{v}). \end{aligned} \tag{5.36}$$

Here we have used the fact that $b(\boldsymbol{v}^3, p^c) = 0$ for all $\boldsymbol{v}^3 \in \boldsymbol{X}^{\circ\circ}$. The second and third equation of (5.32) are also easily verified.

5.4.2 Inf-sup analysis

In the subsequent sections, we consider only the formulation (5.35), and we abandon the super-index 3 to avoid proliferation of indexes. In particular we establish the inf-sup associated to this formulation, see Proposition 5.17. See also Remark

5.18 for the inf-sup of the first and second weak formulations.

Define

$$\boldsymbol{V} = (\boldsymbol{V}_f, \boldsymbol{V}_p) := \{\boldsymbol{v} \in \boldsymbol{X}^{\circ\circ} \ : \ b_\Gamma(\boldsymbol{v}, \mu) = 0 \text{ for all } \ \mu \in \Lambda^\circ\} \tag{5.37}$$

with $\boldsymbol{X}^{\circ\circ}$ and Λ° defined in (5.33) and (5.31), respectively. The space $\boldsymbol{V}$ is closed because the linear map $B_\Gamma : \boldsymbol{X} \to \Lambda'$ defined by $B_\Gamma(\boldsymbol{v})\mu := b_\Gamma(\boldsymbol{v}, \mu)$ is continuous and $\boldsymbol{V} = \text{Ker } B_\Gamma$. It is easy to see that for $\boldsymbol{v} \in \boldsymbol{V}$ we have $\boldsymbol{v}_p\boldsymbol{\eta}_p = \boldsymbol{v}_f\boldsymbol{\eta}_p \in H_{00}^{1/2}(\Gamma)$. Then we can formulate problem (5.35) as:

$$\begin{cases} a(\boldsymbol{u}, \boldsymbol{v}) + b(\boldsymbol{v}, p) & = \ell(\boldsymbol{v}) \quad \text{for all } \boldsymbol{v} \in \boldsymbol{V} \\ b(\boldsymbol{u}, q) & = 0 \qquad \text{for all } q \in M^{\circ\circ}, \end{cases} \tag{5.38}$$

with $M^{\circ\circ}$ defined in (5.34). Since $\boldsymbol{u}_p \cdot \boldsymbol{\eta}_p = \boldsymbol{u}_f \cdot \boldsymbol{\eta}_f \in H_{00}^{1/2}(\Gamma)$, some regularity results on $\boldsymbol{u}_p$ and p_p can be derived which depends on smoothness and convexity properties of ∂D_p. We note however that no extra regularity is used to establish the continuous and discrete inf-sup conditions. Regularity is assumed only in the Section 5.6 where a priori error estimates are established.

Now, define

$$\boldsymbol{Z} = (\boldsymbol{Z}_f, \boldsymbol{Z}_p) := \{\boldsymbol{v} \in \boldsymbol{X}^{\circ\circ} \ : \ b(\boldsymbol{v}, q) = 0 \quad \text{for all } \ q \in M^{\circ\circ}\}\,. \tag{5.39}$$

Then we can also formulate problem (5.35) as:

$$\begin{cases} a(\boldsymbol{u}, \boldsymbol{v}) + b_\Gamma(\boldsymbol{v}, \lambda) & = \ell(\boldsymbol{v}) \quad \text{for all } \boldsymbol{v} \in \boldsymbol{Z} \\ b_\Gamma(\boldsymbol{u}, \mu) & = 0 \qquad \text{for all } \mu \in \Lambda^\circ. \end{cases} \tag{5.40}$$

Remark 5.14 *The Korn inequality implies that the bilinear form a_f defined in (5.14) is $\boldsymbol{X}_f$-*elliptic*; see Braess [2001] and Nečas [1967]. The bilinear for a_p defined in (5.17) is $\boldsymbol{H}(\text{div}^0, D_p)$-*elliptic*, here $\boldsymbol{H}(\text{div}^0, D_p)$ consists of functions in $\boldsymbol{H}(\text{div}, \text{D}_\text{p})$ with vanishing divergence, i.e., the kernel of bilinear form b_p. Then the bilinear form "a" defined in (5.21) is $\boldsymbol{X}_f \times \boldsymbol{H}(\text{div}^0, D_p)$-*elliptic.

Define

$$\boldsymbol{W}_p := \boldsymbol{X}_p \cap H^1(D_p)^2 \quad \text{and } \boldsymbol{W} = (\boldsymbol{X}_f, \boldsymbol{W}_p) \tag{5.41}$$

with

$$\|\boldsymbol{v}\|_{\boldsymbol{W}_p} := \|\boldsymbol{v}_p\|_{H^1(D_p)^2} \quad \text{and} \quad \|\boldsymbol{v}\|^2_{\boldsymbol{W}} := \|\boldsymbol{v}_f\|^2_{\boldsymbol{X}_f} + \|\boldsymbol{v}_p\|^2_{\boldsymbol{W}_p}. \tag{5.42}$$

The use of a subspace $\boldsymbol{W} \cap \boldsymbol{X}^{\circ\circ}$ with a stronger norm $\|\cdot\|_{\boldsymbol{W}} \geq \|\cdot\|_{\boldsymbol{X}}$ is a common strategy in showing continuous and discrete inf-sup conditions without assuming any regularity on the solution of the associated problem Girault and Raviart [1986] and Brezzi and Fortin [1991]; see also Lemmas 5.15 and 5.23, and Proposition 5.21.

From the usual inf-sup condition for the Stokes problem on the whole domain D and since $M^{\circ\circ} \subset M^\circ$, we easily derive the inf-sup condition associated to the formulation (5.38).

Lemma 5.15 *There is a constant $\rho > 0$ such that*

$$\inf_{\substack{q\in M^{\circ\circ}\\ q\neq 0}} \sup_{\substack{\boldsymbol{v}\in \boldsymbol{V}\\ \boldsymbol{v}\neq 0}} \frac{b(\boldsymbol{v},q)}{\|\boldsymbol{v}\|_{\boldsymbol{X}}\|q\|_M} \geq \inf_{\substack{q\in M^{\circ\circ}\\ q\neq 0}} \sup_{\substack{\boldsymbol{v}\in \boldsymbol{V}\cap \boldsymbol{W}\\ \boldsymbol{v}\neq 0}} \frac{b(\boldsymbol{v},q)}{\|\boldsymbol{v}\|_{\boldsymbol{W}}\|q\|_M} \geq \rho > 0.$$

with $\boldsymbol{W}$ and $\|\cdot\|_{\boldsymbol{W}}$ defined in (5.41) and (5.42), respectively.

Lemma 5.15, Remark 5.14 and the fact that (Ker $b\cap \boldsymbol{V}$) $\subset (\boldsymbol{X}_f \times \boldsymbol{H}(\mathrm{div}^0, D_p))$ guarantees stability of the weak formulation (5.38); see Brezzi and Fortin [1991] and Girault and Raviart [1986].

Recall that $\boldsymbol{Z} \subset \boldsymbol{H}(\mathrm{div}^0, D_f) \times \boldsymbol{H}(\mathrm{div}^0, D_p)$; see (5.39). To see that the weak formulation (5.40) is stable, next lemma shows that the inf-sup condition between spaces $\boldsymbol{Z}$ and Λ° holds (see Brezzi and Fortin [1991] and Girault and Raviart [1986]). The proof presented here follows the same ideas as Layton et al. [2002], Lemma 3.4. The main difference is that we are working with the spaces Λ° and $\boldsymbol{Z}$.

Lemma 5.16 *There is a constant $\gamma > 0$, such that*

$$\inf_{\substack{\mu\in\Lambda^\circ\\ \mu\neq 0}} \sup_{\substack{\boldsymbol{v}\in \boldsymbol{Z}\\ \boldsymbol{v}\neq 0}} \frac{b_\Gamma(\boldsymbol{v},\mu)}{\|\boldsymbol{v}\|_{\boldsymbol{X}}|\mu|_{\Lambda^\circ}} \geq \gamma > 0. \tag{5.43}$$

Proof. Fix $\mu \in \Lambda^\circ$, then $\mu \in H^{1/2}(\Gamma)$ and $\int_\Gamma \mu = 0$, in particular if $\mu \neq 0$ then μ is not a constant. From Lemma 5.10 we have that there exists $f_\Gamma \in H^{-1/2}(\Gamma)$ such that $\langle f_\Gamma, 1\rangle_\Gamma = 0$ and:

$$\frac{\langle f_\Gamma, \mu\rangle_\Gamma}{|f_\Gamma|_{H^{-1/2}(\Gamma)}} \geq \frac{1}{2}|\mu|_{H^{1/2}(\Gamma)} = \frac{1}{2}|\mu|_{\Lambda^\circ}. \tag{5.44}$$

From Remark 5.6 we introduce $f \in H^{-1/2}(\partial D_p)$ given by:

$$\langle f, \phi\rangle_{\partial D_p} := \langle f_\Gamma, \phi|_\Gamma\rangle_\Gamma \quad \text{for all } \phi \in H^{1/2}(\partial D_p) \tag{5.45}$$

with

$$|f|_{H^{-1/2}(\partial D_p)} \leq C_1 |f_\Gamma|_{H^{-1/2}(\Gamma)} \tag{5.46}$$

and zero mean on ∂D_p, i.e., $\langle f, 1\rangle_{\partial D_p} = \langle f_\Gamma, 1\rangle_\Gamma = 0$. By using the normal trace theorem, and a continuous Stokes problem (f has zero mean on ∂D_p) we can find $\boldsymbol{v}_p \in \boldsymbol{H}(\mathrm{div}, D_p)$ with $\nabla\cdot\boldsymbol{v}_p = 0$ in D_p such that:

$$\|\boldsymbol{v}_p\|_{\boldsymbol{H}(\mathrm{div},\, D_p)} \leq C|f|_{H^{-1/2}(\partial D_p)} \tag{5.47}$$

$$\boldsymbol{v}_p\cdot\boldsymbol{\eta}_p = f \text{ on } \partial D_p. \tag{5.48}$$

Observe that $\boldsymbol{v}_p \in \boldsymbol{X}_p^\circ$. Indeed, if $\phi \in H_{00}^{1/2}(\Gamma_p)$ then

$$\langle \boldsymbol{v}_p\cdot\boldsymbol{\eta}_p, \phi\rangle_{\partial D_p} = \langle f, \phi\rangle_{\partial D_p} = \langle f_\Gamma, \phi|_\Gamma\rangle_\Gamma = \langle f_\Gamma, 0\rangle_\Gamma = 0$$

and $\langle \boldsymbol{v}_p \cdot \boldsymbol{\eta}_p, 1\rangle_{\partial D_p} = \langle f_\Gamma, 1\rangle_\Gamma = 0$.

Choosing $\boldsymbol{v}_f = 0$, we have $\boldsymbol{v} := (\boldsymbol{v}_f, \boldsymbol{v}_p) \in \boldsymbol{Z}$ and:

$$\begin{aligned}
\frac{b_\Gamma(\boldsymbol{v},\mu)}{\|\boldsymbol{v}\|_{\boldsymbol{X}}} &= \frac{0 + \langle \boldsymbol{v}_p \cdot \boldsymbol{\eta}_p, E_\Gamma^{1/2}\mu\rangle_{\partial D_p}}{\|\boldsymbol{v}_p\|_{\boldsymbol{H}(\mathrm{div},\, D_p)}} && \text{by (5.25)} \\
&\geq \frac{1}{C}\frac{\langle f, E_\Gamma^{1/2}\mu\rangle_{\partial D_p}}{|f|_{H^{-1/2}(\partial D_p)}} && \text{by (5.47) and (5.48)} \\
&= \frac{1}{CC_1}\frac{\langle f_\Gamma, \mu\rangle_\Gamma}{|f_\Gamma|_{H^{-1/2}(\Gamma)}} && \text{by (5.45) and (5.46)} \\
&\geq \frac{1}{CC_1}\frac{1}{2}|\mu|_{H^{1/2}(\Gamma)} && \text{by (5.44).}
\end{aligned}$$

■

For $(q,\mu) \in M^{\circ\circ} \times \Lambda^\circ$ define $|(p,\mu)|^2_{M\times\Lambda^\circ} := \|p\|^2_M + |\mu|^2_{\Lambda^\circ}$. From Lemma 5.15 and Lemma 5.16 we can show:

Proposition 5.17 *There is a constant $\beta > 0$ such that:*

$$\inf_{\substack{(q,\mu)\in M^{\circ\circ}\times\Lambda^\circ \\ (q,\mu)\neq(0,0)}} \sup_{\substack{\boldsymbol{v}\in\boldsymbol{X}^{\circ\circ} \\ \boldsymbol{v}\neq 0}} \frac{b(\boldsymbol{v},q) + b_\Gamma(\boldsymbol{v},\mu)}{\|\boldsymbol{v}\|_{\boldsymbol{X}}|(q,\mu)|_{M\times\Lambda^\circ}} \geq \beta > 0. \tag{5.49}$$

Proof. Given $(q,\mu) \in M^{\circ\circ}\times\Lambda^\circ$, if $q \neq 0$, from Lemma 5.15 there exists $\hat{\boldsymbol{v}} \in \boldsymbol{V}$ such that

$$\frac{b(\hat{\boldsymbol{v}},q)}{\|\hat{\boldsymbol{v}}\|_{\boldsymbol{X}}} \geq \rho\|q\|_M > 0,$$

where ρ independent of q. If $\mu \neq 0$, from Lemma 5.16 there exists $\boldsymbol{z} \in \boldsymbol{Z}$ such that

$$\frac{b_\Gamma(\boldsymbol{z},\mu)}{\|\boldsymbol{z}\|_{\boldsymbol{X}}} \geq \gamma|\mu|_{\Lambda^\circ} > 0,$$

where γ independent of μ.

Observe that, if $q \neq 0$

$$\sup_{\substack{\boldsymbol{v}\in\boldsymbol{X}^{\circ\circ} \\ \boldsymbol{v}\neq 0}} \frac{b(\boldsymbol{v},q) + b_\Gamma(\boldsymbol{v},\mu)}{\|\boldsymbol{v}\|_{\boldsymbol{X}}} \geq \frac{b(\hat{\boldsymbol{v}},q) + b_\Gamma(\hat{\boldsymbol{v}},\mu)}{\|\hat{\boldsymbol{v}}\|_{\boldsymbol{X}}} = \frac{b(\hat{\boldsymbol{v}},q) + 0}{\|\hat{\boldsymbol{v}}\|_{\boldsymbol{X}}} \geq \rho\|q\|_M.$$

Analogously, if $\mu \neq 0$,

$$\sup_{\substack{\boldsymbol{v}\in\boldsymbol{X}^{\circ\circ} \\ \boldsymbol{v}\neq 0}} \frac{b(\boldsymbol{v},q) + b_\Gamma(\boldsymbol{v},\mu)}{\|\boldsymbol{v}\|_{\boldsymbol{X}}} \geq \frac{0 + b_\Gamma(\boldsymbol{z},\mu)}{\|\boldsymbol{z}\|_{\boldsymbol{X}}} \geq \gamma|\mu|_{\Lambda^\circ},$$

then

$$\sup_{\substack{\boldsymbol{v}\in\boldsymbol{X}^{\circ\circ}\\ \boldsymbol{v}\neq 0}} \frac{b(\boldsymbol{v},q)+b_\Gamma(\boldsymbol{v},\mu)}{\|\boldsymbol{v}\|_{\boldsymbol{X}}} \geq \tfrac{\min\{\rho,\gamma\}}{2}\Big(\|q\|_M+|\mu|_{\Lambda^\circ}\Big) \geq \tfrac{\min\{\rho,\gamma\}}{2}|(q,\mu)|_{M\times\Lambda^\circ}.$$

■

Proposition 5.17 permit us to formulate problem (5.35) as

$$\begin{cases} a(\boldsymbol{u},\boldsymbol{v})+c(\boldsymbol{v},(p,\lambda)) &= \ell(\boldsymbol{v}) \quad \text{for all } \boldsymbol{v}\in\boldsymbol{X}^{\circ\circ}\\ c(\boldsymbol{u},(q,\mu)) &= 0 \quad \text{for all } (q,\mu)\in M^{\circ\circ}\times\Lambda^\circ, \end{cases} \tag{5.50}$$

where $c(\boldsymbol{v},(q,\mu)) := b(\boldsymbol{v},q)+b_\Gamma(\boldsymbol{v},\mu)$. Then (5.49) in Proposition 5.17 can be written as: there exists $\beta=\frac{\min\{\rho,\gamma\}}{2}>0$ such that

$$\inf_{\substack{(q,\mu)\in M^{\circ\circ}\times\Lambda^\circ\\ (q,\mu)\neq 0}} \sup_{\substack{\boldsymbol{v}\in\boldsymbol{X}^{\circ\circ}\\ \boldsymbol{v}\neq 0}} \frac{c(\boldsymbol{v},(q,\mu))}{\|\boldsymbol{v}\|_{\boldsymbol{X}}|(q,\mu)|_{M\times\Lambda^\circ}} \geq \beta>0. \tag{5.51}$$

This inf-sup condition, together with the fact that a is $\boldsymbol{X}_f\times\boldsymbol{H}(\mathrm{div}^0,D_p)$-elliptic and a and c are bounded, (according to the abstract saddle point theory) guarantees the well-posedness of the problem (5.50) or (5.35); see Brezzi and Fortin [1991] and Girault and Raviart [1986].

Remark 5.18 *We now obtain the inf-sup condition for the weak formulation (5.32). Consider $\boldsymbol{z}$ introduced in Remark 5.13. Note that in Remark 5.13 we only have required $\boldsymbol{z}\in\boldsymbol{X}^\circ$ and*

$$\int_\Gamma \boldsymbol{z}_f\cdot\boldsymbol{\eta}_f = -\int_\Gamma \boldsymbol{z}_p\cdot\boldsymbol{\eta}_p = \frac{|D_p|}{|D_f|+|D_p|}.$$

Now we also require the divergence of $\boldsymbol{z}$ to be constant on each subdomain and also that $\boldsymbol{z}_f\cdot\boldsymbol{\eta}_f=-\boldsymbol{z}_p\cdot\boldsymbol{\eta}_p$. For instance, we can solve a Stokes problem with constant divergence on the fluid side and a Darcy problem with the corresponding boundary data and constant divergence on the porous side, with divergences values satisfying the subdomain compatibility conditions. Then we have

$$b(\boldsymbol{z},q^3)=0 \quad \textit{for all } q^3\in M^{\circ\circ}, \quad \textit{and} \quad b_\Gamma(\boldsymbol{z},\mu^2)=0 \quad \textit{for all } \mu^2\in\Lambda^\circ. \tag{5.52}$$

We now show show that the inf-sup condition for the weak formulation (5.32) holds. The spaces involved are $\boldsymbol{X}^\circ$ for velocities, and M° and Λ° for pressures and Lagrange multipliers, respectively; see (5.30), (5.24) and (5.31). Take $q^2\in M^\circ$ and $\mu^2\in\Lambda^\circ$ and let $p^c=(1,-\frac{|D_f|}{|D_p|})\in M^\circ$ as in Remark 5.13. We can write $q^2=q^3+\overline{q}p^c$ where $q^3\in M^{\circ\circ}$. Note that

$$\|q^2\|_M \leq \|q^3\|_M+|\bar{q}|\|p^c\|_M.$$

From Proposition 5.25 and a Poincaré inequality, there exists $\boldsymbol{v}^3\in\boldsymbol{X}^{\circ\circ}$ such that

$$b(\boldsymbol{v}^3,q^3)+b_\Gamma(\boldsymbol{v}^3,\mu^2)\geq\tilde{\beta}\|\boldsymbol{v}^3\|_{\boldsymbol{X}}\left\{\|q^3\|_M+\|\mu^2\|_{\Lambda^\circ}\right\},$$

where $\tilde{\beta}$ is a positive constant independent of $\boldsymbol{v}^3$. If $\bar{q} \neq 0$ let

$$\boldsymbol{v}^2 = \boldsymbol{v}^3 + \tilde{\beta}\|\boldsymbol{v}^3\|_{\boldsymbol{X}}\|p^c\|_M \frac{\bar{q}}{|\bar{q}|}\boldsymbol{z} = \boldsymbol{v}^3 + r\boldsymbol{z}, \quad \text{with } r = \tilde{\beta}\|\boldsymbol{v}^3\|_{\boldsymbol{X}}\|p^c\|_M \frac{\bar{q}}{|\bar{q}|}.$$

Observe that $\|\boldsymbol{v}^2\|_{\boldsymbol{X}} \leq (1 + \tilde{\beta}\|\boldsymbol{z}\|_{\boldsymbol{X}}\|p^c\|_M)\|\boldsymbol{v}^3\|_{\boldsymbol{X}}$. We have

$$\begin{aligned} b(\boldsymbol{v}^2, q^2) &= \{b(\boldsymbol{v}^3, q^3) + \bar{q}b(\boldsymbol{v}^3, p^c)\} + r\{b(\boldsymbol{z}, q^3) + \bar{q}b(\boldsymbol{z}, p^c)\} \\ &= \{b(\boldsymbol{v}^3, q^3) + 0\} + r\{0 + \bar{q}\} \qquad \textit{(see (5.52))} \\ &= b(\boldsymbol{v}^3, q^3) + |\bar{q}|\|\boldsymbol{v}^3\|_{\boldsymbol{X}}\|p^c\|_M \end{aligned}$$

and

$$b_\Gamma(\boldsymbol{v}^2, \mu^2) = b_\Gamma(\boldsymbol{v}^3, \mu^2) + rb_\Gamma(\boldsymbol{z}, \mu^2) = b(\boldsymbol{v}^3, \mu^2) + 0.$$

Then

$$\begin{aligned} b(\boldsymbol{v}^2, q^2) + b_\Gamma(\boldsymbol{v}^2, \mu^2) &= b(\boldsymbol{v}^3, q^3) + b_\Gamma(\boldsymbol{v}^3, \mu^2) + |\bar{q}|\|\boldsymbol{v}^3\|_{\boldsymbol{X}}\|p^c\|_M \\ &\geq \tilde{\beta}\|\boldsymbol{v}^3\|_{\boldsymbol{X}}\left\{\|q^3\|_M + \|\mu^2\|_{\Lambda^\circ}\right\} \\ &\quad + |\bar{q}|\|\boldsymbol{v}^3\|_{\boldsymbol{X}}\|p^c\|_M \\ &= \tilde{\beta}\|\boldsymbol{v}^3\|_{\boldsymbol{X}}\left\{\|q^3\|_M + |\bar{q}|\|p^c\|_M + \|\mu^2\|_\Lambda\right\} \\ &\geq \tilde{\beta}\|\boldsymbol{v}^3\|_{\boldsymbol{X}}\left\{\|q^2\|_M + \|\mu^2\|_{\Lambda^\circ}\right\} \\ &\geq \frac{\tilde{\beta}}{1 + \tilde{\beta}\|\boldsymbol{z}\|_{\boldsymbol{X}}\|p^c\|_M}\|\boldsymbol{v}^2\|_{\boldsymbol{X}}\left\{\|q^2\|_M + \|\mu^2\|_{\Lambda^\circ}\right\}. \end{aligned}$$

This gives the inf-sup condition for weak formulation (5.32).

We now obtain the inf-sup condition for the weak formulation (5.28). The spaces are $\boldsymbol{X}$ for velocities, M° for pressures, and Λ defined in (5.26) for Lagrange multipliers. Consider $\boldsymbol{w}$ introduced in Remark 5.12. Note that in Remark 5.12 we have required $\boldsymbol{w} = (\boldsymbol{0}, \boldsymbol{w}_p)$ with

$$\boldsymbol{w}_p \cdot \boldsymbol{\eta}_p = \frac{1}{|\Gamma|} \text{ on } \Gamma \text{ and } \boldsymbol{w}_p \cdot \boldsymbol{\eta}_p = 0 \text{ on } \Gamma_p.$$

Now we also require that the divergence of $\boldsymbol{w}$ be a constant on D_p. Given $\mu^1 \in \Lambda$ and $q^1 \in M^\circ$, we write $\mu^1 = \mu^2 + \bar{\mu}$ where $\int_\Gamma \mu^2 = 0$, i.e., $\mu^2 \in \Lambda^\circ$. From the inf-sup for weak formulation (5.32) deduced above, we can find $\boldsymbol{v}^2 \in \boldsymbol{X}^\circ$ such that

$$b(\boldsymbol{v}^2, q^1) + b_\Gamma(\boldsymbol{v}^2, \mu^2) \geq \hat{\beta}\|v^2\|_{\boldsymbol{X}}\left\{\|q^1\|_M + \|\mu^2\|_{\Lambda^\circ}\right\}$$

If $\bar{\mu} \neq 0$ define $\boldsymbol{v}^1 = \boldsymbol{v}^2 + \hat{\beta}\|\boldsymbol{v}^2\|_{\boldsymbol{X}}|\Gamma|^{\frac{1}{2}}\frac{\bar{\mu}}{|\bar{\mu}|}\boldsymbol{w}$. Note that

$$\|\boldsymbol{v}^1\|_{\boldsymbol{X}} \leq (1 + \hat{\beta}\|\boldsymbol{w}\|_{\boldsymbol{X}}|\Gamma|^{\frac{1}{2}})\|\boldsymbol{v}^2\|_{\boldsymbol{X}} \quad \text{and} \quad \|\mu^1\|_\Lambda \asymp \|\mu^2\|_\Lambda + |\bar{\mu}||\Gamma|^{\frac{1}{2}}.$$

And we proceed as before to obtain the inf-sup condition for the weak formulation (5.28).

5.5 Finite element approximation

In Section 5.3 the problem for the coupling fluid flow with porous media flow in its continuous form was presented, while in Section 5.4 it was presented its variational formulation and well-posedness. Now a two dimensional non-matching grid finite element approximation is discussed. We choose the $P2\backslash P1$ triangular Taylor Hood finite elements for approximating the free fluid side velocity and pressure, while we use the lowest order triangular Raviart-Thomas finite element to approximate the filtration velocity and the porous pressure; see Section 5.5.1 below. In Section 5.5.2 a discrete non-conforming Lagrange multiplier space to couple the Taylor-Hood and Raviart-Thomas spaces is introduced. It is important for the analysis to choose the Stokes side as the mortar side, i.e., to place the discrete Lagrange multiplier on the Darcy side. In this case the discrete map from mortar to non-mortar side is continuous in $L^2(\Gamma)$ norm. Extensions of the results to other than Stokes and Darcy finite element spaces are straightforward; just take the Lagrange multiplier spaces that are used to hybridize mixed finite elements of the Darcy equation; see Brezzi and Fortin [1991]. We establish the discrete inf-sup conditions related to the weak formulation (5.38), (5.40) and (5.35). The extension of the results to the three dimensional case is also straightforward.

5.5.1 Discretization

From now on we assume that D has polygonal boundary. Let $\mathcal{T}_{h_j}$ be a triangulation of D_j, $j = f, p$. We do not assume that they match at the polyhedral interface Γ. We choose $P2\backslash P1$ triangular Taylor-Hood finite elements; see Brenner and Scott [1994], Brezzi and Fortin [1991] and Girault and Raviart [1986]. Define

$$\boldsymbol{X}_{h_f} := \left\{ \boldsymbol{v}_f \in \boldsymbol{X}_f \ : \ \begin{array}{l} \boldsymbol{v}_{fK} = \hat{\boldsymbol{v}}_{fK} \circ F_K^{-1} \text{ on } K \text{ and} \\ \hat{\boldsymbol{v}}_K \in P_2(\hat{K})^2 \end{array} \right\} \cap C^0(\overline{D}_f)^2, \tag{5.53}$$

and

$$\boldsymbol{X}_{h_f}^{\circ} := \{ \boldsymbol{v}_{h_f} \in \boldsymbol{X}_{h_f} : \int_{\Gamma} \boldsymbol{v}_{h_f} \cdot \boldsymbol{\eta}_f = 0 \}, \tag{5.54}$$

where $\boldsymbol{v}_{fK} := \boldsymbol{v}_f|_K$. We also define

$$M_{h_f} := \left\{ q_f \in M_f \ : \ \begin{array}{l} q_{fK} = \hat{q}_{fK} \circ F_K^{-1} \text{ on } K \text{ and} \\ \hat{q}_{fK} \in P_1(\hat{K}) \end{array} \right\} \cap C^0(\overline{D}_f),$$

$$M_{h_f}^{\circ} := \{ q_f \in M_{h_f} : \int_{D_f} q_f = 0 \}. \tag{5.55}$$

We have the following result.

Lemma 5.19 (Taylor-Hood elements) *Suppose that $\mathcal{T}_{h_f}$ is non-degenerate and has no triangle with two edges on ∂D_f. Then, there exists a bounded linear operator $\boldsymbol{I}_{h_f}^{TH} : \boldsymbol{X}_f \to \boldsymbol{X}_{h_f}$ such that*

$$b_f(\boldsymbol{v}_f - \boldsymbol{I}_{h_f}^{TH} \boldsymbol{v}_f, p_{h_f}) = 0 \quad \textit{for all } p_{h_f} \in M_{h_f}^{\circ}$$

and $\|\boldsymbol{I}_{h_f}^{TH}\boldsymbol{v}_f\|_{\boldsymbol{X}_f} \preceq \|\boldsymbol{v}_f\|_{\boldsymbol{X}_f}$*, with constant independent of* h_f*. In addition we have:*

$$\|\boldsymbol{v}_f - \boldsymbol{I}_{h_f}^{TH}\boldsymbol{v}_f\|_{L^2(D_f)^2} \preceq h_f^s |\boldsymbol{v}_f|_{H^s(D_f)^2} \quad s = 1, 2. \tag{5.56}$$

$$|\boldsymbol{v}_f - \boldsymbol{I}_{h_f}^{TH}\boldsymbol{v}_f|_{H^1(D_f)^2} \preceq h_f |\boldsymbol{v}_f|_{H^2(D_f)^2} \tag{5.57}$$

$$\int_\Gamma \boldsymbol{I}_{h_f}^{TH}\boldsymbol{v}_f \cdot \boldsymbol{\eta}_f = \int_\Gamma \boldsymbol{v}_f \cdot \boldsymbol{\eta}_f \quad (\textit{which implies } \ \boldsymbol{I}_{h_f}^{TH} : \boldsymbol{X}_f^\circ \to \boldsymbol{X}_{h_f}^\circ) \tag{5.58}$$

$$|\boldsymbol{I}_{h_f}^{TH}\boldsymbol{v}_f|_{H^{1/2}(\Gamma)^2} \preceq |\boldsymbol{v}_f|_{H^{1/2}(\Gamma)^2}. \tag{5.59}$$

A constructive and apparently new proof using Fortin interpolation is given in Appendix 5.10, or see Brenner and Scott [1994], Brezzi and Fortin [1991] and Girault and Raviart [1986].

A direct consequence of Fortin's criterion and the previous lemma is that, if $\mathcal{T}_{h_f}$ is non-degenerate and has no triangle with two edges on ∂D_f, then $(\boldsymbol{X}_{h_f}^\circ, M_{h_f}^\circ)$ satisfies the inf-sup condition; see (5.54) and (5.55).

For the porous region we are going to use the lowest order Raviart-Thomas finite elements based on triangles. In general the Raviart-Thomas elements in a cell are defined by (see Braess [2001], Brezzi and Fortin [1991] and Girault and Raviart [1986]):

$$RT_k(K) := (P_k(K))^n + P_k(K)\boldsymbol{x},$$

and if $\boldsymbol{v} \in RT_k(K)$ then $\nabla \cdot \boldsymbol{v} \in P_k(K)$ and $\boldsymbol{v} \cdot \boldsymbol{\eta}|_{e_i} \in P_k(e_i)$, for all edge e_i. Then we choose:

$$\boldsymbol{X}_{h_p}^\circ := \left\{ \boldsymbol{v}_p \in \boldsymbol{X}_p \ : \ \boldsymbol{v}_p|_K \in RT_0(K) \text{ and } \int_\Gamma \boldsymbol{v}_p \cdot \boldsymbol{\eta}_p = 0 \right\}, \tag{5.60}$$

and

$$M_{h_p}^\circ := \left\{ p_p \in M_p \ : \ p_p|_K \in P_0(K) \text{ with } \int_{D_p} p_p = 0 \right\}. \tag{5.61}$$

Velocities of lowest order Raviart-Thomas finite elements, $RT_0(K)$, $K \in \mathcal{T}_{h_p}$, are then of the form:

$$\boldsymbol{v}_p(x_1, x_2) = \begin{pmatrix} a \\ b \end{pmatrix} + c \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

We have the following result; see also Braess [2001] and Brezzi and Fortin [1991]. Recall the definition of $\boldsymbol{W}_p$ in (5.41).

Lemma 5.20 (Raviart-Thomas elements) *For* $K \in \mathcal{T}_{h_p}$ *, define the operator* $\boldsymbol{I}_{h_p,K}^{RT} : \boldsymbol{H}(\mathrm{div}, K) \cap H^1(K)^2 \to RT_0(K)$ *by*

$$\boldsymbol{I}_{h_p,K}^{RT}\boldsymbol{v}_p \cdot \boldsymbol{\eta}_p|_e = \frac{1}{|e|}\int_e \boldsymbol{v}_p \cdot \boldsymbol{\eta}_p \tag{5.62}$$

and define $\boldsymbol{I}_{h_p}^{RT} : \boldsymbol{W}_p \to RT_0$ *locally by:* $\boldsymbol{I}_{h_p}^{RT}\boldsymbol{v}_p|_K = \boldsymbol{I}_{h_p,K}^{RT}\boldsymbol{v}_p$*. Then*

$$\int_{D_p} \nabla \cdot (\boldsymbol{v}_p - \boldsymbol{I}_{h_p}^{RT}\boldsymbol{v}_p) q_{h_p} = 0 \quad \textit{for all } q_{h_p} \in M_{h_p}^\circ \tag{5.63}$$

and $\|\boldsymbol{I}_{h_p}^{RT}\boldsymbol{v}_p\|_{\boldsymbol{H}(div,\,D_p)} \preceq \|\boldsymbol{v}_p\|_{\boldsymbol{W}_p}$ *with* $\|\cdot\|_{\boldsymbol{W}_p}$ *defined in (5.42). The property (5.62) implies that* $\boldsymbol{I}_{h_p}^{RT}: \boldsymbol{X}_p^{\circ} \cap \boldsymbol{W}_p \to \boldsymbol{X}_{h_p}^{\circ}$. *In addition, with the property (5.63) we have* $\boldsymbol{I}_{h_p}^{RT}: \boldsymbol{Z}_p \cap \boldsymbol{W}_p \to \boldsymbol{Z}_{h_p}^{\circ}$. *Moreover, if* $\boldsymbol{v}_p \in H^1(D_p)^2$ *then*

$$\|\boldsymbol{v}_p - \boldsymbol{I}_{h_p}^{RT}\boldsymbol{v}_p\|_{L^2(D_p)^2} \preceq h_p|\boldsymbol{v}_p|_{H^1(D_p)^2}, \tag{5.64}$$

and

$$\|\nabla\cdot(\boldsymbol{v}_p - \boldsymbol{I}_{h_p}^{RT}\boldsymbol{v}_p)\|_{L^2(D_p)} \preceq h_p|\nabla\cdot\boldsymbol{v}_p|_{H^1(D_p)}.$$

By using Fortin's idea we can establish the inf-sup condition for the spaces $(X_{h_p}^{\circ}, M_{h_p}^{\circ})$ defined in (5.60) and (5.61), respectively.

5.5.2 Discrete inf-sup condition

Let $\Gamma \cap \mathcal{T}_{h_p}$ be the trace on Γ of the porous side triangulation. We consider piecewise constant Lagrange multiplier space:

$$\Lambda_{h_p}^{\circ} = \left\{ \mu_{h_p} \in L^2(\Gamma) : \begin{array}{l} \mu_{h_p}|_{e_p} \text{ is constant on each edge} \\ e_p \in \Gamma \cap \mathcal{T}_p^{h_p} \text{ and } \int_\Gamma \mu = 0 \end{array} \right\}.$$

We note that this choice leads to non-conforming finite elements associated to Λ° since piecewise constant functions do not belong to $H^{1/2}(\Gamma)$; see (5.31).

We also introduce for later use

$$\Lambda_{h_p} = \left\{ \mu_{h_p} \in L^2(\Gamma) : \quad \mu_{h_p}|_{e_p} \text{is constant on each edge } e_p \in \Gamma \cap \mathcal{T}_p^{h_p} \right\}. \tag{5.65}$$

Define $h = (h_f, h_p)$,

$$\boldsymbol{X}_h^{\circ\circ} := \boldsymbol{X}_{h_f}^{\circ} \times \boldsymbol{X}_{h_p}^{\circ} \subset \boldsymbol{X}^{\circ\circ}, \qquad M_h^{\circ\circ} := M_{h_f}^{\circ} \times M_{h_p}^{\circ} \subset M^{\circ\circ} \tag{5.66}$$

and

$$\boldsymbol{V}_h \;=\; (\boldsymbol{V}_{h_f}, \boldsymbol{V}_{h_p}) := \left\{ \boldsymbol{v}_h \in \boldsymbol{X}_h^{\circ\circ} : \; (\llbracket \boldsymbol{v}_h \rrbracket \cdot \boldsymbol{\eta}_f, \mu_{h_p})_\Gamma = 0 \text{ for all } \mu_{h_p} \in \Lambda_{h_p}^{\circ} \right\},$$

where $\llbracket \boldsymbol{v}_h \rrbracket := \boldsymbol{v}_{h_f} - \boldsymbol{v}_{h_p}$ on Γ for all $\boldsymbol{v}_h \in \boldsymbol{X}_h^{\circ}$. Also define

$$\boldsymbol{Z}_h = (\boldsymbol{Z}_{h_f}, \boldsymbol{Z}_{h_p}) := \{ \boldsymbol{v}_h \in \boldsymbol{X}_h^{\circ\circ} : \; b(\boldsymbol{v}_h, q_h) = 0 \text{ for all } q_h \in M_h^{\circ\circ} \}. \tag{5.67}$$

For $z_{h_p} \in \boldsymbol{X}_{h_p}^{\circ}\cdot\boldsymbol{\eta}_p|_\Gamma = \Lambda_{h_p}^{\circ}$, (i.e., z_{h_p} piecewise constant on Γ relatively to $\mathcal{T}_{h_p}$ and with zero mean on Γ) define $\boldsymbol{E}_{h_p} z_{h_p} \in \boldsymbol{X}_{h_p}^{\circ}$ as the discrete velocity solution of the problem

$$\left\{ \begin{array}{lll} a_p(\boldsymbol{E}_{h_p} z_{h_p}, \boldsymbol{v}_{h_p}) + b_p(\boldsymbol{v}_{h_p}, \hat{p}_{h_p}) & = 0 & \text{for all } \boldsymbol{v}_{h_p} \in \boldsymbol{X}_{h_p}^{\circ} \text{ such} \\ & & \text{that } \boldsymbol{v}_{h_p}\cdot\boldsymbol{\eta}_p = 0 \text{ on } \Gamma_p, \\ b_p(\boldsymbol{E}_{h_p} z_{h_p}, q_{h_p}) & = 0 & \text{for all } q_{h_p} \in M_{h_p}^{\circ} \\ \boldsymbol{E}_{h_p} z_{h_p}\cdot\boldsymbol{\eta}_p & = z_{h_p} & \text{on } \Gamma. \end{array} \right. \tag{5.68}$$

We note that a discrete divergence free Raviart-Thomas vector field is also a divergence free vector field. Therefore, using Mathew [1993] we have

$$\|\boldsymbol{E}_{h_p} z_{h_p}\|^2_{\boldsymbol{L}^2(D)} = \|\boldsymbol{E}_{h_p} z_{h_p}\|_{\boldsymbol{X}_p} \asymp |z_{h_p}|_{H^{-1/2}(\Gamma)}. \tag{5.69}$$

We have the following result.

Proposition 5.21 *Suppose that $\mathcal{T}_{h_f}$ is non-degenerate and has no triangle with two edges on ∂D_f and consider $\boldsymbol{W}$ defined in (5.41). There exists a linear continuous operator*

$$\boldsymbol{\Pi}_h : (\boldsymbol{V} \cap \boldsymbol{W}) \to \boldsymbol{V}_h$$

such that

$$b(\boldsymbol{\Pi}_h \boldsymbol{v} - \boldsymbol{v}, q_h) = 0 \ \text{ for all } q_h \in M_h^{\circ\circ}, \tag{5.70}$$

and

$$\|\boldsymbol{\Pi}_h \boldsymbol{v}\|_{\boldsymbol{X}} \preceq \|\boldsymbol{v}_p\|_{\boldsymbol{W}_p} \leq \|\boldsymbol{v}\|_{\boldsymbol{W}}. \tag{5.71}$$

with $\|\cdot\|_{\boldsymbol{W}}$ defined in (5.42).

Proof. Write $\boldsymbol{\Pi}_h(\boldsymbol{v}) = (\boldsymbol{\Pi}_{h_f}\boldsymbol{v}, \boldsymbol{\Pi}_{h_p}\boldsymbol{v})$ where $\boldsymbol{\Pi}_{h_f}\boldsymbol{v} := \boldsymbol{I}^{TH}_{h_f}\boldsymbol{v}_f$ and

$$\boldsymbol{\Pi}_{h_p}\boldsymbol{v} := \boldsymbol{I}^{RT}_{h_p}\boldsymbol{v}_p + \boldsymbol{E}_{h_p}\left(Q_{h_p}(\boldsymbol{I}^{TH}_{h_f}\boldsymbol{v}_f\cdot\boldsymbol{\eta}_p) - \boldsymbol{I}^{RT}_{h_p}\boldsymbol{v}_p\cdot\boldsymbol{\eta}_p\right),$$

where Q_{h_p} denotes the L^2-projection on Λ^{h_p}, i.e., on the space of piecewise constant functions on Γ.

Let $\mu_{h_p} \in \Lambda^{\circ}_{h_p}$. We have

$$\begin{aligned}
([\![\boldsymbol{\Pi}_h\boldsymbol{v}]\!]\cdot\boldsymbol{\eta}_p, \mu_{h_p})_\Gamma &= (\boldsymbol{\Pi}_{h_p}\boldsymbol{v}\cdot\boldsymbol{\eta}_p, \mu_{h_p})_\Gamma - (\boldsymbol{\Pi}_{h_f}\boldsymbol{v}\cdot\boldsymbol{\eta}_p, \mu_{h_p})_\Gamma \\
&= (Q_{h_p}(\boldsymbol{I}^{TH}_{h_f}\boldsymbol{v}_f\cdot\boldsymbol{\eta}_p), \mu_{h_p})_\Gamma - (\boldsymbol{I}^{TH}_{h_f}\boldsymbol{v}_f\cdot\boldsymbol{\eta}_p, \mu_{h_p})_\Gamma \\
&= 0 \qquad \text{by definition of } Q_{h_p},
\end{aligned}$$

and then obtain $\boldsymbol{\Pi}_h\boldsymbol{v} \in \boldsymbol{V}_h$.

Now we show (5.71). Observe that

$$\begin{aligned}
\|\boldsymbol{\Pi}_h\boldsymbol{v}\|_{\boldsymbol{X}} &\leq \|\boldsymbol{\Pi}_{h_f}\boldsymbol{v}_f\|_{\boldsymbol{X}_f} + \|\boldsymbol{\Pi}_{h_p}\boldsymbol{v}_p\|_{\boldsymbol{X}_p} \\
&\leq \|\boldsymbol{I}^{TH}_{h_f}\boldsymbol{v}_f\|_{\boldsymbol{X}_f} + \|\boldsymbol{I}^{RT}_{h_p}\boldsymbol{v}_p\|_{\boldsymbol{X}_p} \\
&\quad + \|\boldsymbol{E}_{h_p}\left(Q_{h_p}(\boldsymbol{I}^{TH}_{h_f}\boldsymbol{v}_f\cdot\boldsymbol{\eta}_p) - \boldsymbol{I}^{RT}_{h_p}\boldsymbol{v}_p\cdot\boldsymbol{\eta}_p\right)\|_{\boldsymbol{X}_p}.
\end{aligned}$$

The bound (5.71) follows from the boundedness of $\boldsymbol{I}^{TH}_{h_f}$ (Lemma 5.19), $\boldsymbol{I}^{RT}_{h_p}$ (Lemma 5.20), $\boldsymbol{E}_{h_p}$ (Equation (5.69)), and from the following two bounds:

1. From the boundedness of $\boldsymbol{I}^{TH}_{h_f}$ and Q_{h_p}, and from a trace theorem, we have

$$\begin{aligned}
|Q_{h_p}(\boldsymbol{I}^{TH}_{h_f}\boldsymbol{v}_f\cdot\boldsymbol{\eta}_p)|_{H^{-1/2}(\Gamma)} &\preceq \|Q_{h_p}(\boldsymbol{I}^{TH}_{h_f}\boldsymbol{v}_f\cdot\boldsymbol{\eta}_p)\|_{L^2(\Gamma)} \\
&\leq \|\boldsymbol{I}^{TH}_{h_f}\boldsymbol{v}_f\cdot\boldsymbol{\eta}_p\|_{L^2(\Gamma)} \\
&\preceq |\boldsymbol{v}_f\cdot\boldsymbol{\eta}_p|_{H^{1/2}(\Gamma)} \\
&= |\boldsymbol{v}_p\cdot\boldsymbol{\eta}_p|_{H^{1/2}(\Gamma)} \\
&\preceq \|\boldsymbol{v}_p\|_{\boldsymbol{W}_p} \\
&\leq \|\boldsymbol{v}\|_{\boldsymbol{W}}.
\end{aligned}$$

2. From the normal trace theorem and the boundedness of $\boldsymbol{I}_{h_p}^{RT}$ we have

$$|\boldsymbol{I}_{h_p}^{RT}\boldsymbol{v}_p\cdot\boldsymbol{\eta}_p|_{H^{-1/2}(\Gamma)} \preceq \|\boldsymbol{I}_{h_p}^{RT}\boldsymbol{v}_p\|_{\boldsymbol{X}_p} \preceq \|\boldsymbol{v}_p\|_{\boldsymbol{W}_p} \leq \|\boldsymbol{v}\|_{\boldsymbol{W}}.$$

■

Remark 5.22 *We note that when the mesh $\mathcal{T}_{h_f}(D_f)$ restricted to Γ is a refinement of the mesh $\mathcal{T}_{h_p}(D_p)$ restricted to Γ, then by using (5.109) in Appendix B we have $Q_{h_p}(\boldsymbol{I}_{h_f}^{TH}\boldsymbol{v}_f\cdot\boldsymbol{\eta}_p) = Q_{h_p}\boldsymbol{v}_f\cdot\boldsymbol{\eta}_p$. Also from (5.62) we have $\boldsymbol{I}_{h_p}^{RT}\boldsymbol{v}_p\cdot\boldsymbol{\eta}_p = Q_{h_p}\boldsymbol{v}_p\cdot\boldsymbol{\eta}_p$. Hence using that $\boldsymbol{v}_p\cdot\boldsymbol{\eta}_p = \boldsymbol{v}_f\cdot\boldsymbol{\eta}_f \in H_{00}^{1/2}(\Gamma)$ we obtain*

$$\boldsymbol{E}_{h_p}\left(Q_{h_p}(\boldsymbol{v}_f\cdot\boldsymbol{\eta}_p) - \boldsymbol{I}_{h_p}^{RT}\boldsymbol{v}_p\cdot\boldsymbol{\eta}_p\right) = 0. \tag{5.72}$$

In the following result we establish the discrete inf-sup condition using Fortin's Lemma.

Lemma 5.23 *Suppose that $\mathcal{T}_{h_f}$ is non-degenerate and has no triangle with two edges on ∂D_f. Consider $\boldsymbol{V}$ and $M_h^{\circ\circ}$ defined in (5.67) and (5.66), respectively. Then $(\boldsymbol{V}_h, M_h^{\circ\circ})$ satisfies the discrete* inf-sup *condition, i.e., there is a constant $\tilde{\rho} > 0$ independent of h, such that:*

$$\inf_{\substack{q_h\in M_h^{\circ\circ}\\ q_h\neq 0}} \sup_{\substack{\boldsymbol{v}_h\in\boldsymbol{V}_h\\ \boldsymbol{v}_h\neq 0}} \frac{b(\boldsymbol{v}_h, q_h)}{\|\boldsymbol{v}_h\|_{\boldsymbol{X}}\|q_h\|_M} \geq \tilde{\rho} > 0.$$

Proof. Take $q_h \in M_h^{\circ\circ}$. From Lemma 5.15 we can find $\boldsymbol{v} \neq 0 \in \boldsymbol{V}\cap\boldsymbol{W}$ such that

$$\frac{b(\boldsymbol{v}, q_h)}{\|\boldsymbol{v}\|_{\boldsymbol{W}}} \geq \rho\|q_h\|_M.$$

Then from Proposition 5.21 we have

$$\rho\|q_h\|_M \leq \frac{b(\boldsymbol{v}, q_h)}{\|\boldsymbol{v}\|_{\boldsymbol{W}}} = \frac{b(\boldsymbol{\Pi}_h\boldsymbol{v}, q_h)}{\|\boldsymbol{v}\|_{\boldsymbol{W}}} \leq \frac{b(\boldsymbol{\Pi}_h\boldsymbol{v}, q_h)}{\frac{1}{C}\|\boldsymbol{\Pi}_h\boldsymbol{v}\|_{\boldsymbol{X}}},$$

where C is the constant in (5.71). ■

For $\mu_{h_p} \in \Lambda_{h_p}^{\circ}$, define $\widetilde{\boldsymbol{u}}_{h_p} = \widetilde{\boldsymbol{u}}_{h_p}(\mu_{h_p}) \in \boldsymbol{X}_{h_p}^{\circ}$ as velocity solution of the discrete problem

$$\begin{cases} a_p(\widetilde{\boldsymbol{u}}_{h_p}, \boldsymbol{v}_{h_p}) + b_p(\boldsymbol{v}_{h_p}, p_{h_p}) & = -(\boldsymbol{v}_{h_p}\cdot\boldsymbol{\eta}_p, \mu_{h_p})_\Gamma \quad \text{for all } \boldsymbol{v}_{h_p} \in \boldsymbol{X}_{h_p}^{\circ} \\ b_p(\widetilde{\boldsymbol{u}}_{h_p}, q_{h_p}) & = 0 \quad \text{for all } q_{h_p} \in M_{h_p}^{\circ} \end{cases} \tag{5.73}$$

and introduce

$$|\mu_{h_p}|_{\Lambda_{h_p}^{\circ}}^2 := a_p(\widetilde{\boldsymbol{u}}_{h_p}(\mu_{h_p}), \widetilde{\boldsymbol{u}}_{h_p}(\mu_{h_p})). \tag{5.74}$$

and for $\mu_{h_p} \in \Lambda_{h_p}$ (see Remark 5.26),

$$\|\mu_{h_p}\|_{\Lambda_{h_p}^{\circ}} := |\mu_{h_p} - \bar{\mu}_{h_p}|_{\Lambda_{h_p}^{\circ}} + |\Gamma|^{\frac{1}{2}}|\bar{\mu}_{h_p}| \tag{5.75}$$

where $\bar{\mu}_{h_p} := \frac{1}{|\Gamma|}\int_\Gamma \mu_{h_p}$.

In order to see that $|\cdot|_{\Lambda_{h_p}}$ is a norm on $\Lambda^\circ_{h_p}$, observe that if μ_{h_p} is such that $|\mu_{h_p}|_{\Lambda_{h_p}} = 0$, then $\widetilde{\boldsymbol{u}}_{h_p}(\mu_{h_p})$ vanishes. If we take $\boldsymbol{v}_{h_p}$ in (5.73) such that

$$\begin{cases} \boldsymbol{v}_{h_p}\cdot\boldsymbol{\eta}_p & = \mu_{h_p} \\ b_p(\boldsymbol{v}_{h_p}, q_{h_p}) & = 0 \qquad q_{h_p}\in M^\circ_{h_p}, \end{cases}$$

we see that $\|\mu_{h_p}\|_{L^2(\Gamma)} = 0$, that is $\mu_{h_p} = 0$. Then $|\cdot|_{\Lambda_{h_p}}$ is positive.

The norm $\Lambda^\circ_{h_p}$ is the natural discrete version of the norm $H^{1/2}(\Gamma)$ scaled by the factor $\sqrt{\frac{\kappa}{\nu}}$ for the space $\Lambda^\circ_{h_p}$. Indeed, by using (5.68) and (5.73), we have

$$\sup_{z_{h_p}\in \boldsymbol{X}^\circ_{h_p}\cdot\boldsymbol{\eta}_p|_\Gamma = \Lambda^\circ_{h_p}} \frac{(z_{h_p},\mu_{h_p})}{\sqrt{\frac{\nu}{\kappa}}|z_{h_p}|_{H^{-1/2}(\Gamma)}} \asymp \sup_{z_{h_p}\in\Lambda^\circ_{h_p}} \frac{(\boldsymbol{E}_{h_p}z_{h_p}\cdot\boldsymbol{\eta}_p,\mu_{h_p})}{\sqrt{\frac{\nu}{\kappa}}\|\boldsymbol{E}_{h_p}z_{h_p}\|_{\boldsymbol{L}^2(D)}} = |\mu_{h_p}|_{\Lambda^\circ_{h_p}}. \tag{5.76}$$

We have the following result.

Lemma 5.24 *The spaces $(\boldsymbol{Z}_h, \Lambda^\circ_{h_p})$ satisfy the discrete* inf-sup *condition, i.e., there is a constant $\tilde{\gamma} > 0$ such that:*

$$\inf_{\substack{\mu_{h_p}\in\Lambda^\circ_{h_p}\\ \lambda_{h_p}\neq 0}} \sup_{\substack{\boldsymbol{v}_h\in\boldsymbol{Z}_h\\ \boldsymbol{v}_h\neq 0}} \frac{(\llbracket\boldsymbol{v}_h\rrbracket\cdot\boldsymbol{\eta}_f,\mu_{h_p})_\Gamma}{\|\boldsymbol{v}_h\|_{\boldsymbol{X}}|\mu_{h_p}|_{\Lambda^\circ_{h_p}}} \geq \tilde{\gamma} > 0.$$

Proof. Take $\mu_{h_p}\in\Lambda^\circ_{h_p}$ and let $\widetilde{\boldsymbol{u}}_{h_p}(\mu_{h_p})$ be the velocity solution of (5.73). Since $\widetilde{\boldsymbol{u}}_{h_p}(\mu_{h_p})\in\boldsymbol{Z}_{h_p}$ then $\nabla\cdot\widetilde{\boldsymbol{u}}_{h_p} = 0$. Take $\boldsymbol{v}_h = (\boldsymbol{0}, \widetilde{\boldsymbol{u}}_{h_p}(\mu_{h_p}))\in\boldsymbol{Z}_{h_f}\times\boldsymbol{Z}_{h_p}$, then from (5.73)

$$\frac{(\llbracket\boldsymbol{v}_h\rrbracket\cdot\boldsymbol{\eta}_f,\mu_{h_p})_\Gamma}{\|\boldsymbol{v_h}\|_{\boldsymbol{X}}|\mu_{h_p}|_{\Lambda^\circ_{h_p}}} = \frac{a_p(\widetilde{\boldsymbol{u}}_{h_p}(\mu_{h_p}),\widetilde{\boldsymbol{u}}_{h_p}(\mu_{h_p}))}{\|\widetilde{\boldsymbol{u}}_{h_p}(\mu_{h_p})\|_{L^2(D_p)}|\mu_{h_p}|_{\Lambda^\circ_{h_p}}} = \sqrt{\frac{\nu}{\kappa}} > 0.$$

■

For $(q_h,\mu_{h_p})\in M^{\circ\circ}_h\times\Lambda^\circ_{h_p}$ define $|(q_h,\mu_{h_p})|^2_{M\times\Lambda^\circ_{h_p}} := \|q_h\|^2_M + |\mu_{h_p}|^2_{\Lambda^\circ_{h_p}}$. Then using the same argument of Proposition 5.17 we have:

Proposition 5.25 *Under assumptions of Lemmas 5.23 and 5.24 we have that there exists $\tilde{\beta} > 0$ such that*

$$\inf_{\substack{(q_h,\mu_{h_p})\in M^{\circ\circ}_h\times\Lambda^\circ_{h_p}\\ (q_h,\mu_{h_p})\neq(0,0)}} \sup_{\substack{\boldsymbol{v}_h\in\boldsymbol{X}^{\circ\circ}_h\\ \boldsymbol{v}_h\neq 0}} \frac{b(\boldsymbol{v}_h,q_h) + (\llbracket\boldsymbol{v}_h\rrbracket\cdot\boldsymbol{\eta}_f,\mu_{h_p})_\Gamma}{\|\boldsymbol{v}\|_{\boldsymbol{X}}\|(q_h,\mu_{h_p})\|_{M\times\Lambda^\circ_{h_p}}} \geq \tilde{\beta} > 0. \tag{5.77}$$

Remark 5.26 *With the inf-sup condition (5.77) of Proposition 5.25 we can establish the inf-sup conditions corresponding to the discrete versions of the first and the second weak formulations in (5.28) and (5.32), respectively. This is done using similar arguments to those given in Section 5.4.2; see Remark 5.18.*

5.6 Error analysis

We remark that the constants involved in the notation $\preceq$ are all independent, not only of the mesh size but also independent of the parameters ν and κ. In addition, using scaling arguments, it is easy to see that $\frac{1}{\sqrt{\nu}}p_f, \sqrt{\nu}\boldsymbol{u}_f, \sqrt{\frac{\kappa}{\nu}}p_p$ and $\sqrt{\frac{\nu}{\kappa}}\boldsymbol{u}_p$ are all $O(1)$, therefore, we keep those factors on the a priori error estimates.

We introduce the following energy norms

$$|\boldsymbol{v}_f|^2_{a_f} := a_f(\boldsymbol{v}_f, \boldsymbol{v}_f), \tag{5.78}$$

$$\|\boldsymbol{v}_p\|^2_{a_p} := a_p(\boldsymbol{v}_p, \boldsymbol{v}_p), \tag{5.79}$$

and

$$\|\boldsymbol{v}\|^2_a := a(\boldsymbol{v}, \boldsymbol{v}). \tag{5.80}$$

We next establish a priori error estimates for the Stokes and Darcy velocities.

Proposition 5.27 *Suppose that $\mathcal{T}_{h_f}$ is non-degenerate and has no triangle with two edges on ∂D_f. Let $h := \max\{h_f, h_p\}$. Then we have the following estimate*

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_a \preceq h\left(\sqrt{\nu}|\boldsymbol{u}_f|_{H^2(D_f)^2} + \sqrt{\tfrac{\nu}{\kappa}}|\boldsymbol{u}_p|_{H^1(D_p)^2}\right) + h_p\tfrac{1}{\sqrt{\nu}}|p_f|_{H^1(D_f)}.$$

Moreover, if the refinement condition of Remark 5.72 is satisfied then

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_a \preceq h_f\sqrt{\nu}|\boldsymbol{u}_f|_{H^2(D_f)^2} + h_p\sqrt{\tfrac{\nu}{\kappa}}|\boldsymbol{u}_p|_{H^1(D_p)^2}.$$

Proof. From Proposition 5.25 we have that $\boldsymbol{Z}_h \cap \boldsymbol{V}_h$ is not empty, where $\boldsymbol{Z}_h$ and $\boldsymbol{V}_h$ are defined in (5.67) and (5.67), respectively. Then, the discrete problem associated with (5.38) can also be described as: find $\boldsymbol{u}_h \in \boldsymbol{Z}_h \cap \boldsymbol{V}_h$ such that

$$a(\boldsymbol{u}_h, \boldsymbol{v}_h) = \ell(\boldsymbol{v}_h) \quad \boldsymbol{v}_h \in \boldsymbol{Z}_h \cap \boldsymbol{V}_h,$$

where a is *elliptic* in $\boldsymbol{Z}_h \cap \boldsymbol{V}_h$. Furthermore, $\boldsymbol{u}_h$ is also the only velocity solution of

$$\begin{cases} a(\boldsymbol{u}_h, \boldsymbol{v}_h) + b(\boldsymbol{v}_h, p_h) + (\llbracket \boldsymbol{v}_h \rrbracket \cdot \boldsymbol{\eta}_f, \lambda_{h_p}) & = \ell(\boldsymbol{v}_h) \quad \text{for all } \boldsymbol{v}_h \in \boldsymbol{X}_h^{\circ\circ} \\ b(\boldsymbol{u}_h, q_h) & = 0 \quad \text{for all } q_h \in M_h^{\circ\circ} \\ (\llbracket \boldsymbol{u}_h \rrbracket \cdot \boldsymbol{\eta}_f, \mu_{h_p}) & = 0 \quad \text{for all } \mu_{h_p} \in \Lambda_{h_p}^{\circ}. \end{cases} \tag{5.81}$$

For any $\boldsymbol{w}_h \in \boldsymbol{Z}_h \cap \boldsymbol{V}_h$ we have that $\boldsymbol{v}_h := \boldsymbol{u}_h - \boldsymbol{w}_h \in \boldsymbol{Z}_h \cap \boldsymbol{V}_h$ and

$$a(\boldsymbol{v}_h, \boldsymbol{v}_h) = a(\boldsymbol{u}_h, \boldsymbol{v}_h) - a(\boldsymbol{w}_h, \boldsymbol{v}_h) = \ell(\boldsymbol{v}_h) - a(\boldsymbol{w}_h, \boldsymbol{v}_h). \tag{5.82}$$

Let $(\boldsymbol{u}, p, \lambda)$ be the solution of the continuous problem (5.20). Then

$$\ell(\boldsymbol{v}_h) = a(\boldsymbol{u}, \boldsymbol{v}_h) + b(\boldsymbol{v}_h, p) + b_\Gamma(\boldsymbol{v}_h, \lambda)$$

and using (5.82) it follows that

$$a(\boldsymbol{v}_h, \boldsymbol{v}_h) = a(\boldsymbol{u} - \boldsymbol{w}_h, \boldsymbol{v}_h) + b(\boldsymbol{v}_h, p) + b_\Gamma(\boldsymbol{v}_h, \lambda),$$

and

$$\begin{aligned}\|\boldsymbol{u}_h-\boldsymbol{w}_h\|_a=\|\boldsymbol{v}_h\|_a &\leq \|\boldsymbol{u}-\boldsymbol{w}_h\|_a \\ &+ \sup_{\boldsymbol{z}_h\in\boldsymbol{Z}_h\cap\boldsymbol{V}_h}\frac{b(\boldsymbol{z}_h,p)}{\|\boldsymbol{z}_h\|_a}+\sup_{\boldsymbol{z}_h\in\boldsymbol{Z}_h\cap\boldsymbol{V}_h}\frac{b_\Gamma(\boldsymbol{z}_h,\lambda)}{\|\boldsymbol{z}_h\|_a}.\end{aligned}$$

Hence, using

$$\|\boldsymbol{u}-\boldsymbol{u}_h\|_a\leq\|\boldsymbol{u}-\boldsymbol{w}_h\|_a+\|\boldsymbol{u}_h-\boldsymbol{w}_h\|_a,$$

we obtain

$$\begin{aligned}\|\boldsymbol{u}-\boldsymbol{u}_h\|_a &\leq 2\inf_{\boldsymbol{w}_h\in\boldsymbol{Z}_h\cap\boldsymbol{V}_h}\|\boldsymbol{u}-\boldsymbol{w}_h\|_a \\ &+ \sup_{\boldsymbol{z}_h\in\boldsymbol{Z}_h\cap\boldsymbol{V}_h}\frac{b(\boldsymbol{z}_h,p)}{\|\boldsymbol{z}_h\|_a}+\sup_{\boldsymbol{z}_h\in\boldsymbol{Z}_h\cap\boldsymbol{V}_h}\frac{b_\Gamma(\boldsymbol{z}_h,\lambda)}{\|\boldsymbol{z}_h\|_a}.\end{aligned}\tag{5.83}$$

To bound the first term on the right-hand size of (5.83) we let $\boldsymbol{w}_h=\boldsymbol{\Pi}_h\boldsymbol{u}$, where $\boldsymbol{\Pi}_h$ is defined in Proposition 5.21. Proposition 5.21 guarantees that $\boldsymbol{w}_h\in\boldsymbol{V}_h$. In addition, since $b(\boldsymbol{u},q_h)=0$ for all $q_h\in M_h^{\circ\circ}$, (5.70) guaranties that $\boldsymbol{w}_h=\boldsymbol{\Pi}_h\boldsymbol{u}\in\boldsymbol{Z}_h$ and we have

$$\begin{aligned}\|\boldsymbol{u}-\boldsymbol{\Pi}_h\boldsymbol{u}\|_a &\leq \|\boldsymbol{u}_f-\boldsymbol{\Pi}_{h_f}\boldsymbol{u}_f\|_{a_f}+\|\boldsymbol{u}_p-\boldsymbol{\Pi}_{h_p}\boldsymbol{u}_p\|_{a_p} \\ &\leq \|\boldsymbol{u}_f-\boldsymbol{I}_{h_f}^{TH}\boldsymbol{u}_f\|_{a_f}+\|\boldsymbol{u}_p-\boldsymbol{I}_{h_p}^{RT}\boldsymbol{u}_p\|_{a_p} \\ &\quad+\|\boldsymbol{E}_{h_p}\left(Q_{h_p}(\boldsymbol{I}_{h_f}^{TH}\boldsymbol{u}_f\cdot\boldsymbol{\eta}_p)-\boldsymbol{I}_{h_p}^{RT}\boldsymbol{u}_p\cdot\boldsymbol{\eta}_p\right)\|_{a_p}.\end{aligned}$$

From (5.57) in Lemma 5.19 we obtain

$$\|\boldsymbol{u}_f-\boldsymbol{I}_{h_f}^{TH}\boldsymbol{u}_f\|_{a_f}\preceq h_f\sqrt{\nu}|\boldsymbol{u}_f|_{H^2(D_f)^2}\tag{5.84}$$

and from (5.64) in Lemma 5.20 we obtain

$$\|\boldsymbol{u}_p-\boldsymbol{I}_{h_p}^{RT}\boldsymbol{u}_p\|_{a_p}\preceq h_p\sqrt{\frac{\nu}{\kappa}}|\boldsymbol{u}_p|_{H^1(D_p)^2}\tag{5.85}$$

since $\nabla\cdot\boldsymbol{u}_p=0$.

From the boundedness of $\boldsymbol{E}_{h_p}$ in (5.69), we have

$$\begin{aligned}&\|\boldsymbol{E}_{h_p}\left(Q_{h_p}(\boldsymbol{I}_{h_f}^{TH}\boldsymbol{u}_f\cdot\boldsymbol{\eta}_p)-\boldsymbol{I}_{h_p}^{RT}\boldsymbol{u}_p\cdot\boldsymbol{\eta}_p\right)\|_{a_p} \\ &\qquad\preceq \sqrt{\frac{\nu}{\kappa}}|Q_{h_p}(\boldsymbol{I}_{h_f}^{TH}\boldsymbol{u}_f\cdot\boldsymbol{\eta}_p)-\boldsymbol{I}_{h_p}^{RT}\boldsymbol{u}_p\cdot\boldsymbol{\eta}_p|_{H^{-1/2}(\Gamma)}.\end{aligned}$$

Therefore, we need to estimate the following three terms:

$$\begin{aligned}&|Q_{h_p}(\boldsymbol{I}_{h_f}^{TH}\boldsymbol{u}_f\cdot\boldsymbol{\eta}_p)-\boldsymbol{I}_{h_p}^{RT}\boldsymbol{u}_p\cdot\boldsymbol{\eta}_p|_{H^{-1/2}(\Gamma)}\leq \\ &|Q_{h_p}(\boldsymbol{I}_{h_f}^{TH}\boldsymbol{v}_f\cdot\boldsymbol{\eta}_p)-\boldsymbol{I}_{h_f}^{TH}\boldsymbol{u}_f\cdot\boldsymbol{\eta}_p|_{H^{-1/2}(\Gamma)}+|\boldsymbol{I}_{h_f}^{TH}\boldsymbol{u}_f\cdot\boldsymbol{\eta}_p-\boldsymbol{u}_f\cdot\boldsymbol{\eta}_p|_{H^{-1/2}(\Gamma)} \\ &+\|\boldsymbol{u}_f\cdot\boldsymbol{\eta}_p-\boldsymbol{I}_{h_p}^{RT}\boldsymbol{u}_p\cdot\boldsymbol{\eta}_p\|_{H^{-1/2}(\Gamma)}\end{aligned}\tag{5.86}$$

1. Approximation property (5.87), boundedness of $\boldsymbol{I}_{h_f}^{TH}$ in (5.59) and the trace theorem give

$$\begin{aligned}|Q_{h_p}(\boldsymbol{I}_{h_f}^{TH}\boldsymbol{u}_f\cdot\boldsymbol{\eta}_p)-\boldsymbol{I}_{h_f}^{TH}\boldsymbol{u}_f\cdot\boldsymbol{\eta}_p|_{H^{-1/2}(\Gamma)} &\preceq h_p|\boldsymbol{I}_{h_f}^{TH}\boldsymbol{u}_f\cdot\boldsymbol{\eta}_p|_{H^{1/2}(\Gamma)}\\ &\preceq h_p|\boldsymbol{u}_f\cdot\boldsymbol{\eta}_p|_{H^{1/2}(\Gamma)}\\ &= h_p|\boldsymbol{u}_p\cdot\boldsymbol{\eta}_p|_{H^{1/2}(\Gamma)}\\ &\leq h_p|\boldsymbol{u}_p|_{H^{1/2}(\Gamma)^2}\\ &\preceq h_p|\boldsymbol{u}_p|_{H^1(D_p)^2}.\end{aligned}$$

2. The trace theorem and approximation properties of $\boldsymbol{I}_{h_f}^{TH}$ (Lemma 5.19) give

$$\begin{aligned}|\boldsymbol{I}_{h_f}^{TH}\boldsymbol{u}_f\cdot\boldsymbol{\eta}_p-\boldsymbol{u}_f\cdot\boldsymbol{\eta}_p|_{H^{-1/2}(\Gamma)} &\preceq h_f\|\boldsymbol{u}_f\cdot\boldsymbol{\eta}_p\|_{H^{1/2}(\Gamma)}\\ &= h_f\|\boldsymbol{u}_p\cdot\boldsymbol{\eta}_p\|_{H^{1/2}(\Gamma)}\\ &\preceq h_f|\boldsymbol{u}_p|_{H^1(D_p)^2}.\end{aligned}$$

3. The normal trace theorem and the approximation property (5.64) of $\boldsymbol{I}_{h_p}^{RT}$ imply

$$\begin{aligned}|\boldsymbol{u}_p\cdot\boldsymbol{\eta}_p-\boldsymbol{I}_{h_p}^{RT}\boldsymbol{u}_p\cdot\boldsymbol{\eta}_p|_{H^{-1/2}(\Gamma)} &\preceq h_p|\boldsymbol{u}_p\cdot\boldsymbol{\eta}_p|_{H^{1/2}(\Gamma)}\\ &\preceq h_p|\boldsymbol{u}_p|_{H^1(D_p)^2}.\end{aligned}$$

We note that we have used

$$|Q_{h_p}\mu-\mu|_{H^{-1/2}(\Gamma)}\preceq h_p|\mu|_{H^{1/2}(\Gamma)}, \tag{5.87}$$

since by using local arguments we have $\|Q_{h_p}\mu-\mu\|_{L^2(\Gamma)}\preceq h_p^{1/2}|\mu|_{H^{1/2}(\Gamma)}$ and then

$$\begin{aligned}|Q_{h_p}\mu-\mu|_{H^{-1/2}(\Gamma)} &= \sup_{\phi\in H^{1/2}(\Gamma)}\frac{\langle Q_{h_p}\mu-\mu,\phi\rangle_\Gamma}{|\phi|_{H^{1/2}(\Gamma)}}\\ &\leq \sup_{\phi\in H^{1/2}(\Gamma)}\frac{\|Q_{h_p}\mu-\mu\|_{L^2(\Gamma)}\|Q_{h_p}\phi-\phi\|_{L^2(\Gamma)}}{|\phi|_{H^{1/2}(\Gamma)}}\\ &\preceq h_p|\mu|_{H^{1/2}(\Gamma)}.\end{aligned}$$

We now bound the second term on the right-hand size of (5.83). Note that since we are using lowest order Raviart-Thomas elements, the porous side components of $\boldsymbol{Z}_h$ defined in (5.67) are divergence free, i.e., $\boldsymbol{Z}_{h_p}\subset\boldsymbol{Z}_p$, where $\boldsymbol{Z}_p$ is defined in (5.39), therefore, $b_p(\boldsymbol{z_h},q)=0$ for all $q=(q_f,q_p)\in M^{\circ\circ}$. In addition, we have $b(\boldsymbol{z_h},p-q_h)=0$ for $\boldsymbol{z}_h\in\boldsymbol{Z}_h\cap\boldsymbol{V}_h$. In summary, we have

$$|b(\boldsymbol{z}_h,p)|=|b_f(\boldsymbol{z}_{h_f},p_f)|=|b_f(\boldsymbol{z}_{h_f},p_f-Q_fp_f)|\preceq h_f\frac{1}{\sqrt{\nu}}|p_f|_{H^1(D_f)}\|\boldsymbol{z}_h\|_a,$$

where we have used the first order approximation of the L_2-projection operator Q_f on the fluid pressure space $M_{h_f}^\circ$.

To bound the third term on the right-hand size of (5.83) we have

$$\begin{aligned} b_\Gamma(\boldsymbol{z}_h,\lambda) &= \langle\lambda,\boldsymbol{z}_{h_f}\cdot\boldsymbol{\eta}_f\rangle_\Gamma+\langle\lambda,\boldsymbol{z}_{h_p}\cdot\boldsymbol{\eta}_p\rangle_\Gamma \\ &= \langle\lambda,\boldsymbol{z}_{h_f}\cdot\boldsymbol{\eta}_f\rangle_\Gamma+\langle Q_{h_p}\lambda,\boldsymbol{z}_{h_p}\cdot\boldsymbol{\eta}_p\rangle_\Gamma \quad \boldsymbol{z}_{h_p}\cdot\boldsymbol{\eta}_p \text{ is constant in } e \\ &= \langle\lambda-Q_{h_p}\lambda,\boldsymbol{z}_{h_f}\cdot\boldsymbol{\eta}_f\rangle_\Gamma \quad \boldsymbol{z}_{h_p}\in\boldsymbol{Z}_h, \end{aligned}$$

hence,

$$|b_\Gamma(\boldsymbol{z}_h,\lambda)|\preceq h_p\frac{1}{\sqrt{\nu}}|\lambda|_{H^{1/2}(\Gamma)}\sqrt{\nu}|\boldsymbol{z}_{h_f}\cdot\boldsymbol{\eta}_f|_{H^{1/2}(\Gamma)}. \tag{5.88}$$

By using (5.19) on Γ (on the D_f side) and trace theorems we obtain

$$|b_\Gamma(\boldsymbol{z}_h,\lambda)|\preceq h_p\left(\frac{1}{\sqrt{\nu}}|p_f|_{H^1(D_f)}+\sqrt{\nu}|\boldsymbol{u}_f|_{H^2(D_f)^2}\right)\|\boldsymbol{z}_h\|_a \tag{5.89}$$

and the proposition follows. ∎

Remark 5.28 *We note that we could have used the porous media side in (5.19) to bound $|\lambda|_{H^{1/2}(\Gamma)}$ in (5.88). In this case, we would have obtained*

$$|b_\Gamma(\boldsymbol{z}_h,\lambda)|\preceq\frac{h_p}{\sqrt{\nu}}|p_p|_{H^1(D_p)}\|\boldsymbol{z}_h\|_a. \tag{5.90}$$

Even thought we obtain the term h_p multiplying p_p in in (5.90), the bound (5.89) is qualitatively better than the bound (5.90). Note that by using scaling arguments we have $\sqrt{\frac{\kappa}{\nu}}p_p=O(1)$. Therefore, the bound $\frac{h_p}{\sqrt{\nu}}|p_p|_{H^1(D_p)}$ is very pessimist due to the fact that in practice the value of κ is very small.

We next establish a priori error estimates for the Stokes and Darcy pressures.

Proposition 5.29 *Suppose that $\mathcal{T}_{h_f}$ is non-degenerate and has no triangle with two edges on ∂D_f. Let $h:=\max\{h_f,h_p\}$. Then we obtain the following estimate*

$$\begin{aligned} &\frac{1}{\sqrt{\nu}}\|p_f-p_{h_f}\|_{L^2(D_f)}+\sqrt{\frac{\kappa}{\nu}}\|p_p-p_{h_p}\|_{L^2(D_p)} \\ &\quad\preceq\ h\left(\sqrt{\nu}|\boldsymbol{u}_f|_{H^2(D_f)^2}+\sqrt{\frac{\nu}{\kappa}}|\boldsymbol{u}_p|_{H^1(D_p)^2}+\frac{1}{\sqrt{\nu}}|p_f|_{H^1(D_f)}\right)+h_p\sqrt{\frac{\kappa}{\mu}}|p_p|_{H^1(D_p)}. \end{aligned}$$

Moreover, if the refinement condition of Remark 5.72 is satisfied then

$$\begin{aligned} &\frac{1}{\sqrt{\nu}}\|p_f-p_{h_f}\|_{L^2(D_f)}+\sqrt{\frac{\kappa}{\nu}}\|p_p-p_{h_p}\|_{L^2(D_p)} \\ &\quad\preceq\ h_f\left(\sqrt{\nu}|\boldsymbol{u}_f|_{H^2(D_f)^2}+\frac{1}{\sqrt{\nu}}|p_f|_{H^1(D_f)}\right)+h_p\left(\sqrt{\frac{\kappa}{\mu}}|p_p|_{H^1(D_p)}+\sqrt{\frac{\nu}{\kappa}}|\boldsymbol{u}_p|_{H^1(D_p)^2}\right). \end{aligned}$$

Proof. To obtain an expression for the pressure error, observe that for all $\boldsymbol{v}_h\in\boldsymbol{V}_h\cap(H_0^1(D_f)\times\boldsymbol{H}_0(\mathrm{div},D_p))$ (i.e., $\boldsymbol{v}_{h_f}=0$ on ∂D_f and $\boldsymbol{v}_{h_p}\cdot\boldsymbol{\eta}_p=0$ on ∂D_p) and all $q_h\in M_h^{\circ\circ}$

$$b(\boldsymbol{v}_h,p_h-q_h)=a(\boldsymbol{u}-\boldsymbol{u}_h,\boldsymbol{v}_h)+b(\boldsymbol{v}_h,p-q_h). \tag{5.91}$$

This holds true in particular for $\boldsymbol{v}_h = (\boldsymbol{v}_{h_f}, 0)$ and $q_h = (q_{h_f}, 0)$. If we take $q_{h_f} = Q_f p_f$, i.e., the L_2-projection on the discrete fluid pressure space, we obtain

$$b_f(\boldsymbol{v}_{h_f}, p_{h_f} - Q_f p_f) = a_f(\boldsymbol{u}_f - \boldsymbol{u}_{h_f}, \boldsymbol{v}_{h_f}) + b_f(\boldsymbol{v}_{h_f}, p_f - Q_f p_f).$$

Then, using the standard discrete inf-sup condition for the fluid problem, we have

$$\begin{aligned}
&\frac{1}{\sqrt{\nu}}\|p_{h_f} - Q_f p_f\|_{L^2(D_f)} \\
&\quad\preceq \sup_{\boldsymbol{v}_{h_f}\in \boldsymbol{V}_{h_f}\cap H_0^1(D_f)} \frac{a_f(\boldsymbol{u}_f - \boldsymbol{u}_{h_f}, \boldsymbol{v}_{h_f}) + b_f(\boldsymbol{v}_{h_f}, p_f - Q_f p_f)}{\|\boldsymbol{v}_{h_f}\|_{a_f}} \\
&\quad\preceq \|\boldsymbol{u}_f - \boldsymbol{u}_{h_f}\|_{a_f} + \frac{1}{\sqrt{\nu}}\|p_f - Q_f p_f\|_{L^2(D_f)} \\
&\quad\preceq \|\boldsymbol{u}_f - \boldsymbol{u}_{h_f}\|_{a_f} + h_f\frac{1}{\sqrt{\nu}}|p_f|_{H^1(D_f)},
\end{aligned}$$

and from a triangle inequality we obtain

$$\frac{1}{\sqrt{\nu}}\|p_f - p_{h_f}\|_{L^2(D_f)} \preceq \|\boldsymbol{u}_f - \boldsymbol{u}_{h_f}\|_{a_f} + h_f\frac{2}{\sqrt{\nu}}|p_f|_{H^1(D_f)}.$$

Analogously we obtain

$$\sqrt{\frac{\kappa}{\mu}}\|p_p - p_{h_p}\|_{L^2(D_p)} \preceq \|\boldsymbol{u_p} - \boldsymbol{u}_{h_p}\|_{p_f} + 2h_p\sqrt{\frac{\kappa}{\mu}}|p_p|_{H^1(D_p)}.$$

The proposition follows from the bound on velocity error given on Proposition 5.27. ∎

Now we analyze a priori error estimate for λ in the discrete norm $|\cdot|^\circ_{\Lambda_{h_p}}$ defined in (5.74); see also Arbogast et al. [2000]. Note that the norm $\Lambda^\circ_{h_p}$ was defined for piecewise constant functions on the Γ_{h_p} triangulation. For functions $\mu \in L^2(\Gamma)$, we define

$$|\mu|_{\Lambda^\circ_{h_p}} := |Q_{h_p}\mu|_{\Lambda^\circ_{h_p}}, \tag{5.92}$$

where Q_{h_p} is the L^2-projection onto $\Lambda^\circ_{h_p}$. We have the following result:

Proposition 5.30 *Suppose that $\mathcal{T}_{h_f}$ is non-degenerate and has no triangle with two edges on ∂D_f. Let $h := \max\{h_f, h_p\}$. Then we have the following estimates:*

$$|\lambda - \lambda_{h_p}|_{\Lambda^\circ_{h_p}} \preceq h\left(\sqrt{\nu}|\boldsymbol{u}_f|_{H^2(D_f)^2} + \sqrt{\tfrac{\nu}{\kappa}}|\boldsymbol{u}_p|_{H^1(D_p)^2}\right) + h_p\tfrac{1}{\sqrt{\nu}}|p_f|_{H^1(D_f)}, \tag{5.93}$$

and

$$\sqrt{\frac{\kappa}{\nu}}|\lambda - \lambda_{h_p}|_{H^{-1/2}(\Gamma)} \preceq h_p\sqrt{\frac{\kappa}{\nu}}|p_p|_{H^1(D_p)} + |\lambda - \lambda_{h_p}|_{\Lambda^\circ_{h_p}}. \tag{5.94}$$

Moreover, if the refinement condition of Remark 5.72 is satisfied then

$$|\lambda - \lambda_{h_p}|_{\Lambda^\circ_{h_p}} \preceq h_f\sqrt{\nu}|\boldsymbol{u}_f|_{H^2(D_f)^2} + h_p\sqrt{\frac{\nu}{\kappa}}|\boldsymbol{u}_p|_{H^1(D_p)^2}.$$

Proof. Let $\widetilde{\boldsymbol{u}}_{h_p}(Q_{h_p}\lambda)$ and $\widetilde{p}_{h_p}(Q_{h_p}\lambda)$ be the solution of (5.73). Note that the solution of (5.81) satisfies $\boldsymbol{u}_{h_p} = \widetilde{\boldsymbol{u}}_{h_p}(\lambda_{h_p})$ and $p_{h_p} = \widetilde{p}_{h_p}(\lambda_{h_p})$. Then, using the definition of the discrete norm $\Lambda_{h_p}^{\circ}$ we have

$$|\lambda - \lambda_{h_p}|_{\Lambda_{h_p}^{\circ}} = \|\widetilde{\boldsymbol{u}}_{h_p}(Q_{h_p}\lambda) - \boldsymbol{u}_{h_p}\|_{a_p}, \tag{5.95}$$

which can be bounded by

$$\|\widetilde{\boldsymbol{u}}_{h_p}(Q_{h_p}\lambda) - \boldsymbol{u}_{h_p}\|_{a_p} \leq \|\widetilde{\boldsymbol{u}}_{h_p}(Q_{h_p}\lambda) - \boldsymbol{u}_p\|_{a_p} + \|\boldsymbol{u}_p - \boldsymbol{u}_{h_p}\|_{a_p}. \tag{5.96}$$

We use Proposition 5.27 to estimate the second term on the right-hand side of (5.96). We next estimate the first term of the right-hand side of (5.96). Note that

$$a_p(\widetilde{\boldsymbol{u}}_{h_p}(Q_{h_p}\lambda) - \boldsymbol{u}_p, \boldsymbol{v}_{h_p}) + b_p(\boldsymbol{v}_{h_p}, \widetilde{p}_{h_p}(Q_{h_p}\lambda) - p_p) = 0. \tag{5.97}$$

Inserting $\boldsymbol{v}_{h_p} = \widetilde{\boldsymbol{u}}_{h_p}(Q_{h_p}\lambda) - \boldsymbol{u}_{h_p} \in \boldsymbol{Z}_{h_p}$ into (5.97) and recalling that $\boldsymbol{Z}_{h_p} \subset \boldsymbol{Z}_p$ where $\boldsymbol{Z}_{h_p}$ and $\boldsymbol{Z}_p$ are defined in (5.39) and (5.67), respectively, we have

$$a_p(\widetilde{\boldsymbol{u}}_{h_p}(Q_{h_p}\lambda) - \boldsymbol{u}_p, \widetilde{\boldsymbol{u}}_{h_p}(Q_{h_p}\lambda) - \boldsymbol{u}_{h_p}) = 0.$$

Hence,

$$a_p(\widetilde{\boldsymbol{u}}_{h_p}(Q_{h_p}\lambda) - \boldsymbol{u}_p, \widetilde{\boldsymbol{u}}_{h_p}(Q_{h_p}\lambda) - \boldsymbol{u}_p) + a_p(\widetilde{\boldsymbol{u}}_{h_p}(Q_{h_p}\lambda) - \boldsymbol{u}_p, \boldsymbol{u}_p - \boldsymbol{u}_{h_p}) = 0,$$

and by using a Cauchy-Schwarz inequality we obtain

$$\|\widetilde{\boldsymbol{u}}_{h_p}(Q_{h_p}\lambda) - \boldsymbol{u}_p\|_{a_p} \leq \|\boldsymbol{u}_{h_p} - \boldsymbol{u}_p\|_{a_p}$$

and (5.93) follows. To obtain the estimate (5.94), we note that from (5.76) we have

$$\sqrt{\frac{\kappa}{\mu}}\|Q_{h_p}\lambda - \lambda_{h_p}\|_{L^2(\Gamma)} = \sup_{z_{h_p}\in\Lambda_{h_p}^{\circ}} \frac{(z_{h_p}, Q_{h_p}\lambda - \lambda_{h_p})}{\sqrt{\frac{\nu}{\kappa}}|z_{h_p}|_{L^2(\Gamma)}} \preceq |\lambda - \lambda_{h_p}|_{\Lambda_{h_p}^{\circ}},$$

therefore,

$$|\lambda - \lambda_{h_p}|_{H^{-1/2}(\Gamma)} \preceq |\lambda - Q_{h_p}\lambda|_{H^{-1/2}(\Gamma)} + \|Q_{h_p}\lambda - \lambda_{h_p}\|_{L^2(\Gamma)}, \tag{5.98}$$

and (5.94) follows from (5.98) and (5.87). ∎

Remark 5.31 *Note that we are discretizing the third weak formulation (5.35). We have to recover the piecewise constant pressure in each subdomain. Recall the function $\boldsymbol{z}$ of Remark 5.13. Note that we can compute $\boldsymbol{z}_h := \boldsymbol{\Pi}_h(\boldsymbol{z}) = (\boldsymbol{\Pi}_{h_f}\boldsymbol{z}, \boldsymbol{\Pi}_{h_p}\boldsymbol{z})$; see Proposition 5.21. Then*

$$\gamma_h := \ell(\boldsymbol{z}_h) - a(\boldsymbol{u}_h, \boldsymbol{z}_h) - b(\boldsymbol{z}_h, p_h) - (\llbracket \boldsymbol{z}_h \rrbracket \cdot \boldsymbol{\eta}_f, \lambda_{h_p})_\Gamma,$$

and $\gamma_h\, p^c = \gamma_h(p_f^c, p_p^c)$ is the approximation for piecewise constant pressure in each subdomain D_j, $j = f, p$. Observe that

$$|\gamma - \gamma_h| \preceq |a(\boldsymbol{u} - \boldsymbol{u}_h, \boldsymbol{z}_h)| + |b(\boldsymbol{z}_h, p - p_h)| + |(\llbracket \boldsymbol{z}_h \rrbracket \cdot \boldsymbol{\eta}_f, \gamma - \gamma_{h_p})_\Gamma|.$$

These last terms can be estimated using the results of this section. Analogously we can recover the mean value $\bar{\lambda}$ of the Lagrange multiplier. Indeed, we can find $\boldsymbol{w}_h = (0, \boldsymbol{w}_{h_p}) \in \boldsymbol{X}_f \times \boldsymbol{X}_p$ such that

$$\boldsymbol{w}_{h_p} \cdot \boldsymbol{\eta}_p = \frac{1}{|\Gamma|} \text{ on } \Gamma \text{ and } \boldsymbol{w}_{h_p} \cdot \boldsymbol{\eta}_p = 0 \text{ on } \Gamma_p$$

so we can define, see Remark 5.12,

$$\bar{\lambda}_h := \ell(\boldsymbol{w}) - a(\boldsymbol{u}_h, \boldsymbol{w}) - b(\boldsymbol{w}, p_h).$$

In this case

$$|\bar{\lambda} - \bar{\lambda}_h| \preceq |a(\boldsymbol{u} - \boldsymbol{u}_h, w)| + |b(\boldsymbol{w}, p - p_h)|.$$

The last two terms can be estimated using the results of this section.

5.7 Numerical results

In this section we present numerical experiments in order to verify the estimates established in the paper. We consider $D_f = (1,2) \times (0,1)$ and $D_p = (0,1) \times (0,1)$. We consider $\alpha_f = 0$.

The velocity solution for Stokes' equations is given by $\boldsymbol{u}_f(x,y) = (y(1-y), -x+2+2(x-1)y)$ with pressure $p_f(x,y) = -2x - \frac{\nu}{\kappa}y + 5/2 + \frac{5\nu}{12\kappa}$. Note that $\boldsymbol{u}_f$ is not divergence free.
The velocity solution for Darcy's equation is $\boldsymbol{u}_p(x,y) = (1-2x+x^2+y-y^2, -1+x+2y-2xy)$ with pressure $p_p(x,y) = \frac{\nu}{\kappa}((1-x)y(1-y) - x + x^2 - \frac{x^3}{3} + \frac{3}{4} - y) + \frac{1}{2}$.

Note that the normal component of $\boldsymbol{u}_p$ has a parabolic profile on the interface $\Gamma = 1 \times (0,1)$ while its tangential component is zero. Note also that $\boldsymbol{D}\boldsymbol{u}_f = \begin{pmatrix} 0 & 0 \\ 0 & * \end{pmatrix}$ on Γ, and $p_f = p_p$ on Γ. The exact solution is compatible with (5.7) with (5.8) when $\alpha_f = 0$.

A similar example is presented in Burman and Hansbo [2007], where the term $\nabla \cdot \boldsymbol{D}\boldsymbol{u}_f$ is replaced by $\Delta \boldsymbol{u}_f$ in the Stokes equations.

In Figure 5.1 we show the computed solution of the coupled problem. On the porous side we have plotted the velocity in the center of each triangle. In Figure 5.2 we zoom part of the interface and plot the y component of the velocities.

In Figure 5.3 we show the behavior of the error (in the scaled norms defined in (5.78), (5.79) and (5.80)) with respect to the discretization parameters. Here we also show $\|\lambda - \lambda_{h_p}\|_{\Lambda^{h_p}}$, i.e., the Lagrange multiplier approximation error in the discrete norm defined in (5.92).

We observe according to Figure 5.3, the error in the norm $\|\cdot\|_a^2$ defined in (5.80), which is the sum of the fluid velocity and porous velocity errors in the scaled norms, is of linear order. This agree with Proposition 5.27. Analogously, the pressure error is of linear order. This also agree with the result about the pressure error, see Proposition 5.29. We finally observe that the Lagrange multiplier error in the discrete norm defined in (5.92) is also of linear order.

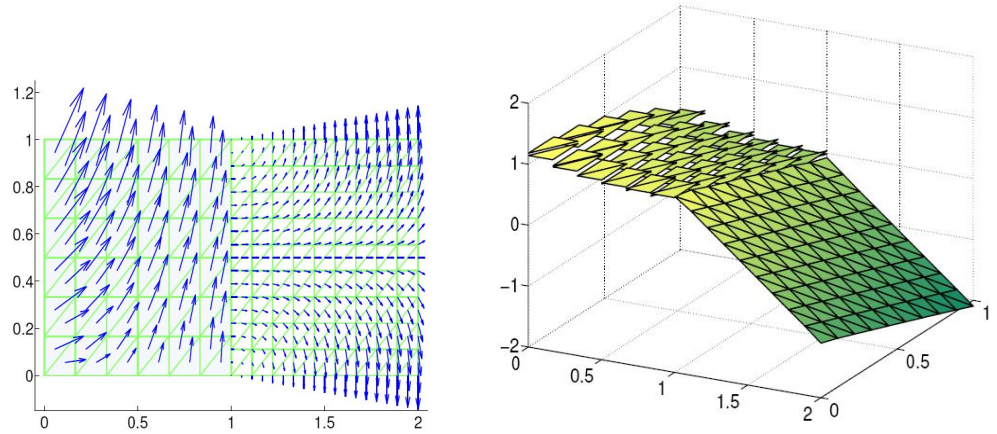

Figure 5.1: Computed velocities (left) and pressures (right). On the porous side we have plot the value of the velocity at the centroid of each triangle.

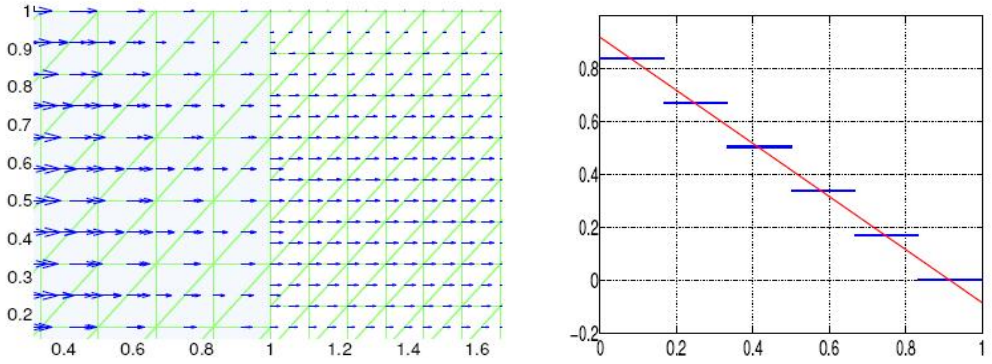

Figure 5.2: The x-component of the discrete velocity (left figure), where on the porous side (left subdomain) we plot the two values of the x component of the velocities at the midpoint of each edge; recall that Raviart-Thomas elements allow discontinuous tangential velocities on interior edges. The discrete (in blue) and the exact (in red) Lagrange multipliers on the interface (right figure).

5.8 Conclusion

We study the coupling across an interface between fluid and porous medium flows, consisting of *Stokes' equations* in the fluid region D_f and *Darcy's law* for the filtration velocity in the porous medium region D_p. After discussing the adequate choice of $H^{1/2}(\Gamma)$, rather than $H_{00}^{1/2}(\Gamma)$, as the Lagrange multiplier space, we present a complete analysis for the inf-sup and approximation results associated with the continuous and discrete formulations of this Stokes/Darcy system. We choose the triangular $P2\backslash P1$ Taylor Hood finite elements and the lower order Raviart-Thomas elements as discrete spaces for the free and porous medium subdomains, respectively. Optimal a priori discrete error estimates do not depend on the coefficients ν and κ and ratio of mesh parameters. Sharper local estimates can also be obtained for the case where the fluid mesh on the interface Γ is a refinement of the porous mesh on Γ. The numerical experiments show good agreements with our theoretical results.

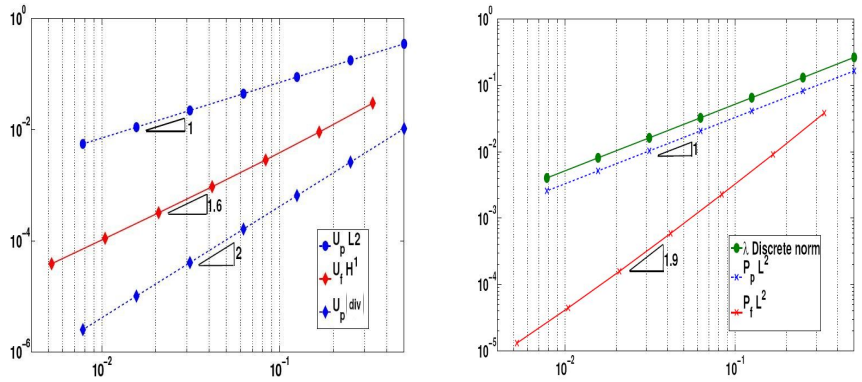

Figure 5.3: Velocities errors (right) and pressures errors (left).

5.9 Appendix A: Non-homogeneous boundary conditions

The non-homogeneous boundary condition can be reduced to the homogeneous case when $\boldsymbol{h}_f \in H^{1/2}(\Gamma_f)^2$ and $h_p \in H^{-1/2}(\Gamma_p)$. First construct $\boldsymbol{\omega}_f \in H^1(D_f)^2$ such that

$$\left\{\begin{array}{rll} -\nabla\cdot T(\boldsymbol{w}_f,\tilde{p}_f) &= 0 & \text{in } D_f \\ \nabla\cdot\boldsymbol{w}_f &= g_f & \text{in } D_f \\ \boldsymbol{w}_f &= \boldsymbol{h}_f & \text{on } \Gamma_f \\ T(\boldsymbol{w}_f,\tilde{p}_f)\cdot\boldsymbol{\eta}_f &= 0 & \text{on } \Gamma \end{array}\right. \tag{5.99}$$

From the divergence theorem

$$\int_{\Gamma_f} \boldsymbol{w}_f\cdot\boldsymbol{\eta}_f = \int_{D_f} g_f - \int_{\Gamma_f} \boldsymbol{h}_f\cdot\boldsymbol{\eta}_f. \tag{5.100}$$

Now put $\boldsymbol{u}_f = \boldsymbol{\omega}_f + \boldsymbol{\zeta}_f$ where $\boldsymbol{u}_f$ satisfies the non-homogeneous system (5.3). So we are looking for $\boldsymbol{\zeta}_f$ that satisfy:

$$\left\{\begin{array}{rll} -\nabla\cdot T(\boldsymbol{\zeta}_f,p_f) &= \boldsymbol{f}_f + \nabla\cdot 2\nu\boldsymbol{D}(\boldsymbol{\omega}_f) & \text{in } D_f \\ \nabla\cdot\boldsymbol{\zeta}_f &= 0 & \text{in } D_f \\ \boldsymbol{\zeta}_f &= 0 & \text{on } \Gamma_f \end{array}\right.$$

.

Analogously, on the porous region, the non-homogeneous case can be reduced to the homogeneous one. In this case $h_p \in H^{-1/2}(\Gamma_p)$. Construct $\boldsymbol{w}_p \in \boldsymbol{H}(\mathrm{div}, D_p)$ such that

$$\left\{\begin{array}{rll} \frac{\nu}{\kappa}\boldsymbol{w}_p + \nabla\tilde{p}_p &= 0 & \text{in } D_p \\ \nabla\cdot\boldsymbol{w}_p &= g_p & \text{in } D_p \\ \boldsymbol{w}_p\cdot\boldsymbol{\eta}_p &= h_p & \text{on } \Gamma_p, \\ \boldsymbol{w}_p\cdot\boldsymbol{\eta}_p &= \boldsymbol{w}_f\cdot\boldsymbol{\eta}_p & \text{on } \Gamma, \end{array}\right. \tag{5.101}$$

with $\boldsymbol{w}_f$ defined in (5.99). This construction is possible since the compatibility condition (5.5) and (5.100) imply that the system (5.101) is compatible.

Put $\boldsymbol{u}_p = \boldsymbol{\omega}_p + \boldsymbol{\zeta}_p$. Then we look for $\boldsymbol{\zeta}_p$ such that:

$$\left\{\begin{array}{rll} \frac{\nu}{\kappa}\boldsymbol{\zeta}_p + \nabla p_p &= -\frac{\nu}{\kappa}\boldsymbol{\omega}_p & \text{in } D_p \\ \nabla\cdot\boldsymbol{\zeta}_p &= 0 & \text{in } D_p \\ \boldsymbol{\zeta}_p\cdot\boldsymbol{\eta}_p &= 0 & \Gamma_p. \end{array}\right.$$

In terms of weak formulation, with $\boldsymbol{\omega} := (\boldsymbol{\omega}_f, \boldsymbol{\omega}_p)$, we have: *find* $(\boldsymbol{\zeta}, p, \lambda) \in \boldsymbol{X} \times M^\circ \times \Lambda$ *satisfying*:

$$\begin{cases} a(\boldsymbol{\zeta}, \boldsymbol{v}) + b(\boldsymbol{v}, p) + b_\Gamma(\boldsymbol{v}, \lambda) & = \ell(\boldsymbol{v}) - a(\boldsymbol{\omega}, \boldsymbol{v}) & \text{for all } \boldsymbol{v} \in \boldsymbol{X} \\ b(\boldsymbol{\zeta}, q) & = 0 & \text{for all } q \in M^\circ \\ b_\Gamma(\boldsymbol{\zeta}, \mu) & = 0 & \text{for all } \mu \in \Lambda, \end{cases}$$

which is the same problem (5.28) with a different right hand side.

5.10 Approximation properties of Taylor-Hood finite elements

In this appendix, the domain of reference is D_f. Recall the definitions of $\boldsymbol{X}_f$ and $M^\circ_{h_f}$ on (5.53) and (5.55), respectively. In order to simplify the notation in some cases we omit the subscript that refers to the domain. In particular, all the operators defined in this section act on velocities defined on D_f.

Let $\boldsymbol{\mathcal{Q}} : \boldsymbol{X}_f \to \boldsymbol{X}_{h_f}$ be Clement interpolation (see Braess [2001], Clément [1975] and Scott and Zhang [1990]). It is know that $\boldsymbol{\mathcal{Q}}$ is bounded, i.e.,

$$|\boldsymbol{\mathcal{Q}}\boldsymbol{v}_f|_{H^1(D_f)^2} \preceq |\boldsymbol{v}|_{H^1(D_f)^2}, \tag{5.102}$$

and we have

$$\begin{aligned} \|\boldsymbol{v}_f - \boldsymbol{\mathcal{Q}}\boldsymbol{v}_f\|_{L^2(D_f)^2} &\preceq h^s |\boldsymbol{v}_f|_{H^s(D_f)^2}, \quad s = 1, 2. && (5.103)\\ |\boldsymbol{v}_f - \boldsymbol{\mathcal{Q}}\boldsymbol{v}_f|_{H^1(D_f)^2} &\preceq h |\boldsymbol{v}_f|_{H^2(D_f)^2}, && (5.104)\\ |\boldsymbol{\mathcal{Q}}\boldsymbol{v}_f|_{H^{1/2}(\Gamma)^2} &\preceq |\boldsymbol{v}_f|_{H^{1/2}(\Gamma)^2}, && (5.105)\\ \|\boldsymbol{v}_f - \boldsymbol{\mathcal{Q}}\boldsymbol{v}_f\|_{L(\Gamma)^2} &\preceq h^{\frac{1}{2}} |\boldsymbol{v}_f|_{H^{1/2}(\Gamma)^2}. && (5.106) \end{aligned}$$

This interpolation is basically a Clement interpolation on Γ, i.e., values zero at the interface relative boundary points and a Clement interpolation at the interior nodes.

Given $K \in \mathcal{T}_{h_f}$ and e edge of K, let $\boldsymbol{\eta}_e^{(K)} = (\eta_e^1, \eta_e^2)$ denote the normal to e exterior to K, $\boldsymbol{\tau}_e^{(K)} = (\tau_e^1, \tau_e^2)$ the tangential vector to e (with ∂K anticlockwise oriented), and x_e the midpoint of the edge e. Each interior edge belongs to two triangles K_1 and K_2. Let $\boldsymbol{\eta}_e$ denote one of the directions $\boldsymbol{\eta}_e^{(K_1)}$ or $\boldsymbol{\eta}_e^{(K_2)}$. For boundary edges $\boldsymbol{\eta}_e$ denotes $\boldsymbol{\eta}_e^{(K)}$. Analogously, for interior edges let $\boldsymbol{\tau}_e$ denote one of the directions $\boldsymbol{\tau}_e^{(K_1)}$ or $\boldsymbol{\tau}_e^{(K_2)}$, and for boundary edges $\boldsymbol{\tau}_e = \boldsymbol{\tau}_e^{(K)}$.

Let $\phi_i^{(K)}$, $i = 1, 2, 3$, be the edge bubble Taylor-Hood basis functions based on the midpoints of the edges of K. Let $\boldsymbol{\psi}_i^{(K)} := \phi_i^{(K)} \boldsymbol{\eta}_{e_i}$, $i = 1, 2, 3$, and $\boldsymbol{\vartheta}_i^{(K)} := \phi_i^{(K)} \boldsymbol{\tau}_{e_i}$, $i = 1, 2, 3$. Observe that:

$$\int_K \boldsymbol{\psi}_i^{(K)} \cdot \boldsymbol{\eta}_{e_i} \neq 0, \quad \boldsymbol{\psi}_i^{(K)} \cdot \boldsymbol{\tau}_{e_i} = 0 \;\; i = 1, 2, 3.$$

$$\boldsymbol{\vartheta}_i^{(K)}(x_{e_i}) \cdot \boldsymbol{\tau}_{e_i} \neq 0, \;\; \boldsymbol{\vartheta}_i^{(K)} \cdot \boldsymbol{\eta}_{e_i} = 0 \;\; i = 1, 2, 3.$$

Now consider the following subspaces of $\boldsymbol{X}_{h_f}$:

$$\boldsymbol{W}_{h_f}^{\boldsymbol{\eta}} := \{\boldsymbol{v}_{h_f} \in \boldsymbol{X}_{h_f} : v|_K \in \text{Span}\{\boldsymbol{\psi}_1^{(K)}, \boldsymbol{\psi}_2^{(K)}, \boldsymbol{\psi}_3^{(K)}\}\} \cap \boldsymbol{X}_{h_f}$$

and

$$\boldsymbol{W}_{h_f}^{\boldsymbol{\tau}} := \{\boldsymbol{v}_{h_f} \in \boldsymbol{X}_{h_f} : v|_K \in \text{Span}\{\boldsymbol{\vartheta}_1^{(K)}, \boldsymbol{\vartheta}_2^{(K)}, \boldsymbol{\vartheta}_3^{(K)}\}\} \cap \boldsymbol{X}_{h_f}.$$

Note that if $\boldsymbol{v}_{h_f} \in \boldsymbol{W}_{h_f}^{\boldsymbol{\eta}}$ then $\boldsymbol{v}_{h_f} \cdot \boldsymbol{\eta}_f|_\Gamma \in H_{00}^{1/2}(\Gamma)$ and $\boldsymbol{v}_{h_f} \cdot \boldsymbol{\tau}_f|_{\partial D_f} = 0$. Also note that if $\boldsymbol{v}_{h_f} \in \boldsymbol{W}_{h_f}^{\boldsymbol{\tau}}$ then $\boldsymbol{v}_{h_f} \cdot \boldsymbol{\tau}_f|_\Gamma \in H_{00}^{1/2}(\Gamma)$ and $\boldsymbol{v}_{h_f} \cdot \boldsymbol{\eta}_f|_{\partial D_f} = 0$.

Let $\boldsymbol{\Pi}_\eta : \boldsymbol{X}_f \to \boldsymbol{W}_{h_f}^{\boldsymbol{\eta}}$ be (locally) defined by :

$$\boldsymbol{\Pi}_\eta \boldsymbol{v}_f \in \text{Span}\{\boldsymbol{\psi}_1^{(K)}, \boldsymbol{\psi}_2^{(K)}, \boldsymbol{\psi}_3^{(K)}\},$$

such that

$$\int_{e_i} \boldsymbol{\Pi}_\eta \boldsymbol{v}_f \cdot \boldsymbol{\eta} = \frac{1}{|e_i|} \int_{e_i} \boldsymbol{v}_f \cdot \boldsymbol{\eta}_{e_i}, \;\; i = 1, 2, 3$$

for all $K \in \mathcal{T}_h$. In other words, $\boldsymbol{\Pi}_\eta \boldsymbol{v}_f = \alpha_1 \boldsymbol{\psi}_1 + \alpha_2 \boldsymbol{\psi}_2 + \alpha_3 \boldsymbol{\psi}_3$, where

$$\alpha_i := \frac{\int_{e_i} \boldsymbol{v}_f \cdot \boldsymbol{\eta}_{e_i}}{\int_{e_i} \boldsymbol{\psi}_i \cdot \boldsymbol{\eta}_{e_i}} = \frac{\int_{e_i} \boldsymbol{v}_f \cdot \boldsymbol{\eta}_{e_i}}{\int_{e_i} \phi_i^{(K)}}.$$

From a trace theorem and a scaling argument we have that:

$$|\alpha_i|^2 \preceq \frac{1}{h_f^2} \|\boldsymbol{v}_f\|_{L^2(K)^2}^2 + |\boldsymbol{v}_f|_{H^1(K)^2}^2.$$

Then

$$|\boldsymbol{\Pi}_\eta \boldsymbol{v}_f|_{H^1(D_f)^2} \preceq \max_{1 \le i \le 3} |\alpha_i|^2 \preceq \frac{1}{h_f^2} \|\boldsymbol{v}_f\|_{L^2(D_f)^2}^2 + |\boldsymbol{v}_f|_{H^1(D_f)^2}^2$$

and

$$\|\boldsymbol{\Pi}_\eta \boldsymbol{v}_f\|_{L^2(D_f)^2}^2 \preceq h_f^2 \max_{1 \le i \le 3} |\alpha_i|^2 \preceq \|\boldsymbol{v}_f\|_{L^2(D_f)^2}^2 + h_f^2 |\boldsymbol{v}_f|_{H^1(D_f)^2}^2. \tag{5.107}$$

Observe that

$$\int_K \nabla \cdot \boldsymbol{\Pi}_\eta \boldsymbol{v}_f = \int_{\partial K} \boldsymbol{\Pi}_\eta(\boldsymbol{v}_f) \cdot \boldsymbol{\eta} = \int_{\partial K} \boldsymbol{v}_f \cdot \boldsymbol{\eta} = \int_K \nabla \cdot \boldsymbol{v}_f.$$

We also have

$$\|\boldsymbol{\Pi}_\eta \boldsymbol{v}_f\|_{L^2(\Gamma)^2} \preceq \|\boldsymbol{v}_f\|_{L^2(\Gamma)^2}. \tag{5.108}$$

Define $\boldsymbol{\Upsilon}_\eta : \boldsymbol{X}_f \to \boldsymbol{X}_{h_f}$ by:

$$\boldsymbol{\Upsilon}_\eta \boldsymbol{v}_f := \boldsymbol{\mathcal{Q}} \boldsymbol{v}_f + \boldsymbol{\Pi}_\eta(\boldsymbol{v}_f - \boldsymbol{\mathcal{Q}} \boldsymbol{v}_f), \tag{5.109}$$

then we obtain the following results.

Lemma 5.32 *The operator $\boldsymbol{\Upsilon}_\eta$ defined in (5.109) is bounded*

$$|\boldsymbol{\Upsilon}_\eta \boldsymbol{v}_f|_{H^1(D_f)^2} \preceq |\boldsymbol{v}_f|_{H^1(D_f)^2}, \tag{5.110}$$

moreover,

$$\|\boldsymbol{v}_f - \boldsymbol{\Upsilon}_\eta \boldsymbol{v}_f\|_{L^2(D_f)^2} \preceq h^s \|\boldsymbol{v}_f\|_{H^s(D_f)^2} \quad s = 1, 2. \tag{5.111}$$

and

$$|\boldsymbol{v}_f - \boldsymbol{\Upsilon}_\eta \boldsymbol{v}_f|_{H^1(D_f)^2} \preceq h |\boldsymbol{v}_f|_{H^2(D_f)^2}. \tag{5.112}$$

We also have

$$|\boldsymbol{\Upsilon}_\eta \boldsymbol{v}_f|_{H^{1/2}(\Gamma)^2} \preceq \|\boldsymbol{v}_f\|_{H^{1/2}(\Gamma)^2}, \tag{5.113}$$

and

$$\int_e \boldsymbol{\Upsilon}_\eta \boldsymbol{v}_f \cdot \boldsymbol{\eta}_e = \int_e \boldsymbol{v}_f \cdot \boldsymbol{\eta}_e \quad \textit{for all edge } e. \tag{5.114}$$

Proof. From (5.107) we have, for $s = 1, 2$,

$$\sum_{K \in \mathcal{T}_h} \|\boldsymbol{\Pi}_\eta(\boldsymbol{v}_f - \boldsymbol{\mathcal{Q}}\boldsymbol{v}_f)\|^2_{L^2(K)^2} \tag{5.115}$$

$$\begin{aligned} &\preceq \sum_{K \in \mathcal{T}_h} \left(\|\boldsymbol{v}_f - \boldsymbol{\mathcal{Q}}\boldsymbol{v}_f\|^2_{L^2(K)^2} + h_f^2 |\boldsymbol{v}_f - \boldsymbol{\mathcal{Q}}\boldsymbol{v}_f|^2_{H^1(K)^2} \right) \\ &\preceq h_f^{2s} |\boldsymbol{v}_f|^2_{H^s(D_f)^2} + h_f^{2(s-1)+2} |\boldsymbol{v}_f|^2_{H^s(D_f)^2} \quad \text{by (5.103) and (5.102).} \\ &\preceq h_f^{2s} |\boldsymbol{v}_f|^2_{H^s(D_f)^2}. \end{aligned} \tag{5.116}$$

Then, using an inverse estimate (see Braess [2001]) and (5.116) we get

$$|\boldsymbol{\Pi}_\eta(\boldsymbol{v}_f - \boldsymbol{\mathcal{Q}}\boldsymbol{v}_f)|_{H^1(K)^2} \preceq \frac{1}{h_f} \|\boldsymbol{\Pi}_\eta(\boldsymbol{v}_f - \boldsymbol{\mathcal{Q}}\boldsymbol{v}_f)\|_{L^2(K)^2} \preceq |\boldsymbol{v}_f|_{H^1(D_f)^2},$$

and hence

$$\begin{aligned} |\boldsymbol{\Upsilon}_\eta \boldsymbol{v}_f|_{H^1(D_f)^2} &\leq |\boldsymbol{\mathcal{Q}}\boldsymbol{v}_f|_{H^1(D_f)^2} + |\boldsymbol{\Pi}_\eta(\boldsymbol{v}_f - \boldsymbol{\mathcal{Q}}\boldsymbol{v}_f)|_{H^1(D_f)^2} \text{ by definition of } \boldsymbol{\Upsilon}_\eta \\ &\preceq |\boldsymbol{v}_f|_{H^1(D_f)^2} + |\boldsymbol{v}_f|_{H^1(D_f)^2} \preceq |\boldsymbol{v}_f|_{H^1(D_f)^2}. \end{aligned}$$

To show (5.111) we have that

$$\begin{aligned} &\|\boldsymbol{v}_f - \boldsymbol{\Upsilon}_\eta \boldsymbol{v}_f\|_{L^2(D_f)^2} \\ &= \|\boldsymbol{v}_f - \boldsymbol{\mathcal{Q}}\boldsymbol{v}_f - \boldsymbol{\Pi}_\eta(\boldsymbol{v}_f - \boldsymbol{\mathcal{Q}}\boldsymbol{v}_f)\|_{L^2(D_f)^2} \text{ by definition of } \boldsymbol{\Upsilon}_\eta. \\ &\leq \|\boldsymbol{v}_f - \boldsymbol{\mathcal{Q}}\boldsymbol{v}_f\|_{L^2(D_f)^2} + \|\boldsymbol{\Pi}_\eta(\boldsymbol{v}_f - \boldsymbol{\mathcal{Q}}\boldsymbol{v}_f)\|_{L^2(D_f)^2} \\ &\preceq h_f^s |\boldsymbol{v}_f|_{H^s(D_f)^2} + h^s |\boldsymbol{v}_f|_{H^s(D_f)^2} \text{ ; } \quad \text{by (5.103) and (5.116)} \\ &\preceq h_f^s |\boldsymbol{v}_f|_{H^s(D_f)^2}, \quad s = 1, 2. \end{aligned}$$

Analogously we get (5.112). To proof (5.113) observe that

$$\begin{aligned} |\boldsymbol{\Upsilon}_\eta \boldsymbol{v}_f|_{H^{1/2}(\Gamma)^2} &\leq |\boldsymbol{\mathcal{Q}}\boldsymbol{v}_f|_{H^{1/2}(\Gamma)^2} + |\boldsymbol{\Pi}_\eta(\boldsymbol{v}_f - \boldsymbol{\mathcal{Q}}\boldsymbol{v}_f)|_{H^{1/2}(\Gamma)^2} \\ &\preceq \|\boldsymbol{\mathcal{Q}}\boldsymbol{v}_f\|_{H^{1/2}(\Gamma)^2} + h_f^{-\frac{1}{2}} |\boldsymbol{\Pi}_\eta(\boldsymbol{v}_f - \boldsymbol{\mathcal{Q}}\boldsymbol{v}_f)|_{L^2(\Gamma)^2} \\ &\preceq |\boldsymbol{v}_f|_{H^{1/2}(\Gamma)^2} + h_f^{-\frac{1}{2}} \|\boldsymbol{v}_f - \boldsymbol{\mathcal{Q}}\boldsymbol{v}_f\|_{L^2(\Gamma)^2} \quad \text{by (5.105) and (5.108)} \\ &\preceq |\boldsymbol{v}_f|_{H^{1/2}(\Gamma)^2} \quad \text{by (5.106).} \end{aligned}$$

The last assertion, (5.114), is straightforward. ∎

Given $q_{h_f} \in M_{h_f}$, define (locally) $\widehat{\mathbf{\Pi}}_\tau q_{h_f} \in \boldsymbol{W}_{h_f}^{\boldsymbol{\tau}}$ by

$$\widehat{\mathbf{\Pi}}_\tau q_{h_f}|_K \in \text{Span}\{\boldsymbol{\vartheta}_1^{(K)}, \boldsymbol{\vartheta}_2^{(K)}, \boldsymbol{\vartheta}_3^{(K)}\}$$

with

$$\widehat{\mathbf{\Pi}}_\tau q_{h_f}(x_e) \cdot \boldsymbol{\eta} = 0 \text{ and } \widehat{\mathbf{\Pi}}_\tau q_{h_f}(x_e) \cdot \boldsymbol{\tau} = \nabla q_{h_f}(x_e) \cdot \boldsymbol{\tau} \tag{5.117}$$

at midpoints x_e of all interior edges e. For edges on Γ_f we define $\widehat{\mathbf{\Pi}}_\tau q|_e = 0$. Note that $\widehat{\mathbf{\Pi}}_\tau q_{h_f}$ is zero at the vertices of all elements of $\mathcal{T}_{h_f}$ and observe that $\widehat{\mathbf{\Pi}}_\tau q_{h_f} \in H^1(D_f)^2$ because the above equation are consistent in neighbor triangles which gives $\widehat{\mathbf{\Pi}}_\tau q_{h_f}$ continuous (see Braess [2001], Chapter II, Theorem 5.2).

Lemma 5.33 *Suppose that $\mathcal{T}_{h_f}$ is non-degenerate and has no triangle with two edges on ∂D_f and consider the operator $\widehat{\mathbf{\Pi}}_\tau$ defined in (5.117). Then*

$$\|\widehat{\mathbf{\Pi}}_\tau q_{h_f}\|_{\boldsymbol{L}^2(D_f)^2} \preceq |q_{h_f}|_{H^1(D_f)} \quad \textit{for all } q_{h_f} \in M_{h_f}^\circ,$$

and there exists a positive constant such that:

$$\int_{D_f} \widehat{\mathbf{\Pi}}_\tau q_{h_f} \cdot \nabla q_{h_f} \succeq |q_{h_f}|^2_{H^1(D_f)} \succeq \|q_{h_f}\|^2_{L^2(D_f)} \quad \textit{for all } q_{h_f} \in M_{h_f}^\circ.$$

From Lemma 5.33 and the boundedness of $\widehat{\mathbf{\Pi}}_\tau$, the spaces $\boldsymbol{W}_{h_f}^{\boldsymbol{\tau}}$ (with the $L^2(D_f)$-norm) and $M_{h_f}^\circ$ (with the $H^1(D_f)$-norm) satisfy the inf-sup condition independent of h_f with respect to the bilinear form defined in (5.15) by:

$$b_f(\boldsymbol{v}_f, q_f) := -(q_f, \nabla \cdot \boldsymbol{v}_f)_{D_f} \quad \text{for all } \boldsymbol{v}_f \in \boldsymbol{X}_f \text{ and } q_f \in M_f^\circ.$$

Also observe that if $\boldsymbol{v}_f \in \boldsymbol{W}_{h_f}^{\boldsymbol{\tau}}$ then $\boldsymbol{v}_f \cdot \boldsymbol{\eta} = 0$ on ∂D_f and then $b_f(\boldsymbol{v}_f, q_f) = \int_{D_f} \boldsymbol{v}_f \cdot \nabla q_f$ by the Green formula. Then, according to the Brezzi's splitting theorem, see Braess [2001] and Girault and Raviart [1986], we can always obtain a stable solution $\boldsymbol{w} \in \boldsymbol{W}_{h_f}^{\boldsymbol{\tau}}$ of:

$$\left\{ \begin{array}{lll} (\boldsymbol{w}, \boldsymbol{v}_f)_{D_f} + b_f(\boldsymbol{v}_f, p_f) & = (\boldsymbol{z}, \boldsymbol{v}_f)_{D_f} & \text{for all } \boldsymbol{v}_f \in \boldsymbol{W}_{h_f}^{\boldsymbol{\tau}} \\ b_f(\boldsymbol{w}, q_f) & = b_f(\boldsymbol{z}, q_f)_{D_f} & \text{for all } q_f \in M_{h_f}^\circ, \end{array} \right. \tag{5.118}$$

where $\boldsymbol{z} \in L^2(D_f)^2$.

Given $\boldsymbol{z}$, denote by $\boldsymbol{\Upsilon}_\tau \boldsymbol{z}$ the solution of (5.118). Then

$$\|\boldsymbol{\Upsilon}_\tau \boldsymbol{z}\|_{L^2(D_f)^2} \preceq \|\boldsymbol{z}\|_{L^2(D_f)^2}, \tag{5.119}$$

and $b_f(\boldsymbol{\Upsilon}_\tau \boldsymbol{z}, q_{h_f}) = b_f(\boldsymbol{z}, q_{h_f})$ for $q_{h_f} \in M_{h_f}^\circ$.

In order to proof Lemma 5.19 define

$$\boldsymbol{I}_{h_f}^{TH} \boldsymbol{v}_f := \boldsymbol{\Upsilon}_\eta \boldsymbol{v}_f + \boldsymbol{\Upsilon}_\tau (\boldsymbol{v}_f - \boldsymbol{\Upsilon}_\eta \boldsymbol{v}_f).$$

Observe that:

$$|\boldsymbol{I}_{h_f}^{TH} \boldsymbol{v}_f|_{H^1(D_f)^2} \leq |\boldsymbol{\Upsilon}_\eta \boldsymbol{v}_f|_{H^1(D_f)^2} + |\boldsymbol{\Upsilon}_\tau (\boldsymbol{v}_f - \boldsymbol{\Upsilon}_\eta \boldsymbol{v}_f)|_{H^1(D_f)^2}$$

$$\begin{aligned}
&\preceq |\boldsymbol{v}_f|_{H^1(D_f)^2} + \frac{1}{h_f}\|\boldsymbol{\Upsilon}_\tau(\boldsymbol{v}_f - \boldsymbol{\Upsilon}_\eta \boldsymbol{v}_f)\|_{L^2(D_f)^2} \quad \text{by (5.110) and inverse estimate.}\\
&\preceq |\boldsymbol{v}_f|_{H^1(D_f)^2} + \frac{1}{h_f}\|\boldsymbol{v}_f - \boldsymbol{\Upsilon}_\eta \boldsymbol{v}_f\|_{L^2(D_f)^2} \quad \text{by (5.119)}\\
&\preceq |\boldsymbol{v}_f|_{H^1(D_f)^2} + |\boldsymbol{v}_f|_{H^1(D_f)^2} \quad \text{by (5.111).}
\end{aligned}$$

Then the operator $\boldsymbol{I}_{h_f}^{TH}$ is bounded (with constant independent of h_f). In addition for $p_{h_f} \in M_{h_f}^\circ$ we get:

$$\begin{aligned}
b_f(\boldsymbol{I}_{h_f}^{TH}\boldsymbol{v}_f, p_{h_f}) &= b_f(\boldsymbol{\Upsilon}_\eta \boldsymbol{v}_f, p_{h_f}) + b_f(\boldsymbol{\Upsilon}_\tau(\boldsymbol{v}_f - \boldsymbol{\Upsilon}_\eta \boldsymbol{v}_f), p_{h_f})\\
&= b_f(\boldsymbol{\Upsilon}_\eta \boldsymbol{v}_f, p_{h_f}) + b_f(\boldsymbol{v}_f - \boldsymbol{\Upsilon}_\eta \boldsymbol{v}_f, p_{h_f}) \quad \text{by definition of } \boldsymbol{\Upsilon}_\tau.\\
&= b_f(\boldsymbol{v}_f, p_{h_f}).
\end{aligned}$$

To obtain (5.56) observe that from definition of $\boldsymbol{I}_{h_f}^{TH}$ we have:

$$\begin{aligned}
&\|\boldsymbol{v}_f - \boldsymbol{I}_{h_f}^{TH}\boldsymbol{v}_f\|_{L^2(D_f)^2}\\
&\quad\leq \|\boldsymbol{v}_f - \boldsymbol{\Upsilon}_\eta \boldsymbol{v}_f\|_{L^2(D_f)^2} + \|\boldsymbol{\Upsilon}_\tau(\boldsymbol{v}_f - \boldsymbol{\Upsilon}_\eta \boldsymbol{v}_f)\|_{L^2(D_f)^2}\\
&\quad\preceq \|\boldsymbol{v}_f - \boldsymbol{\Upsilon}_\eta \boldsymbol{v}_f\|_{L^2(D_f)^2} + \|\boldsymbol{v}_f - \boldsymbol{\Upsilon}_\eta \boldsymbol{v}_f\|_{L^2(D_f)^2} \quad \text{by (5.119)}\\
&\quad\preceq h_f^s |\boldsymbol{v}_f|_{H^s(D_f)^2} \quad s = 1, 2. \quad \text{by (5.111).}
\end{aligned}$$

The proof of (5.57) is similar. Inequality (5.59) is obtained from (5.113).

Acknowledgement

We would like to thank the referees and Professor M. Dryja for their useful comments, which helped to improve the presentation.

Bibliography

Arbogast, T., Cowsar, L. C., Wheeler, M. F., and Yotov, I. (2000). Mixed finite element methods on nonmatching multiblock grids. *SIAM J. Numer. Anal.*, 37(4):1295–1315.

Arbogast, T. and Lehr, H. L. (2006). Homogenization of a Darcy-Stokes system modeling vuggy porous media. *Comput. Geosci.*, 10(3):291–302.

Beavers, G. S. and Joseph, D. D. (1967). Boundary conditions at a naturally permeable wall. *J. Fluid Mech.*, 30:197–207.

Bernardi, C., Maday, Y., and Patera, A. T. (1994). A new nonconforming approach to domain decomposition: the mortar element method. In *Nonlinear partial differential equations and their applications. Collège de France Seminar, Vol. XI (Paris, 1989–1991)*, volume 299 of *Pitman Res. Notes Math. Ser.*, pages 13–51. Longman Sci. Tech., Harlow.

Braess, D. (2001). *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics*. Cambridge University Press, Cambridge. Second Edition.

Brenner, S. C. and Scott, L. R. (1994). *The mathematical theory of finite element methods*, volume 15 of *Texts in Applied Mathematics*. Springer-Verlag, New York.

Brezzi, F. and Fortin, M. (1991). *Mixed and hybrid finite element methods*, volume 15 of *Springer Series in Computational Mathematics*. Springer-Verlag, New York.

Burman, E. and Hansbo, P. (2007). A unified stabilized method for Stokes' and Darcy's equations. *J. Comput. Appl. Math.*, 198(1):35–51.

Clément, P. (1975). Approximation by finite element functions using local regularization. *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. RAIRO Analyse Numérique*, 9(R-2):77–84.

Discacciati, M. (2004). *Domain decomposition methods for the coupling of surface and groundwater flows*. PhD thesis, Ecole Polytechnique Fédérale, Lausanne (Switzerland). Thèse n. 3117.

Discacciati, M., Miglio, E., and Quarteroni, A. (2002). Mathematical and numerical models for coupling surface and groundwater flows. *Appl. Numer. Math.*, 43(1-2):57–74. 19th Dundee Biennial Conference on Numerical Analysis (2001).

Discacciati, M. and Quarteroni, A. (2003). Analysis of a domain decomposition method for the coupling of Stokes and Darcy equations. In Brezzi, F., Buffa, A., Corsaro, S., and Murli, A., editors, *ENUMATH 2001*, Numerical Mathematics and Advanced Applications, pages 3–20. Springer-Verlag.

Discacciati, M. and Quarteroni, A. (2004). Convergence analysis of a subdomain iterative method for the finite element approximation of the coupling of Stokes and Darcy equations. *Comput. Vis. Sci.*, 6(2-3):93–103.

Galvis, J. and Sarkis, M. (2006). Balancing domain decomposition methods for mortar coupling Stokes-Darcy systems. In Keyes, D. and Widlund, O. B., editors, *Domain Decomposition Methods in Science and Engineering XVI*, volume 55 of *Lecture Notes in Computational Science and Engineering*, pages 373–380. Springer.

Girault, V. and Raviart, P.-A. (1986). *Finite element methods for Navier-Stokes equations*, volume 5 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin. Theory and algorithms.

Girault, V., Rivière, B., and Wheeler, M. F. (2005). A discontinuous Galerkin method with nonoverlapping domain decomposition for the Stokes and Navier-Stokes problems. *Math. Comp.*, 74(249):53–84.

Grisvard, P. (1985). *Elliptic problems in nonsmooth domains*, volume 24 of *Monographs and Studies in Mathematics*. Pitman (Advanced Publishing Program), Boston, MA.

Hornung, U., editor (1997). *Homogenization and porous media*, volume 6 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, New York.

Jäger, W. and Mikelić, A. (2000). On the interface boundary condition of Beavers, Joseph, and Saffman. *SIAM J. Appl. Math.*, 60(4):1111–1127.

Layton, W. J., Schieweck, F., and Yotov, I. (2002). Coupling fluid flow with porous media flow. *SIAM J. Numer. Anal.*, 40(6):2195–2218 (2003).

Mardal, K. A., Tai, X.-C., and Winther, R. (2002). A robust finite element method for Darcy-Stokes flow. *SIAM J. Numer. Anal.*, 40(5):1605–1631.

Mathew, T. P. (1993). Schwarz alternating and iterative refinement methods for mixed formulations of elliptic problems. II. Convergence theory. *Numer. Math.*, 65(4):469–492.

Nečas, J. (1967). *Les méthodes directes en théorie des équations elliptiques*. Masson et Cie, Éditeurs, Paris.

Quarteroni, A. and Valli, A. (1999). *Domain decomposition methods for partial differential equations*. Numerical Mathematics and Scientific Computation. The Clarendon Press Oxford University Press, New York. , Oxford Science Publications.

Quarteroni, A., Veneziani, A., and Zunino, P. (2002). A domain decomposition method for advection-diffusion processes with application to blood solutes. *SIAM J. Sci. Comput.*, 23(6):1959–1980.

Rivière, B. and Yotov, I. (2005). Locally conservative coupling of Stokes and Darcy flows. *SIAM J. Numer. Anal.*, 42(5):1959–1977.

Saffman, P. (1971). On the boundary condition at the surface of a porous media. *Stud. Appl. Math.*, 50:93–101.

Scott, L. R. and Zhang, S. (1990). Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comp.*, 54(190):483–493.

Wohlmuth, B. I. (2000). A mortar finite element method using dual spaces for the Lagrange multiplier. *SIAM J. Numer. Anal.*, 38(3):989–1012.

Wohlmuth, B. I., Toselli, A., and Widlund, O. B. (2000). An iterative substructuring method for Raviart-Thomas vector fields in three dimensions. *SIAM J. Numer. Anal.*, 37(5):1657–1676.

Chapter 6

BDD and FETI Methods for Mortar Coupling of Stokes-Darcy Systems

We consider the coupling across an interface of a fluid flow and a porous media flow. The differential equations involve Stokes' equations in the fluid region and Darcy's equations in the porous region, and coupled through an interface with adequate transmission conditions. The discretization consists of $P2/P1$ triangular Taylor-Hood finite elements, the lowest order triangular Raviart-Thomas finite elements, and the mortar piecewise constant Lagrange multipliers on the interface. Nonmatching meshes across the interface are allowed. Due to the small values of the permeability parameter κ of the porous medium, the resulting discrete symmetric saddle point system is very ill conditioned. We design and analyze two preconditioners, one based on Balancing Domain Decomposition (BDD) methods and the other one based on Finite Element by Tearing and Interconnecting (FETI) methods. For both methods, we derive condition number estimates of order $C_1(1+\frac{1}{\kappa})$. In case the fluid discretization is finer than the porous side discretization, we derive a better estimate of order $C_2(\frac{\kappa+1}{\kappa+(h^p)^2})$ for the FETI preconditioner. Here h^p is the mesh size of the porous side triangulation. The constants C_1 and C_2 are independent of the permeability κ, the fluid viscosity ν, and the mesh ratio across the interface. Numerical experiments confirm the sharpness of the theoretical estimates.

6.1 Introduction and problem setting

We consider the coupling across an interface of a fluid flow and a porous media flow. The model consists of Stokes' equations in the fluid region, Darcy's equations for the filtration velocity in the porous medium, and a coupling through an interface with adequate transmission conditions. Such problem appears in several applications like well-reservoir coupling in petroleum engineering, transport of substances across groundwater and surface water, and (bio)fluid-organ interactions. There are some works that address numerical analysis issues of this model. For inf-sup conditions and approximation results associated to the continuous and discrete formulations for Stokes-Darcy systems we refer Galvis [2004], Galvis and

Sarkis [2007b] and Layton et al. [2002], and for Stokes-Laplacian systems we refer Discacciati et al. [2002] and Discacciati and Quarteroni [2003]. For mortar discretization analysis we mention Galvis and Sarkis [2007b] and Rivière and Yotov [2005], while for preconditioning analysis for Stokes-Laplacian systems we refer Discacciati [2004, 2005] and Discacciati and Quarteroni [2004]. Here in this paper, we are interested on preconditioners for *Stokes-Mortar-Darcy* systems with *flux boundary conditions*, therefore, the global system as well as the local systems require flux compatibilities. We propose and analyze two preconditioners based on Balancing Domain Decomposition (BDD) methods and Finite Element by Tearing and Interconnecting (FETI) methods, respectively, and present numerical experiments in order to verify the theory. In the BDD method, the energy of the preconditioner is controlled by the Stokes system, while in the FETI method, the energy is controlled by the Darcy system. For general references, we mention Achdou et al. [1999], Brenner and Sung [2007], Dryja and Proskurowski [2003], Mandel [1993], Mandel and Brezina [1996] and Pavarino and Widlund [2002], Toselli and Widlund [2005] for BDD methods and Brenner and Sung [2007], Dryja and Proskurowski [2003], Farhat and Roux [1991], Klawonn and Widlund [2001], Mandel and Tezaur [1996] and Toselli and Widlund [2005] for FETI methods. Besides the preliminary work presented in Galvis and Sarkis [2006], up to our knowledge, algorithms and theoretical analysis concerning BDD and FETI type preconditioners for Stokes-Darcy coupling are missing.

In this paper we consider Taylor-Hood finite element methods for the free fluid side and lowest order Raviart-Thomas element for the porous side. The BDD and FETI analysis to be developed here can also be straightforwardly extended to the three dimensional case and to other discretizations, e.g., the $P2/P0$ coupled with Raviart-Thomas; see Discacciati et al. [2002], Galvis and Sarkis [2007b], Layton et al. [2002] and Rivière and Yotov [2005]. Here, we consider only the two subdomain case, i.e., when exact local solvers are applied to the Stokes and Darcy regions. For the case of discontinuous pressure finite elements, BDD and FETI type methods based on partitions of the Stokes and Darcy domains can be considered as well. In this case, the most difficult part of designing and analyzing the methods is related to the subdomains that touch the Stokes/Darcy interface, and this is the case studied here. Subdomains that do not touch the Stokes/Darcy interface Γ can be treated as in the classical versions of BDD and FETI type algorithms. We also mention that extension to Balancing Domain Decomposition with Constraints (BDDC) and dual-primal FETI (FETI-DP) methods can also be considered; see Brenner and Sung [2007], Dohrmann [2003], Dryja et al. [2005], Farhat et al. [2000], Li [2005], Li and Widlund [2006, 2007] and Tu [2005].

Let $D^f,\ D^p \subset \mathbb{R}^n$ be polyhedral subdomains, define $D := \text{int}(\overline{D}^f \cup \overline{D}^p)$ and $\Gamma := \partial D^f \cap \partial D^p$, with outward unit normal vectors $\boldsymbol{\eta}^i$ on ∂D^i, $i = f, p$. The tangent vectors on Γ are denoted by $\boldsymbol{\tau}_1$ $(n = 2)$, or $\boldsymbol{\tau}_l$, $l = 1, 2$ $(n = 3)$. The exterior boundaries are $\Gamma^i := \partial D^i \setminus \Gamma$, $i = f, p$. Fluid velocities are denoted by $\boldsymbol{u}^i : D^i \to \mathbb{R}^n$, $i = f, p$, and pressures by $p^i : D^i \to \mathbb{R}$, $i = f, p$.

We consider Stokes equations in the fluid region D^f and Darcy equations for

the filtration velocity in the porous medium D^p. More precisely, we have the following systems of equations in each subdomain:

$$\begin{array}{cc} \text{Stokes' equations} & \text{Darcy's equations} \\ \left\{\begin{array}{rll} -\nabla\cdot T(\boldsymbol{u}^f,p^f) &= \boldsymbol{f}^f & \text{in } D^f \\ \nabla\cdot\boldsymbol{u}^f &= g^f & \text{in } D^f \\ \boldsymbol{u}^f &= \boldsymbol{h}^f & \text{on } \Gamma^f \end{array}\right. & \left\{\begin{array}{rll} \boldsymbol{u}^p &= -\frac{\kappa}{\nu}\nabla p^p & \text{in } D^p \\ \nabla\cdot\boldsymbol{u}^p &= g^p & \text{in } D^p \\ \boldsymbol{u}^p\cdot\boldsymbol{\eta}^p &= h^p & \text{on } \Gamma^p. \end{array}\right. \end{array} \tag{6.1}$$

Here $T(\boldsymbol{v},p) := -pI + 2\nu\boldsymbol{D}\boldsymbol{v}$, where ν is the fluid viscosity, $\boldsymbol{D}\boldsymbol{v} := \frac{1}{2}(\nabla\boldsymbol{v} + \nabla\boldsymbol{v}^T)$ is the linearized strain tensor and κ denotes the rock permeability. For simplicity on the analysis, we assume that κ is a real positive constant. We impose the following conditions:

1. *Interface matching conditions across* Γ; see Discacciati et al. [2002], Discacciati and Quarteroni [2003, 2004] and Layton et al. [2002] and references therein.

   (a) *Conservation of mass across* Γ*:* $\boldsymbol{u}^f\cdot\boldsymbol{\eta}^f + \boldsymbol{u}^p\cdot\boldsymbol{\eta}^p = 0$ on Γ.

   (b) *Balance of normal forces across* Γ*:* $p^f - 2\nu\boldsymbol{\eta}^{fT}\boldsymbol{D}(\boldsymbol{u}^f)\boldsymbol{\eta}^f = p^p$ on Γ.

   (c) *Beavers-Joseph-Saffman condition:* This condition is a kind of empirical law that gives an expression for the component of the Cauchy stress tensor in the tangential direction of the interface Γ; see the works Beavers and Joseph [1967] and Jäger and Mikelić [2000]. It is expressed by:

$$\boldsymbol{u}^f\cdot\boldsymbol{\tau}_l = -\frac{\sqrt{\kappa}}{\alpha^f}2\boldsymbol{\eta}^{fT}\boldsymbol{D}(\boldsymbol{u}^f)\boldsymbol{\tau}_l \quad l = 1, n-1; \text{ on } \Gamma.$$

2. *Compatibility condition:* The divergence and boundary data satisfy (see Galvis and Sarkis [2007b]),

$$\langle g^f, 1\rangle_{D^f} + \langle g^p, 1\rangle_{D^p} - \langle \boldsymbol{h}^f\cdot\boldsymbol{\eta}^f, 1\rangle_{\Gamma^f} - \langle h^p, 1\rangle_{\Gamma^p} = 0.$$

6.2 Weak formulation

In this section we present the weak version of the coupled system of partial differential equations introduced above. Without loss of generality, we consider $\boldsymbol{h}^f = \boldsymbol{0}$, $g^f = 0$, $h^p = 0$ and $g^p = 0$ in (6.1); see Galvis and Sarkis [2007b].

The problem can be formulated as: *Find* $(\boldsymbol{u}, p, \lambda) \in \boldsymbol{X}\times M_0\times\Lambda$ *such that for all* $(\boldsymbol{v}, q, \mu)\in\boldsymbol{X}\times M_0\times\Lambda$:

$$\left\{\begin{array}{ll} a(\boldsymbol{u},\boldsymbol{v}) + b(\boldsymbol{v},p) + b_\Gamma(\boldsymbol{v},\lambda) &= f(\boldsymbol{v}) \\ b(\boldsymbol{u},q) &= 0 \\ b_\Gamma(\boldsymbol{u},\mu) &= 0, \end{array}\right. \tag{6.2}$$

where $\boldsymbol{X} = \boldsymbol{X}^f\times\boldsymbol{X}^p := H_0^1(D^f,\Gamma^f)^n\times\boldsymbol{H}_0(\mathrm{div}, D^p,\Gamma^p)$ and M_0 is the subset of $M := L^2(D^f)\times L^2(D^p)\equiv L^2(D)$ of pressures with zero average value in D. Here

$H_0^1(D^f, \Gamma^f)$ denotes the subspace of $H^1(D^f)$ of functions that vanish on Γ^f. The space $\boldsymbol{H}_0(\mathrm{div}, D^p, \Gamma^p)$ consists of functions in $\boldsymbol{H}(\mathrm{div}, D^p)$ with zero normal trace on Γ^p, where

$$\boldsymbol{H}(\mathrm{div}, D^p) := \left\{ \boldsymbol{v} \in L^2(D^p)^n : \mathrm{div}\, \boldsymbol{v} \in L^2(D^p) \right\}.$$

For the Lagrange multiplier space we consider $\Lambda := H^{1/2}(\Gamma)$. See Galvis and Sarkis [2007b] for a discussion on the choice of the Lagrange multipliers space Λ and how to derive the weak formulation (6.2) and other equivalent weak formulations; see also Layton et al. [2002].

The global bilinear forms are

$$a(\boldsymbol{u}, \boldsymbol{v}) := a^f_{\alpha^f}(\boldsymbol{u}^f, \boldsymbol{v}^f) + a^p(\boldsymbol{u}^p, \boldsymbol{v}^p)$$

and

$$b(\boldsymbol{v}, p) := b^f(\boldsymbol{v}^f, p^f) + b^p(\boldsymbol{v}^p, p^p),$$

with local bilinear forms $a^f_{\alpha^f}, b^f$ and b^p defined by

$$\begin{aligned} a^f_{\alpha^f}(\boldsymbol{u}^f, \boldsymbol{v}^f) \;&:=\; 2\nu(\boldsymbol{D}\boldsymbol{u}^f, \boldsymbol{D}\boldsymbol{v}^f)_{D^f} \\ &\quad + \sum_{\ell=1}^{n-1} \frac{\nu\alpha^f}{\sqrt{\kappa}} \langle \boldsymbol{u}^f \cdot \boldsymbol{\tau}_\ell, \boldsymbol{v}^f \cdot \boldsymbol{\tau}_\ell \rangle_\Gamma, \;\; \boldsymbol{u}^f, \boldsymbol{v}^f \in \boldsymbol{X}^f, \end{aligned} \tag{6.3}$$

$$a^p(\boldsymbol{u}^p, \boldsymbol{v}^p) := (\frac{\nu}{\kappa}\boldsymbol{u}^p, \boldsymbol{v}^p)_{D^p}, \quad \boldsymbol{u}^p, \boldsymbol{v}^p \in \boldsymbol{X}^p, \tag{6.4}$$

$$b^f(\boldsymbol{v}^f, q^f) := -(q^f, \nabla \cdot \boldsymbol{v}^f)_{D^f}, \quad \boldsymbol{v}^f \in \boldsymbol{X}^f, \; q^f \in M^f, \tag{6.5}$$

$$b^p(\boldsymbol{v}^p, p^p) := -(p^p, \nabla \cdot \boldsymbol{v}^p)_{D^p}, \quad \boldsymbol{v}^p \in \boldsymbol{X}^p, \; p^p \in M^p, \tag{6.6}$$

and with weak conservation of mass bilinear form defined by

$$b_\Gamma(\boldsymbol{v}, \mu) := \langle \boldsymbol{v}^f \cdot \boldsymbol{\eta}^f, \mu \rangle_\Gamma + \langle \boldsymbol{v}^p \cdot \boldsymbol{\eta}^p, \mu \rangle_\Gamma, \quad \boldsymbol{v} = (\boldsymbol{v}^f, \boldsymbol{v}^p) \in \boldsymbol{X}, \mu \in \Lambda. \tag{6.7}$$

The second duality pairing of (6.7) is interpreted as

$$\langle \boldsymbol{v}^p \cdot \boldsymbol{\eta}^p, E\boldsymbol{\eta}^p(\mu) \rangle_{\partial D^p}.$$

Here $E\boldsymbol{\eta}^p$ is any continuous lift-in operator from $H^{1/2}(\Gamma)$ to $H^{1/2}(\partial D^p)$; recall that $\Gamma \subset \partial D^p$. See Galvis [2004] and Galvis and Sarkis [2007b].

The functional f in the right hand side of (6.2) is defined by

$$f(\boldsymbol{v}) := f^f(\boldsymbol{v}^f) + f^p(\boldsymbol{v}^p), \quad \text{for all } \boldsymbol{v} = (\boldsymbol{v}^f, \boldsymbol{v}^p) \in \boldsymbol{X},$$

where $f^i(\boldsymbol{v}^i) := (\boldsymbol{f}^i, \boldsymbol{v}^i)_{L^2(D^i)}$ for all $\boldsymbol{v}^i \in \boldsymbol{X}^i$, $i = f, p$.

The bilinear forms $a^f_{\alpha^f}, b^f$ are associated to Stokes' equations and the bilinear forms a^p, b^p to Darcy law. The bilinear form $a^f_{\alpha^f}$ includes interface matching conditions 1.*b* and 1.*c* above. The bilinear form b_Γ is used to impose the weak version of the interface matching condition 1.*a* above. For the analysis of this weak formulation and the well-posedness of the problem; see Galvis and Sarkis [2007b] and Layton et al. [2002].

6.3 Discretization

From now on we consider only the two dimensional case. We note that the ideas developed in the following can be easily extended to case of three dimensional subdomains.

We assume that D^i, $i = f, p$, are *two dimensional* polygonal subdomains. Let $\mathcal{T}^i_{h^i}(D^i)$ be a geometrically conforming shape regular and quasi-uniform triangulation of D^i with mesh size parameter h^i, $i = f, p$. We do not assume that these two triangulations match at the interface Γ. For the fluid region, let $\boldsymbol{X}^f_{h^f}$ and $M^f_{h^f}$ be $P2/P1$ triangular Taylor-Hood finite elements; see Braess [2001], Brenner and Scott [1994] and Brezzi and Fortin [1991]. More precisely,

$$\boldsymbol{X}^f_{h^f} := \left\{ \boldsymbol{u} \in \boldsymbol{X}^f \ : \ \begin{array}{l} \forall K \in \mathcal{T}^f_{h^f}(D_f),\ \boldsymbol{u}_K = \hat{\boldsymbol{u}}_K \circ \\ F_K^{-1} \text{ and } \hat{\boldsymbol{u}}_K \in P_2(\hat{K})^2 \end{array} \right\} \cap C^0(\overline{D}^f)^2, \tag{6.8}$$

where $\boldsymbol{u}_K := \boldsymbol{u}|_K$ and

$$M^f_{h^f} := \left\{ p \in L^2(D^f) \ : \ \begin{array}{l} \forall K \in \mathcal{T}^f_{h^f}(D_f),\ p_K = \hat{p}_K \circ \\ F_K^{-1} \text{ and } \hat{p}_K \in P_1(\hat{K}) \end{array} \right\} \cap C^0(\overline{D}^f).$$

Denote $\mathring{M}^f_{h^f} \subset M^f_{h^f}$ the discrete fluid pressures with zero average value in D^f. For the porous region, let $\boldsymbol{X}^p_{h^p} \subset \boldsymbol{X}^p$ and $M^p_{h^p} \subset L^2(D^p)$ be the lowest order Raviart-Thomas finite elements based on triangles; see Braess [2001] and Brezzi and Fortin [1991]. Let $\mathring{M}^p_{h^p} \subset M^p_{h^p}$ be the subset of pressures in $M^p_{h^p}$ with zero average value in D^p.

Define $\boldsymbol{X}_h := \boldsymbol{X}^f_{h^f} \times \boldsymbol{X}^p_{h^p} \subset \boldsymbol{X}$ and $M_h := M^f_{h^f} \times M^p_{h^p} \subset L^2(D^f) \times L^2(D^p)$. Note that in the definition of the discrete velocities we assume that the boundary conditions are included, i.e., for $\boldsymbol{v}^f_{h^f} \in \boldsymbol{X}^f_{h^f}$ we have $\boldsymbol{v}^f_{h^f} = \boldsymbol{0}$ on Γ^f and for $\boldsymbol{v}^p_{h^p} \in \boldsymbol{X}^p_{h^p}$ we have that $\boldsymbol{v}^p_h \cdot \boldsymbol{\eta}^p = 0$ on Γ^p.

Let $\mathcal{T}^p_{h^p}(\Gamma)$ be the restriction to Γ of the porous side triangulation $\mathcal{T}^p_{h^p}(D^p)$. For the Lagrange multipliers space we choose piecewise constant functions on Γ with respect to the triangulation $\mathcal{T}^p_{h^p}(\Gamma)$,

$$\Lambda_{h^p} := \left\{ \lambda \ : \ \lambda|_{e^p_j} = \lambda_{e^p_j} \text{ is constant in each edge } e^p_j \text{ of } \mathcal{T}^p_{h^p}(\Gamma) \right\}, \tag{6.9}$$

i.e., the *master* is on the fluid region side and the *slave* is on the porous region side; see Ben Belgacem and Maday [1997], Bernardi et al. [1994], Dryja and Proskurowski [2003] and Wohlmuth [2000]. The choice of piecewise constant Lagrange multipliers leads to a nonconforming approximation on Λ_{h^p} since piecewise constant functions do not belong to $H^{1/2}(\Gamma)$. For the analysis of this nonconforming discretization and a priori error estimates we refer to Galvis and Sarkis [2007b].

6.4 Primal and dual formulations

In order to simplify notation and since there is no danger of confusion, we will denote the finite element functions and the corresponding vector representation by the same symbol, i.e., when writing finite element functions we will drop the indices h^i. Recall that we have the pair of spaces $(\boldsymbol{X}_h, M_h)$ associated to the coupled problem and spaces associated to each subproblem: $(\boldsymbol{X}^f_{h^f}, M^f_{h^f})$ and $(\boldsymbol{X}^p_{h^p}, M^p_{h^p})$. We will keep the subscript h^i, $i = f, p$, in the notation for local subspaces $\boldsymbol{X}^f_{h^f}, M^f_{h^f}, \boldsymbol{X}^p_{h^p}$ and $M^p_{h^p}$.

Since we are interested in preconditioning issues we assume $\alpha^f = 0$ in the definition of the fluid side local bilinear form $a^f_{\alpha^f}$ in (6.3). We denote $a^f = a^f_0$; see Remark 6.9 for the case $\alpha^f > 0$.

With the discretization chosen in Section 6.3 we obtain the following symmetric saddle point linear system

$$\left[\begin{array}{cc|cc|c} A^f & B^{fT} & 0 & 0 & C^{fT} \\ B^f & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & A^p & B^{pT} & -C^{pT} \\ 0 & 0 & B^p & 0 & 0 \\ \hline C^f & 0 & -C^p & 0 & 0 \end{array}\right] \left[\begin{array}{c} \boldsymbol{u}^f \\ p^f \\ \hline \boldsymbol{u}^p \\ p^p \\ \hline \lambda \end{array}\right] = \left[\begin{array}{c} \boldsymbol{f}^f \\ g^f \\ \hline \boldsymbol{f}^p \\ g^p \\ \hline 0 \end{array}\right] \tag{6.10}$$

with matrices A^i, B^i, C^i and columns vectors $\boldsymbol{f}^i, g^i$, $i = f, p$, defined by

$$\begin{array}{lcl} a^i(\boldsymbol{u}^i, \boldsymbol{v}^i) & = & \boldsymbol{v}^{iT} A^i \boldsymbol{u}^i, \\ b^i(\boldsymbol{u}^i, q^i) & = & q^{iT} B^i \boldsymbol{u}^i, \\ (\boldsymbol{u}^i \cdot \boldsymbol{\eta}^f, \mu)_\Gamma & = & \mu^T C^i \boldsymbol{u}^i, \\ f^i(\boldsymbol{v}^i) & = & \boldsymbol{v}^{iT} \boldsymbol{f}^i, \\ g^i(q^i) & = & q^{iT} g^i. \end{array} \tag{6.11}$$

Matrix A^f corresponds to ν times the discrete version of the linearized stress tensor on D^f. Note that in the case $\alpha^f > 0$, the bilinear form $a^f_{\alpha^f}$ in (6.3) includes a boundary term; see Remark 6.9. The matrix A^p corresponds to ν/κ times a discrete L^2-norm on D^p. Matrix $-B^i$ is the discrete divergence in D^i, $i = f, p$, and matrices C^f and C^p correspond to the matrix form of the discrete conservation of mass on Γ. Note that ν can be viewed as a scaling factor since it appears in both matrices A^f and A^p. Therefore, it is not relevant for preconditioning issues.

Consider the following partition of the degrees of freedom: For $i = f, p$, let

$$\left[\begin{array}{c} \boldsymbol{u}^i_I \\ p^i_I \\ u^i_\Gamma \\ \bar{p}^i \end{array}\right] \begin{array}{l} \text{Interior displacements + tangential velocities on } \Gamma, \\ \text{Pressures with zero average in } D^i, \\ \text{Interface outward } \textit{normal velocities} \text{ on } \Gamma, \\ \text{Constant pressure in } D^i. \end{array}$$

Then, for $i = f, p$, we have the block structure:

$$A^i = \left[\begin{array}{cc} A^i_{II} & A^{iT}_{\Gamma I} \\ A^i_{\Gamma I} & A^i_{\Gamma\Gamma} \end{array}\right], \quad B^i = \left[\begin{array}{cc} B^i_{II} & B^{iT}_{\Gamma I} \\ 0 & \tilde{B}^{iT} \end{array}\right] \text{ and } \quad C^i = \left[\begin{array}{cccc} 0 & 0 & \tilde{C}^i & 0 \end{array}\right].$$

Note that the $(2,1)$ entry of B^i corresponds to integrating an *interior* velocity against a constant pressure, then it vanishes due to the divergence theorem. We have the following matrix representation of the coupled problem in (6.10):

$$\left[\begin{array}{cccc|cccc|c} A^f_{II} & B^{fT}_{II} & A^{fT}_{\Gamma I} & 0 & 0 & 0 & 0 & 0 & 0 \\ B^f_{II} & 0 & B^{fT}_{\Gamma I} & 0 & 0 & 0 & 0 & 0 & 0 \\ A^f_{\Gamma I} & B^{fT}_{I\Gamma} & A^f_{\Gamma\Gamma} & \bar{B}^{fT} & 0 & 0 & 0 & 0 & \tilde{C}^{fT} \\ 0 & 0 & \bar{B}^f & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & A^p_{II} & B^{pT}_{II} & A^{pT}_{\Gamma I} & 0 & 0 \\ 0 & 0 & 0 & 0 & B^p_{II} & 0 & B^p_{I\Gamma} & 0 & 0 \\ 0 & 0 & 0 & 0 & A^p_{\Gamma I} & B^{pT}_{I\Gamma} & A^p_{\Gamma\Gamma} & \bar{B}^{pT} & -\tilde{C}^{pT} \\ 0 & 0 & 0 & 0 & 0 & 0 & \bar{B}^p & 0 & 0 \\ \hline 0 & 0 & \tilde{C}^f & 0 & 0 & 0 & -\tilde{C}^p & 0 & 0 \end{array}\right] \left[\begin{array}{c} \boldsymbol{u}^f_I \\ p^f_I \\ u^f_\Gamma \\ \bar{p}^f \\ \hline \boldsymbol{u}^p_I \\ p^p_I \\ u^p_\Gamma \\ \bar{p}^p \\ \hline \lambda \end{array}\right] = \left[\begin{array}{c} \boldsymbol{f}^f_I \\ g^f_I \\ f^f_\Gamma \\ \bar{g}^f \\ \hline \boldsymbol{f}^p_I \\ g^p_I \\ f^p_\Gamma \\ \bar{g}^p \\ \hline 0 \end{array}\right]. \tag{6.12}$$

Following Dryja and Proskurowski [2003], Pavarino and Widlund [2002], we choose the following matrix representation in each subdomain D^i, $i = f, p$,

$$\left[\begin{array}{cc|cc} A^i_{II} & B^{iT}_{II} & A^{iT}_{\Gamma I} & 0 \\ B^i_{II} & 0 & B^i_{I\Gamma} & 0 \\ \hline A^i_{\Gamma I} & B^{iT}_{I\Gamma} & A^i_{\Gamma\Gamma} & \bar{B}^{iT} \\ 0 & 0 & \bar{B}^i & 0 \end{array}\right] = \left[\begin{array}{c|c} K^i_{II} & K^{iT}_{\Gamma I} \\ \hline K^i_{\Gamma I} & K^i_{\Gamma\Gamma} \end{array}\right]. \tag{6.13}$$

6.4.1 The primal formulation

From the last equation in (6.12) we see that the mortar condition on Γ (using the Darcy side as the slave side) can be imposed as $u^p_\Gamma = (\tilde{C}^p)^{-1}\tilde{C}^f u^f_\Gamma = \Pi u^f_\Gamma$, where Π is the $L^2(\Gamma)$ projection on the space of piecewise constant functions on each subinterval $e^p \in \mathcal{T}^p_{h^p}(\Gamma)$. We note that $\tilde{C}^p$ is a diagonal matrix for the lowest order Raviart-Thomas elements.

Now we eliminate $\boldsymbol{u}^i_I$, p^i_I, $i = f, p$, and λ, to obtain the following (saddle point) Schur complement

$$S \left[\begin{array}{c} u^f_\Gamma \\ \bar{p}^f \\ \bar{p}^p \end{array}\right] = \left[\begin{array}{c} b_\Gamma \\ \bar{b}^f \\ \bar{b}^p \end{array}\right]. \tag{6.14}$$

Here S is given by

$$\begin{aligned} S: &= \left[\begin{array}{ccc} S^f_\Gamma & \bar{B}^{fT} & 0 \\ \bar{B}^f & 0 & 0 \\ 0 & 0 & 0 \end{array}\right] + \tilde{\Pi}^T \left[\begin{array}{ccc} S^p_\Gamma & 0 & \bar{B}^{pT} \\ 0 & 0 & 0 \\ \bar{B}^p & 0 & 0 \end{array}\right] \tilde{\Pi} \\ &= \tilde{S}^f + \tilde{S}^p \\ &= \left[\begin{array}{c|cc} S^f_\Gamma + \Pi^T S^p_\Gamma \Pi & \bar{B}^{fT} & \Pi^T \bar{B}^{pT} \\ \hline \bar{B}^f & 0 & 0 \\ \bar{B}^p \Pi & 0 & 0 \end{array}\right] = \left[\begin{array}{cc} S_\Gamma & \bar{B}^T \\ \bar{B} & 0 \end{array}\right], \end{aligned}$$

where

$$\tilde{\Pi} := \begin{bmatrix} \Pi & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \bar{B}^T := [\bar{B}^{fT} \;\; \Pi^T \bar{B}^{pT}]. \tag{6.15}$$

Here, we have denoted

$$\tilde{S}^f := \begin{bmatrix} S_\Gamma^f & \bar{B}^{fT} & 0 \\ \bar{B}^f & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \tilde{S}^p := \tilde{\Pi}^T \begin{bmatrix} S_\Gamma^p & 0 & \bar{B}^{pT} \\ 0 & 0 & 0 \\ \bar{B}^p & 0 & 0 \end{bmatrix} \tilde{\Pi}. \tag{6.16}$$

The local matrices S_Γ^i and $\bar{B}^i$ and the local Schur complement S^i are given by

$$S^i = \begin{bmatrix} S_\Gamma^i & \bar{B}^{iT} \\ \bar{B}^i & 0 \end{bmatrix} := K_{\Gamma\Gamma}^i - K_{\Gamma I}^i \left(K_{\Gamma\Gamma}^i\right)^{-1} K_{\Gamma I}^{iT}, \quad i = p, f. \tag{6.17}$$

The right hand side of (6.14) is given by

$$\begin{bmatrix} b_\Gamma \\ \bar{b}^f \\ \bar{b}^p \end{bmatrix} = \left\{ \begin{bmatrix} f_\Gamma^f \\ \bar{g}^f \\ 0 \end{bmatrix} - \begin{bmatrix} K_{\Gamma I}^f \left(K_{\Gamma\Gamma}^f\right)^{-1} \begin{bmatrix} \boldsymbol{f}_I^f \\ g_I^f \end{bmatrix} \\ 0 \end{bmatrix} \right\} + \left\{ \begin{bmatrix} \Pi^T f_\Gamma^p \\ 0 \\ \bar{g}^p \end{bmatrix} - \tilde{\Pi}^T \begin{bmatrix} K_{\Gamma I}^p \left(K_{\Gamma\Gamma}^p\right)^{-1} \begin{bmatrix} \boldsymbol{f}_I^p \\ g_I^p \end{bmatrix} \\ 0 \end{bmatrix} \right\}.$$

We note that the reduced system (6.14), as well as the original system (6.12), is solvable when $\bar{b}^f + \bar{b}^p = 0$, and the solution is unique when we restrict to pressures with zero average value on D.

From now on we only work with functions defined on Γ and extended inside the subdomain using the discrete Stokes and Darcy problems. It is convenient to define the space

$$V_\Gamma := \left\{ v_\Gamma = (v_\Gamma^f, v_\Gamma^p) : v_\Gamma^f = \mathcal{SH}(\boldsymbol{v}^f \cdot \boldsymbol{\eta}^f|_\Gamma) \text{ and } v_\Gamma^p = \mathcal{DH}(\boldsymbol{v}^p \cdot \boldsymbol{\eta}^p|_\Gamma) \right\} \tag{6.18}$$

and

$$M_0 := \left\{ q \in M^h : \begin{array}{l} q^i = \text{piecewise const. in } D^i, i = f, p, \\ \text{and } \int_{D^f} q^f + \int_{D^p} q^p = 0 \end{array} \right\}.$$

Here $\mathcal{SH}$ ($\mathcal{DH}$) is the velocity component of the discrete Stokes (Darcy) harmonic extension operator that maps discrete interface normal velocity $u_\Gamma^f \in H_{00}^{1/2}(\Gamma)$ (respectively $u_\Gamma^p \in (H^{1/2}(\Gamma))'$) to the solution of following problem: Find $\boldsymbol{u}^i \in \boldsymbol{X}_{h^i}^i$ and $p^i \in \mathring{M}_{h^i}^i$ such that for all $\boldsymbol{v}^i \in \boldsymbol{X}_{h^i}^i$ and $q^i \in \mathring{M}_{h^i}^i$, $i = f, p$, we have:

$$\begin{cases} a^f(\mathcal{SH}u^f, \boldsymbol{v}^f) + b^f(\boldsymbol{v}^f, p^f) = 0 \\ b^f(\mathcal{SH}u^f, q^f) = 0 \\ \mathcal{SH}u^f \cdot \boldsymbol{\eta}^f = u_\Gamma^f \quad \text{on } \Gamma \\ \mathcal{SH}u^f = \boldsymbol{0} \quad \text{on } \Gamma^f, \end{cases} \tag{6.19}$$

and

$$\begin{cases} a^p(\mathcal{DH}u^p, \boldsymbol{v}^p) + b^p(\boldsymbol{v}^p, p^p) &= 0 \\ b^p(\mathcal{DH}u^p, q^p) &= 0 \\ \mathcal{DH}u^p \cdot \boldsymbol{\eta}^p &= u_\Gamma^p \quad \text{on } \Gamma \\ \mathcal{DH}u^p \cdot \boldsymbol{\eta}^p &= 0 \quad \text{on } \Gamma^p. \end{cases} \tag{6.20}$$

The degrees of freedom associated with $\mathcal{SH}u^f \cdot \boldsymbol{\tau}^f$ on Γ are free. This corresponds to imposing the natural boundary condition $\boldsymbol{\tau}^T \boldsymbol{D}(\mathcal{SH}u^f)\boldsymbol{\eta}_f = 0$ on Γ.

For $i = f, p$, define the normal trace component of $\boldsymbol{X}^i_{h^i}$ by

$$Z^i_{h^i} = \left\{ \boldsymbol{v}^i \cdot \boldsymbol{\eta}^i|_\Gamma \ : \ \boldsymbol{v}^i \in \boldsymbol{X}^i_{h^i} \right\}. \tag{6.21}$$

Associated with the coupled problem (6.12) we introduce the *balanced subspace*

$$V_{\Gamma,\bar{B}} := \left\{ v^f_\Gamma \in Z^f_{h^f} \ : \ (v^f_\Gamma, \Pi v^f_\Gamma) \in V_\Gamma, \text{ and } \int_\Gamma v^f_\Gamma \cdot \boldsymbol{\eta}_f = 0 \right\}, \tag{6.22}$$

with V_Γ defined in (6.18); see Pavarino and Widlund [2002]. Observe that $V_{\Gamma,\bar{B}} = \mathrm{Ker}\bar{B}$, where $\bar{B}$ is defined in (6.15) and (6.17). Then for $v^f_\Gamma \in V_{\Gamma,\bar{B}}$ we have $\bar{B}v^f_\Gamma = 0$. We will refer to functions $v^f_\Gamma \in V_{\Gamma,\bar{B}}$ as *balanced functions*. If $v^p_\Gamma = \Pi v^f_\Gamma$ and v^f_Γ is a balanced function then we also say that v^p_Γ is a balanced function or the pair $(v^f_\Gamma, \Pi v^f_\Gamma)$ is balanced.

6.4.2 Dual formulation

In the system (6.12), we first eliminate the unknowns $\boldsymbol{u}^f_I, p^f_I$ and $\boldsymbol{u}^p_I, p^p_I$. We obtain

$$\left[\begin{array}{cc|cc|c} S^f_\Gamma & \bar{B}^{fT} & 0 & 0 & \tilde{C}^{fT} \\ \bar{B}^f & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & S^p_\Gamma & \bar{B}^{pT} & -\tilde{C}^{pT} \\ 0 & 0 & \bar{B}^p & 0 & 0 \\ \hline \tilde{C}^f & 0 & -\tilde{C}^p & 0 & 0 \end{array}\right] \left[\begin{array}{c} u^f_\Gamma \\ \bar{p}^f \\ \hline u^p_\Gamma \\ \bar{p}^p \\ \hline \lambda \end{array}\right] = \left[\begin{array}{c} \tilde{b}^f \\ \hline \tilde{b}^p \\ \hline 0 \end{array}\right], \tag{6.23}$$

where S^i_Γ, K^i_{II} and $K^i_{I\Gamma}$, $i = f, p$, are defined in (6.17) and (6.13). The right hand side of (6.23) is given by

$$\left[\begin{array}{c} \tilde{b}^f \\ \hline \tilde{b}^p \\ \hline 0 \end{array}\right] = \left[\begin{array}{c} \left[\begin{array}{c} f^f_\Gamma \\ \bar{g}^f \end{array}\right] - K^f_{\Gamma I}\left(K^f_{\Gamma\Gamma}\right)^{-1} \left[\begin{array}{c} \boldsymbol{f}^f_I \\ g^f_I \end{array}\right] \\ \hline \left[\begin{array}{c} f^p_\Gamma \\ \bar{g}^p \end{array}\right] - K^p_{\Gamma I}\left(K^p_{\Gamma\Gamma}\right)^{-1} \left[\begin{array}{c} \boldsymbol{f}^p_I \\ g^p_I \end{array}\right] \\ \hline 0 \end{array}\right].$$

Let $N_i := \left[\begin{array}{cc} \tilde{C}^i & 0 \end{array}\right]$ and consider S^i, $i = f, p$, defined in (6.17). Then the matrix in the left hand side of (6.23) can be rewritten as

$$\left[\begin{array}{c|c|c} S^f & 0 & N^{fT} \\ \hline 0 & S^p & -N^{pT} \\ \hline N^f & -N^p & 0 \end{array}\right].$$

Now we eliminate the unknowns $u_\Gamma^f, \bar{p}^f$ and $u_\Gamma^p, \bar{p}^p$. We end up with the reduced system

$$F\lambda = c, \tag{6.24}$$

where the operator F is defined by

$$F := N^f (S^f)^{-1} N^{fT} + N^p (S^p)^{-1} N^{pT}, \tag{6.25}$$

and the right hand side c is given by

$$c = N^f (S^f)^{-1} \left\{ \begin{bmatrix} f_\Gamma^f \\ \bar{g}^f \end{bmatrix} - K_{\Gamma I}^f \left(K_{\Gamma\Gamma}^f \right)^{-1} \begin{bmatrix} \boldsymbol{f}_I^f \\ g_I^f \end{bmatrix} \right\} - N^p (S^p)^{-1} \left\{ \begin{bmatrix} f_\Gamma^p \\ \bar{g}^p \end{bmatrix} - K_{\Gamma I}^p \left(K_{\Gamma\Gamma}^p \right)^{-1} \begin{bmatrix} \boldsymbol{f}_I^f \\ g_I^p \end{bmatrix} \right\}.$$

Note that F is positive semidefinite and since a discrete Lagrange multiplier in Λ_{h^p} does not have necessarily zero mean average value on Γ, then, the operator F has one simple zero eigenvalue corresponding to a constant Lagrange multiplier. The above linear system, as well as the original linear system (6.12), is solvable for zero mean right hand side, i.e., for $c^T \cdot (1, \ldots, 1) = 0$.

6.5 BDD preconditioner

In this section we design and analyze a BDD type preconditioner for the Schur complement system (6.14); see Brenner and Sung [2007], Dryja and Proskurowski [2003] and Toselli and Widlund [2005], and also Achdou et al. [1999], Dryja et al. [2005], Mandel [1993], Pavarino and Widlund [2002] and Tu [2005]. For the sake of simplicity on the analysis we assume that $\Gamma = \{1\} \times (0,1)$, $D^f = (1,2) \times (0,1)$ and $D^p = (0,1) \times (0,1)$. We introduce the velocity coarse space on Γ as the span of the normal velocity $v_0 = y(1-y)$ (with v_0 also denoting its vector representation). Define:

$$R_0 := \begin{bmatrix} v_0^T & 0 \\ 0 & I_{2\times 2} \end{bmatrix}, \quad S_0 := R_0 S R_0^T \quad \text{and} \quad Q_0 := R_0^T S_0^\dagger R_0. \tag{6.26}$$

The system (6.14) is solvable when the right hand side satisfy $\bar{b}^f + \bar{b}^p = 0$ with uniqueness of the solution in the space of vectors with pressure component having zero average value on D. Then, we have that S_0 is invertible restricted to vectors with pressure component in M_0. The low dimensionality of the coarse space (which is spanned by ϕ_0^f and a constant pressure per subdomain D^i, $i = f, p,$) and the fact that the functions ϕ_0^f is independent of the triangulation parameters imply stable discrete inf-sup condition for the coarse problem.

Denote $\tilde{S}_0 := v_0^T S_\Gamma v_0$ and $\tilde{S} := \bar{B} v_0 \tilde{S}_0^{-1} v_0^T \bar{B}^T$. We can write (see (6.15) and (6.26))

$$S_0 = \begin{bmatrix} \tilde{S}_0 & (\bar{B}v_0)^T \\ \bar{B}v_0 & 0 \end{bmatrix}.$$

A simple calculation using the formula for the inverse of a saddle point matrix gives

$$Q_0 = \begin{bmatrix} v_0\tilde{S}_0^{-1}v_0^T - v_0\tilde{S}_0^{-1}v_0^T\bar{B}^T\tilde{S}^{-1}\bar{B}v_0\tilde{S}_0^{-1}v_0^T & v_0\tilde{S}_0^{-1}v_0^T\bar{B}^T\tilde{S}^{-1} \\ \tilde{S}^{-1}\bar{B}v_0\tilde{S}_0^{-1}v_0^T & \tilde{S}^{-1} \end{bmatrix},$$

and using (6.15) we obtain

$$Q_0S = \begin{bmatrix} v_0\tilde{S}_0^{-1}v_0^TS_\Gamma - v_0\tilde{S}_0^{-1}v_0^T\bar{B}^T\tilde{S}^{-1}\bar{B}v_0\tilde{S}_0^{-1}v_0^TS_\Gamma + v_0\tilde{S}_0^{-1}v_0^T\bar{B}^T\tilde{S}^{-1}\bar{B} & 0 \\ \tilde{S}^{-1}\bar{B}v_0\tilde{S}_0^{-1}v_0^TS_\Gamma - \tilde{S}^{-1}\bar{B} & I \end{bmatrix},$$

or $Q_0S = \begin{bmatrix} \mathcal{P} & 0 \\ \mathcal{G} & I \end{bmatrix}$. Here, we have defined

$$\begin{aligned} \mathcal{P} &:= \left(v_0\tilde{S}_0^{-1}v_0^TS_\Gamma - v_0\tilde{S}_0^{-1}v_0^T\bar{B}^T\tilde{S}^{-1}\bar{B}v_0\tilde{S}_0^{-1}v_0^TS_\Gamma\right) \\ &\quad + v_0\tilde{S}_0^{-1}v_0^T\bar{B}^T\tilde{S}^{-1}\bar{B} \\ \mathcal{G} &:= \tilde{S}^{-1}\bar{B} - \tilde{S}^{-1}\bar{B}v_0\tilde{S}_0^{-1}v_0^TS_\Gamma. \end{aligned}$$

With this notation we have that $I - Q_0S = \begin{bmatrix} I-\mathcal{P} & 0 \\ \mathcal{G} & 0 \end{bmatrix}$. Elementary calculations show that $\mathcal{P}^2 = \mathcal{P}$ and $\bar{B}(I-\mathcal{P}) = 0$, hence $I-\mathcal{P}$ is a projection and its image is contained on the balanced subspace defined in (6.22); see also Pavarino and Widlund [2002].

Given a residual $r = \begin{bmatrix} f_\Gamma^T & \bar{g}^T \end{bmatrix}^T$, the coarse problem Q_0r, with Q_0 defined in (6.26), is the solution of the coupled problem (6.12) with one velocity degree of freedom (v_0), and a constant pressure per subdomain D^i, $i = f, p$, with mean zero in $D = \text{int}(\overline{D^f} \cup \overline{D^p})$. Note that the matrix S_0 defined in (6.26) can be computed easily and in order to ensure zero mean pressure on D we can use a Lagrange multiplier.

For balanced functions v_Γ^f and u_Γ^f, the S_Γ-inner product is defined by (see (6.15)):

$$\langle u_\Gamma^f, v_\Gamma^f\rangle_{S_\Gamma} := \langle S_\Gamma u_\Gamma^f, v_\Gamma^f\rangle = u_\Gamma^{fT}S_\Gamma v_\Gamma^f.$$

Recall that $\bar{B}u_\Gamma^f{=}0$ when u_Γ^f is balanced. Then, on this subspace of balanced functions, the S_Γ inner product coincides with the S-inner product defined by

$$\left\langle \begin{bmatrix} v_\Gamma^f \\ \bar{q}^f \\ \bar{q}^p \end{bmatrix}, \begin{bmatrix} u_\Gamma^f \\ \bar{p}^f \\ \bar{p}^p \end{bmatrix} \right\rangle_S := \begin{bmatrix} v_\Gamma^f \\ \bar{q}^f \\ \bar{q}^p \end{bmatrix}^T S \begin{bmatrix} u_\Gamma^f \\ \bar{p}^f \\ \bar{p}^p \end{bmatrix} = \begin{bmatrix} v_\Gamma^f \\ \bar{q} \end{bmatrix}^T \begin{bmatrix} S_\Gamma & \bar{B}^T \\ \bar{B} & 0 \end{bmatrix} \begin{bmatrix} u_\Gamma^f \\ \bar{p} \end{bmatrix},$$

where $\bar{p}^T = \begin{bmatrix} \bar{p}^p & \bar{p}^p \end{bmatrix}^T$. Consider the BDD preconditioner operator given by

$$S_N^{-1} := Q_0 + (I - Q_0S)\,(\tilde{S}^f)^\dagger\,(I - SQ_0)\,, \tag{6.27}$$

where $\tilde{S}^f$ is defined in (6.16); see Dryja and Proskurowski [2003] and Pavarino and Widlund [2002]. The notation $(\tilde{S}^f)^\dagger$ stands for the pseudo-inverse of $\tilde{S}^f$, i.e.,

$$(\tilde{S}^f)^\dagger = \begin{bmatrix} (S^f)^{-1} & 0 \\ 0 & 0 \end{bmatrix},$$

with S^f defined in (6.17). The preconditioned operator is given by

$$\begin{aligned} S_N^{-1}S &= Q_0S + (I - Q_0S)\,(\tilde{S}^f)^\dagger S\,(I - Q_0S) \\ &= \begin{bmatrix} \mathcal{P} & 0 \\ \mathcal{G} & I \end{bmatrix} + \begin{bmatrix} I-\mathcal{P} & 0 \\ \mathcal{G} & 0 \end{bmatrix} (\tilde{S}^f)^\dagger \begin{bmatrix} S_\Gamma & \bar{B}^T \\ \bar{B} & 0 \end{bmatrix} \begin{bmatrix} I-\mathcal{P} & 0 \\ \mathcal{G} & 0 \end{bmatrix}. \end{aligned} \tag{6.28}$$

Note that applying $(S^f)^{-1}$ to a vector $\begin{bmatrix} u_\Gamma^f \\ \bar{p} \end{bmatrix}$ is equivalent to solving the linear system

$$\begin{bmatrix} A_{II}^f & B_{II}^{fT} & A_{\Gamma I}^{fT} & 0 \\ B_{II}^f & 0 & B_{I\Gamma}^f & 0 \\ A_{\Gamma I}^f & B_{I\Gamma}^{fT} & A_{\Gamma\Gamma}^f & \bar{B}^{fT} \\ 0 & 0 & \bar{B}^f & 0 \end{bmatrix} \begin{bmatrix} w_I^f \\ s_I^f \\ w_\Gamma^f \\ \bar{s}^f \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ u_\Gamma^f \\ \bar{p}^f \end{bmatrix}.$$

If u_Γ^f is balanced, so is balanced the velocity component of $(S^f)^{-1}\begin{bmatrix} u_\Gamma^f \\ \bar{p}^f \end{bmatrix}$. Then using elementary calculations with the matrices in (6.28) we obtain that

$$\langle S_N^{-1}S \begin{bmatrix} u_\Gamma \\ \bar{p} \end{bmatrix}, \begin{bmatrix} v_\Gamma \\ \bar{q} \end{bmatrix} \rangle_S = \langle (S_\Gamma^f)^{-1} S_\Gamma u_\Gamma, v_\Gamma \rangle_{S_\Gamma},$$

for $u_\Gamma, v_\Gamma \in \text{Range}(I - \mathcal{P})$. In order to bound the condition number of the preconditioned operator $S_N^{-1}S$, we only need to analyze the condition of the operator $(S_\Gamma^f)^{-1}S_\Gamma$. Note that

$$c\langle u_\Gamma^f, u_\Gamma^f\rangle_{S_\Gamma} \leq \langle \left(S^f\right)^{-1} S_\Gamma u_\Gamma^f, u_\Gamma^f\rangle_{S_\Gamma} \leq C\langle u_\Gamma^f, u_\Gamma^f\rangle_{S_\Gamma}$$

is equivalent to

$$c\langle S^f u_\Gamma^f, u_\Gamma^f\rangle \leq \langle S_\Gamma u_\Gamma^f, u_\Gamma^f\rangle \leq C\langle S^f u_\Gamma^f, u_\Gamma^f\rangle. \tag{6.29}$$

The next theorem shows that the condition number estimate for the BDD method introduced in (6.27) is of order $O(1+\frac{1}{\kappa})$, where κ is the permeability of the porous medium; see (6.1).

Theorem 6.1 *If u_Γ^f is a balanced function then*

$$\langle S_\Gamma^f u_\Gamma^f, u_\Gamma^f\rangle \leq \langle S_\Gamma u_\Gamma^f, u_\Gamma^f\rangle \prec \left(1 + \frac{1}{\kappa}\right) \langle S_\Gamma^f u_\Gamma^f, u_\Gamma^f\rangle.$$

Proof. The lower bound follows trivially from $\tilde{S}_\Gamma^f$ and $\tilde{S}_\Gamma^p$ being positive on the subspace of balanced functions. Next we concentrate on the upper bound.

Let v_Γ^f be a balanced function and $v_\Gamma^p = \Pi v_\Gamma^f$. Define $\boldsymbol{v}^p = \mathcal{DH}v_\Gamma^p$; see (6.20). Using properties of the discrete operator $\mathcal{DH}$, see Mathew [1993], we obtain

$$\langle S_\Gamma^p v_\Gamma^p, v_\Gamma^p\rangle = a^p(\boldsymbol{v}^p, \boldsymbol{v}^p) \asymp \frac{\nu}{\kappa}\|v_\Gamma^p\|^2_{(H^{1/2})'(\Gamma)}.$$

Using the L_2-stability property of mortar projection Π, we have

$$\|v_\Gamma^p\|^2_{(H^{1/2})'(\Gamma)} \prec \|v_\Gamma^p\|^2_{L^2(\Gamma)} = \|v_\Gamma^f\|^2_{L^2(\Gamma)} \prec \|v_\Gamma^f\|^2_{H_{00}^{1/2}(\Gamma)}.$$

With $\mathcal{SH}$ defined in (6.19), define $\boldsymbol{v}^f = \mathcal{SH} v_\Gamma^f$. Using properties of $\mathcal{SH}$, see Pavarino and Widlund [2002], we have

$$\nu \|v_\Gamma^f\|^2_{H_{00}^{1/2}(\Gamma)} \asymp a^f(\boldsymbol{v}^f, \boldsymbol{v}^f)$$

and then

$$\langle S_\Gamma^p v_\Gamma^p, v_\Gamma^p \rangle \prec \frac{1}{\kappa} \langle S^f u_\Gamma^f, u_\Gamma^f \rangle. \tag{6.30}$$

This gives the upper bound and finishes the proof. ∎

1. Initialize

$$\begin{aligned} x^{(0)} &= Q_0 b + w \\ d^{(0)} &= b - S x^{(0)} \end{aligned}$$

with $w \in \text{Range}(I - Q_0 S)$. Recall that all vectors have three components, for instance, $x = \begin{bmatrix} x_\Gamma \\ \bar{x}^f \\ \bar{x}^p \end{bmatrix}$ and $b = \begin{bmatrix} b_\Gamma \\ \bar{b}^f \\ \bar{b}^p \end{bmatrix}$.

2. Iterate $k = 1, 2, \ldots$ until convergence

$$\begin{aligned} \text{Precondition: } z^{(k-1)} &= (\tilde{S}^f)^\dagger d^{(k-1)} \\ \text{Project: } y^{(k-1)} &= (I - Q_0 S) z^{(k-1)} \\ \beta^k &= \langle y^{(k-1)}, d^{(k-1)} \rangle / \langle y^{(k-2)}, d^{(k-1)} \rangle \quad [\beta^{(1)} = 0] \\ r^{(k)} &= y^{(k-1)} + \beta^{(k)} r^{(k)} \ [r^{(1)} = y^{(0)}] \\ \alpha^{(k)} &= \langle y^{(k-1)}, d^{(k-1)} \rangle / \langle d^{(k)}, S r^{(k)} \rangle \\ x^{(k)} &= x^{(k-1)} + \alpha^{(k)} r^{(k)} \\ d^{(k)} &= d^{(k-1)} - \alpha^{(k)} S r^{(k)} \end{aligned}$$

Figure 6.1: Implementation of the projected preconditioned conjugate gradient algorithm for the system (6.14) involving the BDD preconditioner (6.27).

Recall that we consider the preconditioned projected conjugate gradient method applied to the Schur complement problem (6.14). We have written the algorithm in Figure 6.1.

6.6 FETI preconditioner

In this section we analyze a FETI preconditioner for the reduced linear system (6.24); see Brenner and Sung [2007], Dryja and Proskurowski [2003] and Toselli and Widlund [2005], and also Farhat and Roux [1991], Klawonn and Widlund [2001] and Mandel and Tezaur [1996]. Recall the definition of F in (6.25). We propose the following preconditioner

$$(N^p)^\dagger (S^p) (N^p)^{\dagger T}, \tag{6.31}$$

where $(N^p)^\dagger$ is the pseudo-inverse $(N^p)^\dagger = \begin{bmatrix} (\tilde{C}^p)^{-1} & 0 \end{bmatrix}$.

Note that after computing the action of $(S^f)^{-1}$ and $(S^p)^{-1}$, in the application of F to a zero average Lagrange multiplier, we end up with balanced functions. Therefore, in order to apply the preconditioned operator $(N^p)^\dagger (S^p) (N^p)^{\dagger T} F$ to a zero mean Lagrange multiplier, we do not need to solve a coarse problem at the

beginning of the CG, neither inside of the CG iteration.

The FETI preconditioner in (6.31) can be considered as the dual preconditioner of the BDD preconditioner defined in (6.27); see the proof of Lemma 6.2 below.

Recall the definition of S^i, $i = f, p$, in (6.17) and the definition of space of balanced functions $V_\Gamma = V_\Gamma^f \times V_\Gamma^p$ in (6.22) and (6.21). We prove the following result.

Lemma 6.2 *Let $\lambda \in \Lambda_{h^p} \cap L_0^2(\Gamma)$ be a zero mean Lagrange multiplier. Then*

$$\langle N^f (S^f)^{-1} N^{fT} \lambda, \lambda \rangle \prec \frac{1}{\kappa} \langle N^p (S^p)^{-1} N^{pT} \lambda, \lambda \rangle.$$

Proof. Consider a zero mean Lagrange multiplier λ. Define $t = (S^p)^{-\frac{1}{2}} N^{pT} \lambda$ and $w^f = N^{fT} \lambda$. Then it is enough to prove that

$$\|(S^f)^{-\frac{1}{2}} w^f\|^2 \prec \|t\|^2.$$

Since w^f is balanced, i.e., $w^f \in V_\Gamma^f$, we have that

$$\begin{aligned}
\|(S^f)^{-\frac{1}{2}} w^f\|^2 &= \sup_{z^f \in Z_{h_f}^f} \frac{\langle (S^f)^{-\frac{1}{2}} w^f, z^f \rangle^2}{\|z^f\|^2} \\
&= \sup_{v^f \,\text{balanced}} \frac{\langle w^f, v^f \rangle^2}{\|(S^f)^{\frac{1}{2}} v^f\|^2} \\
&= \sup_{v^f \,\text{balanced}} \frac{\langle \lambda, N^f v^f \rangle^2}{\|(S^f)^{\frac{1}{2}} v^f\|^2} \\
&= \sup_{v^f \,\text{balanced}} \frac{\langle (S^p)^{-\frac{1}{2}} N^p \lambda, (S^p)^{\frac{1}{2}} (N^p)^{-1} N^f v^f \rangle^2}{\|(S^f)^{\frac{1}{2}} v^f\|^2}.
\end{aligned}$$

Then using the Cauchy-Schwarz inequality and (6.30) in the proof of Theorem 6.1, we have

$$\begin{aligned}
\|(S^f)^{-\frac{1}{2}} w^f\|^2 &= \sup_{v^f \,\text{balanced}} \frac{\langle t, (S^p)^{\frac{1}{2}} (N^p)^{-1} N^f v^f \rangle^2}{\|(S^f)^{\frac{1}{2}} v^f\|^2} \\
&\leq \|t\|^2 \sup_{v^f \,\text{balanced}} \frac{\|(S^p)^{\frac{1}{2}} (N^p)^{-1} N^f v^f\|^2}{\|(S^f)^{\frac{1}{2}} v^f\|^2} \prec \frac{1}{\kappa} \|t\|^2.
\end{aligned}$$

■

Using Lemma 6.2 we can derive the following estimate for the condition number of the FETI preconditioner defined in (6.31).

Theorem 6.3 *Let λ be a zero mean Lagrange multiplier. Then*

$$\langle N^p (S^p)^{-1} N_p^T \lambda, \lambda \rangle \prec \langle F\lambda, \lambda \rangle \prec \left(1 + \frac{1}{\kappa}\right) \langle N^p (S^p)^{-1} N^{pT} \lambda, \lambda \rangle.$$

The condition number estimate $O(\frac{\kappa+1}{\kappa})$ can be improved in the case where the fluid side triangulation is finer than the porous side triangulation. This case has some advantages when κ is small. In order to fix ideas and simplify notation we analyze in detail the case where the triangulation of the fluid side is a *refinement* of the porous side triangulation. In particular, in Theorem 6.6, we will prove that the condition of the FETI preconditioned operator is of order $O(\frac{\kappa+1}{\kappa+(h^p)^2})$ in this simpler situation. The analysis that we will present to prove Theorem 6.6 can be extended easily for the case where the fluid side triangulation is finer than (and not necessarily a refinement of) the porous side triangulation; see Remark 6.7.

We assume that the fluid side discretization on Γ, $\mathcal{T}^f_{h^f}(D^f)|_\Gamma$, is a refinement of the corresponding porous side discretization, $\mathcal{T}^p_{h^p}(D^p)|_\Gamma$. That is, assume that $h^p = rh^f$ for some positive integer r. We will refer to this assumption as the *nested refinement assumption*. For $j = 1, \ldots, m^p$, we introduce the normal fluid velocity ϕ^f_j as the $P2$ bubble function defined on $\mathcal{T}^p_{h^p}(D^p)|_\Gamma$ and with support on the interval $e^p_j = \{0\} \times [(j-1)h^p, jh^p]$. Recall that we are using $P2/P1$ Taylor-Hood discretization on the fluid side. Under the nested refinement assumption we have that $\phi^f_j \in Z^f_{h^f}$ with $Z^f_{h^f}$ defined in (6.21). Denote $Z^f_{h^f,b}$ as the subspace of $Z^f_{h^f}$ spanned by all ϕ^f_j, $j = 1, \ldots, m^p$, and set $Z^f_{h^f,0}$ as the subspace of $Z^f_{h^f}$ spanned by functions with zero average on all edges e^p_j, $j = 1, \ldots, m^p$. Note that $Z^f_{h^f,b}$ and $Z^f_{h^f,0}$ form a direct sum for $Z^f_{h^f}$ and the image $\Pi Z^f_{h^f,0}$ is the zero vector.

Before deriving the condition number estimate of the FETI preconditioner under the nested refinement assumption we first prove a preliminary lemma.

Lemma 6.4 *Assume that $h^p = rh^f$, where r is a positive integer. If $v^f_{\Gamma,b} \in Z^f_{h^f,b}$ and $v^f_{\Gamma,b}$ is a balanced function then*

$$\langle S^f_\Gamma v^f_{\Gamma,b}, v^f_{\Gamma,b}\rangle \prec \frac{\kappa}{(h^p)^2}\langle S^p_\Gamma \Pi v^f_{\Gamma,b}, \Pi v^f_{\Gamma,b}\rangle.$$

Proof. Let $v^f_{\Gamma,b} = \sum_{j=1}^{m^p} \beta_j \phi^f_j \in Z^f_{h^f,b} \subset Z^f_{h^f}$ and note that since the basis functions ϕ^f_j, $j = 1, \ldots, m^p$, do not overlap each other on Γ, they are orthogonal in $L^2(\Gamma)$ and also in $H^1_0(\Gamma)$. Then

$$\|v^f_{\Gamma,b}\|^2_{L^2(\Gamma)} = \sum_{j=1}^{m^p} \beta_j^2 \|\phi^f_j\|^2_{L^2(\Gamma)} \asymp h^p \sum_{j=1}^{m^p} \beta_j^2, \tag{6.32}$$

and

$$|v^f_{\Gamma,b}|^2_{H^1(\Gamma)} = \sum_{j=1}^{m^p} \beta_j^2 |\phi^f_j|^2_{H^1_0(e^p_j)} \asymp \frac{1}{h^p} \sum_{j=1}^{m^p} \beta_j^2. \tag{6.33}$$

Using (6.32), (6.33) and a interpolation estimate we see that

$$\|v^f_{\Gamma,b}\|^2_{H^{1/2}_{00}(\Gamma)} \asymp \sum_{j=1}^{m^p} \beta_j^2 \asymp \frac{1}{h^p}\|v^f_{\Gamma,b}\|^2_{L^2(\Gamma)}.$$

Note also that

$$\langle S^f v^f_{\Gamma,b}, v^f_{\Gamma,b}\rangle \leq a^f(\mathcal{SH}v^f_{\Gamma,b}, \mathcal{SH}v^f_{\Gamma,b}) \asymp \nu \|v^f_{\Gamma,b}\|^2_{H^{1/2}_{00}(\Gamma)}.$$

Denote by $z^p_{\Gamma,b} = \sum_{j=1}^{m^p} \rho_j \chi_{e^p_j}$ the unique piecewise constant function such that $\Pi v^f_{\Gamma,b} = z^p_{\Gamma,b}$. Observe that $|\rho_j| \asymp |\beta_j|$, $j = 1, \ldots, m^p$. We obtain

$$\begin{aligned}\langle S^f_\Gamma v^f_{\Gamma,b}, v^f_{\Gamma,b}\rangle &\prec \frac{\nu}{h^p}\|v^f_{\Gamma,b}\|^2_{L^2(\Gamma)} \asymp \frac{\nu}{h^p}\|z^p_{\Gamma,b}\|^2_{L^2(\Gamma)} \\ &\prec \frac{\nu}{(h^p)^2}\|z^p_{\Gamma,b}\|^2_{(H^{1/2})'(\Gamma)} \asymp \frac{\kappa}{(h^p)^2}\langle S^p_\Gamma z^p_{\Gamma,b}, z^p_{\Gamma,b}\rangle,\end{aligned}$$

where we have used an inverse inequality for piecewise constant functions. ∎

We now translate Lemma 6.4 in a result concerning to our dual preconditioner.

Lemma 6.5 *Assume that $h^p = rh^f$, where r is a positive integer and let λ be a zero mean Lagrange multiplier. Then*

$$\frac{(h^p)^2}{\kappa}\langle N^p (S^p)^{-1} N^{pT}\lambda, \lambda\rangle \prec \langle N^f (S^f)^{-1} N^{fT}\lambda, \lambda\rangle.$$

Proof. We proceed as before. Let $t = (S^f)^{-\frac{1}{2}} N^{fT}\lambda$ and $w = N^p\lambda$. Then

$$\begin{aligned}\|(S^p)^{-\frac{1}{2}} w\|^2 &= \sup_{z^f \in Z^f_{h^f}} \frac{\langle (S^p)^{-\frac{1}{2}} w, z^f\rangle^2}{\|z^p\|^2} \qquad (6.34)\\ &= \sup_{v^p \text{ balanced}} \frac{\langle w, v^f\rangle^2}{\|(S^p)^{\frac{1}{2}} v^p\|^2} \\ &= \sup_{v^p \text{ balanced}} \frac{\langle \lambda, N^p v^p\rangle^2}{\|(S^p)^{\frac{1}{2}} v^p\|^2} \\ &= \sup_{v^f_b \text{ balanced}} \frac{\langle \lambda, N^f v^f_b\rangle^2}{\|(S^p)^{\frac{1}{2}} (N^p)^{-1} N^f v^f_b\|^2} \\ &= \sup_{v^f_b \text{ balanced}} \frac{\langle (S^f)^{-\frac{1}{2}} N^{fT}\lambda, (S^f)^{\frac{1}{2}} v^f_b\rangle^2}{\|(S^p)^{\frac{1}{2}} (N^p)^{-1} N^f v^f_b\|^2} \\ &\leq \|t\|^2 \sup_{v^f_b \text{balanced}} \frac{\|(S^f)^{\frac{1}{2}} v^f_b\|^2}{\|(S^p)^{\frac{1}{2}} (N^p)^{-1} N^f v^f_b\|^2} \\ &\prec \frac{\kappa}{(h^p)^2}\|t\|^2,\end{aligned}$$

where the last step follows from Lemma 6.4. ∎

From Lemmas 6.2 and 6.5, the next theorem follows.

Theorem 6.6 *Assume that $h^p = rh^f$, where r is a positive integer. Let λ be a zero mean Lagrange multiplier, then*

$$\left(1 + \frac{(h^p)^2}{\kappa}\right)\langle N^p (S^p)^{-1} N^{pT}\lambda, \lambda\rangle \prec \langle F\lambda, \lambda\rangle \prec \left(1 + \frac{1}{\kappa}\right)\langle N^p (S^p)^{-1} N^{pT}\lambda, \lambda\rangle.$$

Remark 6.7 *Theorem 6.6 can be extended for the case where $h^f \leq 2h^p$. We only need to extend the argument given in the proof of Lemma 6.4. The basic idea in the proof of Lemma 6.4 is to associate a bubble function $\phi_j^f \in Z_{h^f}^f$ to each porous side element e_j^p, $j = 1, \ldots, m^p$, in such a way that we can construct a one to one and continuous map $v_{\Gamma,b}^f \mapsto z_{\Gamma,b}^p$. The bubble functions ϕ_j^f, $j = 1, \ldots, m^p$, can be chosen orthogonal in $L^2(\Gamma)$ and in $H_0^1(\Gamma)$. This can also be done when $h^f \leq h^p$. The smaller the h^f, the closer is the size of the support of the bubble ϕ_j^f to the size of the element e_j^p since more and more elements e^f can be associated to only one element e^p. This construction can also be carried out in the case $h_p < h^f \leq 2h^p$ where non-orthogonal Taylor-Hood basis functions must be used. This last situation leads to the appearance of an additional constant that depends on the non-orthogonality; see Section 6.7.*

Remark 6.8 *We note that Lemma 6.4 can be used directly to obtain a bound for the balancing domain decomposition preconditioner similar to the one presented in Section 6.5 but with $\tilde{S}^p$ instead of $\tilde{S}^f$ in (6.27); see Proposition 2 of Galvis and Sarkis [2006]. In this case an additional variable elimination is needed. We have to eliminate the component of the normal fluid velocity in the space $Z_{h^f,0}^f$ and work with the Schur complement with respect to the space $Z_{h^f,b}^f$. This is rather difficult to implement (we can use Lagrange multipliers in this case). Then passing to the dual preconditioner permit us to take advantage of the case where the fluid side discretization on Γ is a refinement of the corresponding porous side discretization.*

Remark 6.9 *Theorems 6.1, 6.3 and 6.6 are also valid for the case $\alpha^f > 0$ in (6.3). To see this we need to compare, for different values of α^f, the energy of discrete extensions for a given normal velocity defined on Γ. Given the outward normal velocity v_Γ^f on Γ, let $\mathcal{SH}_{\alpha^f} v_\Gamma^f$ denote the discrete harmonic extension in the sense of $(a_{\alpha^f}^f, b^f)$, that is, the solution of problem (6.19) with a^f replaced by $a_{\alpha^f}^f$. Recall that $a^f = a_0^f$, where $a_0^f = a_{\alpha^f}$ when $\alpha^f = 0$, and therefore, $\mathcal{SH} v_\Gamma^f = \mathcal{SH}_0 v_\Gamma^f$. Note that in (6.19) we have imposed the* natural *boundary condition $\boldsymbol{\tau}^T \boldsymbol{D}(\mathcal{SH} u^f)\boldsymbol{\eta}_f = 0$ on Γ. Now we define another extension denoted by $\widehat{\mathcal{SH}} v_\Gamma^f$. Given the outward normal velocity v_Γ^f on Γ, let $\widehat{\mathcal{SH}} v_\Gamma^f$ be the (a^f, b^f)-discrete harmonic extension given by the solution of (6.19) with the boundary condition $\widehat{\mathcal{SH}} v_\Gamma^f \cdot \boldsymbol{\tau} = 0$. For both $\mathcal{SH}$ and $\widehat{\mathcal{SH}}$ are imposed essential boundary condition v_Γ^f for the normal component on Γ. The difference between them is in how the boundary condition is imposed for the tangential component on Γ. For the $\mathcal{SH}$ is imposed homogeneous* natural *boundary condition, while for $\widehat{\mathcal{SH}}$ is imposed homogeneous* essential *boundary condition. Note that both extensions, $\mathcal{SH}_{\alpha^f}$ and $\widehat{\mathcal{SH}}$, satisfy the zero discrete divergence condition, i.e., the second equation of (6.19).*
Using the minimization property of the $(a_{\alpha^f}^f, b^f)$-discrete harmonic extension $\mathcal{SH}_{\alpha^f}$ and the (a^f, b^f)-discrete harmonic extension $\widehat{\mathcal{SH}}$ we get

$$
\begin{aligned}
& a^f(\mathcal{SH} v_\Gamma^f, \mathcal{SH} v_\Gamma^f) \\
&\quad = \; a_0^f(\mathcal{SH} v_\Gamma^f, \mathcal{SH} v_\Gamma^f) \quad \textit{(by definition)}
\end{aligned}
$$

1. Initialize

$$\begin{aligned} x^{(0)} &= 0 \quad \text{(No coarse problem)} \\ \lambda^{(0)} &= c \end{aligned}$$

2. Iterate $k = 1, 2, \ldots$ until convergence

$$\text{Precondition: } y^{(k-1)} = (N^p)^\dagger (S^p)(N^{pT})^\dagger d^{(k-1)}$$

$$\begin{aligned} \beta^k &= \langle y^{(k-1)}, d^{(k-1)} \rangle / \langle y^{(k-2)}, d^{(k-1)} \rangle \quad [\beta^{(1)} = 0] \\ r^{(k)} &= y^{(k-1)} + \beta^{(k)} r^{(k)} \; [r^{(1)} = y^{(0)}] \\ \alpha^{(k)} &= \langle y^{(k-1)}, d^{(k-1)} \rangle / \langle d^{(k)}, F r^{(k)} \rangle \\ x^{(k)} &= x^{(k-1)} + \alpha^{(k)} r^{(k)} \\ d^{(k)} &= d^{(k-1)} - \alpha^{(k)} F r^{(k)} \end{aligned}$$

Figure 6.2: Implementation of the preconditioned conjugate gradient algorithm for the system (6.24) involving the FETI preconditioner (6.31).

$$\begin{aligned} &\leq a_0^f(\mathcal{SH}_{\alpha^f} v_\Gamma^f, \mathcal{SH}_{\alpha^f} v_\Gamma^f) \quad \textit{(by the minimization property of } \mathcal{SH}\textit{)} \\ &\leq a_{\alpha^f}^f(\mathcal{SH}_{\alpha^f} v_\Gamma^f, \mathcal{SH}_{\alpha^f} v_\Gamma^f) \quad (\alpha^f > 0) \\ &\leq a_{\alpha^f}^f(\widehat{\mathcal{SH}} v_\Gamma^f, \widehat{\mathcal{SH}} v_\Gamma^f) \quad \textit{(by the minimization property of } \mathcal{SH}_{\alpha^f}\textit{)} \\ &= a_0^f(\widehat{\mathcal{SH}}_0 v_\Gamma^f, \widehat{\mathcal{SH}}_0 v_\Gamma^f) \quad \textit{(because } \widehat{\mathcal{SH}} u^f \cdot \boldsymbol{\tau}^f = 0 \textit{ on } \Gamma\textit{)} \\ &\asymp \nu \|v_\Gamma^f\|^2_{H_{00}^{1/2}(\Gamma)} \\ &\asymp a^f(\mathcal{SH}_0 v_\Gamma^f, \mathcal{SH} v_\Gamma^f). \end{aligned}$$

The last two equivalences follow from properties of the (a^f, b)-discrete harmonic extensions $\mathcal{SH}$ and $\widehat{\mathcal{SH}}$ (which coincides with the discrete Stokes harmonic extension); see Girault and Raviart [1986] and Pavarino and Widlund [2002]. The two equivalences appearing above are independent of the permeability, fluid viscosity and mesh sizes. Then, the energy of the $(a_{\alpha^f}^f, b)$-discrete harmonic extensions is equivalent to the energy of the (a^f, b)-discrete harmonic extension, i.e., the discrete Stokes harmonic extension. This equivalence guarantees the extensions of Theorems 6.1, 6.3 and 6.6 to the case $\alpha^f > 0$.

We solve the system (6.24) using preconditioned conjugate gradient. We have written the algorithm in Figure 6.2.

6.7 Numerical results

In this section we present numerical tests in order to verify the estimates in Theorems 6.1, 6.3 and 6.6. We consider $D^f = (1, 2) \times (0, 1)$ and $D^p = (0, 1) \times (0, 1)$; see Burman and Hansbo [2007] and Galvis [2004] for examples of exact solutions and compatible divergence and boundary data. Note that the reduced systems (6.14) and (6.24) involve only degrees of freedom on the interface Γ. In our test problems we compute the eigenvalues of the preconditioned operators.

To solve both reduced systems (6.14) and (6.24) we can use the PCG algorithms described in Figures 6.1 and 6.2. Recall that the original system (6.10) is a "three times" saddle point problem. Note that since the finite element basis of $M_{h^f}^f \times M_{h^p}^p$ and Λ^{h^p} have non-zero mean, the finite element matrix in (6.12)

has the kernel composed by constant pressures in $D = \text{int}(\overline{D^f} \cup \overline{D^p})$ and constant Lagrange multipliers on Γ. The corresponding system is solved up to a constant pressure and a constant Lagrange multiplier. These constants can be recovered when imposing the zero average pressure constraint; see Galvis and Sarkis [2007b].

$h^f \downarrow h^p \rightarrow$	$3^{-1} * 2^{-0}$	$3^{-1} * 2^{-1}$	$3^{-1} * 2^{-2}$	$3^{-1} * 2^{-3}$	$3^{-1} * 2^{-4}$
$2^{-1} * 2^{-0}$	1, 1.0189	1, 1.0198	1, 1.0194	1, 1.0193	1, 1.0193
$2^{-1} * 2^{-1}$	1, 1.0209	1, 1.0200	1, 1.0197	1, 1.0196	1, 1.0196
$2^{-1} * 2^{-2}$	1, 1.0217	1, 1.0205	1, 1.0202	1, 1.0201	1, 1.0201
$2^{-1} * 2^{-3}$	1, 1.0220	1, 1.0208	1, 1.0204	1, 1.0203	1, 1.0203
$2^{-1} * 2^{-4}$	1, 1.0221	1, 1.0209	1, 1.0205	1, 1.0204	1, 1.0204

Table 6.1: Minimum and maximum eigenvalues for the BDD preconditioned operator. Here $\kappa = 1$ and $\alpha^f = 0$.

6.7.1 BDD preconditioner

In the case of the BDD preconditioner (6.27) for (6.14), we solve a coarse problem before reducing the system to ensure balanced velocities at the beginning of the CG iterations.

$h^f \downarrow h^p \rightarrow$	$3^{-1} * 2^{-0}$	$3^{-1} * 2^{-1}$	$3^{-1} * 2^{-2}$	$3^{-1} * 2^{-3}$	$3^{-1} * 2^{-4}$
$2^{-1} * 2^{-0}$	1, 27.7647	1, 21.0147	1, 20.6035	1, 20.3686	1, 20.2893
$2^{-1} * 2^{-1}$	1, 28.1390	1, 21.3303	1, 20.8549	1, 20.6550	1, 20.5836
$2^{-1} * 2^{-2}$	1, 28.8104	1, 22.0017	1, 21.3392	1, 21.1424	1, 21.0735
$2^{-1} * 2^{-3}$	1, 29.0687	1, 22.2367	1, 21.6045	1, 21.3626	1, 21.2955
$2^{-1} * 2^{-4}$	1, 29.1810	1, 22.3479	1, 21.7006	1, 21.4666	1, 21.3929

Table 6.2: Minimum and maximum eigenvalues for the BDD preconditioned operator. Here $\kappa = 10^{-3}$ and $\alpha^f = 0$.

$h^f \downarrow h^p \rightarrow$	$3^{-1} * 2^{-0}$	$3^{-1} * 2^{-1}$	$3^{-1} * 2^{-2}$	$3^{-1} * 2^{-3}$	$3^{-1} * 2^{-4}$
$2^{-1} * 2^{-0}$	1, 1891.43	1, 1977.08	1, 1945.05	1, 1932.10	1, 1928.32
$2^{-1} * 2^{-1}$	1, 2095.91	1, 1997.27	1, 1972.77	1, 1961.34	1, 1957.88
$2^{-1} * 2^{-2}$	1, 2168.17	1, 2053.57	1, 2021.03	1, 2010.27	1, 2006.90
$2^{-1} * 2^{-3}$	1, 2201.81	1, 2079.68	1, 2044.05	1, 2032.42	1, 2029.13
$2^{-1} * 2^{-4}$	1, 2215.58	1, 2090.10	1, 2054.33	1, 2042.26	1, 2038.90

Table 6.3: Minimum and maximum eigenvalues for the BDD preconditioned operator. Here $\kappa = 10^{-5}$ and $\alpha^f = 0$.

We consider $\alpha^f = 0$ and $\nu = 1$, and different values of h^f and h^p with non-matching grids across the interface Γ; see Table 6.1 for the results when $\kappa = 1$, Table 6.2 for $\kappa = 10^{-3}$ and Table 6.3 for the case $\kappa = 10^{-5}$. These three tables reveal growth of order $O(1+\frac{1}{\kappa})$ in κ and hence, verify the sharpness of the estimate in Theorem 6.1.

$h^f \downarrow h^p \rightarrow$	$3^{-1} * 2^{-1}$	$3^{-1} * 2^{-2}$	$3^{-1} * 2^{-3}$	$3^{-1} * 2^{-5}$
$2^{-1} * 2^{-0}$	1.0000, 1.0208	1.0000, 1.0194	1.0000, 1.0193	1.0000, 1.0193
$2^{-1} * 2^{-1}$	1.0017, 1.0200	1.0000, 1.0197	1.0000, 1.0196	1.0000, 1.0196
$2^{-1} * 2^{-2}$	1.0026, 1.0205	1.0004, 1.0202	1.0000, 1.0200	1.0000, 1.0201
$2^{-1} * 2^{-3}$	1.0027, 1.0208	1.0007, 1.0204	1.0001, 1.0203	1.0000, 1.0203
$2^{-1} * 2^{-4}$	1.0028, 1.0209	1.0007, 1.0205	1.0002, 1.0204	1.0000, 1.0204
$2^{-1} * 2^{-5}$	1.0028, 1.0209	1.0007, 1.0206	1.0002, 1.0205	1.0000, 1.0204

Table 6.4: Minimum and maximum eigenvalues of the FETI preconditioned operator. Here $\kappa = 1$ and $\alpha^f = 0$.

$h_f \downarrow h_p \rightarrow$	$3^{-1} * 2^{-1}$	$3^{-1} * 2^{-2}$	$3^{-1} * 2^{-3}$	$3^{-1} * 2^{-4}$
$2^{-1} * 2^{-0}$	1.000, 20.7608	1.000, 20.4405	1.000, 20.3110	1.000, 20.2732
$2^{-1} * 2^{-1}$	2.707, 20.9627	1.000, 20.7177	1.000, 20.6034	1.000, 20.5688
$2^{-1} * 2^{-2}$	3.634, 21.5257	1.425, 21.2003	1.000, 21.0927	1.000, 21.0590
$2^{-1} * 2^{-3}$	3.714, 21.7868	1.651, 21.4305	1.106, 21.3142	1.000, 21.2813
$2^{-1} * 2^{-4}$	3.760, 21.891	1.663, 21.5333	1.162, 21.4126	1.026, 21.3790
$2^{-1} * 2^{-5}$	3.771, 21.937	1.673, 21.5768	1.164, 21.4561	1.040, 21.4220

Table 6.5: Minimum and maximum eigenvalues of the FETI preconditioned operator. Here $\kappa = 10^{-3}$ and $\alpha^f = 0$.

6.7.2 FETI preconditioner

In the case of the FETI preconditioner (6.31), we solve the reduced system (6.24) up to a constant Lagrange multiplier and a constant pressure. These constants are recovered after enforcing zero mean pressure on $D = \text{int}\,(\overline{D}^f \cup \overline{D}^p)$; see Galvis and Sarkis [2007b]. We note that the FETI method can be viewed as the dual preconditioner counterpart of the BDD preconditioner. We repeat the same experiments mentioned above for this preconditioner.

We consider $\alpha^f = 0$, $\nu = 1$ and different values of h^f and h^p with nonmatching grids across the interface Γ; see Table 6.4 for the results when $\kappa = 1$, Table 6.5 for $\kappa = 10^{-3}$ and Table 6.6 for the case $\kappa = 10^{-5}$. Note that in Tables 6.4, 6.5 and 6.6 the minimum eigenvalues strictly greater than one when $h^f \leq 2h^p$, and the value of the minimum eigenvalues seem to stabilize very quick for smaller h^f with fixed h^p. This confirms the extension of Theorem 6.6 for the case where $h^f \leq 2h^p$; see Remark 6.7. In Table 6.7 we present the numerical results where one of the meshes on the interface is a refinement of the other side triangulation on the interface. We observe a behavior similar to the behavior of Table 6.6 with a bigger value for

$h^f \downarrow h^p \rightarrow$	$3^{-1} * 2^{-1}$	$3^{-1} * 2^{-2}$	$3^{-1} * 2^{-3}$	$3^{-1} * 2^{-4}$
$2^{-1} * 2^{-0}$	1.00, 1977.08	1.00, 1945.05	1.00, 1932.10	1.00, 1928.32
$2^{-1} * 2^{-1}$	171.72, 1997.27	1.00, 1972.77	1.00, 1961.34	1.00, 1957.88
$2^{-1} * 2^{-2}$	264.44, 2053.57	43.45, 2021.03	1.00, 2010.27	1.00, 2006.90
$2^{-1} * 2^{-3}$	272.35, 2079.68	66.10, 2044.05	11.58, 2032.42	1.00, 2029.13
$2^{-1} * 2^{-4}$	276.95, 2090.10	67.29, 2054.33	17.20, 2042.26	3.64, 2038.90
$2^{-1} * 2^{-5}$	278.09, 2094.70	68.32, 2058.68	17.42, 2046.61	5.04, 2043.20

Table 6.6: Minimum and maximum eigenvalues of the FETI preconditioned operator. Here $\kappa = 10^{-5}$ and $\alpha^f = 0$.

the minimum eigenvalue when $h_f \leq h_p$. This verifies the estimates of Theorem 6.6. This shows that the FETI preconditioner is scalable for the parameters faced in practice, i.e., the fluid side mesh finer than the porous side mesh and a small permeability κ. We conclude that the numerical experiments concerning the FETI preconditioner reveal the sharpness of the results obtained in Theorems 6.3 and 6.6 and Remark 6.7.

$h^f \downarrow h^p \rightarrow$	$2^{-1} * 2^{-1}$	$2^{-1} * 2^{-2}$	$2^{-1} * 2^{-3}$	$2^{-1} * 2^{-4}$
$2^{-1} * 2^{-0}$	1.00, 2002.47	1.00, 1961.35	1.00, 1937.86	1.00, 1929.93
$2^{-1} * 2^{-1}$	690.43, 2034.03	1.00, 1986.49	1.00, 1966.50	1.00, 1959.36
$2^{-1} * 2^{-2}$	627.36, 2101.17	176.56, 2034.92	1.00, 2015.24	1.00, 2008.35
$2^{-1} * 2^{-3}$	639.68, 2124.67	151.62, 2061.45	44.91, 2037.26	1.00, 2030.55
$2^{-1} * 2^{-4}$	642.44, 2135.79	154.45, 2071.06	38.04, 2047.66	11.98, 2040.29
$2^{-1} * 2^{-5}$	643.47, 2140.73	154.86, 2075.43	38.73, 2051.91	10.20, 2044.66

Table 6.7: Minimum and Maximum eigenvalues of the FETI preconditioned operator. Here $\kappa = 10^{-5}$ and $\alpha^f = 0$. The refinement condition of Theorem 6.6 is satisfied under the diagonal.

Recall that we have assumed $\alpha^f = 0$. Now consider $\alpha^f > 0$. Numerical experiment were performed with $\alpha^f > 0$ reveling results similar to the ones presented above for the case $\alpha^f = 0$. We only include Table 6.8 which shows the extreme eigenvalues of the FETI preconditioned operator for the case $\alpha^f = 1$, $\nu = 1$ and $\kappa = 10^{-5}$. This table presents a similar behavior to the one with $\alpha^f = 0$ in Table 6.6 and hence, confirms Remark 6.9 which says that the parameter α^f does not play much role for preconditioning.

$h_f \downarrow h_p \rightarrow$	$3^{-1} * 2^{-1}$	$3^{-1} * 2^{-2}$	$3^{-1} * 2^{-3}$	$3^{-1} * 2^{-4}$
$2^{-1} * 2^{-0}$	1.00, 1705.47	1.00, 1678.07	1.00, 1666.84	1.00, 1663.55
$2^{-1} * 2^{-1}$	162.74, 1814.26	1.00, 1787.53	1.00, 1776.50	1.00, 1773.22
$2^{-1} * 2^{-2}$	251.56, 1843.50	41.65, 1812.69	1.00, 1801.61	1.00, 1798.29
$2^{-1} * 2^{-3}$	267.47, 1849.46	63.63, 1816.43	11.24, 1804.66	1.00, 1801.34
$2^{-1} * 2^{-4}$	272.29, 1850.65	66.82, 1817.38	16.75, 1805.30	3.58, 1801.91
$2^{-1} * 2^{-5}$	273.34, 1851.08	67.99, 1817.68	17.37, 1805.57	4.97, 1802.14

Table 6.8: Minimum and maximum eigenvalues of the FETI preconditioned operator. Here $\kappa = 10^{-5}$ and $\alpha^f = 1$.

6.8 Conclusions and final comments

We consider the problem of coupling fluid flows with porous media flows with Beavers-Joseph-Saffman condition on the interface. We choose a discretization consisting of Taylor-Hood finite elements of order two on the free fluid side and the lowest order Raviart-Thomas finite element on the porous fluid side. The meshes are allowed to be nonmatching across the interface.

We design and analyze two preconditioners for the resulting symmetric linear system. We note that the original linear system is symmetric indefinite and involves three Lagrange multipliers: one for each subdomain pressure and a third

one to impose the weak conservation of mass across the interface Γ; see Section 6.1.

One preconditioner is based on BDD methods and the other one is based on FETI methods. In the case of the BDD preconditioner, the energy is controlled by the Stokes side, while in the FETI preconditioner, the energy is controlled by the Darcy system; see Theorems 6.1 and 6.3. In both cases a bound $C_1(\frac{\kappa+1}{\kappa})$ is derived. Furthermore, under the assumption that the fluid side mesh on the interface is finer than the corresponding porous side mesh, we derive the better bound $C_2(\frac{\kappa+1}{\kappa+(h^p)^2})$ for the FETI preconditioner; see Theorem 6.6 and Remark 6.7. This better bound also shows that the FETI preconditioner is more scalable for parameters faced in practice, e.g., problems with small permeability κ and where the fluid side mesh is finer than the porous side mesh. The constants C_1 and C_2 above are independent of the fluid viscosity ν, the mesh ratio across the interface, and the permeability κ.

Bibliography

Achdou, Y., Maday, Y., and Widlund, O. B. (1999). Iterative substructuring preconditioners for mortar element methods in two dimensions. *SIAM J. Numer. Anal.*, 36(2):551–580.

Beavers, G. S. and Joseph, D. D. (1967). Boundary conditions at a naturally permeable wall. *J. Fluid Mech.*, 30:197–207.

Ben Belgacem, F. and Maday, Y. (1997). The mortar element method for three-dimensional finite elements. *RAIRO Modél. Math. Anal. Numér.*, 31(2):289–302.

Bernardi, C., Maday, Y., and Patera, A. T. (1994). A new nonconforming approach to domain decomposition: the mortar element method. In *Nonlinear partial differential equations and their applications. Collège de France Seminar, Vol. XI (Paris, 1989–1991)*, volume 299 of *Pitman Res. Notes Math. Ser.*, pages 13–51. Longman Sci. Tech., Harlow.

Braess, D. (2001). *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics*. Cambridge University Press, Cambridge. Second Edition.

Brenner, S. C. and Scott, L. R. (1994). *The mathematical theory of finite element methods*, volume 15 of *Texts in Applied Mathematics*. Springer-Verlag, New York.

Brenner, S. C. and Sung, L.-Y. (2007). BDDC and FETI-DP without matrices or vectors. *Comput. Methods Appl. Mech. Engrg.*, 196(8):1429–1435.

Brezzi, F. and Fortin, M. (1991). *Mixed and hybrid finite element methods*, volume 15 of *Springer Series in Computational Mathematics*. Springer-Verlag, New York.

Burman, E. and Hansbo, P. (2007). A unified stabilized method for Stokes' and Darcy's equations. *J. Comput. Appl. Math.*, 198(1):35–51.

Discacciati, M. (2004). *Domain decomposition methods for the coupling of surface and groundwater flows*. PhD thesis, Ecole Polytechnique Fédérale, Lausanne (Switzerland). Thèse n. 3117.

Discacciati, M. (2005). Iterative methods for Stokes/Darcy coupling. In *Domain decomposition methods in science and engineering*, volume 40 of *Lect. Notes Comput. Sci. Eng.*, pages 563–570. Springer, Berlin.

Discacciati, M., Miglio, E., and Quarteroni, A. (2002). Mathematical and numerical models for coupling surface and groundwater flows. *Appl. Numer. Math.*, 43(1-2):57–74. 19th Dundee Biennial Conference on Numerical Analysis (2001).

Discacciati, M. and Quarteroni, A. (2003). Analysis of a domain decomposition method for the coupling of Stokes and Darcy equations. In Brezzi, F., Buffa, A., Corsaro, S., and Murli, A., editors, *ENUMATH 2001*, Numerical Mathematics and Advanced Applications, pages 3–20. Springer-Verlag.

Discacciati, M. and Quarteroni, A. (2004). Convergence analysis of a subdomain iterative method for the finite element approximation of the coupling of Stokes and Darcy equations. *Comput. Vis. Sci.*, 6(2-3):93–103.

Dohrmann, C. R. (2003). A preconditioner for substructuring based on constrained energy minimization. *SIAM J. Sci. Comput.*, 25(1):246–258 (electronic).

Dryja, M., Kim, H. H., and Widlund, O. B. (2005). A BDDC algorithm for problems with mortar discretization. Technical Report TR2005-873, Courant Institute of Mathematical Sciences, Computer Science Department. http://cs.nyu.edu/web/Research/TechReports/.

Dryja, M. and Proskurowski, W. (2003). On preconditioners for mortar discretization of elliptic problems. *Numer. Linear Algebra Appl.*, 10(1-2):65–82. Dedicated to the 60th birthday of Raytcho Lazarov.

Farhat, C., Lesoinne, M., and Pierson, K. (2000). A scalable dual-primal domain decomposition method. *Numer. Linear Algebra Appl.*, 7(7-8):687–714. Preconditioning techniques for large sparse matrix problems in industrial applications (Minneapolis, MN, 1999).

Farhat, C. and Roux, F.-X. (1991). A method of finite element tearing and interconnecting and its parallel solution algorithm. *Internat. J. Numer. Methods Engrg.*, 32:1205 – 1227.

Galvis, J. (2004). Finite elements for well-reservoir coupling. Master's thesis, Instituto Nacional de Matemática Pura e Aplicada. TR-B011 / 2005, http://www.preprint.impa.br/cgi-bin/MMMsearch.cgi.

Galvis, J. and Sarkis, M. (2006). Balancing domain decomposition methods for mortar coupling Stokes-Darcy systems. In Keyes, D. and Widlund, O. B., editors, *Domain Decomposition Methods in Science and Engineering XVI*, volume 55 of *Lecture Notes in Computational Science and Engineering*, pages 373–380. Springer.

Galvis, J. and Sarkis, M. (2007). Non-matching mortar discretization analysis for the coupling Stokes-Darcy equations. *Electron. Trans. Numer. Anal.*, 26:350–384.

Girault, V. and Raviart, P.-A. (1986). *Finite element methods for Navier-Stokes equations*, volume 5 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin. Theory and algorithms.

Jäger, W. and Mikelić, A. (2000). On the interface boundary condition of Beavers, Joseph, and Saffman. *SIAM J. Appl. Math.*, 60(4):1111–1127.

Klawonn, A. and Widlund, O. B. (2001). FETI and Neumann-Neumann iterative substructuring methods: connections and new results. *Comm. Pure Appl. Math.*, 54(1):57–90.

Layton, W. J., Schieweck, F., and Yotov, I. (2002). Coupling fluid flow with porous media flow. *SIAM J. Numer. Anal.*, 40(6):2195–2218 (2003).

Li, J. (2005). A dual-primal FETI method for incompressible Stokes equations. *Numer. Math.*, 102(2):257–275.

Li, J. and Widlund, O. (2006). BDDC algorithms for incompressible Stokes equations. *SIAM J. Numer. Anal.*, 44(6):2432–2455 (electronic).

Li, J. and Widlund, O. (2007). A BDDC preconditioner for saddle point problems. In *Domain decomposition methods in science and engineering XVI*, volume 55 of *Lect. Notes Comput. Sci. Eng.*, pages 413–420. Springer, Berlin.

Mandel, J. (1993). Balancing domain decomposition. *Comm. Numer. Methods Engrg.*, 9(3):233–241.

Mandel, J. and Brezina, M. (1996). Balancing domain decomposition for problems with large jumps in coefficients. *Math. Comp.*, 65(216):1387–1401.

Mandel, J. and Tezaur, R. (1996). Convergence of a substructuring method with Lagrange multipliers. *Numer. Math.*, 73(4):473–487.

Mathew, T. P. (1993). Schwarz alternating and iterative refinement methods for mixed formulations of elliptic problems. II. Convergence theory. *Numer. Math.*, 65(4):469–492.

Pavarino, L. F. and Widlund, O. B. (2002). Balancing Neumann-Neumann methods for incompressible Stokes equations. *Comm. Pure Appl. Math.*, 55(3):302–335.

Rivière, B. and Yotov, I. (2005). Locally conservative coupling of Stokes and Darcy flows. *SIAM J. Numer. Anal.*, 42(5):1959–1977.

Toselli, A. and Widlund, O. (2005). *Domain decomposition methods—algorithms and theory*, volume 34 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin.

Tu, X. (2005). A BDDC algorithm for a mixed formulation of flow in porous media. *Electron. Trans. Numer. Anal.*, 20:164–179.

Wohlmuth, B. I. (2000). A mortar finite element method using dual spaces for the Lagrange multiplier. *SIAM J. Numer. Anal.*, 38(3):989–1012 (electronic).

Chapter 7

Balancing Domain Decomposition Methods for Discontinuous Galerkin Discretization

A Discontinuous Galerkin (DG) discretization of a Dirichlet problem for second order elliptic equations with discontinuous coefficients in two dimensions is considered. The problem is considered in a polygonal region D which is a union of disjoint polygonal substructures D_i of size $O(H_i)$. Inside each substructure D_i, a triangulation $\mathcal{T}_{h_i}(D_i)$ with a parameter h_i and a conforming finite element method are introduced. To handle nonmatching meshes across ∂D_i, a DG method, symmetric and with interior penalty terms on the ∂D_i, is considered. In this paper we design and analyze Balancing Domain Decomposition algorithms for solving the resulting discrete systems. Under certain assumptions on the coefficients and the mesh sizes across ∂D_i, a condition number estimate $C(1 + \max_i \log^2 \frac{H_i}{h_i})$ is established with C independent of h_i, H_i and the jumps of the coefficients. The algorithm is well suited for parallel computations and can be straightforwardly extended to three-dimensional problems. Results of numerical tests are included which confirm the theoretical results and the imposed assumption.

7.1 Introduction

DG methods are becoming more and more popular for approximation of PDEs since they are well suited for dealing with complex geometries, discontinuous coefficients and local or patch refinements; see Arnold et al. [2002], Dryja [2003] and the references therein. A goal of this paper is to design and analyze Balancing Domain Decomposition (BDD) algorithms for the resulting discrete problem; see Mandel [1993], Dryja and Widlund [1995] and also Toselli and Widlund [2005]. There are also several papers devoted to algorithms for solving DG discrete problems. In particular in connection with domain decomposition methods, we can mention Feng and Karakashian [2001], Lasser and Toselli [2003], Antonietti and Ayuso [2005] where overlapping Schwarz methods were proposed and analyzed for DG discretization of elliptic problems with continuous coefficients. In Dryja [2003] a non optimal multilevel additive Schwarz method is designed and analyzed for the discontinuous coefficient case. In Brenner and Wang [2005] a two-level

ASM is proposed and analyzed for DG discretization of fourth order problems. In Dryja et al. [2007b] we have also successfully extended these preconditioners to the Balancing Domain Decomposition (BDDC) with constraints method. Up to our knowledge BDD algorithms for DG discretization of elliptic problems with continuous and discontinuous coefficients have not been analyzed in literature.

The paper is organized as follows. In Section 7.2, the differential problem and its DG discretization are formulated. In Section 7.3, the problem is reduced to a Schur complement problem with respect to the unknowns on ∂D_i, and discrete harmonic functions defined in a special way are introduced. In Section 7.4, the BDD algorithm is designed and analyzed for the Schur complement problem using the general theory of hybrid methods; see Toselli and Widlund [2005]. The local problems are defined on ∂D_i and on faces of ∂D_j common to D_i, while the coarse space, restriction and prolongation operators are defined via a special partitioning of unity on the ∂D_i. Sections 7.5 and 7.6 are devoted to numerical experiments and final remarks, respectively.

7.2 Differential and discrete problems

Consider the following problem: Find $u^* \in H_0^1(D)$ such that

$$a(u^*, v) = f(v) \quad \text{for all } v \in H_0^1(D) \tag{7.1}$$

where $a(u,v) = \sum_{i=1}^{N} \int_{D_i} \rho_i \nabla u \nabla v dx$ and $f(v) = \int_D f v dx$.
We assume that $\bar{D} = \cup_{i=1}^{N} \bar{D}_i$ and the substructures D_i are disjoint shape regular polygonal subregions of diameter $O(H_i)$ that form a geometrically conforming partition of D, i.e., for all $i \neq j$ the intersection $\partial D_i \cap \partial D_j$ is empty, or a common vertex or face of ∂D_i and ∂D_j. We assume $f \in L^2(D)$ and for simplicity of presentation let ρ_i be a positive constant, $i = 1, \ldots, N$.

Let us introduce a shape regular triangulation in each D_i with triangular elements and the mesh parameter h_i . The resulting triangulation on D is in general nonmatching across ∂D_i. Let $X_i(D_i)$ be a finite element (FE) space of piecewise linear continuous functions in D_i. Note that we do not assume that the functions in $X_i(D_i)$ vanish on $\partial D_i \cap \partial D$. Define

$$X_h(D) = X_1(D_1) \times \cdots \times X_N(D_N).$$

The discrete problem obtained by the DG method, see Arnold et al. [2002], Dryja [2003], is of the form:

Find $u_h^* \in X_h(D)$ such that

$$a_h(u_h^*, v) = f(v) \quad \text{for all } v \in X_h(D) \tag{7.2}$$

where

$$a_h(u,v) \equiv \sum_{i=1}^{N} b_i(u,v) \quad \text{and} \quad f(v) \equiv \sum_{i=1}^{N} \int_{D_i} f v_i dx, \tag{7.3}$$

$$b_i(u,v) \equiv a_i(u,v) + s_i(u,v) + p_i(u,v), \tag{7.4}$$

$$a_i(u,v) \equiv \int_{D_i} \rho_i \nabla u_i \nabla v_i dx, \tag{7.5}$$

$$s_i(u,v) \equiv \sum_{F_{ij} \subset \partial D_i} \int_{F_{ij}} \frac{\rho_{ij}}{l_{ij}} \left(\frac{\partial u_i}{\partial n}(v_j - v_i) + \frac{\partial v_i}{\partial n}(u_j - u_i) \right) ds, \tag{7.6}$$

$$p_i(u,v) \equiv \sum_{F_{ij} \subset \partial D_i} \int_{F_{ij}} \frac{\rho_{ij}}{l_{ij}} \frac{\delta}{h_{ij}} (u_j - u_i)(v_j - v_i) ds, \tag{7.7}$$

$$d_i(u,v) \equiv a_i(u,v) + p_i(u,v), \tag{7.8}$$

with $u = \{u_i\}_{i=1}^N \in X_h(D)$ and $v = \{v_i\}_{i=1}^N \in X_h(D)$. We set $l_{ij} = 2$ when $F_{ij} \equiv \partial D_i \cap \partial D_j$ is a common face of ∂D_i and ∂D_j, and define $\rho_{ij} = 2\rho_i\rho_j/(\rho_i+\rho_j)$ as the harmonic average of ρ_i and ρ_j, and $h_{ij} = 2h_ih_j/(h_i+h_j)$. In order to simplify the notation we include the index $j = 0$ and set $l_{i0} = 1$ when $F_{i0} \equiv \partial D_i \cap \partial D$ has a positive measure, and set $u_0 = 0$ and $v_0 = 0$, and define $\rho_{i0} = \rho_i$ and $h_{i0} = h_i$. The outward normal derivative on ∂D_i is denoted by $\frac{\partial}{\partial n}$ and δ is the positive penalty parameter.

It is known that there exists a $\delta_0 = O(1) > 0$ such that for $\delta \geq \delta_0$, we obtain $2|s_i(u,u)| \leq d_i(u,u)$ and therefore, the problem (7.2) is elliptic and has a unique solution. An error bound of this method is given in Arnold et al. [2002] for continuous and in Dryja [2003] for discontinuous coefficients.

7.3 Schur complement problem

In this section we derive a Schur complement problem for the problem (7.2).

Define $\overset{o}{X}_i(D_i)$ as the subspace of $X_i(D_i)$ of functions that vanish on ∂D_i. Let $u = \{u_i\}_{i=1}^N \in X_h(D)$. For each $i = 1, \ldots, N$, the function $u_i \in X_i(D)$ can be represented as

$$u_i = \hat{\mathcal{P}}_i u + \hat{\mathcal{H}}_i u, \tag{7.9}$$

where $\hat{\mathcal{P}}_i u$ is the projection of u into $\overset{o}{X}_i(D_i)$ in the sense of $b_i(.,.)$. Note that since $\hat{\mathcal{P}}_i u$ and v_i belong to $\overset{o}{X}_i(D_i)$, we have

$$a_i(\hat{\mathcal{P}}_i u, v_i) = b_i(\hat{\mathcal{P}}_i u, v_i) = a_h(u, v_i). \tag{7.10}$$

The $\hat{\mathcal{H}}_i u$ is the discrete harmonic part of u in the sense of $b_i(.,.)$, where $\hat{\mathcal{H}}_i u \in X_i(D_i)$ is the solution of

$$b_i(\hat{\mathcal{H}}_i u, v_i) = 0 \qquad v_i \in \overset{o}{X}_i(D_i), \tag{7.11}$$

with boundary data given by

$$u_i \text{ on } \partial D_i \quad \text{ and } \quad u_j \text{ on } F_{ji} = \partial D_i \cap \partial D_j. \tag{7.12}$$

We point out that for $v_i \in \overset{o}{X}_i(D_i)$ we have

$$b_i(\hat{\mathcal{H}}_i u, v_i) = (\rho_i \nabla \hat{\mathcal{H}}_i u, \nabla v_i)_{L^2(D_i)} + \sum_{F_{ij} \subset \partial D_i} \frac{\rho_{ij}}{l_{ij}} (\frac{\partial v_i}{\partial n}, u_j - u_i)_{L^2(F_{ij})}. \tag{7.13}$$

Note that $\hat{\mathcal{H}}_i u$ is the classical discrete harmonic except at nodal points close to ∂D_i. We will sometimes call $\hat{\mathcal{H}}_i u$ discrete harmonic in a special sense, i.e., in the sense of $b_i(.,.)$ or $\hat{\mathcal{H}}_i$. Hence $\hat{\mathcal{H}} u = \{\hat{\mathcal{H}}_i u\}_{i=1}^N$ and $\hat{\mathcal{P}} u = \{\hat{\mathcal{P}}_i u\}_{i=1}^N$ are orthogonal in the sense of $a_h(.,.)$. The discrete solution of (7.2) can be decomposed as $u_h^* = \hat{\mathcal{P}} u_h^* + \hat{\mathcal{H}} u_h^*$ where for all $v \in X_h(D)$, $a_h(\hat{\mathcal{P}} u_h^*, \hat{\mathcal{P}} v) = f(\hat{\mathcal{P}} v)$ and

$$a_h(\hat{\mathcal{H}} u_h^*, \hat{\mathcal{H}} v) = f(\hat{\mathcal{H}} v). \tag{7.14}$$

Define $\Gamma \equiv (\cup_i \partial D_{ih_i})$ where ∂D_{ih_i} is the set of nodal points of ∂D_i. We note that the nodes on both side of $\cup_i \partial D_i$ belong to Γ. We denote the space $V = V_h(\Gamma)$ as the set of all functions v_h in $X_h(D)$ such that $\hat{\mathcal{P}} v_h = 0$, i.e., the space of discrete harmonic functions in the sense of $\hat{\mathcal{H}}_i$. The equation (7.14) is the Schur complement problem associated to (7.2).

7.4 Balancing domain decomposition

We design and analyze a BDD method Mandel [1993], Toselli and Widlund [2005] for solving (7.14) and use the general framework of balancing domain decomposition methods; see Toselli and Widlund [2005]. For $i = 1, \ldots, N$, let V_i be auxiliary spaces and I_i prolongation operators from V_i to V, and define the operators $\tilde{T}_i : V \to V_i$ as

$$b_i(\tilde{T}_i u, v) = a_h(u, I_i v) \quad \text{for all } v \in V_i.$$

and set $T_i = I_i \tilde{T}_i$. The coarse problem is defined as

$$a_h(P_0 u, v) = a_h(u, v) \quad \text{for all } v \in V_0.$$

Then the BDD method is defined as

$$T = P_0 + (I - P_0) \left(\sum_{i=1}^{N} T_i \right) (I - P_0). \tag{7.15}$$

We next define the prolongation operators I_i and the local spaces V_i for $i = 1, ..., N$, and the coarse space V_0. The bilinear forms b_i and a_h are given by (7.4) and (7.3), respectively.

7.4.1 Local problems

Let us denote by Γ_i the set of all nodes on ∂D_i and on neighboring faces $\bar{F}_{ji} \subset \partial D_j$. We note that the nodes of ∂F_{ji} (which are vertices of D_j) are included in Γ_i. Define V_i as the vector space associated to the nodal values on Γ_i and extended via $\hat{\mathcal{H}}_i$ inside D_i. We say that $u \in V_i$ if it can be represented as $u := \{u_l^{(i)}\}_{l \in \#(i)}$, where $\#(i) = \{i \text{ and } \cup j : F_{ij} \subset \partial D_i\}$. Here $u_i^{(i)}$ and $u_j^{(i)}$ stand for the nodal value of u on ∂D_i and $\bar{F}_{ji}$. We write $u = \{u_l^{(i)}\} \in V_i$ to refer to a function defined on Γ_i, and $u = \{u_i\} \in V$ to refer to a function defined on all Γ. Let us define the regular zero extension operator $\tilde{I}_i : V_i \to V$ as follows: Given $u \in V_i$, let $\tilde{I}_i u$ be equal to u on the nodes of Γ_i and zero on $\Gamma \backslash \Gamma_i$. Then we associate with each D_k, $k = 1, \cdots, N$, the discrete harmonic function u_k inside each D_k in the sense of $\hat{\mathcal{H}}_k$.

A face across D_i and D_j has two sides, the side inside $\bar{D}_i$, denoted by F_{ij}, and the side inside $\bar{D}_j$, denoted by F_{ji}. In addition, we assign to each face one master side $m(i,j) \in \{i,j\}$ and one slave side $s(i,j) \in \{i,j\}$. Then, using the *interface condition*, see below, we show that Theorem 7.1 holds, see below, with a constant C independent of the ρ_i, h_i and H_i.

The Interface Condition. We say that the coefficients $\{\rho_i\}$ and the local mesh sizes $\{h_i\}$ satisfy the *interface condition* if there exist constants C_0 and C_1, of order $O(1)$, such that for any face $F_{ij} = F_{ji}$ the following condition holds

$$h_{s(i,j)} \leq C_0 h_{m(i,j)} \quad \text{and} \quad \rho_{s(i,j)} \leq C_1 \rho_{m(i,j)}. \tag{7.16}$$

We associate with each D_i, $i = 1, \cdots, N$, the weighting diagonal matrices $D^{(i)} = \{D_l^{(i)}\}_{l \in \#(i)}$ on Γ_i defined as follows:

- On ∂D_i $(l = i)$

$$D_i^{(i)}(x) = \begin{cases} 1 & \text{if } x \text{ is a vertex of } \partial D_i, \\ 1 & \text{if } x \text{ is an interior node of a master face } F_{ij} \\ 0 & \text{if } x \text{ is an interior node of a slave face } F_{ij} \end{cases} \tag{7.17}$$

- On ∂D_j $(l = j)$

$$D_j^{(i)}(x) = \begin{cases} 0 & \text{if } x \text{ is an end point of } F_{ji}, \\ 1 & \text{if } x \text{ is an interior node of a slave face } F_{ji} \\ 0 & \text{if } x \text{ is an interior node of a master face } F_{ji} \end{cases} \tag{7.18}$$

- For $x \in F_{i0}$ we set $D_i^{(i)}(x) = 1$

The prolongation operators $I_i : V_i \to V$, $i = 1, \ldots, N$, are defined as $I_i = \tilde{I}_i D^{(i)}$ and they form a partition of unity on Γ described as

$$\sum_{i=1}^{N} I_i \tilde{I}_i^T = I_\Gamma, \tag{7.19}$$

where I_Γ is an identity operator.

7.4.2 Coarse problem

We define the coarse space $V_0 \subset V$ as

$$V_0 \equiv \text{Span}\{I_i \Phi^{(i)}, i = 1, ..., N\} \tag{7.20}$$

where $\Phi^{(i)} \in V_i$ denotes the function equal to one at every node of Γ_i.

Theorem 7.1 *If the interface condition (7.16) holds then there exists a positive constant C independent of h_i, H_i and the jumps of ρ_i such that*

$$a_h(u,u) \leq a_h(Tu,u) \leq C(1 + \log^2 \frac{H}{h}) a_h(u,u) \qquad \forall u \in V, \tag{7.21}$$

where T is defined in (7.15). Here $\log \frac{H}{h} = \max_i \log \frac{H_i}{h_i}$.

For the proof see Dryja et al. [2008].

7.5 Numerical experiments

In this section, we present numerical results for the preconditioner introduced in (7.15) and show that the bounds of Theorem 7.1 are reflected in the numerical tests. In particular we show that the interface condition (7.16) is necessary and sufficient.

We consider the domain $D = (0,1)^2$ divided into $N = M \times M$ squares subdomains D_i which are unions of fine elements, with $H = 1/M$. Inside each subdomain D_i we generate a structured triangulation with n_i subintervals in each coordinate direction and apply the discretization presented in Section 7.2 with $\delta = 4$. In the numerical experiments we use a red and black checkerboard type of subdomain partition. On the black subdomains we let $n_i = 2 * 2^{L_b}$ and on the red subdomains we let $n_i = 3 * 2^{L_r}$, where L_b and L_r are integers denoting the number of refinements inside each subdomain D_i. Hence the mesh sizes are $h_b = \frac{2^{-L_b}}{2N}$ and $h_r = \frac{2^{-L_r}}{3N}$, respectively. We solve the second order elliptic problem $-\mathrm{div}(\rho(x)\nabla u^*(x)) = 1$ in D with homogeneous Dirichlet boundary conditions. In the numerical experiments, we run PCG until the l_2 initial residual is reduced by a factor of 10^6.

In the first test we consider the constant coefficient case $\rho = 1$. We consider different values of $M \times M$ coarse partitions and different values of local refinements $L_b = L_r$, therefore keeping constant the mesh ratio $h_b/h_r = 3/2$. We place the master on the black subdomains. Table 7.1 lists the number of PCG iterations and in parenthesis the condition number estimate of the preconditioned system. We note that the interface condition (7.16) is satisfied. As expected from the analysis, the condition numbers appear to be independent of the number of subdomains and grow by a logarithmical factor when the size of the local problems increases. Note that in the case of continuous coefficients the Thereom 7.1 is valid without any assumption on h_b and h_r if the master sides are chosen on the larger meshes.

M↓ L_r →	0	1	2	3	4	5
2	13 (6.86)	17 (8.97)	18 (12.12)	19 (16.82)	21 (22.23)	22 (28.25)
4	18 (8.39)	22 (11.30)	26 (14.74)	30 (19.98)	33 (26.64)	36 (34.19)
8	20 (8.89)	24 (11.57)	28 (14.82)	32 (20.03)	37 (26.64)	42 (34.04)
16	19 (9.02)	24 (11.63)	27 (14.83)	32 (20.05)	37 (26.67)	42 (34.06)

Table 7.1: PCG/BDD iterations count and condition numbers for different sizes of coarse and local problems and constant coefficients ρ_i.

We now consider the discontinuous coefficient case where we set $\rho_i = 1$ on the black subdomains and $\rho_i = \mu$ on the red subdomains. The subdomains are kept fixed to 4×4. Table 7.2 lists the results on runs for different values of μ and for different levels of refinements on the red subdomains. On the black subdomains $n_i = 2$ is kept fixed. The masters are placed on the black subdomains. It is easy to see that the interface condition (7.16) holds if and only if μ is not large, which it is in agreement with the results in Table 7.2.

$\mu \downarrow L_r \rightarrow$	0	1	2	3	4
1000	90 (2556)	133 (3744)	184 (5362)	237 (7178)	303 (9102)
10	33 (29.16)	40 (42.31)	47 (58.20)	52 (75.55)	57 (94.59)
0.1	17 (8.28)	19 (8.70)	19 (9.21)	19 (9.50)	19 (9.65)
0.001	18 (8.83)	18 (8.95)	18 (9.46)	18 (9.83)	18 (10.08)

Table 7.2: PCG/BDD iterations count and condition numbers for different values of the coefficients and the local mesh sizes on the red subdomains only. The coefficients and the local mesh sizes on the black subdomains are kept fixed. The subdomains are also kept fixed to 4×4.

7.6 Final remarks

We end this paper by mentioning extensions and alternative Neumann-Neumann methods for DG discretizations where the Theorem 7.1 holds: 1) The BDD algorithms can be straightforwardly extended to three-dimensional problems; 2) Additive Schwarz versions and inexact local Neumann solvers can be considered; see Dryja et al. [2008]; 3) BDDC methods can be designed and analyzed, see Dryja et al. [2007b]; 4) On faces F_{ij} where h_i and h_j are of the same order, the values of (7.17) and (7.18) at interior nodes x of the faces F_{ij} and F_{ji} can be replaced by $\frac{\sqrt{\rho_i}}{\sqrt{\rho_i}+\sqrt{\rho_j}}$. 5) Similarly, on faces F_{ij} where ρ_i and ρ_j are of the same order, we can replace (7.17) and (7.18) at interior nodes x of the faces F_{ij} and F_{ji} by $\frac{h_i}{h_i+h_j}$. Finally, we remark the conditioning of the preconditioned systems deteriorates as we increase the penalty parameter δ to large values.

Bibliography

Antonietti, P. F. and Ayuso, B. (2005). Schwarz domain decomposition preconditioners for discontinuous Galerkin approximations of elliptic problems: non-overlapping case. Technical Report 20-VP, IMATI-CNR.

Arnold, D. N., Brezzi, F., Cockburn, B., and Martin, D. (2002). Unified analysis of discontinuous Galerkin method for elliptic problems. *SIAM J. Numer. Anal.*, 39(5):1749–1779.

Brenner, S. C. and Wang, K. (2005). Two-level additive Schwarz preconditioners for C^0 interior penalty methods. *Numer. Math.*, 102(2):231–255.

Dryja, M. (2003). On discontinuous Galerkin methods for elliptic problems with discontinuous coefficients. *Comput. Methods Appl. Math.*, 3(1):76–85.

Dryja, M., Galvis, J., and Sarkis, M. (2008). Neumann-Neumann methods for DG discretization of elliptic problems. In preparation.

Dryja, M. and Widlund, O. B. (1995). Schwarz methods of Neumann-Neumann type for three-dimensional elliptic finite elements problems. *Comm. Pure Appl. Math.*, 48(2):121–155.

Feng, X. and Karakashian, O. A. (2001). Two-level additive Schwarz methods for a discontinuous Galerkin approximation of second order elliptic problems. *SIAM J. Numer. Anal.*, 39(4):1343–1365.

Lasser, C. and Toselli, A. (2003). An overlapping domain decomposition preconditioners for a class of discontinuous Galerkin approximations of advection-diffusion problems. *Math. Comp.*, 72(243):1215–1238.

Mandel, J. (1993). Balancing domain decomposition. *Comm. Numer. Methods Engrg.*, 9(3):233–241.

Toselli, A. and Widlund, O. (2005). *Domain decomposition methods—algorithms and theory*, volume 34 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin.

Chapter 8

BDDC Methods for Discontinuous Galerkin Discretization of Elliptic Problems

A Discontinuous Galerkin (DG) discretization of Dirichlet problem for second order elliptic equations with discontinuous coefficients in 2-D is considered. For this discretization, Balancing Domain Decomposition with Constraints (BDDC) algorithms are designed and analyzed as an additive Schwarz method (ASM). The coarse and local problems are defined using special partitions of unity and edge constrains. Under assumption on the coefficients and mesh sizes across ∂D_i, where D_i are subregions of the original region D, a condition number estimate $C(1+\max_i \log(H_i/h_i))^2$ is established with C independent of h_i, H_i and the jumps of the coefficients. The algorithm is well suited for parallel computations and can be straightforwardly extended to the 3-D problems. The results of numerical tests are enclosed which confirm the theoretical results and the imposed assumption.

8.1 Introduction

In this paper, a Discontinuous Galerkin approximation of elliptic problems with discontinuous coefficients is considered. The problem is considered in a polygonal region D which is a union of disjoint polygonal subregions D_i. The discontinuities of the coefficients occur across ∂D_i. The problem is approximated by a conforming finite element method (FEM) on matching triangulation in each D_i and nonmatching one across ∂D_i. This kind of triangulation and composite discretization are motivated first of all by the regularity of the solution of the problem being discussed. Discrete problems are formulated using DG methods, symmetric and with interior penalty terms on the ∂D_i; see Arnold [1982], Arnold et al. [2002], Dryja [2003]. A goal of this paper is to design and analyze Balancing Domain Decomposition with Constraints (BDDC) algorithms for the resulting discrete problem; see Dohrmann [2003], Mandel et al. [2005] and Li and Widlund [2006] for conforming finite elements. In the first step, the problem is reduced to the Schur complement

problem with respect to unknowns on ∂D_i for $i = 1, \ldots, N$. For that, discrete harmonic functions defined in a special way are used. The method is designed and analyzed for the Schur complement problem using the general theory of ASMs; see Toselli and Widlund [2005]. The local problems are defined on D_i and faces of ∂D_j which are common to D_i plus zero average values constrains on edges of D_i or/and faces of D_j. The coarse spaces are defined using a special partitioning of unity with respect to D_i and introducing master and slave sides of substructures. A side $F_{ij} = \partial D_i \cap \partial D_j$ is master when $\rho_i \geq \rho_j$, otherwise it is slave, so if $F_{ij} \subset \partial D_i$ is master then $F_{ji} \subset \partial D_j$, $F_{ij} = F_{ji}$, is slave. The h_i- and h_j- triangulations on F_{ij} and F_{ji}, respectively, are built in a way that $h_i \geq h_j$ if $\rho_i \geq \rho_j$ where h_i and h_j are the parameters of these triangulations. It is proved that the algorithm is almost optimal and its rate of convergence is independent of h_i and h_j, the number of subdomains D_i and the jumps of coefficients. The algorithm is well suited for parallel computations and it can be straightforwardly extended to the problems in the 3-D cases.

DG methods are becoming more and more popular for approximation of PDEs; see Arnold [1982], Arnold et al. [2002] and literature therein. There are also some papers devoted to algorithms for solving the resulting discrete problem, in particular domain decomposition methods. We first mention Feng and Karakashian [2001] and Lasser and Toselli [2003] where overlapping Schwarz methods were proposed and analyzed for DG discretization of elliptic problems with continuous coefficients. In Dryja [2003] for the considered discrete problem, a multilevel ASM is designed and analyzed but it is not optimal. In Brenner and Wang [2005] a two-level ASM is proposed and analyzed for DG discretization of fourth order problems. Up to our knowledge BDDC algorithms for DG discretization of elliptic problems with continuous and discontinuous coefficients have not been analyzed in literature. We note that part of the analysis presented here has previously appeared as a technical report for analyzing several DG preconditioners of Neumann-Neumann type; see Dryja and Sarkis [2006]. In Dryja et al. [2007a] we have also successfully extended these preconditioners to the Balancing Domain Decomposition (BDD) method.

The paper is organized as follows. In Section 8.2 the differential problem and its DG discretization are formulated. In Section 8.3 the Schur complement problem is derived using discrete harmonic function in a special way. Some technical tools are presented in Section 8.4. Sections 8.5 and 8.6 are devoted to designing a BDDC algorithm while Section 8.7 and 8.8 are devoted to the proof of the main result Theorem 8.5. In Section 8.9 are introduced coarse spaces of dimension twice less than those defined in Section 8.6. Finally in Section 8.10 some numerical experiments are presented which confirm the theoretical results. The enclosed numerical results show that the introduced assumption on the coefficients and the parameter steps are sufficient and necessary.

8.2 Differential and discrete problems

8.2.1 Differential problem

Consider the following problem: Find $u^* \in H_0^1(D)$ such that

$$a(u^*, v) = f(v), \quad \forall v \in H_0^1(D) \tag{8.1}$$

where

$$a(u, v) := \sum_{i=1}^{N} \int_{D_i} \rho_i \nabla u \nabla v dx \quad \text{and} \quad f(v) := \int_D f v dx.$$

We assume that $\bar{D} = \cup_{i=1}^N \bar{D}_i$ and the substructures D_i are disjoint shaped regular polygonal subregions of diameter $O(H_i)$ and form a geometrical conforming partition of D, i.e., $\forall i \neq j$ the intersection $\partial D_i \cap \partial D_j$ is empty or is a common vertex or face of ∂D_i and ∂D_j. We assume $f \in L^2(D)$ and for simplicity of presentation let ρ_i be a positive constant.

8.2.2 Discrete problem

Let us introduce the shape regular triangulation in each D_i with triangular elements and h_i as mesh parameter. The resulting triangulation on D is in general nonmatching across ∂D_i. Let $X_i(D_i)$ be the regular finite element (FE) space of piecewise linear continuous functions in D_i. Note that we do not assume that functions in $X_i(D_i)$ vanish on $\partial D_i \cap \partial D$. Define

$$X_h(D) := X_1(D_1) \times \cdots \times X_N(D_N).$$

A discrete problem obtained by DG method, see Arnold et al. [2002], Dryja [2003], is of the form:

Find $u_h^* \in X_h(D)$ such that

$$a_h(u_h^*, v_h) = f(v_h), \quad \forall v_h \in X_h(D) \tag{8.2}$$

where

$$a_h(u, v) = \sum_{i=1}^{N} \hat{a}_i(u, v) \quad \text{and} \quad f(v) = \sum_{i=1}^{N} \int_{D_i} f v_i dx, \tag{8.3}$$

$$\hat{a}_i(u, v) := a_i(u, v) + s_i(u, v) + p_i(u, v), \tag{8.4}$$

$$a_i(u, v) := \int_{D_i} \rho_i \nabla u_i \nabla v_i dx, \tag{8.5}$$

$$s_i(u, v) := \sum_{F_{ij} \subset \partial D_i} \int_{F_{ij}} \frac{\rho_{ij}}{l_{ij}} \left(\frac{\partial u_i}{\partial n}(v_j - v_i) + \frac{\partial v_i}{\partial n}(u_j - u_i) \right) ds, \tag{8.6}$$

$$p_i(u, v) := \sum_{F_{ij} \subset \partial D_i} \int_{F_{ij}} \frac{\rho_{ij}}{l_{ij}} \frac{\delta}{h_{ij}} (u_j - u_i)(v_j - v_i) ds, \tag{8.7}$$

and $u = \{u_i\}_{i=1}^N \in X_h(D)$, $v = \{v_i\}_{i=1}^N \in X_h(D)$. We set $l_{ij} = 2$ when $F_{ij} = \partial D_i \cap \partial D_j$ is a common face of ∂D_i and ∂D_j, and define $\rho_{ij} := 2\rho_i\rho_j/(\rho_i + \rho_j)$ as the harmonic average of ρ_i and ρ_j, and $h_{ij} := 2h_ih_j/(h_i + h_j)$. In order to simplify

notations we include the index $j = \partial$ and put $l_{i\partial} := 1$ when $F_{i\partial} := \partial D_i \cap \partial D$ has positive measure. We also set $u_\partial = 0$, $v_\partial = 0$ and define $\rho_{i\partial} := \rho_i$ and $h_{i\partial} := h_i$. The $\frac{\partial}{\partial n}$ denotes the outward normal derivative on ∂D_i, and δ is the penalty positive parameter. We note that when ρ_{ij} is given by harmonic average, $\min\{\rho_i, \rho_j\} \leq \rho_{ij} \leq 2\min\{\rho_i, \rho_j\}$.

We also define

$$d_i(u, v) := a_i(u, v) + p_i(u, v), \tag{8.8}$$

and

$$d_h(u, v) := \sum_{i=1}^{N} d_i(u, v). \tag{8.9}$$

It is known that there exists a $\delta_0 = O(1) > 0$ such that for $\delta \geq \delta_0$, we obtain $2|s_i(u, u)| \leq d_i(u, u)$ and therefore, the problem (8.2) is elliptic and has a unique solution. An a priori error estimates of the method are optimal for the continuous coefficient, see Arnold [1982], Arnold et al. [2002], but is not for discontinuous coefficients; see Dryja [2003]. In the later case the error is $O(h^{1/2})$ only in the H^1- broken norm if the solution of (8.1) $u^* \in H^{3/2+\epsilon}(D)$, with $\epsilon > 0$. On the other hand we cannot expect more regularity of u^* in the case of discontinuous coefficients in a general case.

We use the d_h-norm, also called broken norm in $X_h(D)$ with weights given by ρ_i and $\frac{\delta}{l_{ij}}\frac{\rho_{ij}}{h_{ij}}$. For $u = \{u_i\} \in X_h(D)$ we note that

$$d_h(u, u) = \sum_{i=1}^{N} \{\rho_i \parallel \nabla u_i \parallel^2_{L^2(D_i)} + \sum_{F_{ij} \subset \partial D_i} \frac{\delta}{l_{ij}} \frac{\rho_{ij}}{h_{ij}} \int_{F_{ij}} (u_i - u_j)^2 ds\}. \tag{8.10}$$

Lemma 8.1 *There exists $\delta_0 > 0$ such that for $\delta \geq \delta_0$ we have that for all $u \in X_h(D)$ holds*

$$\gamma_0 d_i(u, u) \leq \hat{a}_i(u, u) \leq \gamma_1 d_i(u, u), \qquad i = 1, \ldots, N, \tag{8.11}$$

and

$$\gamma_0 d_h(u, u) \leq a_h(u, u) \leq \gamma_1 d_h(u, u) \tag{8.12}$$

where γ_0 and γ_1 are positive constants independent of the ρ_i, h_i and H_i.

The proof essentially follows from (8.38), see below, or refer to Dryja [2003].

8.3 Schur complement problem

In this section we derive a Schur complement version for the problem (8.2). We first introduce some auxiliary notations.

Let $u = \{u_i\} \in X_h(D)$ be given. We can represent u_i as

$$u_i = \mathcal{H}_i u_i + \mathcal{P}_i u_i \tag{8.13}$$

where $\mathcal{H}_i u_i$ is the discrete harmonic part of u_i in the sense of $a_i(.,.)$, see (8.5), i.e.,

$$a_i(\mathcal{H}_i u_i, v_i) = 0 \qquad \forall v_i \in \overset{o}{X}_i(D_i) \tag{8.14}$$

$$\mathcal{H}_i u_i = u_i \quad \text{on} \quad \partial D_i, \tag{8.15}$$

while $\mathcal{P}_i u_i$ is the projection of u_i on $\overset{o}{X}_i(D_i)$ in the sense of $a_i(.,.)$, i.e.

$$a_i(\mathcal{P}_i u_i, v_i) = a_i(u_i, v_i), \quad \forall v_i \in \overset{o}{X}_i(D_i). \tag{8.16}$$

Here $\overset{o}{X}_i(D_i)$ is a subspace of $X_i(D_i)$ of functions which vanish on ∂D_i, and $\mathcal{H}_i u_i$ is the classical discrete harmonic part of u_i. Let us denote by $\overset{o}{X}_h(D)$ the subspace of $X_h(D)$ defined by $\overset{o}{X}_h(D) := \{\overset{o}{X}_i(D_i)\}_{i=1}^N$ and consider the global projections $\mathcal{H}u := \{\mathcal{H}_i u_i\}_{i=1}^N$ and $\mathcal{P}u := \{\mathcal{P}_i u_i\}_{i=1}^N : X_h(D) \to \overset{o}{X}_h(D)$ in the sense of $\sum_{i=1}^N a_i(.,.)$. Hence, a function $u \in X_h(D)$ can therefore be decomposed as

$$u = \mathcal{H}u + \mathcal{P}u. \tag{8.17}$$

The function $u \in X_h(D)$ can also be represented as

$$u = \hat{\mathcal{H}}u + \hat{\mathcal{P}}u \tag{8.18}$$

where $\hat{\mathcal{P}}u = \{\hat{\mathcal{P}}_i u_i\}_{i=1}^N : X_h(D) \to \overset{o}{X}_h(D)$ is the projection in the sense of $a_h(.,.)$, the original bilinear form of (8.2), see (8.3). Since $\hat{\mathcal{P}}_i u_i \in \overset{o}{X}_i(D_i)$ and $v_i \in \overset{o}{X}_i(D_i)$, we have

$$a_i(\hat{\mathcal{P}}_i u, v_i) = a_h(u, v_i).$$

The discrete solution of (8.2) can be decomposed as $u_h^* = \hat{\mathcal{H}}u_h^* + \hat{\mathcal{P}}u_h^*$. To find $\hat{\mathcal{P}}u_h^*$ we need to solve the following set of usual discrete Dirichlet problems:

Find $\hat{\mathcal{P}}_i u_h^* \in \overset{o}{X}_i(D)$ such that

$$a_i(\hat{\mathcal{P}}_i u_h^*, v_i) = f(v_i), \qquad \forall v_i \in \overset{o}{X}_i(D_i) \tag{8.19}$$

for $i = 1, \cdots, N$. Note that these problems are local and independent, so they can be solved in parallel. This is a precomputational step.

We now formulate the problem for $\hat{\mathcal{H}}u_h^*$. Let $\hat{\mathcal{H}}_i u$ be the discrete harmonic part of u in the sense of $\hat{a}_i(.,.)$, see (8.4), where $\hat{\mathcal{H}}_i u \in X_i(D_i)$ is the solution of

$$\hat{a}_i(\hat{\mathcal{H}}_i u, v_i) = 0 \qquad \forall v_i \in \overset{o}{X}_i(D_i), \tag{8.20}$$

$$u_i \quad \text{on} \quad \partial D_i \quad \text{and} \quad u_j \quad \text{on} \quad F_{ji} \subset \partial D_j \quad \text{are given} \tag{8.21}$$

where u_j are given on $F_{ji} = \partial D_i \cap \partial D_j$. We point out that for $v_i \in \overset{o}{X}_i(D_i)$ we have

$$\hat{a}_i(u_i, v_i) = (\rho_i \nabla u_i, \nabla v_i)_{L^2(D_i)} + \sum_{F_{ij} \subset \partial D_i} \frac{\rho_{ij}}{l_{ij}} (\frac{\partial v_i}{\partial n}, u_j - u_i)_{L^2(F_{ij})}. \tag{8.22}$$

Note that (8.20) - (8.21) has a unique solution. To see this, let us rewrite (8.20) in the form

$$\rho_i(\nabla \hat{\mathcal{H}}_i u, \nabla \varphi_i^k)_{L^2(D_i)} = - \sum_{F_{ij} \subset \partial D_i} \frac{\rho_{ij}}{l_{ij}} (\frac{\partial \varphi_i^k}{\partial n}, u_j - u_i)_{L^2(F_{ij})} \tag{8.23}$$

where φ_i^k are nodal basis functions of $\overset{o}{X}_i\,(D_i)$ associated with interior nodal points x_k of the h_i-triangulation of D_i. Note that $\frac{\partial\varphi_i^k}{\partial n}$ does not vanish on ∂D_i when x_k is a node of an element touching ∂D_i. We see that $\hat{\mathcal{H}}_i u$ is a special extension into D_i where u is given on ∂D_i and on all the F_{ji}, and therefore, it depends on the values of u_j given on $F_{ji} = \partial D_i \cap \partial D_j$ and on $F_{\partial i}$ (we already have assumed $u_\partial = 0$ for $j = \partial$). Note that $\hat{\mathcal{H}}_i u$ is the discrete harmonic except at nodal points close to ∂D_i. We will call sometimes $\hat{\mathcal{H}}_i u$ as discrete harmonic in special sense, i.e., in the sense of $\hat{a}_i(.,.)$ or $\hat{\mathcal{H}}_i$. We set that $\hat{\mathcal{H}}u = \{\hat{\mathcal{H}}_i u\}_{i=1}^N \in X_h(D)$.

Note that (8.20) is obtained from

$$a_h(\hat{\mathcal{H}}u, v) = 0 \tag{8.24}$$

for $u \in X_h(D)$ and when taking $v = \{v_i\}_{i=1}^N \in \overset{o}{X}_h(D)$. It is easy to see that $\hat{\mathcal{H}}u = \{\hat{\mathcal{H}}_i u\}_{i=1}^N$ and $\hat{\mathcal{P}}u = \{\hat{\mathcal{P}}_i u_i\}_{i=1}^N$ are orthogonal in the sense of $a_h(.,.)$, i.e.

$$a_h(\hat{\mathcal{H}}u, \hat{\mathcal{P}}v) = 0, \qquad u, v \in X^h(D). \tag{8.25}$$

In addition,

$$\mathcal{H}\hat{\mathcal{H}}u = \mathcal{H}u, \qquad \hat{\mathcal{H}}\mathcal{H}u = \hat{\mathcal{H}}u \tag{8.26}$$

since $\hat{\mathcal{H}}u$ and $\mathcal{H}u$ do not change the values of u on all the nodes on boundaries of the subdomais D_i also denoted by

$$\Gamma := (\cup_i \partial D_{ih_i}), \tag{8.27}$$

where ∂D_{ih_i} is the set of nodal points of ∂D_i. We note that definition of Γ includes the nodes on both sides of $\cup_i \partial D_i$.

We are now in the position to be able to derive a Schur complement problem for (8.2). Let us apply the decomposition (8.18) in (8.2). We get

$$a_h(\hat{\mathcal{H}}u_h^* + \hat{\mathcal{P}}u_h^*, \hat{\mathcal{H}}v_h + \hat{\mathcal{P}}v_h) = f(\hat{\mathcal{H}}v_h + \hat{\mathcal{P}}v_h)$$

or

$$a_h(\hat{\mathcal{H}}u_h^*, \hat{\mathcal{H}}v_h) + 2a_h(\hat{\mathcal{H}}u_h^*, \hat{\mathcal{P}}v_h) + a_h(\hat{\mathcal{P}}u_h^*, \hat{\mathcal{P}}v_h) = f(\hat{\mathcal{H}}v_h) + f(\hat{\mathcal{P}}v_h).$$

Using (8.19) and (8.24) we have

$$a_h(\hat{\mathcal{H}}u_h^*, \hat{\mathcal{H}}v_h) = f(\hat{\mathcal{H}}v_h), \qquad \forall v_h \in X_h(D). \tag{8.28}$$

This is the Schur complement problem for (8.2). We denote the space $V_h(\Gamma)$ or in short notation V, which we will use later, as the set of all functions v_h in $X_h(D)$ such $\hat{\mathcal{P}}v_h = 0$, i.e., the space of discrete harmonic functions in the sense of the $\hat{\mathcal{H}}_i$. We rewrite the Schur complement problem as:

Find $u_h^* \in V_h(\Gamma)$ such that

$$\mathcal{S}(u_h^*, v_h) = g(v_h), \qquad \forall v_h \in V_h(\Gamma) \tag{8.29}$$

where here and below $u_h^* \equiv \hat{\mathcal{H}}u_h^*$, and

$$\mathcal{S}(u_h, v_h) = a_h(\hat{\mathcal{H}}u_h, \hat{\mathcal{H}}v_h), \qquad g(v_h) = f(\hat{\mathcal{H}}v_h). \tag{8.30}$$

This problem has a unique solution.

8.4 Technical tools

Our main goal is to design and analyze a BDDC method for solving (8.29). This will be done in the next section. We now introduce some notations and facts used later. Let $u = \{u_i\}_{i=1}^N \in X_h(D)$ and $v = \{v_i\}_{i=1}^N \in X_h(D)$. Let $d_i(.,.)$ and $d_h(.,.)$ be the bilinear forms defined in (8.8) and (8.9).

Note that for $u, v \in \overset{o}{X}_h(D)$

$$d_i(u,v) = a_i(u,v) = \rho_i(\nabla u_i, \nabla v_i)_{L^2(D_i)} \tag{8.31}$$

and for $u \in X_h(D)$

$$\gamma_0 d_h(u,u) \leq a_h(u,u) \leq \gamma_1 d_h(u,u) \tag{8.32}$$

in view of Lemma 8.1, where γ_0 and γ_1 are positive constants independent of h_i, H_i and ρ_i. The next lemma shows the equivalence between discrete harmonic functions in the sense $\mathcal{H}$ and in the sense $\hat{\mathcal{H}}$, and therefore we can take advantage of all the discrete Sobolev results known for $\mathcal{H}$ discrete harmonic extensions.

Lemma 8.2 *For $u \in X_h(D)$ we have*

$$d_i(\mathcal{H}u, \mathcal{H}u) \leq d_i(\hat{\mathcal{H}}u, \hat{\mathcal{H}}u) \leq C d_i(\mathcal{H}u, \mathcal{H}u), \quad i = 1, \ldots, N, \tag{8.33}$$

and

$$d_h(\mathcal{H}u, \mathcal{H}u) \leq d_h(\hat{\mathcal{H}}u, \hat{\mathcal{H}}u) \leq C d_h(\mathcal{H}u, \mathcal{H}u) \tag{8.34}$$

where $\mathcal{H}u = \{\mathcal{H}_i u_i\}_{i=1}^N$ and $\hat{\mathcal{H}}u = \{\hat{\mathcal{H}}_i u\}_{i=1}^N$ are defined by (8.14) - (8.15) and (8.20) - (8.21) respectively, and C is a positive constant independent of h_i, u, ρ_i and H_i.

Proof. We note that $\mathcal{P}$ and $\mathcal{H}$ are projections in the sense of $\sum_i a_i(.,.)$ while $\hat{\mathcal{P}}$ and $\hat{\mathcal{H}}$ are projections in the sense of $a_h(.,.)$. Therefore, the left hand side of (8.34) follows from properties of minimum energy of discrete harmonic extensions in the $\sum_i a_i(.,.)$ sense. To prove the right hand side of (8.34) note that

$$\begin{aligned} d_h(\hat{\mathcal{H}}u, \hat{\mathcal{H}}u) &= d_h(\hat{\mathcal{H}}u, \mathcal{H}\hat{\mathcal{H}}u + \mathcal{P}\hat{\mathcal{H}}u) \\ &= d_h(\hat{\mathcal{H}}u, \mathcal{H}u)) + d_h(\hat{\mathcal{H}}u, \mathcal{P}\hat{\mathcal{H}}u) \end{aligned} \tag{8.35}$$

in view of (8.26). The first term is estimated as

$$d_h(\hat{\mathcal{H}}u, \mathcal{H}u) \leq \varepsilon d_h(\hat{\mathcal{H}}u, \hat{\mathcal{H}}u) + \frac{1}{4\varepsilon} d_h(\mathcal{H}u, \mathcal{H}u), \tag{8.36}$$

with arbitrary $\varepsilon > 0$. To estimate the second term in the right hand side of (8.35) note that for $v := \mathcal{P}\hat{\mathcal{H}}u \in \overset{o}{X}(D)$ and using (8.23), we get

$$\begin{aligned} d_h(\hat{\mathcal{H}}u, v) &= \sum_{i=1}^N \rho_i(\nabla \hat{\mathcal{H}}_i u_i, \nabla v_i)_{L^2(D_i)} \\ &= -\sum_{i=1}^N \sum_{F_{ij} \subset \partial D_i} \frac{\rho_{ij}}{l_{ij}} (\frac{\partial v_i}{\partial n}, u_j - u_i)_{L^2(F_{ij})}. \end{aligned} \tag{8.37}$$

The term in the right hand side of (8.37) are estimated as

$$\begin{aligned}|\rho_{ij}(\frac{\partial v_i}{\partial n}, u_j - u_i)_{L^2(F_{ij})}| \quad &\leq \rho_{ij} \parallel \frac{\partial v_i}{\partial n} \parallel_{L^2(F_{ij})} \parallel u_i - u_j \parallel_{L^2(F_{ij})} \\ &\leq C\frac{\rho_{ij}}{h_i^{1/2}} \parallel \nabla v_i \parallel_{L^2(D_i)} \parallel u_i - u_j \parallel_{L^2(F_{ij})} \\ &\leq C\frac{\rho_{ij}}{h_{ij}^{1/2}} \parallel \nabla v_i \parallel_{L^2(D_i)} \parallel u_i - u_j \parallel_{L^2(F_{ij})} \\ &\leq C\{\varepsilon\rho_{ij} \parallel \nabla v_i \parallel^2_{L^2(D_i)} + \frac{\rho_{ij}}{4\varepsilon h_{ij}} \parallel u_i - u_j \parallel^2_{L^2(F_{ij})}\} \\ &\leq C\{2\varepsilon\rho_i \parallel \nabla v_i \parallel^2_{L^2(D_i)} + \frac{\rho_{ij}}{4\varepsilon h_{ij}} \parallel u_i - u_j \parallel^2_{L^2(F_{ij})}\},\end{aligned} \tag{8.38}$$

where we have used that $h_{ij} \leq 2h_i$ and $\rho_{ij} \leq 2\rho_i$. Substituting this into (8.37), we get

$$d_h(\hat{\mathcal{H}}u, v) \leq C\sum_{i=1}^{N}\{2\varepsilon\rho_i \parallel \nabla\mathcal{P}_i\hat{\mathcal{H}}_i u_i \parallel^2_{L^2(D_i)} + \frac{\rho_{ij}}{4h_{ij}\varepsilon}\sum_{F_{ij}\subset\partial D_i} \parallel u_i - u_j \parallel^2_{L^2(F_{ij})}\}, \tag{8.39}$$

and using

$$\parallel \nabla\mathcal{P}_i\hat{\mathcal{H}}_i u_i \parallel_{L^2(D_i)} \leq \parallel \nabla\hat{\mathcal{H}}_i u_i \parallel_{L^2(D_i)},$$

we obtain

$$d_h(\hat{\mathcal{H}}u, v) \leq C\{\varepsilon d_h(\hat{\mathcal{H}}u, \hat{\mathcal{H}}u) + \frac{1}{4\varepsilon}d_h(\mathcal{H}u, \mathcal{H}u)\}, \tag{8.40}$$

and then

$$d_h(\hat{\mathcal{H}}u, \hat{\mathcal{H}}u) \leq C\{\varepsilon d_h(\hat{\mathcal{H}}u, \hat{\mathcal{H}}u) + \frac{1}{4\varepsilon}d_h(\mathcal{H}u, \mathcal{H}u)\}.$$

Choosing a sufficiently small ε, the right hand side of (8.34) follows.

8.5 Balancing domain decomposition with constraints

We design and analyze BDDC methods for solving the Schur complement problem (8.29); see Dohrmann [2003], Mandel et al. [2005], Li and Widlund [2006] for conforming elements. We follow the general framework of ASM as stated below in Lemma 8.3; see Toselli and Widlund [2005]. For $i = 0, \ldots, N$, let V_i be auxiliary spaces and I_i prolongation operators from V_i to V, and define the operators $\tilde{T}_i : V \to V_i$ as

$$b_i(\tilde{T}_i u, v) = a_h(u, I_i v) \quad \forall v \in V_i$$

and set $T_i = I_i\tilde{T}_i$. Then the ASMs, in particular the BDDC method, are defined as

$$T = \sum_{i=0}^{N} T_i. \tag{8.41}$$

The bilinear form a_h is defined in (8.3). The bilinear forms b_i, the operators I_i, and the spaces V_i, $i = 0, \ldots, N$, are defined in the next subsections.

Lemma 8.3 *Suppose the following three assumptions hold:*

i) There exists a constant C_0 such that for all $u \in V$ there exists a decomposition $u = \sum_{i=0}^{N} I_i u^{(i)}$ with $u^{(i)} \in V_i$, $i = 0, \ldots, N$, and

$$\sum_{i=0}^{N} b_i(u^{(i)}, u^{(i)}) \leq C_0^2 a_h(u, u).$$

ii) There exist constants $\epsilon_{ij}, i, j = 1, \ldots, N$, such that for all $u^{(i)} \in V_i$, $u^{(j)} \in V_j$,

$$a_h(I_i u^{(i)}, I_j u^{(j)}) \leq \epsilon_{ij} a_h(I_i u^{(i)}, I_i u^{(i)})^{1/2} a_h(I_j u^{(j)}, I_j u^{(j)})^{1/2}.$$

iii) There exists a constant ω such that

$$a_h(I_i u, I_i u) \leq \omega b_i(u, u) \quad \forall u \in V_i, \quad i = 0, \ldots, N.$$

Then, T is invertible and

$$C_0^2 a_h(u, u) \leq a_h(Tu, u) \leq (\rho(\epsilon) + 1)\omega a_h(u, u), \quad \forall u \in V.$$

Here, $\rho(\epsilon)$ is the spectral radius of the matrix $\epsilon = \{\epsilon\}_{i,j=1}^{N}$.

8.5.1 Notations and the interfacing condition

Let us denote by Γ_i the set of all nodes on ∂D_i and on neighboring faces $F_{ji} \subset \partial D_j$. We note that the nodes of ∂F_{ji} (which are vertices of D_j) are included in Γ_i. Define W_i as the vector space associated to the nodal values on Γ_i and extended via $\hat{\mathcal{H}}_i$ inside D_i. We say that $u^{(i)} \in W_i$ if $u^{(i)}$ is represented as $u^{(i)} := \{u_l^{(i)}\}_{l \in \#(i)}$, where $\#(i) = \{i \text{ and } \cup j : F_{ij} \subset \partial D_i\}$. Here $u_i^{(i)}$ and the $u_j^{(i)}$ stand for the nodal values of $u^{(i)}$ on ∂D_i and the F_{ji}, respectively. We write $u = \{u_i\} \in V$ to refer to a function defined on all Γ with each u_i defined (only) on ∂D_i. We point out that F_{ij} and F_{ji} are geometrically the same even though the mesh on F_{ij} is inherited from the D_i mesh while the mesh on F_{ji} corresponds to the D_j mesh.

Denote by $\Lambda_i := \{F_{ij} : F_{ij} \subset \partial D_i\} \cup \{F_{ji} : F_{ji} = F_{ij},\ F_{ji} \subset \partial D_j\}$ the set of all faces of D_i and all faces of D_j touching D_i. Given $u^{(i)} \in W_i$ and $F_{\ell k} \in \Lambda_i$ we use the notation

$$\overline{u}_{\ell k}^{(i)} = \frac{1}{|F_{\ell k}|} \int_{F_{\ell k}} u^{(i)} ds.$$

Let us define the regular zero extension operator $\tilde{I}_i : W_i \to V$ as follows: Given $u^{(i)} \in W_i$, let $\tilde{I}_i u^{(i)}$ be equal to $u^{(i)}$ on nodes Γ_i and zero on $\Gamma \backslash \Gamma_i$.

A face across D_i and D_j has two sides, the side contained in ∂D_i, denoted by F_{ij}, and the side contained in ∂D_j, denoted by F_{ji}. In addition, we assign to each pair $\{F_{ij}, F_{ji}\}$ a master (mortar) and a slave (nonmortar) side. If F_{ij} is a slave side then F_{ji} is a master side and vice versa. If F_{ij} is a slave side we will use the notation δ_{ij} (instead of F_{ij}) to recall this fact while if F_{ij} is mortar side we will use the notation γ_{ij}. The choice of slave-master sides are such that the *interfacing condition*, stated next, can be satisfied. In this case Theorem 8.5 below holds with constant C independent of the ρ_i, h_i and H_i.

Assumption 8.4 (The interfacing condition) *We say that the coefficients* $\{\rho_i\}$ *and the local mesh sizes* $\{h_i\}$ *satisfy the interfacing condition if exist constants* C_0 *and* C_1*, of order* $O(1)$*, such that for any face* F_{ij} *the following conditions hold*

$$\begin{cases} h_i \leq C_0 h_j \quad \textit{and} \quad \rho_i \leq C_1 \rho_j & \textit{if } F_{ij} \textit{ is a slave side, or} \\ h_j \leq C_0 h_i \quad \textit{and} \quad \rho_j \leq C_1 \rho_i & \textit{if } F_{ij} \textit{ is a master side.} \end{cases} \tag{8.42}$$

We associate with each D_i, $i = 1, \cdots, N$, the weighting diagonal matrices $D^{(i)} = \{D_l^{(i)}\}_{l \in \#(i)}$ on Γ_i defined as follows:

- On ∂D_i $(l = i)$

$$D_i^{(i)}(x) = \begin{cases} 1 & \text{if } x \text{ is a vertex of} \partial D_i, \\ 1 & \text{if } x \text{ is interior node of a master face } F_{ij}, \\ 0 & \text{if } x \text{ is interior node of a slave face } F_{ij}, \end{cases} \tag{8.43}$$

- On F_{ji} $(l = j)$

$$D_j^{(i)}(x) = \begin{cases} 0 & \text{if } x \text{ is end point of the face } F_{ji}, \\ 1 & \text{if } x \text{ is interior of a slave face } F_{ji}, \\ 0 & \text{if } x \text{ is interior node of a master face } F_{ji}, \end{cases} \tag{8.44}$$

- For $x \in F_{i\partial}$ we set $D_i^{(i)}(x) = 1$.

The prolongation operators $I_i : W_i \to V$, $i = 1, \ldots, N$, are defined as

$$I_i = \tilde{I}_i D^{(i)}, \tag{8.45}$$

and they form a partition of unity on Γ described as

$$\sum_{i=1}^{N} I_i \tilde{I}_i^T = I_\Gamma. \tag{8.46}$$

8.6 Local and global spaces

The local spaces $V_i = V_i(\Gamma_i)$, $i = 1, \ldots, N$, are defined as the subspaces of W_i of functions with zero average faces values on all faces F_{ij} and F_{ij} associated to the subdomain D_i, i.e., for all $F_{\ell k} \in \Lambda_i$.

For $u^{(i)}, v^{(i)} \in V_i(\Gamma_i)$ we define the local bilinear form b_i as

$$b_i(u^{(i)}, v^{(i)}) := \hat{a}_i(\tilde{I}_i u^{(i)}, \tilde{I}_i v^{(i)}), \tag{8.47}$$

where the bilinear form $\hat{a}_i$ was defined in (8.4).

Now we define a BDDC coarse space. As in BDDC methods, here we define the coarse space using local bases and imposing continuity condition with respect to the primal variables; see Dohrmann [2003], Mandel et al. [2005], Li and Widlund [2006].

Recall that $\Lambda_i := \{F_{ij} : F_{ij} \subset D_i\} \cup \{F_{ji} : F_{ji} = F_{ij},\ F_{ji} \subset D_j\}$ is the set of all faces of D_i and all faces of D_j touching D_i. For $F_{\ell k} \in \Lambda_i$ define the local coarse basis function $\Phi^{(i)}_{F_{\ell k}} \in W_i$ by

$$b_i(\Phi^{(i)}_{F_{\ell k}}, v) = 0, \quad \forall v \in V_i(\Gamma_i) \tag{8.48}$$

with

$$\frac{1}{|F_{\ell k}|}\int_{F_{\ell k}} \Phi^{(i)}_{F_{\ell k}} = 1$$

and

$$\int_{F_{\ell' k'}} \Phi^{(i)}_{F_{\ell k}} = 0, \quad \forall F_{\ell' k'} \neq F_{\ell k} \text{ with } F_{\ell' k'} \in \Lambda_i.$$

Note that $\Phi^{(i)}_{F_{k\ell}} \neq \Phi^{(i)}_{F_{\ell k}}$.

Define $V_{0i} = V_{0i}(\Gamma_i) := \mathrm{Span}\{\Phi^{(i)}_{F_{\ell k}} : F_{\ell k} \in \Lambda_i\} \subset W_i$. Then (8.48) implies that V_i is $\hat{\mathcal{H}}_i-$orthogonal to V_{0i}, and W_i is a direct sum of V_{0i} and V_i, i.e., $V_{0i} \oplus V_i = W_i$.

The global coarse space V_0 is defined as the set of all $u_0 := \{u_0^{(i)}\} \in \prod_{i=1}^N V_{0i}(\Gamma_i)$ such that for $i, j = 1, \ldots, N$, we have

$$\overline{u}^{(i)}_{0\ell k} = \overline{u}^{(j)}_{0\ell k} \quad \forall F_{\ell k} \in \Lambda_i \cap \Lambda_j. \tag{8.49}$$

The coarse prolongation operator $I_0 : V_0 \to V$ is defined as $I_0 u_0 = \sum_{i=1}^N I_i u_0^{(i)}$ and the bilinear form b_0 is of the form

$$b_0(u_0, v_0) := \sum_{i=1}^N b_i(u_0^{(i)}, v_0^{(i)}).$$

8.7 Main result

In this section we state and proof our main result.

Theorem 8.5 *Let the Assumption 8.4 be satisfied. Then there exists a positive constant C independent of h_i, H_i and the jumps of ρ_i such that*

$$a_h(u,u) \leq a_h(Tu,u) \leq C\left(1 + \log\frac{H}{h}\right)^2 a_h(u,u) \qquad \forall u \in V, \tag{8.50}$$

where T is defined in (8.41). Here $\log\frac{H}{h} = \max_i \log\frac{H_i}{h_i}$.

Proof. By the general theorem of ASMs we need to check the three key assumptions of Lemma 8.3.

Assumption(i) We prove that for $u = \{u_i\}_{i=1}^N \in V$ there exist $u_0 \in V_0$ and $u^{(i)} \in V_i$ such that

$$I_0 u_0 + \sum_{i=1}^N I_i u^{(i)} = u \tag{8.51}$$

and

$$b_0(u_0,u_0)+\sum_{i=1}^{N} b_i(u^{(i)},u^{(i)})=a(u,u). \tag{8.52}$$

Let $u=\{u_i\}_{i=1}^N \in V(\Gamma)$. Define $u_0^{(i)}\in V_{0i}(\Gamma_i)$ as

$$u_0^{(i)}=\sum_{F_{\ell k}\in\Lambda_i}\Big(\frac{1}{|F_{\ell k}|}\int_{F_{\ell k}} u_\ell ds\Big)\Phi^{(i)}_{F_{\ell k}} \tag{8.53}$$

where functions $\Phi^{(i)}_{F_{ik}}$ were defined in (8.48). Note that $u_0^{(i)}$ and u have the same average faces values on all faces $F_{\ell k}\in\Lambda_i$, i.e.,

$$\begin{cases} \dfrac{1}{|F_{\ell k}|}\displaystyle\int_{F_{\ell k}} u_\ell ds=\frac{1}{|F_{\ell k}|}\int_{F_{\ell k}} u_0^{(i)} ds=\overline{u}^{(i)}_{0\ell k} \\ \dfrac{1}{|F_{\ell k}|}\displaystyle\int_{F_{\ell k}} u_\ell ds=\frac{1}{|F_{\ell k}|}\int_{F_{\ell k}} u_0^{(j)} ds=\overline{u}^{(j)}_{0\ell k}, \end{cases} \tag{8.54}$$

then for all $F_{\ell k}\in\Lambda_i\cap\Lambda_j$ we have

$$\overline{u}^{(i)}_{0\ell k}=\overline{u}^{(j)}_{0\ell k}. \tag{8.55}$$

Define $u_0\in V_0$ by $u_0=\{u_0^{(i)}\}_{i=1}^N$ and set $w=u-I_0u_0$. Then we can write

$$w=\sum_{i=1}^{N} I_i(\tilde{I}_i^T u-u_0^{(i)})=\sum_{i=1}^{N} I_i u^{(i)},$$

where we have defined $u^{(i)}=\tilde{I}_i^T u-u_0^{(i)}\in V_i$. Since the prolongation operators I_i form a partition of unity, (8.51) holds.

To check (8.52) observe that $u^{(i)}$ has zero edge average values on all faces $F_{\ell k}\in\Lambda_i$, hence it is $\hat{\mathcal{H}}_i-$orthogonal to $u_0^{(i)}$; see (8.48). Then from the definition of b_0 we have

$$\begin{aligned} b_0(u_0,u_0)+\sum_{i=1}^{N} b_i(u^{(i)},u^{(i)}) &= \sum_{i=1}^{N} b_i(u_0^{(i)},u_0^{(i)})+b_i(u^{(i)},u^{(i)}) \\ &= \sum_{i=1}^{N} b_i(u_0^{(i)}+u^{(i)},u_0^{(i)}+u^{(i)}) \\ &= \sum_{i=1}^{N} b_i(\tilde{I}_i^T u,\tilde{I}_i^T u)=a_h(u,u). \end{aligned}$$

This ends the proof of Assumption(i).

Assumption(ii) We need to prove that

$$a_h(I_iu^{(i)},I_ju^{(j)})\le C\varepsilon_{ij}a_h^{1/2}(I_iu^{(i)},I_iu^{(i)})\;a_h^{1/2}(I_ju^{(j)},I_ju^{(j)}) \tag{8.56}$$

for $u^{(i)} \in V_i$ and $u^{(j)} \in V_j$, $\quad i,j = 1,\cdots,N$, and the spectral radius $\varrho(\varepsilon)$ of $\varepsilon = \{\varepsilon_{ij}\}_{i,j=1}^{N}$ is bounded. In our case $\varrho(\varepsilon) \leq C$ with constant independent of h_i and H_i. This follows from coloring arguments and the fact that $u^{(i)}$ and $u^{(j)}$ are different from zero only on D_i and D_j and their neighboring substructures.

Assumption(iii). We need to prove that for $i = 1,\cdots,N$,

$$a_h(I_i u^{(i)}, I_i u^{(i)}) \leq \omega b_i(u^{(i)}, u^{(i)}), \quad \forall u^{(i)} \in V_i \tag{8.57}$$

and

$$a_h(I_0 u_0, I_0 u_0) \leq \omega b_0(u_0, u_0), \quad \forall u_0 \in V_0 \tag{8.58}$$

with $\omega \leq C(1+\log \frac{H}{h})^2$ where C is a positive constant independent of h_i, H_i and the jumps of ρ_i.

For the proof of (8.57) see Lemma 8.6, and for the proof of (8.58) see Lemma 8.7 in the next section.

8.8 Auxiliary lemmas

In this section we complete the proof of Theorem 8.5 by proving two auxiliary lemmas associated with (8.57) and (8.58).

Lemma 8.6 *Assume that Assumption 8.4 holds. Then for $u^{(i)} \in V_i$, $i = 1,\ldots,N$, we have*

$$a_h(I_i u^{(i)}, I_i u^{(i)}) \leq C\left(1+\log\frac{H}{h}\right)^2 b_i(u^{(i)}, u^{(i)}), \tag{8.59}$$

where C is independent of h_i, H_i and the jumps of ρ_i.

Proof. In order to prove (8.59) we can replace $a_h(\hat{\mathcal{H}}u, \hat{\mathcal{H}}u)$ by $d_h(\mathcal{H}u, \mathcal{H}u)$ in the left hand side of (8.59) and in its right hand side we can put $d_i(\mathcal{H}\tilde{I}_i u^{(i)}, \mathcal{H}\tilde{I}_i u^{(i)})$ instead of $b_i(u^{(i)}, u^{(i)})$; see Lemma 8.1 and Lemma 8.2.

In order to simplify notations, all the functions are considered as harmonic extensions in the $\mathcal{H}$ sense. Hence, we denote $\mathcal{H}I_i u$ by $I_i u$ and let $u = \{u_l^{(i)}\}_{l\in\#(i)} \in V_i$. Using (8.8), (8.9) and (8.45) we obtain

$$d_h(I_i u^{(i)}, I_i u^{(i)}) = d_i(\tilde{I}_i D^{(i)} u^{(i)}, \tilde{I}_i D^{(i)} u^{(i)}) + \sum_j d_j(\tilde{I}_i D^{(i)} u^{(i)}, \tilde{I}_i D^{(i)} u^{(i)}) \tag{8.60}$$

where the sum is taken over D_j with common faces to D_i. The first term of the right hand side of (8.60) can be estimated as

$$\begin{aligned} & d_i(\tilde{I}_i D^{(i)} u^{(i)}, \tilde{I}_i D^{(i)} u^{(i)}) \\ &= \rho_i \int_{D_i} |\nabla D_i^{(i)} u_i^{(i)}|^2 dx + \sum_{F_{ij}\subset\partial D_i} \frac{\delta}{l_{ij}} \frac{\rho_{ij}}{h_{ij}} \int_{F_{ij}} (D_i^{(i)} u_i^{(i)} - D_j^{(i)} u_j^{(i)})^2 dx. \end{aligned} \tag{8.61}$$

To bound the first term of (8.61) we use

$$\begin{aligned} \rho_i \parallel \nabla D_i^{(i)} u_i^{(i)} \parallel^2_{L^2(D_i)} \;\leq\; & 2\rho_i \{ \parallel \nabla (D_i^{(i)} u_i^{(i)} - u_i^{(i)}) \parallel^2_{L^2(D_i)} \\ & + \parallel \nabla u_i^{(i)} \parallel^2_{L^2(D_i)} \} \end{aligned}$$

therefore,

$$\rho_i \parallel \nabla(D_i^{(i)}u_i^{(i)} - u_i^{(i)}) \parallel^2_{L^2(D_i)} \leq C \sum_{\delta_{ij}\subset\partial D_i} \rho_i \parallel \tilde{u}_i^{(i)} \parallel^2_{H^{1/2}_{00}(\delta_{ij})} .$$

Here $\tilde{u}_i^{(i)} = u_i^{(i)}$ at the interior nodal points of δ_{ij} and $\tilde{u}_i^{(i)} = 0$ on $\partial\delta_{ij}$. Recall that δ_{ij} denotes F_{ij} when F_{ij} is a slave side. It can be proved, see for example Toselli and Widlund [2005], that

$$\rho_i \parallel \tilde{u}_i^{(i)} \parallel^2_{H^{1/2}_{00}(\delta_{ij})} \leq C \left(1 + \log \frac{H_i}{h_i}\right)^2 \rho_i |u_i^{(i)}|^2_{H^1(D_i)}. \tag{8.62}$$

Here we have used the fact that $u_i^{(i)}$ has zero average face values.

We now estimate the second term of (8.61) and (8.67), see below. Note that for $F_{i\partial}$, i.e. for faces on ∂D, the estimates of the terms corresponding to $F_{i\partial}$ follow straightforwardly. On a slave face F_{ij} of ∂D_i, i.e. where $h_i \leq C_0 h_j$ and $\rho_i \leq C_1\rho_j$, or on $F_{i\partial}$, we have

$$\parallel D_i^{(i)}u_i^{(i)} - D_j^{(i)}u_j^{(i)} \parallel^2_{L^2(F_{ij})} \leq C h_i \max_{F_{ij}} |u_i^{(i)}|^2 \tag{8.63}$$

and

$$\begin{aligned} \frac{\rho_{ij}}{h_{ij}} \parallel D_i^{(i)}u_i^{(i)} - D_j^{(i)}u_j^{(i)} \parallel^2_{L^2(F_{ij})} &\leq C\rho_i \max_{F_{ij}} |u_i|^2 \\ &\leq C \left(1 + \log \frac{H_i}{h_i}\right) \rho_i |u_i^{(i)}|^2_{H^1(D_i)}, \end{aligned}$$

where we have used $\rho_{ij} \leq 2\rho_i$ and $h_i \leq C h_{ij}$ since $h_i < C_0 h_j$. We also have used that $u^{(i)}$ has zero average face value on any face of Λ_i, therefore, Poincaré inequality has be used to bound the $H^1(D_i)-$norm by the seminorm.

On a master side F_{ij} of ∂D_i, i.e. where $h_j \leq C_0 h_i$ and $\rho_j \leq C_1\rho_i$, we have

$$\begin{aligned} \parallel D_i^{(i)}u_i^{(i)} - D_j^{(i)}u_j^{(i)} \parallel_{L^2(F_{ij})} &\leq \parallel u_i^{(i)} - u_j^{(i)} \parallel_{L^2(F_{ij})} \\ &\quad + \parallel \sum_{x_v^j \in \partial F_{ij}} u_j^{(i)}(x_v^j)\varphi_v^j \parallel_{L^2(F_{ij})}, \end{aligned} \tag{8.64}$$

and using a triangular inequality we obtain

$$\parallel u_j^{(i)}(x_v^j)\varphi_v^j \parallel_{L^2(F_{ij})} \leq \parallel u_i^{(i)}(x_v^i)\varphi_v^i \parallel_{L^2(F_{ij})} + \parallel u_i^{(i)}(x_v^i)\varphi_v^i - u_j^{(i)}(x_v^j)\varphi_v^j \parallel_{L^2(F_{ij})} \tag{8.65}$$

Where φ_v^j is the nodal basis functions corresponding to x_v^j. The first term of (8.65) can be estimated as

$$\parallel u_i^{(i)}\varphi_v^i \parallel^2_{L^2(F_{ij})} \leq C \max_{F_{ij}} |u_i^{(i)}|^2 h_i \leq C h_i \left(1 + \log \frac{H_i}{h_i}\right) |u_i^{(i)}|^2_{H^1(D_i)},$$

while the second term of (8.65) can be bounded as in (8.81), see below. Using these estimates in (8.61) and Lemma 8.1 we get

$$d_i(I_i u^{(i)}, I_i u^{(i)}) \leq C \left(1 + \log \frac{H_i}{h_i}\right)^2 b_i(u^{(i)}, u^{(i)}). \tag{8.66}$$

We estimate the second term of (8.60) by bounding $d_j(\tilde{I}_i D^{(i)} u^{(i)}, \tilde{I}_i D^{(i)} u^{(i)})$ by the term $b_i(u^{(i)}, u^{(i)})$. For $u = \{u_l^{(i)}\} \in V_i$ we have

$$\begin{aligned} & d_j(\tilde{I}_i D^{(i)} u^{(i)}, \tilde{I}_i D^{(i)} u^{(i)}) \\ & = \rho_j \parallel \nabla D_j^{(i)} u_j^{(i)} \parallel^2_{L^2(D_j)} + \frac{\delta}{l_{ij}} \frac{\rho_{ij}}{h_{ij}} \int_{F_{ij}} (D_i^{(i)} u_i^{(i)} - D_j^{(i)} u_j^{(i)})^2 dx. \end{aligned} \tag{8.67}$$

We need only to estimate the first term of (8.67) since the second term has been already estimated; see (8.63), (8.64) and (8.65). If F_{ij} is a slave side of ∂D_i then $D_j^{(i)}$ vanishes, and so vanishes $\parallel \nabla D_j^{(i)} u_j^{(i)} \parallel^2_{L^2(D_j)}$. We now estimate the case where F_{ij} is a master side of ∂D_i and it is not equal to $F_{i\partial}$. On F_{ji} we decompose $u_j^{(i)} = w_j^{(i)} + \sum_{x_v^j \in \partial F_{ji}} u_j^{(i)}(x_v^j) \varphi_v^j$, where $w_j^{(i)} = D_j^{(i)} u_j^{(i)}$. We have

$$\begin{aligned} \parallel \nabla w_j^{(i)} \parallel^2_{L^2(D_j)} & \leq C \parallel w_j^{(i)} \parallel^2_{H^{1/2}_{00}(F_{ji})} \\ & = C\{|w_j^{(i)}|^2_{H^{1/2}(F_{ji})} + \int_{F_{ji}} \frac{(w_j^{(i)})^2}{dist(s, \partial F_{ji})} \, ds\}. \end{aligned} \tag{8.68}$$

We now estimate the first term of (8.68). Let Q_j be the L_2- projection on the h_j-triangulation of F_{ji}. Then

$$\begin{aligned} |w_j^{(i)}|^2_{H^{1/2}(F_{ji})} & \leq 2\{|w_j^{(i)} - Q_j u_i^{(i)}|^2_{H^{1/2}(F_{ij})} + |Q_j u_i^{(i)}|^2_{H^{1/2}(F_{ij})}\} \\ & \leq C\{\frac{1}{h_j} \parallel w_j^{(i)} - u_i^{(i)} \parallel^2_{L^2(F_{ij})} + \parallel \nabla u_i^{(i)} \parallel^2_{L^2(D_i)}\} \end{aligned} \tag{8.69}$$

and

$$\begin{aligned} & \parallel w_j^{(i)} - u_i^{(i)} \parallel^2_{L^2(F_{ij})} \\ & \quad \leq 2 \parallel u_j^{(i)} - u_i^{(i)} \parallel^2_{L^2(F_{ij})} + 2 \parallel \sum_{x_v^j \in \partial F_{ij}} u_j^{(i)}(x_v^j) \varphi_v^i \parallel^2_{L^2(F_{ij})} \end{aligned} \tag{8.70}$$

where the second term of (8.70) can be bounded as before and using the fact that $\rho_j \leq C_1 \rho_i$.

It remains to estimate the second term of (8.68). In order to simplify notation, we take F_{ij} as the interval $[0, H]$. Note that

$$\int_{F_{ji}} \frac{(w_j^{(i)})^2}{dist(s, \partial F_{ji})} ds \leq C\{\int_0^{H/2} \frac{(w_j^{(i)})^2}{s} ds + \int_{H/2}^{H} \frac{(w_j^{(i)})^2}{(H-s)} ds\}. \tag{8.71}$$

Let us estimate the first term in the right hand side of (8.71). We have

$$\begin{aligned} & \int_0^{H/2} \frac{(w_j^{(i)})^2}{s} ds = \int_0^{h_j} \frac{(w_j^{(i)})^2}{s} ds + \int_{h_j}^{H/2} \frac{(u_j^{(i)})^2}{s} ds \\ & \leq C\{u_j^{(i)}(h_j)^2 + \int_{h_j}^{H/2} \frac{(u_i^{(i)})^2 - (u_j^{(i)})^2}{s} ds + \int_{h_j}^{H/2} \frac{(u_i^{(i)})^2}{s} ds\} \end{aligned}$$

$$\leq C\{u_j^{(i)}(h_j)^2 + \frac{1}{h_j} \parallel u_i^{(i)} - u_j^{(i)} \parallel^2_{L^2(F_{ji})} + \left(1 + \log \frac{H_j}{h_j}\right) \max_{F_{ij}} |u_i^{(i)}|^2\}$$
$$\leq C\{\frac{1}{h_j} \parallel u_i^{(i)} - u_j^{(i)} \parallel^2_{L^2(F_{ij})} + \left(1 + \log \frac{H_i}{h_i}\right)\left(1 + \log \frac{H_j}{h_j}\right) \parallel u_i^{(i)} \parallel^2)_{H^1(D_i)}\}.$$

The second term of (8.71) is estimated similarly. Substituting these estimates to (8.71) and using that $u_i^{(i)}$ has zero average faces values we get

$$\int_{F_{ji}} \frac{(u_j^{(i)})^2}{dist(s, \delta F_{ji})} ds \leq C\{\left(1 + \log \frac{H}{h}\right)^2 (\parallel \nabla u_i^{(i)} \parallel^2_{L^2(D_i)} + \tag{8.72}$$
$$+ \frac{1}{H_i^2} \parallel u_i^{(i)} \parallel^2_{L^2(D_i)}) + \frac{1}{h_j} \parallel u_i^{(i)} - u_j^{(i)} \parallel^2_{L^2(F_{ij})}\}.$$

In turn, substituting (8.69) and (8.72) into (8.68), and the resulting estimate into (8.67), plus using Lemma 8.1, we get

$$d_j(\tilde{I}_i D^{(i)} u^{(i)}, \tilde{I}_i D^{(i)} u^{(i)}) \leq C \left(1 + \log \frac{H}{h}\right)^2 b_i(u^{(i)}, u^{(i)}). \tag{8.73}$$

Using (8.66) and (8.73) in (8.60), we get

$$d_h(I_i u^{(i)}, I_i u^{(i)}) \leq C \left(1 + \log \frac{H}{h}\right)^2 b_i(u^{(i)}, u^{(i)}).$$

The proof of Lemma 8.6 is complete.

Lemma 8.7 *Assume that Assumption 8.4 holds. Then for $u_0 \in V_0$, V_0 defined by (8.49), we have the following inequality*

$$a_h(I_0 u_0, I_0 u_0) \leq C \left(1 + \log \frac{H}{h}\right)^2 b_0(u_0, u_0) \tag{8.74}$$

where C is independent of h_i, H_i and the jumps of ρ_i.

Proof. By Lemma 8.1 and Lemma 8.2

$$a_h(\hat{\mathcal{H}} u, \hat{\mathcal{H}} u) \leq C d_h(\hat{\mathcal{H}} u, \hat{\mathcal{H}} u) \leq C d_h(\mathcal{H} u, \mathcal{H} u), \tag{8.75}$$

where $d_h(.,.)$ is defined by (8.9). Hence, to prove the result (8.74) we can replace $a_h(\hat{\mathcal{H}} u, \hat{\mathcal{H}} u)$ by $d_h(\mathcal{H} u, \mathcal{H} u)$ in the left hand side of (8.74).

In order to simplify the notation we write u instead of u_0 and put $I_0 u_0 = I_0 u = \sum_{i=1}^N I_i u^{(i)}$. We have

$$d_i(I_0 u, I_0 u) = \rho_i \parallel \nabla\{(I_i u^{(i)})_i + \sum_{F_{ij} \subset \partial D_i} (I_j u^{(j)})_i\} \parallel^2_{L^2(D_i)} \tag{8.76}$$
$$+ \sum_{F_{ij} \subset \partial D_i} \int_{F_{ij}} \frac{\rho_{ij}}{l_{ij}} \frac{\delta}{h_{ij}} (\{(I_i u^{(i)})_i + (I_j u^{(j)})_i\} - \{(I_i u^{(i)})_j + (I_j u^{(j)})_j\})^2 ds.$$

To bound the second term of the right hand side of (8.76) let us consider the case where F_{ij} is a mortar face. The proof for the case where F_{ij} is a slave side is

similar; see also the arguments given in (8.63) and afterwards. Then using the definition of I_i and $D^{(i)}$ we obtain

$$\begin{aligned} J &= \int_{F_{ij}} \frac{\rho_{ij}}{l_{ij}} \frac{\delta}{h_{ij}} (\{(I_i u^{(i)})_i + (I_j u^{(j)})_i\} - \{(I_i u^{(i)})_j + (I_j u^{(j)})_j)^2 ds \qquad (8.77)\\ &= \int_{F_{ij}} \frac{\rho_{ij}}{l_{ij}} \frac{\delta}{h_{ij}} \left(\{D_i^{(i)} u_i^{(i)} - D_j^{(i)} u_j^{(i)}\} - \{D_j^{(j)} u_j^{(j)} - D_i^{(j)} u_i^{(j)}\}\right)^2 ds \\ &= \int_{F_{ij}} \frac{\rho_{ij}}{l_{ij}} \frac{\delta}{h_{ij}} \left(\{D_i^{(i)} u_i^{(i)} - D_j^{(i)} u_j^{(i)}\} - \{D_j^{(j)} u_j^{(j)} - 0\}\right)^2 ds \\ &= \int_{F_{ij}} \frac{\rho_{ij}}{l_{ij}} \frac{\delta}{h_{ij}} \left(\{D_i^{(i)} u_i^{(i)} - (D_j^{(i)} + D_j^{(j)}) u_j^{(i)}\} + D_j^{(j)} \{u_j^{(i)} - u_j^{(j)}\}\right)^2 ds \\ &= \int_{F_{ij}} \frac{\rho_{ij}}{l_{ij}} \frac{\delta}{h_{ij}} \left(\{u_i^{(i)} - u_j^{(i)}\} - \sum_{x_v^j \in \partial F_{ji}} \{u_j^{(i)}(x_v^j) - u_j^{(j)}(x_v^j)\} \varphi_v^j\right)^2 ds \end{aligned}$$

Where φ_v^j is the nodal basis function corresponding to x_v^j. Then

$$\begin{aligned} J &\leq C \int_{F_{ij}} \frac{\rho_{ij}}{l_{ij}} \frac{\delta}{h_{ij}} \{u_i^{(i)} - u_j^{(i)}\}^2 ds \\ &\quad + C h_j \frac{\rho_{ij}}{l_{ij}} \frac{\delta}{h_{ij}} \max_{x_v \in \partial F_{ji}} \{u_j^{(i)}(x_v^j) - u_j^{(j)}(x_v^j)\}^2 \qquad (8.78) \end{aligned}$$

It remains to estimate the second term of (8.78), we estimate a bound for the difference $|u_j^{(i)}(x_v^j) - u_j^{(j)}(x_v^j)|$ with $x_v^j \in \partial F_{ji}$. First note that $\overline{u}_{ji}^{(i)} = \overline{u}_{ji}^{(j)}$ since there are primal variables associated to the faces $F_{ji} \in \Lambda_i$ and $F_{ji} \in \Lambda_j$; see (8.49). Therefore

$$\begin{aligned} |u_j^{(i)}(x_v^j) - u_j^{(j)}(x_v^j)| &\leq |u_j^{(j)}(x_v^j) - \overline{u}_{ji}^{(j)}| + |u_j^{(i)}(x_v^j) - \overline{u}_{ji}^{(i)}| \qquad (8.79)\\ &\leq C \left(1 + \log \frac{H_j}{h_j}\right)^{\frac{1}{2}} \| \nabla u_j^{(j)} \|_{L^2(D_i)} + |u_j^{(i)}(x_v^j) - \overline{u}_{ji}^{(i)}|. \end{aligned}$$

To get the estimate of the first term of the right hand side of (8.79) we have used a Poincaré inequality and a L^∞ bound for FEM functions, see Toselli and Widlund [2005]. The second term of (8.79) is estimated as

$$\begin{aligned} |u_j^{(i)}(x_v^j) - \overline{u}_{ji}^{(i)}| &\leq |u_j^{(i)}(x_v^j) - u_i^{(i)}(x_v^i)| + |u_i^{(i)}(x_v^i) - \overline{u}_{ij}^{(i)}| + |\overline{u}_{ij}^{(i)} - \overline{u}_{ji}^{(i)}| \\ &\leq C\{|u_j^{(i)}(x_v^j) - u_i^{(i)}(x_v^i)| + \left(1 + \log \frac{H_i}{h_i}\right)^{\frac{1}{2}} \| \nabla u_i^{(i)} \|_{L^2(D_i)} \qquad (8.80)\\ &\quad + h_j^{-\frac{1}{2}} \| u_i^{(i)} - u_j^{(i)} \|_{L^2(F_{ij})}\}, \end{aligned}$$

where we have used a Poncaré inequality and a L^∞ bound for FEM functions to obtain the second term in the right hand side of (8.80) and a Cauchy-Schwarz inequality to obtain the third term of (8.80). To estimate the first term of (8.80), let $Q_j u_i^{(i)}$ be the L^2-projection of $u^{(i)}$ on the h_j triangulation of F_{ji}. We obtain

$$\begin{aligned} |u_j^{(i)}(x_v^j) - u_i^{(i)}(x_v^i)| &\leq |u_j^{(i)}(x_v^j) - Q_j u_i^{(i)}(x_v^i)| + |Q_j u_i^{(i)}(x_v^i) - u_i^{(i)}(x_v^i)| \\ &\leq C\{h_j^{-\frac{1}{2}} \| u_j^{(i)} - u_i^{(i)} \|_{L^2(F_{ij})} + \left(1 + \log \frac{H_j}{h_j}\right)^{\frac{1}{2}} \| \nabla u_i^{(i)} \|_{L^2(D_i)}\}, \qquad (8.81) \end{aligned}$$

where the first estimate has followed from a inverse inequality and the second from the approximation properties of the L^2 projection and a L^∞ bound for FEM functions.

By Lemma 8.1 and Lemma 8.2 we can bound $d_i(\mathcal{H}\tilde{I}_i u^{(i)}, \mathcal{H}\tilde{I}_i u^{(i)})$ by $b_i(\hat{\mathcal{H}}_i u^{(i)}, \hat{\mathcal{H}}_i u^{(i)})$. Then we conclude that J of (8.77) can be estimated as

$$J \leq C\left(1+\log\frac{H}{h}\right)\{b_i(u^{(i)},u^{(i)})+b_j(u^{(j)},u^{(j)}\}, \tag{8.82}$$

since $\rho_{ij} \leq C\rho_i$ and $h_j \leq Ch_{ij}$.

It remains to estimate the first term in (8.76). We have

$$\begin{aligned}
&\| \nabla\{(I_i u^{(i)})_i + \sum_{F_{ij}\subset\partial D_i} (I_j u^{(j)})_i\} \|^2_{L^2(D_i)} \\
&= \| \nabla\{(D_i^{(i)} + \sum_{F_{ij}\subset\partial D_i} D_i^{(j)}) u_i^{(i)} + \sum_{F_{ij}\subset\partial D_i} D_i^{(j)}(u_i^{(j)} - u_i^{(i)})\} \|^2_{L^2(D_i)} \\
&\leq C\{\| \nabla u_i^{(i)} \|^2_{L^2(D_i)} + \sum_{\delta_{ij}\subset\partial D_i} |D_i^{(j)}(u_i^{(j)} - u_i^{(i)})|^2_{H^{1/2}_{00}(\delta_{ij})}\},
\end{aligned} \tag{8.83}$$

where the sum in (8.83) reduces to the slaves sides F_{ij}. From (8.49) we obtain

$$\begin{aligned}
&|D_i^{(j)}(u_i^{(j)} - u_i^{(i)})|^2_{H^{1/2}_{00}(F_{ij})} \\
&\quad \leq 2\{|D_i^{(j)}(u_i^{(j)} - \overline{u}_{ij}^{(j)})|^2_{H^{1/2}_{00}(F_{ij})} + |D_i^{(j)}(u_i^{(i)} - \overline{u}_{ij}^{(i)})|^2_{H^{1/2}_{00}(F_{ij})}\}
\end{aligned} \tag{8.84}$$

and therefore the first term of (8.84) is estimated as

$$\begin{aligned}
&\rho_i |D_i^{(j)}(u_i^{(j)} - \overline{u}_{ij}^{(j))})|^2_{H^{1/2}_{00}(F_{ij})} \\
&\leq 2\rho_i\{|D_i^{(j)}(u_i^{(j)} - u_j^{(j)})|^2_{H^{1/2}_{00}(F_{ij})} + |D_i^{(j)}(u_j^{(j)} - \overline{u}_{ji}^{(j)})|^2_{H^{1/2}_{00}(F_{ij})} \\
&\quad + |D_i^{(j)}(\overline{u}_{ji}^{(j)} - \overline{u}_{ij}^{(j)})|^2_{H^{1/2}_{00}(F_{ij})}\} \\
&\leq C\rho_i\{\frac{1}{h_j} \| u_i^{(j)} - u_j^{(j)} \|^2_{L^2(F_{ji})} + \left(1+\log\frac{H_j}{h_j}\right)^2 \| \nabla u_j^{(j)} \|^2_{L^2(D_j)}\} \\
&\leq C\left(1+\log\frac{H_j}{h_j}\right)^2 b_j(u^{(j)},u^{(j)})
\end{aligned} \tag{8.85}$$

since when F_{ij} is a slave side $\rho_i \leq C_1\rho_j$ and in view of Lemma 8.1. The second term of (8.84) is bounded by

$$\begin{aligned}
\rho_i |D^{(j)}(u_i^{(i)} - \overline{u}_{ij}^{(i)})|^2_{H^{1/2}_{00}(F_{ij})} &\leq C\rho_i\left(1+\log\frac{H_i}{h_i}\right)^2 \| \nabla u_i^{(i)} \|_{L^2(D_i)} \\
&\leq \left(1+\log\frac{H_i}{h_i}\right)^2 b_i(u^{(i)},u^{(i)}).
\end{aligned} \tag{8.86}$$

Using (8.85) and (8.86) in (8.84) and the resulting inequality in (8.83) and (8.77) we see that

$$\rho_i \parallel \nabla\{(I_i u^{(i)})_i + \sum_{F_{ij}\subset\partial D_i} (I_j u^{(j)})_i\} \parallel^2_{L^2(D_i)} \\ \leq C\left(1+\log\frac{H}{h}\right)^2 \{b_i(u^{(i)},u^{(i)}) + b_j(u^{(j)},u^{(j)})\},$$

this estimate and (8.82) imply that

$$d_i(I_0u_0, I_0u_0) \leq C\left(1+\log\frac{H}{h}\right)^2 \{b_i(u^{(i)},u^{(i)}) + b_j(u^{(j)},u^{(j)})\}.$$

Summing this over i and using Lemma 8.1 and Lemma 8.2 we get (8.74).

8.9 Smaller global spaces

In Section 8.6 we have defined the coarse space with a primal variable associated to each face $F_{\ell k} \in \Lambda_i$. In this case the number of constrains per subdomain is twice the number of edges of ∂D_i for floating subdomains D_i. In this section we discuss choices of subsets of Λ_i which imply smaller coarse problems and still maintain the bound (8.50) of Theorem 8.5.

Recall that a face across D_i and D_j has two sides, the side contained in ∂D_i, denoted by F_{ij}, and the side contained in ∂D_j, denoted by F_{ji}. Let $\tilde{\Lambda}_i$, $i = 1,\ldots,N$, be such that for all pair of neighboring subdomains D_i and D_j the subset $\tilde{\Lambda}_i \cap \tilde{\Lambda}_j$ contains one and only one face from each pair $\{F_{ij}, F_{ji}\}$, i.e., F_{ij} or F_{ji}. We denote the chosen face by $\lambda_{ij} = \lambda_{ji}$. For instance, we can choose $\tilde{\Lambda}_i$ as the set of mortar faces λ_{ij} associated to D_i.

After choosing $\tilde{\Lambda}_i$, the local spaces $V_i = V_i(\Gamma_i)$, $i = 1,\ldots,N$, are defined as the subspaces of W_i of functions with zero average faces values on all faces $\lambda_{\ell k} \in \tilde{\Lambda}_i$ while the spaces V_{0i} are defined as $V_{0i} = V_{0i}(\Gamma_i) = \text{Span}\{\Phi^{(i)}_{\lambda_{\ell k}} : \lambda_{\ell k} \in \tilde{\Lambda}_i\} \subset W_i$ where the functions $\Phi^{(i)}_{\lambda_{ik}}$ are defined as in Section 8.6 replacing Λ_i by $\tilde{\Lambda}_i$ in each subdomain; see (8.48).

From now on we will use the notation

$$\overline{u}^{(i)}_{\lambda_{\ell k}} = \frac{1}{|\lambda_{\ell k}|}\int_{\lambda_{\ell k}} u^{(i)} ds,$$

where $u^{(i)} \in W_i$. The global coarse space V_0 is now defined as the set of all $u_0 = \{u_0^{(i)}\} \in \prod_{i=1}^N V_{0i}(\Gamma_i)$ such that for $i = 1,\ldots,N$, we have

$$\overline{u}^{(i)}_{0\lambda_{ij}} = \overline{u}^{(j)}_{0\lambda_{ij}} \quad \forall \lambda_{ij} \in \tilde{\Lambda}_i. \tag{8.87}$$

Recall that $u_0^{(i)}$ is defined locally. Then we have the following possible cases of continuity with respect to the primal variables:

Case 1 $\lambda_{ij} = \lambda_{ji} = F_{ij}$. This case imposes continuity of the average face values of $u_0^{(i)}$ and $u_0^{(j)}$ on F_{ij}; see (8.87).

Case 2 $\lambda_{ij} = \lambda_{ji} = F_{ji}$. This case imposes continuity of the average face value on F_{ji}.

Example 8.8 *Consider the domain $D = (0,1)^2$ and divide it into $N = M \times M$ squares subdomains D_i which are unions of fine elements, with $H = 1/M$. We note that for floating subdomains D_i, Λ_i has eight coarse basis functions while $\tilde{\Lambda}_i$ has only four coarse basis functions.*

The bilinear forms a_h, b_i and the operators I_i, $i = 1, \ldots, N$, and the operator I_0 are defined in Section 8.5 and Section 8.6.

We now show that with these new local and global spaces Theorem 8.5 still holds. The proof is basically the same as the one given in Section 8.7 and Section 8.8 with some minor modifications depending on which of the above cases is considered and also on a modification of the Poincaré inequality.

Theorem 8.9 *If the Assumption 8.4 holds, then there exists a positive constant C independent of h_i, H_i and the jumps of ρ_i such that*

$$a_h(u,u) \leq a_h(Tu,u) \leq C\left(1 + \log\frac{H}{h}\right)^2 a_h(u,u) \qquad \forall u \in V, \tag{8.88}$$

where T is defined in (8.41), the local spaces V_i, $i = 1, \ldots, N$, are defined above in this section and the global space V_0 is defined using (8.87). Here $\log\frac{H}{h} = \max_i \log\frac{H_i}{h_i}$.

Proof. We now mention the main modifications of the proof of the three key assumptions of Lemma 8.3.

Assumption(i) Let $u = \{u_i\}_{i=1}^N \in V(\Gamma)$. Define $u_0^{(i)} \in V_{0i}(\Gamma_i)$ by

$$u_0^{(i)} = \sum_{\lambda_{\ell k} \in \tilde{\Lambda}_i} \Big(\frac{1}{|\lambda_{\ell k}|}\int_{\lambda_{\ell k}} uds\Big)\Phi_{\lambda_{\ell k}}^{(i)} \tag{8.89}$$

and proceed as in the proof of Theorem 8.5.

Assumption(ii) Same argument given to verify Assumption(ii) in the proof of Theorem 8.5.

Assumption(iii) We modify the proof of Lemma 8.7 and Lemma 8.6 as follows:

For the proof of Lemma 8.7 we consider the following cases to obtain a bound for the left hand side of (8.79),

Case 1 $\lambda_{ij} = \lambda_{ji} = F_{ji}$. In this case we use the same argument as in the proof of in Lemma 8.7 to estimate the left hand side of (8.79).

Case 2 $\lambda_{ij} = \lambda_{ji} = F_{ij}$. In this case we estimate

$$|u_j^{(i)}(x_v^j) - u_j^{(j)}(x_v^j)| \leq$$
$$|u_j^{(i)}(x_v^j) - \overline{u}_{F_{ji}}^{(i)}| + |u_j^{(j)}(x_v^j) - \overline{u}_{F_{ji}}^{(j)}| + |\overline{u}_{F_{ji}}^{(i)} - \overline{u}_{F_{ji}}^{(j)}|. \tag{8.90}$$

The first and second term of (8.90) can be bounded as in **Case 1**. The third term of (8.90) is bounded as follows. Since $\lambda_{ij} = \lambda_{ji} = F_{ij}$ we have that $\overline{u}_{F_{ij}}^{(i)} = \overline{u}_{F_{ij}}^{(j)}$; see (8.87). Then

$$|\overline{u}_{F_{ji}}^{(i)} - \overline{u}_{F_{ji}}^{(j)}| \leq |\overline{u}_{F_{ji}}^{(i)} - \overline{u}_{F_{ij}}^{(i)}| + |\overline{u}_{F_{ij}}^{(j)} - \overline{u}_{F_{ji}}^{(j)}| \tag{8.91}$$

and we obtain

$$|\overline{u}_{F_{ji}}^{(i)} - \overline{u}_{F_{ij}}^{(i)}| \leq C H_j^{-\frac{1}{2}} \parallel u_{F_{ji}}^{(i)} - u_{F_{ij}}^{(i)} \parallel_{L^2(F_{ij})} \leq C h_j^{-\frac{1}{2}} \parallel u_{ji}^{(i)} - u_{ij}^{(i)} \parallel_{L^2(F_{ij})} .$$

An analogous bound holds also for the second term of (8.91); see (8.79).

For the proof of Lemma 8.6 we can apply Poincaré inequality only in the case $\lambda_{ij} = F_{ij} \subset \partial D_i$. If this is not the case, i.e., if $\lambda_{ij} = F_{ji} \subset D_j$, we still can bound the $H^1(D_i)$ norm by the seminorm using the following argument: If $u^{(i)} \in V_i$ and $\lambda_{ij} = F_{ji}$ then $u^{(i)}$ has zero average value on F_{ji}. Therefore,

$$\begin{aligned} \parallel u_i \parallel_{L^2(D_i)} &\leq \parallel u_i - \overline{u}_{F_{ij}}^{(i)} \parallel_{L^2(D_i)} + \parallel \overline{u}_{F_{ij}}^{(i)} - \overline{u}_{F_{ji}}^{(i)} \parallel_{L^2(F_{ij})} \\ &\leq \parallel \nabla u_i \parallel_{L^2(D_i)} + \frac{1}{H_i^{1/2}} \parallel u_{ij}^{(i)} - u_{ji}^{(i)} \parallel_{L^2(F_{ij})} . \end{aligned}$$

Having modified the proof of Lemma 8.7 and Lemma 8.6, then Assumption(iii) follows.

8.10 Numerical experiments

In this section, we present numerical results for the preconditioner introduced in (8.41) and show that the bounds of Theorem 8.5 and Theorem 8.9 are reflected on the numerical tests. In particular we show that the Assumption 8.4, see (8.42), is sufficient and necessary.

We consider the domain $D = (0,1)^2$ and divide into $N = M \times M$ squares subdomains D_i which are unions of fine elements, with $H = 1/M$. Inside each subdomain D_i we generate a structured triangulation with n_i subintervals in each coordinate direction and apply the discretization presented in Section 8.2 with $\delta = 4$. This value $\delta = 4$ was chosen because numerically it was observed that the L^2 approximation error seem to stabilize when δ becomes larger. The minimum value of δ that gives a positive definite system is $\delta_{\min} = 1.565$. In the numerical experiments we use the black and white checkerboard type of subdomain partition. On the white subdomains we let $n_i = 2 * 2^{L_w}$ and on the black subdomains we let $n_i = 3 * 2^{L_b}$, where L_w and L_b are integers denoting the number of refinements inside each subdomain D_i. Hence, the mesh sizes are $h_w = \frac{2^{-L_w}}{2M}$ and $h_b = \frac{2^{-L_b}}{3M}$, respectively. We solve the second order elliptic problem $-\text{div}(\rho(x)\nabla u^*(x)) = 1$ in D

with homogeneous Dirichlet boundary conditions. In the numerical experiments, we run PCG until the l_2 initial residual is reduced by a factor of 10^{-6}.

In the first test we consider the constant coefficient case $\rho = 1$. We consider different values of $M \times M$ coarse partitions and different values of local refinements $L_w = L_b$, therefore, keeping constant the mesh ratio $h_w/h_b = 3/2$. We place the mortars on the white subdomains. We note that the interfacing condition (8.42) is satisfied. Table 8.1 lists the number of PCG iterations and in parenthesis the condition number estimate of the preconditioned system in the case we choose eight coarse functions per subdomain. As expected from the analysis, the condition numbers appear to be independent of the number of subdomains and seems to grow by a logarithmic factor when the size of the local problems increases. Note that in the case of continuous coefficients Theorem 8.5 and Theorem 8.9 are valid without any assumption on h_w and h_b if the mortar sides are chosen on the larger meshes.
Table 8.2 is the same as before however, now we have chosen $\tilde{\Lambda}_i$ as being the set of mortar faces of D_i. In this case we have four coarse basis functions in each subdomain. We note that even though the coarse problems are smaller the results are very similar to the ones presented in Table 8.1 where coarse problems are larger. As in the case of Table 8.2 the smallest eigenvalue of the preconditioned operator is 1.

We now consider the discontinuous coefficient case where we set $\rho_i = 1$ on the white subdomains and $\rho_i = \mu$ on the black subdomains. The subdomains are kept fixed to 4×4, i.e., 16 subdomains. Table 8.3 lists the results on runs for different values of μ and for different levels of refinements on the black subdomains. On the white subdomains $n_i = 2$ is kept fixed. The mortars are placed on the white subdomains. It is easy to see that the interfacing condition (8.42) holds if and only if μ is not large, which it is in agreement with the results seem in Table 8.3. We repeat the same experiment of Table 8.3 but this time with four coarse local basis functions associated to the mortar side of the subdomain. The results are presented in Table 8.4.

$M \downarrow L_b \rightarrow$	0	1	2	3	4	5
2	12 (5.7)	14 (6.7)	15 (7.5)	18 (10.6)	19 (14.5)	19 (19.0)
4	14 (5.8)	18 (8.5)	21 (11.7)	24 (15.2)	27 (19.2)	29 (23.9)
8	15 (5.9)	20 (9.1)	24 (12.3)	27 (15.8)	31 (19.6)	34 (24.0)
16	15 (6.0)	20 (9.4)	25 (12.8)	28 (16.3)	31 (20.1)	35 (24.5)
32	15 (6.0)	20 (9.3)	25 (12.8)	28 (16.3)	32 (20.2)	

Table 8.1: PCG/BDDC iterations count and condition numbers for different sizes of coarse and local problems and constant coefficients ρ_i with 8 coarse basis functions per subdomain.

8.11 Conclusions

In this paper several BDDC methods with different coarse spaces for DG discretization of second order elliptic equations with discontinuous coefficients have been designed and analyzed. It has been proved that the methods are almost optimal and very well suited for parallel computations. Their rate of convergence

$M \downarrow L_b$	0	1	2	3	4	5
2	13 (5.7)	15 (6.7)	16 (7.5)	18 (10.7)	19 (14.5)	19 (18.9)
4	15 (5.8)	19 (8.5)	22 (11.7)	24 (15.1)	27 (19.2)	29 (23.8)
8	17 (6.1)	21 (9.1)	25 (12.3)	28 (15.7)	31 (19.6)	34 (24.0)
16	18 (6.1)	23 (9.4)	27 (12.8)	30 (16.3)	32 (20.1)	
32	18 (6.1)	24 (9.4)	27 (12.8)	30 (16.3)		

Table 8.2: PCG/BDDC iterations count and condition numbers for different sizes of coarse and local problems and constant coefficients ρ_i with 4 coarse basis functions per subdomain associated to its mortar faces.

$\mu \downarrow L_b \rightarrow$	1	2	3	4	5
1000	165(2822.6)	263(3746.6)	282(4758.9)	287(5922.3)	310(7168.88)
10	37(32.9)	43(42.3)	47(52.8)	51(64.8)	53(77.7)
0.1	17(6.8)	16(6.8)	17(6.8)	17(6.9)	17(6.9)
0.001	16(7.12)	16(7.16)	16(7.25)	17(7.38)	18(7.50)

Table 8.3: PCG/BDDC iterations count and condition numbers for different values of coefficients and the local mesh sizes on the black subdomains only. The coefficients and the local mesh sizes on the white subdomains are kept fixed. The subdomains are also kept fixed to 4×4 and 8 coarse basis functions in each subdomain are used.

are independent of the parameter of triangulations, the number of substructures and the jumps of coefficients. The numerical tests confirm theoretical results. The methods can be straightforwardly extended to 3-D cases. Finally, we remark that the condition of the preconditioned systems deteriorates as we increase the penalty parameter δ to large values.

Bibliography

Arnold, D. N. (1982). An interior penalty finite element method with discontinuous elements. *SIAM J. Numer. Anal.*, 19(4):742–760.

Arnold, D. N., Brezzi, F., Cockburn, B., and Martin, D. (2002). Unified analysis of discontinuous Galerkin method for elliptic problems. *SIAM J. Numer. Anal.*, 39(5):1749–1779.

Brenner, S. C. and Wang, K. (2005). Two-level additive Schwarz preconditioners for C^0 interior penalty methods. *Numer. Math.*, 102(2):231–255.

$\mu \downarrow L_b \rightarrow$	1	2	3	4	5
1000	133(2905.1)	188(3827.7)	254(4838.4)	326(5980.0)	384(7205.4)
10	40(33.4)	45(43.0)	49(53.5)	53(65.3)	54(78.0)
0.1	16 (6.8)	16(6.8)	17 (6.8)	17 (6.9)	17 (7.0)
0.001	15 (7.3)	16 (7.2)	17 (7.3)	17 (7.42)	18 (7.52)

Table 8.4: PCG/BDDC iterations count and condition numbers for different values of coefficients and the local mesh sizes on the black subdomains only. The coefficients and the local mesh sizes on the white subdomains are kept fixed. The subdomains are also kept fixed to 4×4 and 4 coarse basis functions in each subdomain are used. Mortar faces are chosen.

Dohrmann, C. R. (2003). A preconditioner for substructuring based on constrained energy minimization. *SIAM J. Sci. Comput.*, 25(1):246–258 (electronic).

Dryja, M. (2003). On discontinuous Galerkin methods for elliptic problems with discontinuous coefficients. *Comput. Methods Appl. Math.*, 3(1):76–85.

Dryja, M. and Sarkis, M. (2006). A Neumann-Neumann method for DG discretization of elliptic problems. Technical Report Serie A 456, Intituto de Mathemática Pura e Aplicada. http://www.preprint.impa.br/Shadows/SERIE_A/2006/456.html.

Feng, X. and Karakashian, O. A. (2001). Two-level additive Schwarz methods for a discontinuous Galerkin approximation of second order elliptic problems. *SIAM J. Numer. Anal.*, 39(4):1343–1365.

Lasser, C. and Toselli, A. (2003). An overlapping domain decomposition preconditioners for a class of discontinuous Galerkin approximations of advection-diffusion problems. *Math. Comp.*, 72(243):1215–1238.

Li, J. and Widlund, O. (2006). FETI-DP, BDDC, and block Cholesky methods. *Internat. J. Numer. Methods Engrg.*, 66(2):250–271.

Mandel, J., Dohrmann, C. R., and Tezaur, R. (2005). An algebraic theory for primal and dual substructuring methods by constraints. *Appl. Numer. Math.*, 54(2):167–193.

Toselli, A. and Widlund, O. (2005). *Domain decomposition methods—algorithms and theory*, volume 34 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin.

Chapter 9

A Priori Error Estimates for Wiener-Chaos Finite Element Approximations of the Darcy's Equation in Random Porous Media

We consider the white noise analysis constructed from a Hilbert space and an operator to define and characterize adequate spaces for solving the ordinary (rather than Wick) product stochastic pressure equation. A weak form of this equation involves different spaces for the solution and test functions and we establish a continuous inf-sup condition and well-posedness of the problem. We generalize the numerical approximations proposed in Benth and Theting [Stochastic Anal. Appl., 20 (2002), pp. 1191–1223] for Wick stochastic partial differential equations, and in Roman and S. [Discrete Contin. Dyn. Syst. Ser. B, 6 (2006), pp. 941–955] for the ordinary product stochastic pressure equation, and establish discrete inf-sup conditions and provide a priori error estimates for a wide class of norms. The proposed numerical approximation is based on Wiener-Chaos finite element methods and leads to the solution of a positive symmetric linear system. We also improve and generalize the approximation results of Benth and Gjerde [Stochastics Stochastics Rep., 63 (1998), pp. 313–326] and Cao [Stochastics, 78 (2006), pp. 179–187] when a (generalized) process is truncated by a finite Wiener-Chaos expansion. Finally, we present numerical experiments to validade the results.

9.1 Introduction

In the mathematical studies of transport of pollutants in groundwater and oil recovery processes, we have to approximate the solution of a system of stochastic partial differential equations which models the two-phase flow in a porous medium. The system is composed of a transport equation for the saturation (the relative volume of one of the two fluids) coupled with an equation for the velocity field given by the Darcy's Law and the incompressibility condition of the flow. The randomness enters the problem through the unknown properties of the rocks,

especially the permeability tensor. In this paper we deal with one of the equations derived from this system, specifically, we consider an equation of the form:

$$\begin{cases} -\nabla_x \cdot (\kappa(x,\cdot)\nabla_x u(x,\cdot)) &= f(x,\cdot), \text{ for all } x \in D \\ u(x,\cdot) &= 0, \text{ for all } x \in \partial D, \end{cases} \tag{9.1}$$

where $\log \kappa(x,\cdot)$ is a Gaussian field and f is a (possible random) forcing term; see Ghanem and Spanos [1991], Furtado and Pereira [2003], Roman and Sarkis [2006], Babuška et al. [2007] and references therein. We emphasize that the assumption of the stochastic structure of the permeability function $\kappa(x,\cdot)$ is due, more than anything else, to lack of data and accuracy in the measurements of the media. One approach that has been studied, when the parameters in the equation are not completely known, is to replace the true values of these parameters by some kind of average. By replacing the stochastic coefficient $\kappa(x,\cdot)$ in the equations by the average $\bar{\kappa}(x)$ we obtain some information about the solution, but usually this information is not enough to make more precise predictions of nonlinear functionals of the solution or to know what effect the small fluctuations in the parameter values actually have on the solution.

One way to evaluate nonlinear functionals of the solution involves Monte-Carlo approximations. Briefly speaking, Monte-Carlo approximations require knowing the solution, or approximations of the solution, for many paths or realizations in the space of outcomes. In principle, if we know the distribution of the processes modeling the coefficients in the equation, we can simulate many trajectories of the coefficients, and for each trajectory, we can apply a finite element method (FEM) to obtain an approximation of the solution for that particular realization. Then, Monte-Carlo approximations of a nonlinear functional g of the solution are of the form $\mathbb{E}[g(\hat{u}(x,\cdot)] \approx \frac{1}{M}\sum_{i=1}^{M} g(\hat{u}^{(a)}(x,\omega_i))$, where $\mathbb{E}$ denotes the expectation operator, M is the number of realizations, and $\hat{u}^{(a)}(x,\omega_i)$ is a finite element approximation of the solution at x for the ith-trajectory ω_i. However, this procedure is very expensive and time consuming as it involves assembling and solving large linear systems as many times as trajectories are simulated. We also mention that the Monte Carlo approach gives relevant information about the average value of the solution only when the variance of the solution is of moderated size. The variance of the solution depends on the size and correlation width of the introduced noise; see Holden et al. [1996].

Alternatively to the Monte-Carlo approach, methods that somehow "separate" the stochastic part from the deterministic part are very attractive to researchers. A point of practical importance of these methods is that they can be used in numerical simulation schemes to obtain numerical realizations of a random process. To illustrate the advantage of these methods, let us suppose that the solution of (9.1) can be represented as

$$\hat{u}(x,\omega) = \sum_{\alpha \in \mathcal{I}} \hat{u}_\alpha(x) Y_\alpha(\omega), \tag{9.2}$$

where $\mathcal{I}$ is a countable index set and $\{Y_\alpha\}_{\alpha\in\mathcal{I}}$ is a collection of random variables with known probability distributions. Let us say we have an approximation of the

solution of the form

$$\hat{u}(x,\omega) \approx \hat{u}^{(a)}(x,\omega) = \sum_{\alpha\in\widetilde{\mathcal{I}}} \hat{u}_\alpha^{(a)}(x) Y_\alpha(\omega), \tag{9.3}$$

where $\widetilde{\mathcal{I}}$ is a finite index set with $\widehat{\mathcal{I}} \subset \mathcal{I}$. Assume that we compute and store the deterministic functions $\{\hat{u}_\alpha^{(a)}\}_{\alpha\in\widetilde{\mathcal{I}}}$. Then, when a simulation is required, all we need to do is to generate values for the random variables $\{Y_\alpha\}_{\alpha\in\widetilde{\mathcal{I}}}$, and assemble the solution according to (9.3). In this way, we will need to solve a very large linear system only once in order to compute the deterministic coefficients $\hat{u}_\alpha^{(a)}(x)$. Since

$$\hat{u}(x,\omega) - \hat{u}^{(a)}(x,\omega) = \sum_{\alpha\in\mathcal{I}\setminus\widehat{\mathcal{I}}} \hat{u}_\alpha(x) Y_\alpha(\omega) + \sum_{\alpha\in\widetilde{\mathcal{I}}} \left(\hat{u}_\alpha(x) - \hat{u}_\alpha^{(a)}(x)\right) Y_\alpha(\omega)$$

we conclude that relevancy of the information we obtain using the procedure described above depends mainly on:

1. The kind of expansion used in (9.2) and (9.3). (In this paper we use the *Wiener-Chaos* expansion).
2. The finite dimensional problem involved in the computation of the coefficients $\{\hat{u}_\alpha^{(a)}\}_{\alpha\in\widetilde{\mathcal{J}}}$ in (9.3). Usually this problem is a Galerkin type problem that uses the original coefficient κ or an approximation of it, for instance, a truncated *Karhúnen-Loève* or *Wiener-Chaos* expansion.

The Karhúnen-Loève (KL) expansion of a stochastic process with continuous covariance function is well-used in many engineering applications as an efficient tool to store a random process. This expansion is optimal in the Fourier sense, as it minimizes the mean square error resulting from truncation after a finite number terms. A fundamental issue inherited from the KL expansion is the excess of oscillations of the terms of the expansion near the boundary ∂D. Using the KL expansion also requires the computation of the eigenvalues and eigenfunctions of the *Covariance* operator associated to the coefficient; see Babuška and Chatzipantelidis [2002], Babuška et al. [2007, 2004], Frauenfelder et al. [2005], Ghanem [1999a,b], Ghanem and R. [1999], Ghanem and Spanos [1991], Jin et al. [2007], Keese [2003], Matthies and Keese [2005], Nobile et al. [2007], Schwab and Todor [2003] and references therein. We note that for an approximation of a log $-$normal coefficient κ it is possible to use its truncated KL expansion or the exponential of the truncated KL expansion of $\log\kappa$.

In this paper we will propose a finite dimensional problem for computing the coefficient of an approximation of the type (9.3) that uses the original log $-$normal coefficient κ. We avoid the use of an approximated KL expansion for the coefficient κ of (9.1) and we do not assume finite dimensionality of the noise. In this work the expansion (9.3) is of the form of a truncated Wiener-Chaos expansion. The Wiener-Chaos expansion is the orthogonal expansion, in terms of Fourier-Hermite stochastic polynomials, of random processes defined in the *white noise space*. The use of Wiener-Chaos expansion gives us some freedom when defining the white

noise space. For instance, we can use the construction of the white noise space based on $L^2(\mathbb{R}^d)$ and the Hermite functions that have several interesting analytical and algebraic properties. Such properties can be explored to manipulate products of functions, to compute norms, and to improve the complexity and stability of the algorithms. We note that using Wiener-Chaos expansion for approximating the solution has an additional advantage since it is possible to deduce, based on Sobolev, Kodratiev, Hida and others norms on the white noise space, convergence rates of the procedure of truncating the Wiener-Chaos expansion; see Da Prato [2006], Hida et al. [1993], Holden et al. [1996], Kuo [1996] and references therein.

The approximation of solutions of partial differential equations based on Wiener-Chaos expansion has been considered extensively in the literature. We mention Benth and Theting [2002] and Theting [2000] where they analyze the stochastic pressure equation interpreting the ordinary product as Wick product. These works also consider other Wick stochastic partial differential equations. They use the white noise calculus and propose and approximation by truncating the Wiener-Chaos expansion of the solution. They present a priori error estimates based in the work of Benth and Gjerde [1998] on convergence rates of the procedure of truncating the Wiener-Chaos expansion. Here we also mention Cao [2006] where the estimates in Benth and Gjerde [1998] are improved. In Roman and Sarkis [2006] they present and explain several features which show the advantages of using the white noise calculus as a natural framework for the study of the stochastic pressure equation without replacing the ordinary product by Wick product. We note however that they consider a permeability process $\kappa(x,\omega;\phi) = \rho_0 + e^{W_\phi(x,\omega)}$, where $W_\phi(x,\omega)$ is the 1-dimensional smoothed white noise process defined on the 1-dimensional white noise probability space $\big(\mathcal{S}'(\mathbb{R}^d), \mathcal{B}(\mathcal{S}'(\mathbb{R}^d)), \mu\big)$; see Section 9.7.1 below. The constant $\rho_0 > 0$ is added to obtain uniform ellipticity of the associated bilinear form. Observe that for different $\phi \in L^2(\mathbb{R}^d)$ there is a different permeability function $\kappa(\cdot,\cdot,\phi)$ associated to it. Applying properly the Wiener-Chaos decomposition they obtain a symmetric positive definite linear system of equations whose solutions are the coefficients of a Galerkin-type approximation to the solution of the original equation. They do not provide a priori error estimates.

In this paper we use the white noise theory in a general setup to construct and characterize adequate spaces to prove the existence and uniqueness of the solution of the ordinary product stochastic pressure equation. Here, when writing the weak form of the equation we choose different spaces for the solution and test functions. This infinite dimensional problem is of Petrov-Galerkin type and we can establish the inf-sup condition and the well-posedness of the problem. This generalizes the approximation proposed in Benth and Theting [2002] and Roman and Sarkis [2006] to the spaces introduced to obtain existence and uniqueness of the weak solution of (9.1). For the finite dimensional problem we also use different spaces for the solution and the test functions, however it leads to the solution of a positive symmetric linear system. We prove the inf-sup stability of this approximation and provide a priory error estimates for a wide class of norms that depends on the choice of a sequence of weights. We also generalize and improve the results of Cao [2006] and Benth and Gjerde [1998] on approximation of a (generalized) process

by its truncated Wiener-Chaos expansion. We present numerical experiments in order to confirm the theory.

From the modeling point of view, the general setup considered, permit us several choices which result in different numerical methods and different a priori error estimates.

In order to circumvent the ellipticity and to develop a priori error estimates, we consider two types of norms. The first type, introduced to define the spaces for the solution of the stochastic pressure equation, measures the exponential decay of the solution in the white noise probability space implied by the same type of decay in the forcing term for our particular choice of the permeability κ which is of the form $\kappa(x,\omega;\phi) = e^{W_\phi(x,\omega)}$ as in Roman and Sarkis [2006] without the constant term added to obtain uniform ellipticity. The second type of norms are used to derive a priori error estimates, are of Hida-Kuo-Kondrative-Streit type and depend on the choice of sequence of weights; see Hida [1980], Kuo [1996], Holden et al. [1996], Hida et al. [1993], Da Prato [2006] and references therein. For some particular choices, these norms measure the *regularity* of the process in the white noise probability space variable just as the Sobolev norms measure regularity of functions. We point out that we consider a general setup which permits a unified analysis for several modelling choices such as smoothed white noise process and genelized KL expansions. For regularity result of the pressure equation for this type of norms we refer to Theting [2000] and Galvis and Sarkis [2008].

We mention that a different approach to ours is considered in Babuška et al. [2007]. By using L_p spaces for the probability space, they show the well-posedness for problems where the permeability coefficient is finite dimensional lognormal; see Lemma 1.2 and Example 1 in Babuška et al. [2007]. They consider a coefficient of the form of $\kappa(x,\omega) = e^{\sum_{j=1}^K a_j(x)Y_j(\omega)}$ and assume that each deterministic function a_j is bounded in D and their result depend on the quantity $\sum_{j=1}^K \sup_{x\in D} a_j(x)^2$. Using our approach, the assumption required to show well-posednes of the problem for this particular form of the permeability κ is $\sup_{x\in D}\sum_{j=1}^K \lambda_j^{2\theta} a_j(x)^2 < \infty$. For a coefficient of the form

$$\kappa(x,\omega) = e^{\sum_{j=1}^\infty a_j(x)Y_j(\omega)}$$

an extension of their result would require $\sum_{j=1}^\infty \sup_{x\in D} a_j(x)^2 < \infty$ and our assumption for this infinite dimensional noise case is that $\sup_{x\in D}\sum_{j=1}^\infty \lambda_j^{2\theta} a_j(x)^2 < \infty$. Here $\theta > 0$ and $\{\lambda_j\}_{j=1}^\infty$ is any sequence with $1 < \lambda_1 \leq \lambda_2 \leq \cdots$ and $\sum_{j=1}^\infty \lambda_j^{-2\theta} < \infty$; see Theorem 9.6 below.

This paper is structured as follows. In Section 9.2 we introduce the white noise calculus framework to be used in the rest of the paper. Section 9.3 is dedicated to describe the problem we are dealing with and to introduce the adequate spaces for the solution of the stochastic pressure equation. These spaces are characterized in Section 9.4 where additional norms are introduced in order to measure de regularity in the ω variable. Two examples of such norms are presented. In Section 9.5 we consider a Galerkin approximation and deduce a priori error estimates. The resulting linear system is studied in Section 9.6. Section 9.7 discusses some

modeling choices and finally in Section 9.8 we present a one dimensional numerical experiment.

9.2 Framework: White Noise Analysis

Let H be a real Hilbert space with inner product $(\cdot,\cdot)_H$ and norm $\|\cdot\|_H$. Let A be an operator on H such that there exists an H-orthonormal basis $\{\eta_j\}_{j=1}^{\infty}$ with

1. $A\eta_j = \lambda_j \eta_j$, $j = 1, 2, \ldots$.
2. $1 < \lambda_1 \leq \lambda_2 \leq \cdots$.
3. $\sum_{j=1}^{\infty} \lambda_j^{-2\theta} < \infty$ for some constant $\theta > 0$.

For $p > 0$ define $\mathcal{S}_p := \{\xi \in H; \quad \|\xi\|_p < \infty\}$ where

$$\|\xi\|_p^2 := \|A^p \xi\|_H^2 = \sum_{j=0}^{\infty} \lambda_j^{2p} (\xi, \eta_j)_H^2$$

and for $p < 0$ let $\mathcal{S}_p$ be the dual space of $\mathcal{S}_{-p}$. It is easy to see that $\|\cdot\|_p = \|A^p \cdot\|_H$ and the duality pairing between $\mathcal{S}_p$ and $\mathcal{S}_{-p}$ is an extension of the H inner product. We also define

$$\mathcal{S} = \cap_{p \geq 0} \mathcal{S}_p \text{ (with the projective limit topology)}$$

and $\mathcal{S}'$ as the dual space of $\mathcal{S}$, i.e., we use the standard countable Hilbert space constructed from (H, A); see Kuo [1996] and Obata [1994].

Let $\mathcal{S}'$ be the probability space with the sigma-field $\mathcal{B}(\mathcal{S}')$ of Borel subsets of $\mathcal{S}'$. The probability measure μ is given by the Bochner-Minlos theorem and characterized by

$$E_\mu e^{i\langle \cdot, \xi \rangle} := \int_{\mathcal{S}'} e^{i\langle \omega, \xi \rangle} d\mu(\omega) = e^{-\frac{1}{2}\|\xi\|_H^2}, \text{ for all } \xi \in \mathcal{S}. \tag{9.4}$$

Here, the pairing $\langle \omega, \xi \rangle = \omega(\xi)$ is the action of $\omega \in \mathcal{S}'$ on $\xi \in \mathcal{S}$, and E_μ denotes the expectation with respect to the measure μ; see [Obata- 1994, Chapter 1-3], [Holden et al.- 1996, Chapter 2], [Hida- 1980, Chapter 3], Hida et al. [1993], Kuo [1996] and Berezanskiĭ [1986]. The measure μ is often called the (normalized) *Gaussian measure* on $\mathcal{S}'$. The reason for this can be seen from the following remark:

Remark 9.1 *Equation (9.4) says that: for any test function $\xi \in \mathcal{S}$, the random variable $\langle \cdot, \xi \rangle$, is normally distributed with zero mean and variance $\|\xi\|_H^2$. If $\xi_1, \ldots, \xi_j \in \mathcal{S}$ are orthonormal in H then the random variables $\langle \cdot, \xi_1 \rangle, \ldots, \langle \cdot, \xi_j \rangle$ are independent and normally distributed with mean zero and variance equal to one; see Holden et al. [1996], Kuo [1996] and Obata [1994].*

The following particular case of Fernique's Theorem will be used throughout this paper; see Shigekawa [2004], Bogachev [1998], Kuo [1975], Da Prato [2006] and Da Prato and Zabczyk [1992].

Lemma 9.2 *We have*

$$\int_{\mathcal{S}'} e^{s\|\omega\|^2_{-\theta}} d\mu(\omega) = \begin{cases} \prod_{j=1}^{\infty} \left(1 - \frac{2s}{\lambda_j^{2\theta}}\right)^{-\frac{1}{2}}, & s < \frac{\lambda_1^{2\theta}}{2} \\ +\infty & s \geq \frac{\lambda_1^{2\theta}}{2} \end{cases}$$

Proof. Note that $\|\omega\|^2_{-\theta} = \sum_{j=1}^{\infty} \lambda_j^{-2\theta} \langle \omega, \eta_j \rangle^2$. Using the monotone convergence theorem when $s > 0$ or the dominated convergence theorem when $s < 0$ we have

$$\begin{aligned} \int_{\mathcal{S}'} e^{s\|\omega\|^2_{-\theta}} d\mu(\omega) &= \lim_{J\to\infty} \int_{\mathcal{S}'} e^{s\sum_{j=i}^{J} \lambda_j^{-2\theta} \langle \omega, \eta_j \rangle^2} d\mu(\omega) \\ &= \lim_{J\to\infty} \prod_{j=1}^{J} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{s\lambda_j^{-2\theta} y^2} e^{-\frac{1}{2}y^2} dy \\ &= \lim_{J\to\infty} \prod_{j=1}^{J} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp\left(-\frac{1}{2}\left(1 - \frac{2s}{\lambda_j^{2\theta}}\right) y^2\right) dy \\ &= \prod_{j=1}^{\infty} \left(1 - \frac{2s}{\lambda_j^{2\theta}}\right)^{-\frac{1}{2}}. \end{aligned}$$

In view of the assumption $\sum_{j=1}^{\infty} \lambda_j^{-2\theta} < \infty$, the infinite product above converges when $s < \frac{\lambda_1^{2\theta}}{2}$, and goes to $+\infty$ when s approaches $\frac{\lambda_1^{2\theta}}{2}$ from below. Because the integral above is monotonically increasing with respect to s, the lemma follows. ∎

We note that Lemma 9.2 implies that $\int_{\mathcal{S}'} \|\omega\|^2_{-\theta} d\mu < \infty$ which in turn implies that $\mu(\mathcal{S}_{-\theta}) = 1$. To see this note that if $\mathcal{S}' \setminus \mathcal{S}_{-\theta} = \{\omega : \|\omega\|^2_{-\theta} = \infty\}$ then $\mu(\mathcal{S}' \setminus \mathcal{S}_{-\theta}) > 0$ would imply that $\int_{S'} \|\omega\|^2_{-\theta} d\mu = \infty$. Without further comments, we use that $\mu(\mathcal{S}_{-\theta}) = 1$ throughout this paper.

Definition 9.3 *The 1-dimensional smoothed white noise associated to H and A is the map $w : \mathcal{S} \times$*
$mathcal S' \longrightarrow \mathbb{R}$ given by $w(\xi) = w(\xi, \omega) = \langle \omega, \xi \rangle$ for $\omega \in \mathcal{S}'$, $\xi \in \mathcal{S}$.

It is not difficult to prove that when $\xi \in H$ and we choose $\xi_n \in \mathcal{S}$ such that $\xi_n \to \xi$ in H, then $\langle \omega, \xi \rangle := \lim_{n\to+\infty} \langle \omega, \xi_n \rangle$ exists in $L^2(\mu)$, and is independent of the choice of $\{\xi_n\}_{n=1}^{\infty}$. Thus, the definition of smoothed white noise can be extended to functions in H. In what follows we use the notation (L^2) for the space $L^2(\mu)$. We always interpretate properties in the "almost everywhere" or "almost surely" sense, therefore to make notation and formula less cumbersome, we will sometimes omit such words.

Definition 9.4 *Let $D \subset \mathbb{R}^d$. Using the map w of Definition 9.3 we can construct a stochastic process, called the* smoothed white noise process *$W_\phi(x, \omega)$, as follows:*

$$W_\phi(x, \omega) := w(\phi_x, \omega) = \langle \omega, \phi_x \rangle, \ x \in D, \ \omega \in \mathcal{S}',$$

where $\phi_x \in H$, for all $x \in D$. For examples and properties of ϕ_x, see Section 9.7.

Remark 9.5 *Note that the process $\{W_\phi(x, \cdot)\}_{x\in D}$ has the following properties:*

i) *For each $x \in D$, $W_\phi(x,\cdot)$ is normally distributed with zero mean and variance $\|\phi_x\|_H$.*

ii) *For each $x, \hat{x} \in D$ we have that $E_\mu W_\phi(x,\cdot)W_\phi(\hat{x},\cdot) = (\phi_x, \phi_{\hat{x}})_H$.*

9.3 The Problem and Variational Formulation

Given $\phi_x \in \mathcal{S}_\theta$ for all $x \in D$ we consider the following problem:

$$\begin{cases} -\nabla_x \cdot (\kappa(x,w;\phi)\nabla_x u(x,w;\phi)) &= f(x,\omega), \text{ for all } x \in D \\ u(x,\cdot;\phi) &= 0, \text{ for all } x \in \partial D \end{cases} \tag{9.5}$$

for all $w \in \mathcal{S}'$, where

$$\kappa(x,\omega;\phi) := e^{W_\phi(x,\omega)} = e^{\langle \omega, \phi_x \rangle} \tag{9.6}$$

and the exponent $W_\phi(x,\omega)$ is the 1-dimensional smoothed white noise process of Definition 9.4. Thus, κ is log-normal random process. Observe that for different maps $x \mapsto \phi_x \in \mathcal{S}_\theta$ there is a different permeability function $\kappa(\cdot,\cdot,\phi)$ associated to it. We will omit, whenever there is no danger of confusion, the dependence of κ on the map $x \mapsto \phi_x$ just to make the notation less cumbersome.

To motivate the definition of the spaces for the solution of (9.5), observe that since $\mu(\mathcal{S}_{-\theta}) = 1$ and $\phi_x \in \mathcal{S}_\theta$ for all $x \in D$, we can write

$$|\langle \omega, \phi_x \rangle| \le \|\omega\|_{-\theta} \sup_{x \in D} \|\phi_x\|_\theta \qquad \omega\text{-a.s. in } \mathcal{S}'.$$

Denote $C_\theta = C_\theta(\phi) := \sup_{x\in D} \|\phi_x\|_\theta$. Then we have for all $\epsilon > 0$

$$-\frac{\epsilon}{2}\|\omega\|^2_{-\theta} - \frac{C_\theta^2}{2\epsilon} \le -\|\omega\|^2_{-\theta} C_\theta \le \langle \omega, \phi_x \rangle \le \|\omega\|^2_{-\theta} C_\theta \le \frac{\epsilon}{2}\|\omega\|^2_{-\theta} + \frac{C_\theta^2}{2\epsilon}$$

and

$$\kappa_{\min}(w) := e^{-\frac{C_\theta^2}{2\epsilon}} e^{-\frac{\epsilon}{2}|\omega|^2_{-\theta}} \le \kappa(x,w) \le e^{\frac{C_\theta^2}{2\epsilon}} e^{\frac{\epsilon}{2}\|\omega\|^2_{-\theta}} =: \kappa_{\max}(w). \tag{9.7}$$

When $u(\cdot,\omega)$ is the weak solution of (9.5) for almost all $\omega \in \mathcal{S}'$, then from the Lax-Millgram Lemma we should have

$$|u(\cdot,\omega)|^2_{H_0^1(D)} \le \frac{1}{\kappa_{\min}(\omega)^2}\|f(\cdot,\omega)\|^2_{H^{-1}(D)} = e^{\frac{C_\theta^2}{\epsilon}} e^{\epsilon\|\omega\|^2_{-\theta}} \|f(\cdot,\omega)\|^2_{H^{-1}(D)}.$$

Then for $s \in \mathbb{R}$ we can write

$$|u(\cdot,\omega)|^2_{H_0^1(D)} e^{s\|\omega\|^2_{-\theta}} \le e^{\frac{C_\theta^2}{\epsilon}} \|f(\cdot,\omega)\|^2_{H^{-1}(D)} e^{(s+\epsilon)|\omega|^2_{-\theta}} \qquad \text{a.s. in } \mathcal{S}'$$

and integrating both sides we obtain

$$\int_{\mathcal{S}'} |u(\cdot,\omega)|^2_{H_0^1(D)} e^{s\|\omega\|^2_{-\theta}} d\mu(\omega) \le e^{\frac{C_\theta^2}{\epsilon}} \int_{\mathcal{S}'} \|f(\cdot,\omega)\|^2_{H^{-1}(D)} e^{(s+\epsilon)\|\omega\|^2_{-\theta}} d\mu(\omega). \tag{9.8}$$

This last inequality gives us an idea of the spaces where we can seek the solutions and choose the test functions. For the solution space we use the left-hand side norm given in (9.8), while for the test functions spaces we take right-hand side

dual norm.

Define $\mathcal{U}_s^m$ as the space of functions $u : D \times \mathcal{S}' \to \mathbb{R}$ such that

$$\int_{\mathcal{S}'} \|u(\cdot,\omega)\|^2_{H^m(D)} e^{s\|\omega\|^2_{-\theta}} d\mu(\omega) < +\infty \tag{9.9}$$

with norm

$$\|u\|^2_{\mathcal{U}_s^m} := \int_{\mathcal{S}'} \|u(\cdot,\omega)\|^2_{H^m(D)} e^{s\|\omega\|^2_{-\theta}} d\mu(\omega) \tag{9.10}$$

and seminorm

$$|u|^2_{\mathcal{U}_s^m} := \int_{\mathcal{S}'} |u(\cdot,\omega)|^2_{H^m(D)} e^{s\|\omega\|^2_{-\theta}} d\mu(\omega). \tag{9.11}$$

Note that $\mathcal{U}_0^0 = L^2(D) \otimes (L^2)$ and in general $\mathcal{U}_s^m = H^m(D) \otimes (L^2)_s$ where

$$(L^2)_s := L^2(\mathcal{S}', e^{s\|\omega\|^2_{-\theta}} d\mu(\omega)) \tag{9.12}$$

with norm $\|v\|^2_{(L^2)_s} := \int_{\mathcal{S}'} |v(\omega)|^2 e^{s\|\omega\|^2_{-\theta}} d\mu$. We also define $\widehat{\mathcal{U}}_s^1 := H_0^1(D) \otimes (L^2)_s \subset \mathcal{U}_s^1$, i.e., the functions in $\mathcal{U}_s^1$ which vanish on ∂D almost sure in ω. By using a Poincaré inequality, the seminorm $|\cdot|_{\mathcal{U}_s^1}$ is a norm equivalent to $\|\cdot\|_{\mathcal{U}_s^1}$ in $\widehat{\mathcal{U}}_s^1$. Since the space $(L^2)_s$ is the dual of $(L^2)_{-s}$ and the $H^{-1}(D)$ is the dual of $H_0^1(D)$, we can identify the dual space of $\widehat{\mathcal{U}}_{-s}^1$ with $\mathcal{U}_s^{-1}$ where the duality pairing is given by

$$\langle f, v\rangle := \int_{D\times\mathcal{S}'} f(x,\omega)v(x,\omega)dxd\mu \quad \text{for all } v \in \widehat{\mathcal{U}}_{-s}^1, \quad f \in \mathcal{U}_s^{-1}. \tag{9.13}$$

We also define the bilinear form $a : \widehat{\mathcal{U}}_s^1 \times \widehat{\mathcal{U}}_{-s+\epsilon}^1 \to \mathbb{R}$ by

$$a(u,v) := \int_{D\times\mathcal{S}'} \kappa(x,\omega)\nabla u(x,\omega)\nabla v(x,\omega)dxd\mu. \tag{9.14}$$

The *weak formulation* of problem (9.5) is introduced as follows:

$$\begin{cases} \text{Find } \hat{u} \in \widehat{\mathcal{U}}_s^1 \text{ such that} \\ a(\hat{u}, v) = \langle f, v\rangle \text{ for all } v \in \widehat{\mathcal{U}}_{-s+\epsilon}^1, \end{cases} \tag{9.15}$$

where the duality pairing between $f \in \mathcal{U}_{s-\epsilon}^1$ and $v \in \widehat{\mathcal{U}}_{-s+\epsilon}$ is given by

$$\langle f, v\rangle = \int_{D\times\mathcal{S}'} f(x,\omega)v(x,\omega)dxd\mu.$$

Theorem 9.6 (Inf-sup condition) *Let $\epsilon > 0$ and assume that $C_\theta = \sup_{x\in D} \|\phi_x\|_\theta < \infty$. Then the following results follow:*

1. *The bilinear form $a : \widehat{\mathcal{U}}_s^1 \times \widehat{\mathcal{U}}_{-s+\epsilon}^1 \to \mathbb{R}$ is continuous and $||a|| \le e^{\frac{C_\theta^2}{2\epsilon}}$.*
2. *The bilinear form "a" satisfies the following inf-sup condition:*

$$\inf_{u\in\widehat{\mathcal{U}}_s^1\setminus\{0\}} \sup_{v\in\widehat{\mathcal{U}}_{-s-\epsilon}^1\setminus\{0\}} \frac{a(u,v)}{|u|_{\mathcal{U}_s^1}|v|_{\mathcal{U}_{-s-\epsilon}^1}} \ge e^{-\frac{C_\theta^2}{2\epsilon}}. \tag{9.16}$$

3. *For any* $v \in \widehat{\mathcal{U}}^1_{-s-\epsilon} \setminus \{0\}$ *there exists* $u(v) \in \widehat{\mathcal{U}}^1_{s+2\epsilon}$ *such that* $a(u,v) \neq 0$.
4. *For any* $f \in \mathcal{U}^{-1}_{s+\epsilon} \subset \mathcal{U}^{-1}_{s-\epsilon}$ *there exists a unique solution* $\hat{u} \in \widehat{\mathcal{U}}^1_s$ *of problem (9.15) and*

$$\|\hat{u}\|_{\mathcal{U}^1_s} \leq C e^{\frac{C_\theta^2}{2\epsilon}} \|f\|_{\mathcal{U}^{-1}_{s+\epsilon}}, \tag{9.17}$$

where C *is the Poincaré inequality constant which is independent of* ϵ *and* θ.

Proof. We proceed as follows:

1. From (9.7) we have

$$\begin{aligned} a(u,v) &= \int_{D\times\mathcal{S}'} \kappa(x,\omega)\nabla u(x,\omega)\nabla v(x,\omega)dxd\mu \\ &\leq e^{\frac{C_\theta^2}{2\epsilon}} \int_{\mathcal{S}'} e^{\frac{\epsilon}{2}|\omega|^2_{-\theta}} |u(\cdot,\omega)|_{H^1_0(D)} |v(\cdot,\omega)|_{H^1_0(D)} d\mu \leq e^{\frac{C_\theta^2}{2\epsilon}} |u|_{\mathcal{U}^1_s} |v|_{\mathcal{U}^1_{-s+\epsilon}}. \end{aligned}$$

2. Given $u \in \widehat{\mathcal{U}}^1_s \setminus \{0\}$ define

$$v_r(x,\omega) := \begin{cases} u(x,\omega)e^{(s+\frac{\epsilon}{2})\|\omega\|^2_{-\theta}}, & \text{if } x \in D \text{ and } \|w\|_{-\theta} \leq r \\ 0, & \text{if } x \in D \text{ and } \|w\|_{-\theta} > r. \end{cases}$$

Denote $B(r) := \{\omega \in \mathcal{S}' : \|w\|_{-\theta} \leq r\}$. From (9.7) we see that

$$\begin{aligned} a(u,v_r) &= \int_{D\times B(r)} \kappa(x,\omega)|\nabla u(x,\omega)|^2 e^{(s+\frac{\epsilon}{2})\|\omega\|^2_{-\theta}} dxd\mu(\omega) \\ &\leq e^{\frac{C_\theta^2}{2\epsilon}} \int_{D\times B(r)} |\nabla u(x,\omega)|^2 e^{(s+\epsilon)\|\omega\|^2_{-\theta}} dxd\mu(\omega) \\ &\leq e^{\frac{C_\theta^2}{2\epsilon}} e^{\epsilon r^2} \int_{D\times B(r)} |\nabla u(x,\omega)|^2 e^{s\|\omega\|^2_{-\theta}} dxd\mu(\omega) \leq e^{\frac{C_\theta^2}{2\epsilon}+\epsilon r^2} |u|^2_{\mathcal{U}^1_s} < \infty, \end{aligned}$$

and therefore, $a(u,v_r)$ is well defined for all r. We also have

$$\begin{aligned} |v_r|^2_{\mathcal{U}^1_{-s-\epsilon}} &= \int_{D\times B(r)} |\nabla u(x,\omega)|^2 e^{2(s+\frac{\epsilon}{2})\|\omega\|^2_{-\theta}} e^{-(s+\epsilon)\|\omega\|^2_{-\theta}} dxd\mu(\omega) \\ &= \int_{D\times B(r)} |\nabla u(x,\omega)|^2 e^{s\|\omega\|^2_{-\theta}} dxd\mu(\omega) \leq |u|^2_{\mathcal{U}^1_s} \end{aligned}$$

and using (9.7) we obtain

$$\begin{aligned} a(u,v_r) &= \int_{D\times B(r)} \kappa(x,\omega)|\nabla u(x,\omega)|^2 e^{(s+\frac{\epsilon}{2})\|\omega\|^2_{-\theta}} dxd\mu(\omega) \\ &\geq e^{-\frac{C_\theta^2}{2\epsilon}} \int_{D\times B(r)} e^{-\frac{\epsilon}{2}\|\omega\|^2_{-\theta}} |\nabla u(x,\omega)|^2 e^{(s+\frac{\epsilon}{2})\|\omega\|^2_{-\theta}} dxd\mu(\omega) \\ &= e^{-\frac{C_\theta^2}{2\epsilon}} \int_{D\times B(r)} |\nabla u(x,\omega)|^2 e^{s\|\omega\|^2_{-\theta}} dxd\mu(\omega). \end{aligned}$$

For any arbitrary $\delta > 0$, take $R(\delta) > 0$ such that $\int_{D\times B(R)} |\nabla u(x,\omega)|^2 e^{s\|\omega\|^2_{-\theta}} dx d\mu(\omega) > (1-\delta)|u|^2_{\mathcal{U}^1_s}$. We obtain $a(u, v_R) > (1-\delta)e^{-\frac{C_\theta^2}{2\epsilon}}|u|^2_{\mathcal{U}^1_s}$ and then,

$$\sup_{v\in\mathcal{U}^1_{-s-\epsilon}} \frac{a(u,v)}{|v|_{\mathcal{U}^1_{-s-\epsilon}}} \geq \frac{a(u,v_R)}{|v_R|_{\mathcal{U}^1_{-s-\epsilon}}} > (1-\delta)e^{-\frac{C_\theta^2}{2\epsilon}} \frac{|u|^2_{\mathcal{U}^1_s}}{|u|_{\mathcal{U}^1_s}} = (1-\delta)e^{-\frac{C_\theta^2}{2\epsilon}}|u|_{\mathcal{U}^1_s}.$$

Because $\delta > 0$ is arbitrary, we conclude that the inf-sup condition (9.16) holds.

3. Given $v \in \widehat{\mathcal{U}}^1_{-s-\epsilon} \setminus \{0\}$ we can take u_r defined by

$$u_r(x,\omega) := \begin{cases} v(x,\omega)e^{(-s-\frac{\epsilon}{2})\|\omega\|^2_{-\theta}}, & \text{if } x \in D \text{ and } \|w\|_{-\theta} \leq r \\ 0, & \text{if } x \in D \text{ and } \|w\|_{-\theta} > r. \end{cases}$$

Note that

$$\begin{aligned} |u_r|^2_{\mathcal{U}^1_{s+2\epsilon}} &= \int_{D\times B(r)} |\nabla v(x,\omega)|^2 e^{2(-s-\frac{\epsilon}{2})\|\omega\|^2_{-\theta}} e^{(s+2\epsilon)\|\omega\|^2_{-\theta}} dx d\mu(\omega) \\ &\leq e^{2\epsilon r^2} \int_{D\times B(r)} |\nabla v(x,\omega)|^2 e^{(-s-\epsilon)\|\omega\|^2_{-\theta}} dx d\mu(\omega) \leq e^{2\epsilon r^2}|v|^2_{\mathcal{U}^1_{-s-\epsilon}} < \infty, \end{aligned}$$

and take R large enough to have

$$a(u_R, v) \geq e^{-\frac{C_\theta^2}{2\epsilon}} \int_{D\times B(R)} e^{-\frac{\epsilon}{2}\|\omega\|^2_{-\theta}} |\nabla v(x,\omega)|^2 e^{(-s-\frac{\epsilon}{2})\|\omega\|^2_{-\theta}} dx d\mu(\omega) > 0. \tag{9.18}$$

4. Let $T_a : \widehat{\mathcal{U}}^1_s \to \mathcal{U}^{-1}_{s-\epsilon}$ be the linear continuous operator defined by

$$a(u,v) = \langle T_a u, v\rangle \quad \text{for all } u \in \widehat{\mathcal{U}}^1_s, \quad v \in \widehat{\mathcal{U}}^1_{-s+\epsilon},$$

and let $\mathcal{R}(T_a)$ be the range of T_a. Now we show that $\mathcal{R}(T_a) \cap \mathcal{U}^{-1}_{s+\epsilon}$ is closed in $\mathcal{U}^{-1}_{s+\epsilon}$. Indeed, let $\{u_n\} \subset \widehat{\mathcal{U}}^1_s$ be a sequence such that $\{T_a u_n\} \subset \mathcal{U}^{-1}_{s+\epsilon}$ converge to f in $\mathcal{U}^{-1}_{s+\epsilon}$. From the inf-sup (9.16), for all integers m and n we have

$$\begin{aligned} |u_m - u_n|_{\mathcal{U}^1_s} &\leq e^{\frac{C_\theta^2}{2\epsilon}} \sup_{v\in\widehat{\mathcal{U}}^1_{-s-\epsilon}\setminus\{0\}} \frac{a(u_m-u_n,v)}{|v|_{\mathcal{U}^1_{-s-\epsilon}}} \\ &\leq e^{\frac{C_\theta^2}{2\epsilon}} \sup_{v\in\widehat{\mathcal{U}}^1_{-s-\epsilon}\setminus\{0\}} \frac{\langle T_a(u_m-u_n),v\rangle}{|v|_{\mathcal{U}^1_{-s-\epsilon}}} \leq e^{\frac{C_\theta^2}{2\epsilon}} \|T_a(u_m-u_n)\|_{\mathcal{U}^{-1}_{s+\epsilon}} \end{aligned}$$

which implies that $\{u_n\}$ is a Cauchy sequence in $\widehat{\mathcal{U}}^1_s$, and hence has a limit $u \in \widehat{\mathcal{U}}^1_s$. By continuity we have that $T_a u = f$, then $f \in \mathcal{R}(T_a)$ and since $f \in \mathcal{U}^{-1}_{s+\epsilon}$ we have that $f \in \mathcal{R}(T_a) \cap \mathcal{U}^{-1}_{s+\epsilon}$. So we have $\mathcal{R}(T_a) \cap \mathcal{U}^{-1}_{s+\epsilon}$ is closed in $\mathcal{U}^{-1}_{s+\epsilon}$. Now we show $\mathcal{R}(T_a) \cap \mathcal{U}^{-1}_{s+\epsilon} = \mathcal{U}^{-1}_{s+\epsilon}$. Assume by contradiction that there exists $v \in (\mathcal{U}^{-1}_{s+\epsilon})^* = \widehat{\mathcal{U}}^1_{-s-\epsilon}$ such that $v \neq 0$ and $\langle f, v\rangle = 0$ for all $f \in \mathcal{R}(T_a) \cap \mathcal{U}^{-1}_{s+\epsilon}$. With u_R introduced in 3. above, we have from (9.7)

$$\langle T_a u_R, z\rangle = a(u_R, z) \leq e^{\frac{C_\theta^2}{2\epsilon}} |u_R|_{\mathcal{U}^1_{s+2\epsilon}} |z|_{\mathcal{U}^1_{-s-\epsilon}} \text{ for all } z \in \widehat{\mathcal{U}}^1_{-s-\epsilon},$$

which implies that $T_a u_R \in (\widehat{\mathcal{U}}^1_{-s-\epsilon})^* = \mathcal{U}^{-1}_{s+\epsilon}$. Taking $f = T_a u_R$ implies $0 = \langle T_a u_R, v\rangle = a(u_R, v)$ and using (9.18) we conclude $v = 0$ which gives a contradiction. So $\mathcal{R}(T_a) \cap \mathcal{U}^{-1}_{s+\epsilon} = \mathcal{U}^{-1}_{s+\epsilon}$. Therefore, problem (9.15) has a unique solution when the right hand side $f \in \mathcal{U}^{-1}_{s+\epsilon}$. Now (9.17) follows directly from the inf-sup condition (9.16). ∎

Remark 9.7 *From Theorem 9.6 , when $f \in \mathcal{U}_0^{-1} = H^{-1}(D) \otimes (L^2)$, the solution $u \in \widehat{\mathcal{U}}^1_s$ for every $s < 0$ (take $\epsilon = -s$). In order that $u \in \widehat{\mathcal{U}}^1_0 = H^1_0(D) \otimes (L^2)$ we need $f \in \mathcal{U}^{-1}_\epsilon$ for some $\epsilon > 0$. When the right hand side f is given by a finite sum of Fourier-Hermite polynomials we have that the solution $u \in \widehat{\mathcal{U}}^1_s$ for every s with $s < \frac{\lambda_1^{2\theta}}{2}$; see Definition 9.9 and Theorem 9.10.*

Remark 9.8 *If $\widetilde{U} \subset \widehat{\mathcal{U}}^1_{s+\epsilon} \subset \widehat{\mathcal{U}}^1_s$ for some $\epsilon > 0$, then the pair of spaces $\widetilde{U}$ and $\widetilde{V}$, where $\widetilde{V}$ is defined by $\widetilde{V} := \left\{ue^{(s+\frac{\epsilon}{2})\|\cdot\|^2_{-\theta}}; u \in \widetilde{U}\right\}$, also satisfies the inf-sup condition. This will be useful when constructing finite element spaces in Section 9.5; see Remark 9.19.*

9.4 Characterization of the Spaces $(L^2)_s$ and $\mathcal{U}^m_s$

In the following we characterize the space $(L^2)_s$ defined in (9.12). This is enough for characterizing the tensor product space $\mathcal{U}^m_s$. Since there is no danger of confusion we denote functions in $\mathcal{U}^m_s$ and $(L^2)_s$ by the same symbols.

We need to consider multi-index of arbitrary length. To simplify the notation, we regard multi-indices as elements of the space $(\mathbb{N}_0^{\mathbb{N}})_c$ of all sequences $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots)$ with elements $\alpha_j \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$ and with compact support, i.e., with only finitely many $\alpha_j \neq 0$. We write $\boldsymbol{\mathcal{J}} = (\mathbb{N}_0^{\mathbb{N}})_c$. Given $\boldsymbol{\alpha} \in \boldsymbol{\mathcal{J}}$, define the order and length of $\boldsymbol{\alpha}$, denoted by $d(\boldsymbol{\alpha})$ and $|\boldsymbol{\alpha}|$ respectively, by

$$d(\boldsymbol{\alpha}) := \max\{j : \alpha_j \neq 0\} \tag{9.19}$$

and

$$|\boldsymbol{\alpha}| := \alpha_1 + \alpha_2 + \ldots + \alpha_{d(\boldsymbol{\alpha})}.$$

We also introduce the σ-Hermite polynomials, $h_{\sigma^2,n}$, where $\sigma > 0$ and $n = 0, 1, 2, \ldots$. These polynomials can be defined by the generating function identity

$$e^{tx - \frac{1}{2}\sigma^2 t^2} = \sum_{n=0}^{\infty} \frac{t^n}{n!} h_{\sigma^2,n}(x). \tag{9.20}$$

When $\sigma^2 = 1$ we denote $h_{1,n}$ simply by h_n. The σ-Hermite polynomials are an orthogonal basis for $L^2(\mathbb{R}, e^{-\frac{1}{2\sigma^2}x^2}dx)$.

For $s < \frac{\lambda_1^{2\theta}}{2}$, define $\sigma_j = \sigma_j(s) := \left(1 - \frac{2s}{\lambda_j^{2\theta}}\right)^{-\frac{1}{2}}$, $j = 1, 2, \ldots$, and for $\boldsymbol{\alpha} \in \boldsymbol{\mathcal{J}}$ denote

$$\sigma^{\boldsymbol{\alpha}} = \sigma^{\boldsymbol{\alpha}}(s) := \prod_{j=1}^{d(\boldsymbol{\alpha})} \sigma_j^{\alpha_j}(s)$$

and define

$$\sigma_* = \sigma_*(s) := \int_{\mathcal{S}'} e^{s\|\omega\|^2_{-\theta}} d\mu(\omega). \tag{9.21}$$

From Lemma 9.2, $\sigma_* = \prod_{j=1}^{\infty} \sigma_j < \infty$ when $s < \frac{\lambda_1^{2\theta}}{2}$. Now we define the $\sigma(s)$-Fourier-Hermite polynomials.

Definition 9.9 *Let $s < \frac{\lambda_1^{2\theta}}{2}$, $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots) \in \boldsymbol{\mathcal{J}}$ and $\sigma = \sigma(s)$ as before. Define*

$$H_{\sigma^2,\boldsymbol{\alpha}}(\omega) := \frac{1}{\sqrt{\sigma_*}} \prod_{j=1}^{d(\boldsymbol{\alpha})} h_{\sigma_j^2,\alpha_j}(\langle \omega, \eta_j \rangle); \quad \omega \in \mathcal{S}'.$$

We now state the Wiener-Itô Wiener-Chaos expansion theorem; see Da Prato [2006], Hida [1980], [Holden et al.- 1996, Theorem 2], Hida et al. [1993] and Obata [1994]. For completeness we include the proof of the orthogonality of the $\sigma(s)$-Fourier-Hermite polynomials.

Theorem 9.10 *For $s < \frac{\lambda_1^{2\theta}}{2}$ the $\sigma(s)$-Fourier-Hermite polynomials are orthogonal in $(L^2)_s$. Moreover,*

$$\|H_{\sigma^2(s),\boldsymbol{\alpha}}\|^2_{(L^2)_s} = \boldsymbol{\alpha}!\sigma(s)^{2\boldsymbol{\alpha}}.$$

Then, polynomials in ω are elements of in $(L)_s$ and every $u \in (L^2)_s$ is of the form

$$u = \sum_{\boldsymbol{\alpha}\in\boldsymbol{\mathcal{J}}} u_{\boldsymbol{\alpha},s} H_{\sigma(s)^2,\boldsymbol{\alpha}}$$

with $\|u\|^2_{(L^2)_s} = \sum_{\boldsymbol{\alpha}\in\boldsymbol{\mathcal{J}}} \boldsymbol{\alpha}!\sigma(s)^{2\boldsymbol{\alpha}} u^2_{\boldsymbol{\alpha},s}$.

Proof. Take $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \boldsymbol{\mathcal{J}}$ and let $M = \max\{d(\boldsymbol{\alpha}), d(\boldsymbol{\beta})\}$. Note that the random variables

$$\prod_{j=1}^{M} h_{\sigma_j^2,\alpha_j}(\langle \cdot, \eta_j \rangle) h_{\sigma_j^2,\beta_j}(\langle \cdot, \eta_j \rangle) e^{s\lambda_j^{-2\theta}\langle \cdot, \eta_j \rangle} \text{ and } \prod_{j=M+1}^{\infty} e^{s\lambda_j^{-2\theta}\langle \cdot, \eta_j \rangle}$$

are independent (see Remark 9.1). Then

$$\begin{aligned}
& \frac{1}{\sigma_*} \int_{\mathcal{S}'} H_{\sigma^2,\boldsymbol{\alpha}}(\omega) H_{\sigma^2\boldsymbol{\beta}}(\omega) e^{s\|\omega\|^2_{-\theta}} d\mu(\omega) \\
= & \int_{\mathcal{S}'} \prod_{j=1}^{\infty} h_{\sigma_j^2,\alpha_j}(\langle \cdot, \eta_j \rangle) h_{\sigma_j^2,\beta_j}(\langle \cdot, \eta_j \rangle) \frac{e^{s\lambda_j^{-2\theta}\langle \cdot, \eta_j \rangle}}{\sigma_j(s)} d\mu \\
= & \int_{\mathcal{S}'} \prod_{j=1}^{M} h_{\sigma_j^2,\alpha_j}(\langle \cdot, \eta_j \rangle) h_{\sigma_j^2,\beta_j}(\langle \cdot, \eta_j \rangle) \frac{e^{s\lambda_j^{-2\theta}\langle \cdot, \eta_j \rangle}}{\sigma_j(s)} d\mu \int_{\mathcal{S}'} \prod_{j=M+1}^{\infty} \frac{e^{s\lambda_j^{-2\theta}\langle \cdot, \eta_j \rangle}}{\sigma_j(s)} d\mu \\
= & \prod_{j=1}^{M} \sigma_j^{2\alpha_j} \alpha_j! \delta_{\alpha_j,\beta_j} = \boldsymbol{\alpha}!\sigma^{2\boldsymbol{\alpha}} \delta_{\boldsymbol{\alpha},\boldsymbol{\beta}}.
\end{aligned}$$

■

Remark 9.11 *The corresponding tensor product norm for $u \in \mathcal{U}_s^m$ with $s < \frac{\lambda_1^{2\theta}}{2}$ is given by*

$$\|u\|_{\mathcal{U}_s^m}^2 = \sum_{\boldsymbol{\alpha}\in\boldsymbol{\mathcal{J}}} \boldsymbol{\alpha}!\sigma(s)^{2\boldsymbol{\alpha}}\|u_{\boldsymbol{\alpha},s}\|_{H^m(D)}^2,$$

where $u = \sum_{\boldsymbol{\alpha}\in\boldsymbol{\mathcal{J}}} u_{\boldsymbol{\alpha},s}H_{\sigma(s)^2,\boldsymbol{\alpha}}$ with $u_{\boldsymbol{\alpha},s} \in H^m(D)$ for all $\boldsymbol{\alpha} \in \boldsymbol{\mathcal{J}}$.

In $(L^2)_s$ with $s < \frac{\lambda_1^{2\theta}}{2}$ we introduce the system of Hilbert norms

$$||u||_{p;\rho,s}^2 := \sum_{\boldsymbol{\alpha}\in\boldsymbol{\mathcal{J}}} \rho(\boldsymbol{\alpha},p)^2\boldsymbol{\alpha}!\sigma(s)^{2\boldsymbol{\alpha}}u_{\boldsymbol{\alpha},s}^2, \tag{9.22}$$

where $u = \sum_{\boldsymbol{\alpha}\in\boldsymbol{\mathcal{J}}} u_{\boldsymbol{\alpha},s}H_{\sigma(s)^2,\boldsymbol{\alpha}}$. We assume that $\rho(\boldsymbol{\alpha},q) \geq \rho(\boldsymbol{\alpha},p) > 0$ and $\rho(\boldsymbol{\alpha},0) = 1$ for all $q > p \geq 0$ and $\boldsymbol{\alpha} \in \boldsymbol{\mathcal{J}}$. Usually the *weights* $\rho(\boldsymbol{\alpha},s)$ are the eigenvalues of some nonnegative operator in $(L^2)_s$ with the $\sigma(s)$-Fourier-Hermite polynomials as eigenfunctions; see Benth and Gjerde [1998], Holden et al. [1996], Hida et al. [1993], Obata [1994], Kuo [1996], Bogachev [1998], Cochran et al. [1998] and Benth and Theting [2002].

For $p > 0$ define the spaces $(\mathcal{S}_p)_{\rho,s}$ by

$$(\mathcal{S}_p)_{\rho,s} := \{v \in (L^2)_s : \|v\|_{p;\rho,s} < \infty\}. \tag{9.23}$$

For $p < 0$ define $(\mathcal{S}_p)_{\rho,s}$ as the dual space of $(\mathcal{S}_{-p})_{\rho,s}$. We have $(\mathcal{S}_0)_{\rho,s} = (L^2)_s$ and the inclusion $(\mathcal{S}_q)_{\rho,s} \subset (\mathcal{S}_p)_{\rho,s}$ holds for all $q > p$.

Let $N, K \in \mathbb{N}_0$ and define

$$\boldsymbol{\mathcal{J}}^{N,K} := \{\boldsymbol{\alpha} \in \boldsymbol{\mathcal{J}} : d(\boldsymbol{\alpha}) \leq K, \text{ and, } |\boldsymbol{\alpha}| \leq N\} \tag{9.24}$$

and

$$\mathcal{P}^{N,K} := \operatorname{span}\left\{H_{\sigma(s)^2,\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \boldsymbol{\mathcal{J}}^{N,K}\right\} = \operatorname{span}\left\{\prod_{j=1}^{d(\boldsymbol{\alpha})}\langle\omega,\eta_j\rangle^{\alpha_j} : \boldsymbol{\alpha} \in \boldsymbol{\mathcal{J}}\right\},$$

i.e., $\mathcal{P}^{N,K}$ consists of polynomials in $\langle\omega,\eta_1\rangle,\ldots,\langle\omega,\eta_K\rangle$ of total degree at most N. Let $Q_s^{N,K}$ be the orthogonal projection on $\mathcal{P}^{N,K}$ in the $(L^2)_s$-norm. This projection is equivalent to the procedure of truncating the expansion in terms of $\sigma(s)$-Fourier-Hermite polynomials by eliminating the coefficients corresponding to multi-indices outside $\boldsymbol{\mathcal{J}}^{N,K}$. We also define the $\mathbb{R}^K$ approximation of $(L^2)_s$ by

$$\mathcal{A}_s^K := \operatorname{span}\left\{H_{\sigma(s)^2,\boldsymbol{\alpha}} : d(\boldsymbol{\alpha}) \leq K\right\}$$

and denote by Q_s^K the orthogonal projection on $\mathcal{A}_s^K$ in the $(L^2)_s$-norm.

We have the following approximation results:

Theorem 9.12 *Assume that $s < \frac{\lambda_1^{2\theta}}{2}$. Then for all $v \in (\mathcal{S}_q)_{\rho,s}$ and $p < q$ we have*

$$\|v - Q_s^{N,K}v\|_{p;\rho,s}^2 \leq M_1^2\|Q_s^Kv - Q_s^{N,K}v\|_{q;\rho,s}^2 + M_2^2\|v - Q_s^Kv\|_{q;\rho,s}^2$$

with

$$M_1 = M_1(\rho, p, q) := \max_{d(\boldsymbol{\alpha})\leq K, |\boldsymbol{\alpha}|>N} \frac{\rho(\boldsymbol{\alpha}, p)}{\rho(\boldsymbol{\alpha}, q)} \tag{9.25}$$

and

$$M_2 = M_2(\rho, p, q) := \max_{d(\boldsymbol{\alpha})>K} \frac{\rho(\boldsymbol{\alpha}, p)}{\rho(\boldsymbol{\alpha}, q)}. \tag{9.26}$$

Proof. Fix $s < \frac{\lambda_1^{2\theta}}{2}$ and note that $Q_s^{N,K} v = \sum_{\boldsymbol{\alpha}\in\boldsymbol{\mathcal{J}}^{N,K}} v_{\boldsymbol{\alpha},s} H_{\sigma^2,\boldsymbol{\alpha}}$. Then recalling the definition of $\boldsymbol{\mathcal{J}}^{N,K}$ in (9.24) we see that

$$\begin{aligned}
||v - Q_s^{N,K} v||_{p;\rho,s}^2 &= \sum_{\boldsymbol{\alpha}\notin\boldsymbol{\mathcal{J}}^{N,K}} \rho(\boldsymbol{\alpha}, p)^2 \boldsymbol{\alpha}! \sigma(s)^{2\boldsymbol{\alpha}} v_{\boldsymbol{\alpha},s}^2 \\
&= \sum_{\boldsymbol{\alpha}\notin\boldsymbol{\mathcal{J}}^{N,K}} \rho(\boldsymbol{\alpha}, q)^2 \boldsymbol{\alpha}! \sigma(s)^{2\boldsymbol{\alpha}} v_{\boldsymbol{\alpha},s}^2 \frac{\rho(\boldsymbol{\alpha}, p)^2}{\rho(\boldsymbol{\alpha}, q)^2} \\
&= \sum_{d(\boldsymbol{\alpha})\leq K, |\boldsymbol{\alpha}|>N} \rho(\boldsymbol{\alpha}, q)^2 \boldsymbol{\alpha}! \sigma(s)^{2\boldsymbol{\alpha}} v_{\boldsymbol{\alpha},s}^2 \frac{\rho(\boldsymbol{\alpha}, p)^2}{\rho(\boldsymbol{\alpha}, q)^2} + \\
&\qquad \sum_{d(\boldsymbol{\alpha})>K} \rho(\boldsymbol{\alpha}, q)^2 \boldsymbol{\alpha}! \sigma(s)^{2\boldsymbol{\alpha}} v_{\boldsymbol{\alpha},s}^2 \frac{\rho(\boldsymbol{\alpha}, p)^2}{\rho(\boldsymbol{\alpha}, q)^2} \\
&\leq \left(\max_{d(\boldsymbol{\alpha})\leq K, |\boldsymbol{\alpha}|>N} \frac{\rho(\boldsymbol{\alpha}, p)^2}{\rho(\boldsymbol{\alpha}, q)^2} \right) \|Q_s^K v - Q_s^{N,K} v\|_{q;\rho,s}^2 + \\
&\qquad \left(\max_{d(\boldsymbol{\alpha})>K} \frac{\rho(\boldsymbol{\alpha}, p)^2}{\rho(\boldsymbol{\alpha}, q)^2} \right) \|v - Q_s^K v\|_{q;\rho,s}^2 \\
&\leq M_1^2 \|Q_s^K v - Q_s^{N,K} v\|_{q;\rho,s}^2 + M_2^2 \|v - Q_s^K v\|_{q;\rho,s}^2,
\end{aligned}$$

where M_1 and M_2 are defined in (9.25) and (9.26), respectively. ∎

Corollary 9.13 *Assume that $s < \frac{\lambda_1^{2\theta}}{2}$. Then for all $v \in (\mathcal{S}_q)_{\rho,s}$ and $p < q$ we have*

$$\|v - Q_s^{N,K} v\|_{p;\rho,s} \leq \max\{M_1, M_2\} \|v\|_{q;\rho,s}$$

with M_1 and M_2 defined in (9.25) and (9.26), respectively.

If $v \in (\mathcal{S}_q)_{\rho,s}$ is of finite dimensional noise type, i.e., is such that

$$v = \sum_{d(\boldsymbol{\alpha})\leq K} v_{\boldsymbol{\alpha},s} H_{\sigma(s)^2,\boldsymbol{\alpha}}(\omega)$$

then, for all $p < q$

$$\|v - Q_s^{N,K} v\|_{p;\rho,s} \leq M_1(\rho, p, q) \|v\|_{q;\rho,s}.$$

Several examples of weights $\rho(\boldsymbol{\alpha}, p)$ can be considered. We next mention two examples.

Example 9.14 *See Cao [2006], Benth and Gjerde [1998], Holden et al. [1996], Kuo [1996] and Obata [1994]. Take $\nu \in [0, 1)$ and*

$$\rho(\boldsymbol{\alpha}, p)^2 = (\boldsymbol{\alpha}!)^\nu \lambda^{2p\boldsymbol{\alpha}}, \quad \boldsymbol{\alpha} \in \boldsymbol{\mathcal{J}}. \tag{9.27}$$

Note that we can write

$$||u||^2_{p;\rho,s} = ||\Gamma_{\otimes,\nu}(A)^p u||^2_{(L^2)_s} = \int_{\mathcal{S}'} |\Gamma_{\otimes,\nu}(A)^p u(\omega)|^2 e^{s\|\omega\|^2_{-\theta}} d\mu(\omega),$$

where $\Gamma_{\otimes,\nu}(A)$ *is the operator defined by* $\Gamma_{\otimes,\nu}(A)H_{\sigma^2,\boldsymbol{\alpha}} = (\alpha!)^\nu \lambda^{\boldsymbol{\alpha}} H_{\sigma^2,\boldsymbol{\alpha}}$. *Note also that* $\Gamma_{\otimes,0}(A^p) = \Gamma_{\otimes,0}(A)^p$. *In the case of* $\nu = 0$ *and* $s = 0$, $\Gamma_{\otimes,0}(A)$ *is called the Second Quantization of A; Hida et al. [1993].*

Corollary 9.15 *Assume that* $s < \frac{\lambda_1^{2\theta}}{2}$ *and consider* ρ *defined in (9.27). Then for every* $p < q$ *we have*

$$\|v - Q_s^{N,K} v\|_{p;\rho,s} \leq \max\left\{\frac{1}{\lambda_1^{N+1}}, \frac{1}{\lambda_{K+1}}\right\}^{q-p} \|v\|_{q;\rho,s}.$$

Proof. Recalling the fact $1 < \lambda_1 \leq \lambda_2 \leq \ldots$, we see that for all $q > p$,

$$M_1(\rho,p,q) = \max_{d(\boldsymbol{\alpha})\leq K, |\boldsymbol{\alpha}|>N} \prod_{i=1}^{d(\boldsymbol{\alpha})} \frac{1}{\lambda_i^{\alpha_i(q-p)}} = \frac{1}{\lambda_1^{(N+1)(q-p)}}$$

and

$$M_2(\rho,p,q) = \max_{d(\boldsymbol{\alpha})>K} \prod_{i=1}^{d(\boldsymbol{\alpha})} \frac{1}{\lambda_i^{\alpha_i(q-p)}} = \frac{1}{\lambda_{K+1}^{(q-p)}},$$

and the lemma follows. ∎

Remark 9.16 *We note that Corollary 9.15 is valid for any choice of the sequence* $\{\lambda_j\}_{j=1}^\infty$ *with* $1 < \lambda_1 \leq \lambda_2 \leq \ldots$ *such that* $\sum_{j=1}^\infty \lambda_j^{-2\theta} < \infty$*; see also Remark 9.23 and Section 9.7.1. For instance, Corollary 9.15 applied to the sequence* $\{\lambda_j = 2j\}_{j=1}^\infty$ *with* $\theta = 1$ *(*$s < 2$*) gives for all* $p \in \mathbb{R}$ *and* $t > 0$,

$$\|v - Q_s^{N,K} v\|_{p;\rho,s} \leq \frac{1}{2^t} \max\left\{\frac{1}{2^{tN}}, \frac{1}{(K+1)^t}\right\} \|v\|_{p+t;\rho,s}.$$

Benth and Gjerde [1998] and Cao [2006] consider the norms of Example 9.14 with weight ρ *defined in (9.27) for the special case* $s = 0$ *and the sequence* $\{\lambda_j = 2j\}_{j=1}^\infty$. *They derive approximation estimates valid only for* $p < 0$ *and* $t > 1$. *Using our notation, Theorem 2 in Cao [2006], which substantially improves the result of Benth and Gjerde [1998], reads as follows: Let* $p < 0$ *and assume that* $t > 1$. *Then for any* $v \in (\mathcal{S})_{\rho,p+t,0}$

$$\|v - Q_s^{N,K} v\|_{p;\rho,0} \leq \sqrt{B(t)\frac{1}{2^{tN}} + A(t)\frac{1}{K^{t-1}}}\ \|v\|_{p+t;\rho,0}, \tag{9.28}$$

where

$$A(t) = e^{\frac{2}{t-1}} \frac{t}{t-1} \quad \text{and} \quad B(t) = e^{\frac{1}{2^{t-1}(t-1)}} \frac{1}{2^t(t-1)}.$$

It is easy to see that

$$\frac{1}{2^t} \max\left\{\frac{1}{2^{tN}}, \frac{1}{(K+1)^t}\right\} \leq \frac{1}{2^t\sqrt{B(t)}} \sqrt{B(t)\frac{1}{2^{tN}} + A(t)\frac{1}{K^{t-1}}}$$

and then, our estimate is sharper that the one given in Cao [2006] and moreover, the proof is simpler and is valid for all $p \in \mathbb{R}$ and $t > 0$.

Example 9.17 *See Bogachev [1998], Da Prato [2006], Hida et al. [1993], Malliavin [1995], Shigekawa [2004]. Given a multi-index $\boldsymbol{\alpha}$ we denote $\langle \boldsymbol{\alpha}, \lambda \rangle := \sum_{i=1}^{d(\boldsymbol{\alpha})} \alpha_i \lambda_i$. As an alternative to the weight ρ introduced in Example 9.14 we can define*

$$\rho(\boldsymbol{\alpha}, p)^2 = 1 + \langle \boldsymbol{\alpha}, \lambda \rangle^{2p}, \quad p > 0, \text{ and } \rho(\boldsymbol{\alpha}, 0) = 1, \quad \boldsymbol{\alpha} \in \boldsymbol{\mathcal{J}}. \tag{9.29}$$

In this case we can write

$$\begin{aligned} ||u||^2_{p;\rho,s} &= ||u||^2_{(L^2)_s} + ||\Gamma_\oplus(A)^p u||^2_{(L^2)_s} \\ &= \int_{\mathcal{S}'} \left(|u(\omega)|^2 + |\Gamma_\oplus(A)^p u(\omega)|^2 \right) e^{s\|\omega\|^2_{-\theta}} d\mu(\omega), \end{aligned}$$

where $\Gamma_\oplus(A)$ is the operator defined by $\Gamma_\oplus(A) H_{\sigma^2,\boldsymbol{\alpha}} = \langle \boldsymbol{\alpha}, \lambda \rangle H_{\sigma^2,\boldsymbol{\alpha}}$. Note also that in this case $\Gamma(A^p) \neq \Gamma(A)^p$. It is easy to see that $||\Gamma_\oplus(A)^p \cdot||^2_{(L^2)_s}$ is a seminorm in the space of function in $(L^2)_s$ with $u_{\mathbf{0}} = 0$ in its $\sigma(s)$-Fourier-Hermite expansion. This seminorm can be computed using directional derivatives in the ω variable; see Da Prato [2006].

Corollary 9.18 *Assume that $s < \frac{\lambda_1^{2\theta}}{2}$ and consider ρ defined in (9.29). Then for every $p < q$ we have*

$$\|v - Q_s^{N,K} v\|_{p;\rho,s} \leq \max\left\{ \frac{1}{1 + (N+1)\lambda_1}, \frac{1}{1 + \lambda_{K+1}} \right\}^{q-p} \|v\|_{q;\rho,s}.$$

Proof. Recalling that $1 < \lambda_1 \leq \lambda_2 \leq \ldots$, we have for all $q > p$,

$$M_1(\rho, p, q) = \max_{d(\boldsymbol{\alpha}) \leq K, |\boldsymbol{\alpha}| > N} \frac{1}{\left(1 + \sum_{i=1}^{d(\boldsymbol{\alpha})} \alpha_i \lambda_i\right)^{q-p}} = \frac{1}{(1 + (N+1)\lambda_1)^{q-p}}$$

and

$$M_2(\rho, p, q) = \max_{d(\boldsymbol{\alpha}) > K} \frac{1}{\left(1 + \sum_{i=1}^{d(\boldsymbol{\alpha})} \alpha_i \lambda_i\right)^{q-p}} = \frac{1}{(1 + \lambda_{K+1})^{q-p}}.$$

This finishes the proof. ∎

9.5 The Galerkin Approximation and a Priori Error Estimates

Recall that when $s < \frac{\lambda_1^{2\theta}}{2}$, polynomials in ω are functions in $(L^2)_s$. Let $X_0^h(D) \subset H_0^1(D)$ be the finite element space of piecewise linear functions with respect to a triangulation of D. For $N, K \in \mathbb{N}_0$ and $h > 0$ define the following discrete spaces:

$$\mathcal{X}_s^{N,K,h} := X_0^h(D) \otimes \mathcal{P}^{N,K} \subset \widehat{\mathcal{U}}_s^1 \subset \mathcal{U}_s^1 = H^1(D) \otimes (L^2)_s \tag{9.30}$$

and

$$\mathcal{Y}_s^{N,K,h} := \left\{ v : v(x,\omega) = \tilde{v}(x,\omega) e^{(s+\frac{\epsilon}{2})\|P_K \omega\|^2_{-\theta}},\ \tilde{v} \in \mathcal{X}_s^{N,K,h} \right\} \subset \widehat{\mathcal{U}}^1_{-(s+\epsilon)}, \tag{9.31}$$

where the (H-orthogonal) projection on the span$\{\eta_1, \ldots, \eta_K\}$, denoted by P_K, is defined by

$$P_K\omega := \sum_{j=1}^{K} \langle \omega, \eta_j \rangle \eta_j, \text{ for all } \omega \in \mathcal{S}'. \tag{9.32}$$

Remark 9.19 *Alternatively to Remark 9.8, instead of multiplying $\tilde{v}(x,\omega)$ by the weight $e^{(s+\frac{\epsilon}{2})\|\omega\|^2_{-\theta}}$, we have defined $\mathcal{Y}_s^{N,K,h}$ in (9.31) by multiplying by the weight $e^{(s+\frac{\epsilon}{2})\|P_K\omega\|^2_{-\theta}}$. This is done in order to avoid computations of infinite series when assembling the resulting linear system; see Section 9.6. We note also the Remark 9.8 case would require the assumption $s+\epsilon < \lambda_1^{2\theta}/2$ in order to establish the inf-sup condition.*

The discrete version of problem (9.15) is introduced as:

$$\begin{cases} \text{Find } \hat{u}_s^{N,K,h} \in \mathcal{X}_s^{N,K,h} \text{ such that} \\ a(\hat{u}_s^{N,K,h}, v) = \langle f, v \rangle \text{ for all } v \in \mathcal{Y}_s^{N,K,h}. \end{cases} \tag{9.33}$$

Remark 9.20 *Observe that the above Galerkin approximation $\hat{u}_s^{N,K,h}$ satisfies a variational equation with the original permeability κ defined in (9.6).*

Since functions in $\mathcal{X}_s^{N,K,h}$ depend only on $\langle \omega, \eta_j \rangle$, $j = 1, \ldots, K$, and not on $\langle \omega, \eta_j \rangle$, $j = K+1, \ldots$, then for all $u \in \mathcal{X}_s^{N,K,h}$ and $v \in \mathcal{Y}_s^{N,K,h}$ (i.e., $v(x,\omega) = \tilde{v}(x,\omega)e^{(s+\frac{\epsilon}{2})\|P_K\omega\|^2_{-\theta}}$ with $\tilde{v} \in \mathcal{X}_s^{N,K,h}$ and P_K defined in (9.32)) we have

$$\begin{aligned} a(u,v) &= \int_{D\times\mathcal{S}'} e^{\langle \omega, \phi_x \rangle} \nabla u(x,\omega) \nabla \tilde{v}(x,\omega) e^{(s+\frac{\epsilon}{2})\|P_K\omega\|^2_{-\theta}} dx d\mu \\ &= \int_{D\times\mathcal{S}'} e^{\sum_{j=1}^{\infty} (\phi_x, \eta_j)\langle \omega, \eta_j \rangle} \nabla u(x,\omega) \nabla \tilde{v}(x,\omega) e^{(s+\frac{\epsilon}{2})\|P_K\omega\|^2_{-\theta}} dx d\mu \\ &= \int_D e^{\frac{1}{2}\sum_{j=K+1}^{\infty} a_j(x)^2} \int_{\mathcal{S}'} e^{\sum_{j=1}^{K} a_j(x)\langle \cdot, \eta_j \rangle} \nabla u \nabla \tilde{v} e^{(s+\frac{\epsilon}{2})\|P_K\cdot\|^2_{-\theta}} d\mu dx \\ &= \int_D e^{\frac{1}{2}(\|\phi_x\|_H^2 - \sum_{j=1}^{K} a_j(x)^2)} \int_{\mathcal{S}'} e^{\sum_{j=1}^{K} a_j(x)\langle \cdot, \eta_j \rangle} \nabla u \nabla \tilde{v} e^{(s+\frac{\epsilon}{2})\|P_K\cdot\|^2_{-\theta}} d\mu dx, \end{aligned} \tag{9.34}$$

where we have used the formula $\int_{\mathbb{R}} e^{a_j(x)y_j} \frac{1}{\sqrt{2\pi}} e^{-\frac{y_j^2}{2}} dy_j = e^{\frac{a_j(x)^2}{2}}$ and the notation

$$a_j(x) := (\phi_x, \eta_j)_H. \tag{9.35}$$

We have the following result:

Lemma 9.21 *Let $\epsilon > 0$ and $s \in \mathbb{R}$ be such that $s < \frac{\lambda_{K+1}^{2\theta}}{2}$ and $-s-\epsilon < \frac{\lambda_{K+1}^{2\theta}}{2}$. The bilinear form "a" and the spaces $(\mathcal{X}_s^{N,K,h}, \mathcal{Y}_s^{N,K,h})$ satisfy the following inf-sup condition:*

$$\inf_{u \in \mathcal{X}_s^{N,K,h}\setminus\{0\}} \sup_{v \in \mathcal{Y}_s^{N,K,h}\setminus\{0\}} \frac{a(u,v)}{|u|_{\mathcal{U}_s^1} |v|_{\mathcal{U}^1_{-(s+\epsilon)}}} \geq \frac{e^{-\frac{C_\theta^2}{2\epsilon}}}{\prod_{j=K+1}^{\infty} \sigma_j(-s-\epsilon)}. \tag{9.36}$$

Proof. Let $u \in \mathcal{X}_s^{N,K,h} \setminus \{0\}$. If $v(x,\omega) := u(x,\omega)e^{(s+\frac{\epsilon}{2})\|P_K\omega\|^2_{-\theta}}$ then $v \in \mathcal{Y}_s^{N,K,h}$

and

$$|v|^2_{\mathcal{U}^1_{-(s+\epsilon)}} = \int_{D\times\mathcal{S}'} |\nabla u(x,\omega)|^2 e^{2(s+\frac{\epsilon}{2})\|P_K\omega\|^2_{-\theta}} e^{-(s+\epsilon)\|\omega\|^2_{-\theta}} dxd\mu$$
$$= \int_{D\times\mathcal{S}'} |\nabla u(x,\omega)|^2 e^{2(s+\frac{\epsilon}{2})\|P_K\omega\|^2_{-\theta}} e^{-(s+\epsilon)(\|P_K\omega\|^2_{-\theta}+\|(I-P_K)\omega\|^2_{-\theta})} dxd\mu.$$

As in Lemma 9.2 for $-(s+\epsilon) < \frac{\lambda^{2\theta}_{K+1}}{2}$ we have

$$\int_{\mathcal{S}'} e^{-(s+\epsilon)\|(I-P_K)\omega\|^2_{-\theta}} d\mu = \prod_{j=K+1}^{\infty} \sigma_j(-s-\epsilon) < \infty.$$

Analogous computation holds for $\int_{S'} e^{s\|(I-P_K)\omega\|^2_{-\theta}} d\mu$ when $s < \frac{\lambda^{2\theta}_{K+1}}{2}$. Then,

$$\begin{aligned}
|v|^2_{\mathcal{U}^1_{-(s+\epsilon)}} &= \prod_{j=K+1}^{\infty} \sigma_j(-s-\epsilon) \int_{D\times\mathcal{S}'} |\nabla u(x,\omega)|^2 e^{s\|P_K\omega\|^2_{-\theta}} dxd\mu \\
&= \frac{\prod_{j=K+1}^{\infty} \sigma_j(-s-\epsilon)}{\prod_{j=K+1}^{\infty} \sigma_j(s)} \int_{D\times\mathcal{S}'} |\nabla u(x,\omega)|^2 e^{s\|\omega\|^2_{-\theta}} dxd\mu \\
&= \frac{\prod_{j=K+1}^{\infty} \sigma_j(-s-\epsilon)}{\prod_{j=K+1}^{\infty} \sigma_j(s)} |u|^2_{\mathcal{U}^1_s},
\end{aligned}$$

and from (9.34) and the fact that $e^{\frac{1}{2}\sum_{j=K+1}^{\infty} a_j(x)^2} \geq 1$, we have

$$\begin{aligned}
a(u,v) &\geq \int_{D\times\mathcal{S}'} e^{\sum_{j=1}^{K} a_j(x)\langle\omega,\eta_j\rangle} |\nabla u(x,\omega)|^2 e^{(s+\frac{\epsilon}{2})\|P_K\omega\|^2_{-\theta}} dxd\mu(\omega) \\
&\geq e^{-\frac{C^2_\theta}{2\epsilon}} \int_{D\times\mathcal{S}'} e^{-\frac{\epsilon}{2}|P_K\omega|^2_{-\theta}} |\nabla u(x,\omega)|^2 e^{(s+\frac{\epsilon}{2})\|P_K\omega\|^2_{-\theta}} dxd\mu(\omega) \\
&= e^{-\frac{C^2_\theta}{2\epsilon}} \int_{D\times\mathcal{S}'} |\nabla u(x,\omega)|^2 e^{s\|P_K\omega\|^2_{-\theta}} dxd\mu(\omega) \\
&= \frac{1}{\prod_{j=K+1}^{\infty} \sigma_j(s)} e^{-\frac{C^2_\theta}{2\epsilon}} |u|^2_{\mathcal{U}^1_s}.
\end{aligned}$$

Then

$$\frac{a(u,v)}{|v|_{\mathcal{U}^1_{-(s+\epsilon)}}} \geq \frac{\frac{1}{\prod_{j=K+1}^{\infty} \sigma_j(s)}}{\frac{\prod_{j=K+1}^{\infty} \sigma_j(-s-\epsilon))}{\prod_{j=K+1}^{\infty} \sigma_j(s)}} e^{-\frac{C^2_\theta}{2\epsilon}} \frac{|u|^2_{\mathcal{U}^1_s}}{|u|_{\mathcal{U}^1_s}} \geq \frac{e^{-\frac{C^2_\theta}{2\epsilon}}}{\prod_{j=K+1}^{\infty} \sigma_j(-s-\epsilon)} |u|_{\mathcal{U}^1_s}.$$

We conclude that the discrete inf-sup condition holds. ∎

Lemma 9.22 *For $v = \sum_{\boldsymbol{\alpha}\in\boldsymbol{\mathcal{J}}} v_{\boldsymbol{\alpha},s} H_{\sigma^2(s),\boldsymbol{\alpha}} \in \widehat{\mathcal{U}}^1_s \cap \mathcal{U}^m_s$ let $v^h = \sum_{\boldsymbol{\alpha}\in\boldsymbol{\mathcal{J}}} v^h_{\boldsymbol{\alpha},s} H_{\sigma^2(s),\boldsymbol{\alpha}}$ where $v^h_{\boldsymbol{\alpha},s}$ is the Clement finite element interpolation of $v_{\boldsymbol{\alpha},s}$ on the space $X^h_0(D)$. Then*

$$\|v - v^h\|_{\mathcal{U}^1_s} \leq \hat{C} h^{\ell-1} \|v\|_{\mathcal{U}^\ell_s}, \quad \ell = 1, 2,$$

with the constant $\hat{C}$ independent of s and h.

Define the tensor product spaces $\mathcal{U}^m_{p;\rho,s} := H^m(D) \otimes (\mathcal{S}_p)_{\rho,s}$ with $(\mathcal{S}_p)_{\rho,s}$ defined in (9.22) and (9.23). The tensor product norm is given by

$$||u||^2_{\mathcal{U}^m_{p;\rho,s}} := \sum_{\boldsymbol{\alpha}\in\boldsymbol{\mathcal{J}}} \rho(\boldsymbol{\alpha},p)^2 \boldsymbol{\alpha}!\sigma(s)^{2\boldsymbol{\alpha}} \|u_{\boldsymbol{\alpha},s}\|^2_{H^m(D)}, \tag{9.37}$$

and the seminorm is

$$|u|^2_{\mathcal{U}^m_{p;\rho,s}} := \sum_{\boldsymbol{\alpha}\in\boldsymbol{\mathcal{J}}} \rho(\boldsymbol{\alpha},p)^2 \boldsymbol{\alpha}!\sigma(s)^{2\boldsymbol{\alpha}} |u_{\boldsymbol{\alpha},s}|^2_{H^m(D)}, \tag{9.38}$$

where the weights $\{\rho(\boldsymbol{\alpha},p)\}$ were introduced in Section 9.4.

Remark 9.23 *As in Section 9.4 we have $\mathcal{U}^m_{0;\rho,s} = \mathcal{U}^m_s$ and $\mathcal{U}^m_{q;\rho,s} \subset \mathcal{U}^m_{p;\rho,s}$ for all $q > p$. Theorem 9.12 and and Corollaries 9.13, 9.15 and 9.18 extend trivially to the tensor product norm and seminorm defined in (9.37) and (9.38).*

Using the tensor product norm (9.37) we can easily deduce the following a priori error estimates:

Theorem 9.24 *Let $s \in \mathbb{R}$ and $\epsilon > 0$ be such that $s + 2\epsilon < \frac{\lambda_1^{2\theta}}{2}$ and $-s-\epsilon < \frac{\lambda_{K+1}^{2\theta}}{2}$. We have the following estimate:*

$$|\hat{u} - \hat{u}^{N,K,h}_s|_{\mathcal{U}^1_s} \le \left(1 + e^{\frac{C_\theta^2}{\epsilon}} \prod_{j=K+1}^{\infty} \sigma_j(-s-\epsilon)\right) \inf_{z\in\mathcal{X}^{N,K,h}_s} |\hat{u} - z|_{\mathcal{U}^1_{s+2\epsilon}}. \tag{9.39}$$

Moreover, for all $q > 0$

$$|\hat{u} - \hat{u}^{N,K,h}_s|_{\mathcal{U}^1_s} \le$$

$$C_*(s,\epsilon)\Big\{ \max\left\{M_1(\rho,0,q), M_2(\rho,0,q)\right\} |\hat{u}|_{\mathcal{U}^1_{q;\rho,s+2\epsilon}} + \hat{C}h^{\ell-1}\|\hat{u}\|_{\mathcal{U}^\ell_{s+2\epsilon}}\Big\}$$

where $C_*(s,\epsilon) = 1 + e^{\frac{C_\theta^2}{\epsilon}} \prod_{j=K+1}^{\infty} \sigma_j(-s-\epsilon)$, M_1 and M_2 are defined in Theorem 9.13, and $\hat{C}$ is the constant of Lemma 9.22.

Proof. First note that for all $v \in \mathcal{Y}^{N,K,h}_s$ we have

$$a(\hat{u} - \hat{u}^{N,K,h}_s, v) = 0$$

and therefore, for all $z \in \mathcal{X}^{N,K,h}_s$

$$a(\hat{u}^{N,K,h}_s - z, v) = a(\hat{u} - z, v).$$

Using the continuity 2. in Theorem 9.6, with $s + 2\epsilon$ instead of s, we obtain

$$a(\hat{u}^{N,K,h}_s - z, v) \le e^{\frac{C_\theta^2}{2\epsilon}} |\hat{u} - z|_{\mathcal{U}^1_{s+2\epsilon}} |v|_{\mathcal{U}^1_{-s-\epsilon}}.$$

From the discrete inf-sup of Lemma 9.21 we have

$$|\hat{u} - \hat{u}^{N,K,h}_s|_{\mathcal{U}^1_s} \le |\hat{u} - z|_{\mathcal{U}^1_s} + |\hat{u}^{N,K,h}_s - z|_{\mathcal{U}^1_s}$$

$$
\begin{aligned}
&\leq |\hat{u}-z|_{\mathcal{U}_s^1} + e^{\frac{C_\theta^2}{2\epsilon}} \prod_{j=K+1}^{\infty} \sigma_j(-s-\epsilon) \sup_{v\in\mathcal{Y}_s^{N,K,h}\setminus\{0\}} \frac{a(\hat{u}_s^{N,K,h}-z,v)}{|v|_{\mathcal{U}_{-s-\epsilon}^1}} \\
&\leq |\hat{u}-z|_{\mathcal{U}_s^1} + e^{\frac{C_\theta^2}{2\epsilon}} e^{\frac{C_\theta^2}{2\epsilon}} \prod_{j=K+1}^{\infty} \sigma_j(-s-\epsilon) |\hat{u}-z|_{\mathcal{U}_{s+2\epsilon}^1} \\
&\leq \left(1 + e^{\frac{C_\theta^2}{\epsilon}} \prod_{j=K+1}^{\infty} \sigma_j(-s-\epsilon)\right) |\hat{u}-z|_{\mathcal{U}_{s+2\epsilon}^1}
\end{aligned}
$$

which gives (9.39). We need only to bound the second term of (9.39). This can be done as follows: take the polynomial $z = \zeta_{N,K}^h \in \mathcal{X}_s^{N,K,h}$ where $\zeta_{N,K} = (Id \otimes Q_{s+2\epsilon}^{N,K})\hat{u}$. Note that since $s+2\epsilon < \frac{\lambda_1^{2\theta}}{2}$, polynomials in w are in $(L^2)_{s+2\epsilon}$ and then $\zeta_{N,K} \in \widehat{\mathcal{U}}_{s+2\epsilon}^1$ is well defined. We have

$$
|\hat{u}-\zeta_{N,K}^h|_{\mathcal{U}_{s+2\epsilon}^1} \leq |\hat{u}-\zeta_{N,K}|_{\mathcal{U}_{s+2\epsilon}^1} + |\zeta_{N,K}-\zeta_{N,K}^h|_{\mathcal{U}_{s+2\epsilon}^1}. \tag{9.40}
$$

Apply Theorem 9.13 (see Remark 9.23) with $p=0$ and $q>0$ to get

$$
\begin{aligned}
|\hat{u}-\zeta_{N,K}|_{\mathcal{U}_{s+2\epsilon}^1} &= |\hat{u}-\zeta_{N,K}|_{\mathcal{U}_{0;\rho,s+2\epsilon}^1} \\
&\leq \max\left\{M_1(\rho,0,q), M_2(\rho,0,q)\right\} |\hat{u}|_{\mathcal{U}_{q;\rho,s+2\epsilon}^1}.
\end{aligned} \tag{9.41}
$$

From Lemma 9.22 we get

$$
|\zeta_{N,K}-\zeta_{N,K}^h|_{\mathcal{U}_{s+2\epsilon}^1} \leq Ch^{\ell-1}\|\hat{u}\|_{\mathcal{U}_{s+2\epsilon}^\ell}. \tag{9.42}
$$

Inserting (9.41) and (9.42) into (9.40) we get the result. ∎

The following result follows from Corollary 9.15:

Corollary 9.25 *Consider the weight ρ defined in (9.27). Under the assumptions of Theorem 9.24 we have for all $q>0$*

$$
\begin{aligned}
&|\hat{u}-\hat{u}_s^{N,K,h}|_{\mathcal{U}_s^1} \leq \\
&\qquad C_*(s,\epsilon)\Big\{ \max\left\{\frac{1}{\lambda_1^{N+1}}, \frac{1}{\lambda_{K+1}}\right\}^q |\hat{u}|_{\mathcal{U}_{q;\rho,s+2\epsilon}^1} + \hat{C}h^{\ell-1}\|\hat{u}\|_{\mathcal{U}_{s+2\epsilon}^\ell}\Big\},
\end{aligned}
$$

where $C_*(s,\epsilon) = 1 + e^{\frac{C_\theta^2}{\epsilon}} \prod_{j=K+1}^{\infty} \sigma_j(-s-\epsilon)$ and $\hat{C}$ is the constant of Lemma 9.22.

From Corollary 9.18 we have the following a priori error estimate:

Corollary 9.26 *Consider ρ defined in (9.29). Under the assumptions of Theorem 9.24 we have for all $q>0$*

$$
\begin{aligned}
&|\hat{u}-\hat{u}_s^{N,K,h}|_{\mathcal{U}_s^1} \leq \\
&\qquad C_*(s,\epsilon)\Big\{ \max\left\{\tfrac{1}{1+(N+1)\lambda_1}, \tfrac{1}{1+\lambda_{K+1}}\right\}^q |\hat{u}|_{\mathcal{U}_{q;\rho,s+2\epsilon}^1} + \hat{C}h^{\ell-1}\|\hat{u}\|_{\mathcal{U}_{s+2\epsilon}^\ell}\Big\}.
\end{aligned}
$$

where $C_*(s,\epsilon) = 1 + e^{\frac{C_\theta^2}{\epsilon}} \prod_{j=K+1}^{\infty} \sigma_j(-s-\epsilon)$ and $\hat{C}$ is the constant of Lemma 9.22.

9.6 The Resulting Linear System

We now analyze the properties of the resulting linear system for the discrete spaces $\mathcal{X}_s^{N,K,h} \subset \widehat{\mathcal{U}}_s^1$ and $\mathcal{Y}_s^{N,K,h} \subset \widehat{\mathcal{U}}_{-(s+\epsilon)}^1$ defined in (9.30) and (9.31), respectively.

From (9.34), we see that for all functions $u \in \mathcal{X}_s^{N,K,h}$ and $v \in \mathcal{Y}_s^{N,K,h}$, i.e., $v(x,\omega) = \tilde{v}(x,\omega)e^{(s+\frac{\epsilon}{2})\|P_K\omega\|_{-\theta}^2}$ with $\tilde{v} \in \mathcal{X}_s^{N,K,h}$ and P_K defined in (9.32), we have

$$a(u,v) = \int_D b_K(x) \int_{\mathcal{S}'} e^{\sum_{j=1}^K a_j(x)\langle\omega,\eta_j\rangle} \nabla u(x,\omega)\nabla\tilde{v}(x,\omega)e^{(s+\frac{\epsilon}{2})\|P_K\omega\|_{-\theta}^2} d\mu dx, \quad (9.43)$$

where

$$b_K(x) := e^{\frac{1}{2}(\|\phi_x\|_H^2 - \sum_{j=1}^K a_j(x)^2)}. \quad (9.44)$$

For every $u \in \mathcal{X}_s^{N,K,h}$ we introduce the function $\underline{u} : D \times \mathbb{R}^K \to \mathbb{R}$ such that

$$\underline{u}(x, \langle\omega,\eta_1\rangle, \ldots, \langle\omega,\eta_j\rangle) = u(x,\omega), \text{ for all } \omega \in \mathcal{S}',$$

and we denote $\underline{\mathcal{X}}_s^{N,K,h} := \{\underline{u} : \quad u \in \mathcal{X}_s^{N,K,h}\}$. We also introduce

$$\kappa_K(x,y) := e^{\sum_{j=1}^K a_j(x)y_j}, \qquad D_K := \operatorname{diag}(\lambda_1^{-\theta}, \ldots, \lambda_K^{-\theta})$$

and define the bilinear form $\underline{a}$ by

$$\underline{a}(\underline{u},\underline{\tilde{v}}) := a(u,v), \text{ for all } \underline{u},\underline{\tilde{v}} \in \underline{\mathcal{X}}_s^{N,K,h}. \quad (9.45)$$

Here $v(x,\omega) = \tilde{v}(x,\omega)e^{(s+\frac{\epsilon}{2})\|P_K\omega\|_{-\theta}^2}$. With this notation and using (9.43) we have

$$\underline{a}(\underline{u},\underline{\tilde{v}}) = \int_D b_K(x) \int_{\mathbb{R}^K} \kappa_K(x,y)\nabla_x\underline{u}(x,y)\nabla_x\underline{\tilde{v}}(x,y)e^{(s+\frac{\epsilon}{2})|D_K y|^2} d\mu_K(y)dx, \quad (9.46)$$

where μ_K denotes the standard Gaussian measure in $\mathbb{R}^K$.

To simplify notation we set $\check{\sigma} = \sigma(s+\frac{\epsilon}{2})$, i.e., let $\check{\sigma}_j = \sigma_j(s+\frac{\epsilon}{2})$, $j = 1,2,\ldots$. Let $\{\psi_\ell\}_{\ell=1}^L$ be the standard hat basis functions for $X_0^h(D)$, then, the collection

$$\left\{\underline{\Psi}_{\ell,\check{\sigma}^2,\boldsymbol{\alpha}} : \quad \underline{\Psi}_{\ell,\check{\sigma}^2,\boldsymbol{\alpha}}(x,y) = \psi_\ell(x)\underline{H}_{\check{\sigma}^2,\boldsymbol{\alpha}}(y), \quad \ell = 1,\ldots,L; \quad \boldsymbol{\alpha} \in \boldsymbol{\mathcal{J}}^{N,K}\right\}$$

is a basis of $\underline{\mathcal{X}}_s^{N,K,h}$. Denote by $\{A_{(\ell,\boldsymbol{\alpha}),(m,\boldsymbol{\beta})}\}$ the matrix associated to the bilinear form $\underline{a}$ defined in (9.45). From (9.46) we have

$$\begin{aligned} A_{(\ell,\boldsymbol{\alpha}),(m,\boldsymbol{\beta})} &= \\ &= \int_D b_K(x) \int_{\mathbb{R}^K} \kappa_K(x,y)\underline{\Psi}_{\ell,\check{\sigma}^2,\boldsymbol{\alpha}}(x,y)\underline{\Psi}_{m,\check{\sigma}^2,\boldsymbol{\beta}}(x,y)e^{(s+\frac{\epsilon}{2})|D_K y|^2} d\mu_K(y)dx \quad (9.47) \\ &= \int_D b_K(x)\kappa^\star_{K,\boldsymbol{\alpha},\boldsymbol{\beta}}(x)\nabla\psi_\ell(x)\nabla\psi_k(x)dx, \quad (9.48) \end{aligned}$$

where we have defined

$$\kappa^\star_{K,\boldsymbol{\alpha},\boldsymbol{\beta}}(x) := \int_{\mathbb{R}^K} \kappa_K(x,y)\underline{H}_{\check{\sigma}^2,\boldsymbol{\alpha}}(y)\underline{H}_{\check{\sigma}^2,\boldsymbol{\beta}}(y)e^{(s+\frac{\epsilon}{2})|D_K y|^2} d\mu_K(y). \quad (9.49)$$

Now we compute the integral in (9.49). From the definition of the $\check{\sigma}$-Fourier-Hermite polynomials we see that

$$\begin{aligned}\kappa^{\star}_{K,\boldsymbol{\alpha},\boldsymbol{\beta}}(x) &= \prod_{j=1}^{K}\int_{\mathbb{R}} h_{\check{\sigma}_j^2,\alpha_j}(y_j)h_{\check{\sigma}_j^2,\beta_j}(y_j)e^{a_j(x)y_j}e^{(s+\frac{\epsilon}{2})\lambda_j^{-2\theta}y_j^2}d\mu_1\\ &= \prod_{j=1}^{K}\check{\sigma}_j\int_{\mathbb{R}} h_{\check{\sigma}_j^2,\alpha_j}(y_j)h_{\check{\sigma}_j^2,\beta_j}(y_j)e^{a_j(x)y_j}\frac{e^{-\frac{1}{2\check{\sigma}_j^2}y_j^2}}{\sqrt{2\pi}\check{\sigma}_j}dy_j.\\ &= \prod_{j=1}^{K}\check{\sigma}_j\kappa^{\star}(x;\alpha_j,\beta_j), \end{aligned} \tag{9.50}$$

where

$$\kappa^{\star}(x;\alpha_j,\beta_j) := \int_{\mathbb{R}} h_{\check{\sigma}_j^2,\alpha_j}(y_j)h_{\check{\sigma}_j^2,\beta_j}(y_j)e^{a_j(x)y_j}\frac{e^{-\frac{1}{2\check{\sigma}_j^2}y_j^2}}{\sqrt{2\pi}\check{\sigma}_j}dy_j.$$

From the generating function identity (9.20) one can easily deduce

$$h_{\check{\sigma}_j^2,\alpha_j}(t)h_{\check{\sigma}_j^2,\beta_j}(t) = \sum_{m=0}^{\min\{\alpha_j,\beta_j\}} m!\binom{\alpha_j}{m}\binom{\beta_j}{m}\check{\sigma}_j^{2m}h_{\check{\sigma}_j^2,\alpha_j+\beta_j-2m}(t)$$

and

$$e^{a_j(x)t} = e^{\frac{1}{2}\check{\sigma}_j^2a_j(x)^2}\sum_{m=0}^{\infty}\frac{1}{m!}a_j(x)^m h_{\check{\sigma}_j^2,m}(t).$$

Then, we have

$$\kappa^{\star}(x;\alpha_j,\beta_j) = e^{\frac{1}{2}\check{\sigma}_j^2a_j(x)^2}\sum_{m=0}^{\min\{\alpha_j,\beta_j\}} m!\binom{\alpha_j}{m}\binom{\beta_j}{m}a_j(x)^{\alpha_j+\beta_j-2m}\check{\sigma}_j^{2(\alpha_j+\beta_j-m)}. \tag{9.51}$$

Summarizing, we have that $A_{(\ell,\boldsymbol{\alpha}),(m,\boldsymbol{\beta})}$ defined in (9.47) can be easily computed using (9.48). We only need to compute b_K defined in (9.44) and $\kappa^{\star}_K$ defined in (9.49). The later computation is reduced to the finite product in (9.50) where each factor is given by (9.51).

Denote by $\{g_{\ell,\boldsymbol{\alpha}}\}$ the load vector. Each entry is given by

$$g_{\ell,\boldsymbol{\alpha}} = \int_{D\times\mathcal{S}'} f(x,\omega)\psi_\ell(x)\prod_{j=1}^{d(\boldsymbol{\alpha})} h_{\check{\sigma}_j^2,\alpha_j}(\langle\omega,\eta_j\rangle)e^{(s+\frac{\epsilon}{2})\|P_K\omega\|^2_{-\theta}}dxd\mu(\omega).$$

The integral with respect the ω variable is exactly the computation of the $\boldsymbol{\alpha}-$th coefficient of the expansion of $f(x,\cdot)$ in terms of $\sigma(s+\frac{\epsilon}{2})$-Fourier-Hermite polynomials. In particular if f does not depend on ω we have that $g_{\ell,\boldsymbol{\alpha}} = 0$ when $\boldsymbol{\alpha}\neq\mathbf{0}$.

Remark 9.27 *It is easy to see that the matrix $\{A_{(\ell,\boldsymbol{\alpha}),(m,\boldsymbol{\beta})}\}$ associated to the bilinear form $\underline{a}$ is symmetric and positive definite. It is a block square matrix of*

dimension $\binom{K+N}{K}$ *where each block* $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ *is the usual finite element matrix of the discretization of a elliptic equation with coefficient given by* $b_K(x)\kappa_K^\star(x; \boldsymbol{\alpha}, \boldsymbol{\beta})$ *where* b_K *is defined in (9.44) and* $\kappa_K^\star$ *defined in (9.49) is computed using (9.50) and (9.51). This corresponds to a discretization of a coupled system of elliptic equations.*

Remark 9.28 *Recall that we have set* $\check{\sigma}_j = \sigma_j(s + \frac{\epsilon}{2})$. *The coefficients computed after solving the resulting linear system are the coefficients of the approximated solution in terms of* $\sigma(s + \frac{\epsilon}{2})$*-Fourier-Hermite polynomials. We can change these coefficients to the coefficients of the solution in terms of the* $\sigma(s)$*-Fourier-Hermite polynomials using the following formula easily deduced from the generating function identity (9.20):*

$$\widehat{H}_{\check{\sigma}^2,\boldsymbol{\alpha}}(\omega) = \sum_{\boldsymbol{\gamma} \leq \boldsymbol{\alpha}/2} \frac{\boldsymbol{\alpha}!}{\boldsymbol{\gamma}!(\boldsymbol{\alpha} - 2\boldsymbol{\gamma})!} \left(\frac{\sigma^2 - \check{\sigma}^2}{2} \right)^{\boldsymbol{\gamma}} \widehat{H}_{\sigma^2,\boldsymbol{\alpha}-2\boldsymbol{\gamma}}(\omega)$$

where

$$\widehat{H}_{\sigma^2,\boldsymbol{\alpha}}(\omega) := \prod_{j=1}^{d(\boldsymbol{\alpha})} h_{\sigma_j^2,\alpha_j}(\langle \omega, \eta_j \rangle); \quad \omega \in \mathcal{S}'.$$

This is a postprocessing step. Note that this formula can also be used to deduce the expansion of the right-hand side term f *in* $\sigma(s + \frac{\epsilon}{2})$*-Fourier-Hermite polynomials from the standard Wiener-Chaos expansion (i.e., in in terms of Fourier-Hermite polynomials).*

9.7 On the Choice of H, A and ϕ_x

Several choices for the Hilbert space H and the operator A are possible. We mention three possible choices of (H, A, ϕ_x). We will first review some known results:

9.7.1 Known Results

The Schwartz Space and the Operator $-\frac{d^2}{dx^2} + x^2 + 1$

Consider the densely defined differential operator

$$A_1 = -\frac{d^2}{dx^2} + x^2 + 1. \tag{9.52}$$

We have an $L^2(\mathbb{R})$ orthonormal system of eigenfunctions of A_1 which are the Hermite functions

$$e_n(x) := \frac{1}{\sqrt{\sqrt{\pi}(n-1)!}} e^{-\frac{1}{2}x^2} h_{n-1}(\sqrt{2}x), \quad n = 1, 2, \ldots, \tag{9.53}$$

where h_n is the nth degree Hermite polynomial. We have $A_1 e_n = (2n)e_n$, $n = 1, 2, \ldots$.

The family of tensors products

$$e_{\boldsymbol{n}} := e_{(n_1,\ldots,n_d)} := e_{n_1} \otimes \ldots \otimes e_{n_d}, \quad \boldsymbol{n} = (n_1, \ldots, n_d) \in \mathbb{N}^d$$

form an orthonormal basis for $L^2(\mathbb{R}^d)$. Let $\boldsymbol{n}^{(j)} = (n_1^{(j)}, \ldots, n_d^{(j)})$ be the j-th multi-index in some fixed ordering of all d-dimensional multi-indices $\boldsymbol{n} = (n_1, \ldots, n_d) \in \mathbb{N}^d$. We assume that this ordering has the property that

$$i < j \Longrightarrow \prod_{k=1}^{d} (2n_k^{(i)}) \leq \prod_{k=1}^{d} (2n_k^{(j)}). \tag{9.54}$$

Now define

$$\eta_j := e_{\boldsymbol{n}^{(j)}} = e_{n_1^{(j)}} \otimes \ldots \otimes e_{n_d^{(j)}}, \ j = 1, 2, \ldots. \tag{9.55}$$

We have $A_1^{\otimes d} \eta_j = \lambda_j \eta_j$ where

$$\lambda_j := \prod_{k=1}^{d} (2n_k^{(j)}), j = 1, 2, \ldots. \tag{9.56}$$

Note that $\boldsymbol{n}^{(1)} = (1, \ldots, 1) \in \mathbb{R}^d$, $\lambda_1 = 2^d$ and that $1 < \lambda_1 \leq \lambda_2 \leq \ldots$.

For the next result, see [Holden et al.- 1996, Lemma 2.3.3].

Lemma 9.29 (Zhang) *With $d(\boldsymbol{\alpha})$ defined in (9.19) we have that*

$$\sum_{\boldsymbol{\alpha} \in \boldsymbol{\mathcal{J}}} \prod_{k=1}^{d(\boldsymbol{\alpha})} (2k)^{-q\alpha_k} < \infty$$

if and only if $q > 1$.

Corollary 9.30 *For $\{\lambda_j\}_{j=1}^{\infty}$ defined in (9.56) we have that $\sum_{j=1}^{\infty} \lambda_j^{-q} < \infty$ for all $q > 1$.*

Proof. For $\boldsymbol{\alpha} \in \boldsymbol{\mathcal{J}}$ define $N_Z(\boldsymbol{\alpha}) = \#\{j : \alpha_j \neq 0\}$. Observe that

$$\sum_{j=1}^{\infty} \lambda_j^{-p} = \sum_{\boldsymbol{n} \in \mathbb{N}^d} \prod_{k=1}^{d} (2n_k)^{-q} = \sum_{N_Z(\boldsymbol{\alpha}) \leq d} \prod_{k=1}^{d(\boldsymbol{\alpha})} (2k)^{-q\alpha_k} < \sum_{\boldsymbol{\alpha} \in \boldsymbol{\mathcal{J}}} \prod_{k=1}^{d(\boldsymbol{\alpha})} (2k)^{-q\alpha_k} < \infty,$$

where $d(\boldsymbol{\alpha})$ is defined in (9.19) and d is the dimension of $\mathbb{R}^d$. ∎

The covariance integral operator on $L^2(D)$ and the Mercer's theorem

Consider the covariance operator $Q : L^2(D) \to L^2(D)$ associated to $W(x, \omega)$, i.e., the integral operator with kernel the symmetric positive function $C(x, \hat{x}) = E_\mu W(x, \cdot) W(\hat{x}, \cdot)$. This is a compact operator in $L^2(D)$ when its kernel is square integrable and symmetric. We denote by $\{\mu_j\}_{j=1}^{\infty}$ and $\{\zeta_j\}_{j=1}^{\infty}$ the eigenvalues and eigenfunction of Q. We have $\sum_j \mu_j^2 < \infty$.

We recall that from Mercer's theorem (see Riesz and Sz.-Nagy [1990]) we can write

$$C(x,\hat{x}) = \sum_{j=1}^{\infty} \mu_j \zeta_j(x)\zeta_j(\hat{x}). \tag{9.57}$$

For results on the decay of the eigenvalues, see [Frauenfelder et al.- 2005, Proposition 2.3, 2.5 and 2.6].

9.7.2 Three modeling choices

With the notation and review of Section 9.7.1 we now mention three modeling choices. The use of one or another depends on the information and computations available for the problem.

First choice

The first modeling choice we propose is:

1. The Hilbert space $H = L^2(\mathbb{R}^d)$.
2. The operator $A = A_1^{\otimes d}$ with the sequence $\{\lambda_j\}_{j=1}^{\infty}$ in (9.56) and the eigenfunctions $\{\eta_j\}_{j=1}^{\infty}$ in (9.55).
3. For all $x \in D$ we define $\phi_x(\hat{x}) := \phi(\hat{x} - x)$, $\hat{x} \in \mathbb{R}^d$, where $\phi \in L^2(\mathbb{R}^d)$.

From Corollary 9.30 we can take any $\theta > \frac{1}{2}$ (independent of the dimension d). In order to verify the assumption of Theorem 9.6 it is enough to take $\phi \in \mathcal{S}_\theta$. The functions a_j defined in (9.35) are $a_j(x) = (\phi(\cdot - x), \eta_j)$ and in general can be computed using numerical integration. In explicit applications the test function or *window* ϕ can be chosen such that the diameter of the support of ϕ is the maximal distance within which $W_\phi(x_1, \cdot)$ and $W_\phi(x_2, \cdot)$ are correlated; see Holden et al. [1996]. We also recall that the map $x \mapsto \phi_x$ may be chosen to match covariance function; see Roman and Sarkis [2006].

It is easy to see that in this case, $\mathcal{S} = \mathcal{S}(\mathbb{R}^d)$ is the Schwartz space of rapidly decreasing functions and then $\mathcal{S}' = \mathcal{S}'(\mathbb{R}^d)$ is the space of *tempered distributions*. The triplet $(\mathcal{S}'(\mathbb{R}^d), \mathcal{B}(\mathcal{S}'(\mathbb{R}^d)), \mu)$ is called 1-*dimensional white noise probability space*. The smoothed white noise of Definition 9.3 is called the 1-dimensional d-parameter smoothed white noise. In Section 9.8 we present a numerical experiment using this setup for the case of $d = 1$.

Second choice

The second modeling choice we propose is:

1. The Hilbert space $H = L^2(\mathbb{R}^d)$.
2. The operator $A = A_1^{\otimes d}$ with the sequence $\{\lambda_j\}_{j=1}^{\infty}$ in (9.56) and the eigenfunctions $\{\eta_j\}_{j=1}^{\infty}$ in (9.55).
3. For all $x \in D$, $\phi_x = \sum_j^{\infty} a_j(x)\eta_j$ where $a_j(x) := \sqrt{\mu_j}\zeta_j(x)$, $j = 1, \ldots,$ where $\{\mu_j\}_{j=1}^{\infty}$ and $\{\zeta_j\}_{j=1}^{\infty}$ are the eigenvalues and eigenfunctions of the operator Q introduced in Section 9.7.1.

From (9.57) we see that $C(x,\hat{x}) = \sum_{j=1}^{\infty} a_j(x)a_j(\hat{x}) = (\phi_x, \phi_{\hat{x}})_H$, $x, \hat{x} \in D$. Note that

$$\|\phi_x\|_\theta^2 = \sum_{j=1}^{\infty} \lambda_j^{2\theta} \mu_j \zeta_j(x)^2, \text{ for all } x \in D.$$

The assumption $\phi_x \in \mathcal{S}_\theta$ for all $x \in D$ must be checked in each case, and the convergence of the series depends on the decay of the eigenvalues and the $L^\infty(D)$-norm of the eigenfunction which in turns depend on the regularity of the function $C(x,\hat{x})$; see [Frauenfelder et al.- 2005, Proposition 2.3, 2.5 and 2.6]. For the numerical computation of the eigenfunctions and eigenvalues of Q, **?**] and also Frauenfelder et al. [2005]. We also note that ϕ_x defined in 3. above can be used with any choice of Hilbert space H and operator A. A possible choice is the one describe next and which can be viewed as a generalization of the Karhúnen-Loève expansion.

Third choice

The third modeling choice we propose is:

1. The Hilbert space $H = L^2(D)$.
2. $A = Q^{-1}$ with $\lambda_j := \frac{1}{\mu_j}$, $j = 1, 2, \ldots$, and $\eta_j := \zeta_j$, $j = 1, 2, \ldots$, where the μ_j and ζ_j are the eigenvalues and eigenfunctions of the integral operator Q.
3. For all $x \in D$, $\phi_x = \sum_j^{\infty} a_j(x)\eta_j$ where $a_j(x) := \sqrt{\mu_j}\zeta_j(x)$, $j = 1, \ldots$.

According to Section 9.7.1 we can take any $\theta \geq 1$. In this case the expansion of $W(x,\omega)$ in terms of $\sigma(s)$-Fourier-Hermite polynomials coincides with its Karhúnen-Loève expansion for the case $s = 0$. We mention that in order to make calculations, such as writing the expansion of the right-hand side $f(x,\omega)$ in terms of $\sigma(s)$-Fourier-Hermite polynomials, we need to know the eigenfunctions of Q. The assumption $\phi_x \in \mathcal{S}_\theta$ for all $x \in D$ must be checked for each particular problem. Observe also that $\|\phi_x\|_\theta^2 = \sum_{j=1}^{\infty} \frac{1}{\mu_j^{2\theta-1}} \zeta_j(x)^2$ for all $x \in D$.

9.8 Numerical Experiments

In this section we present numerical experiments with $D = [0,1]$, $H = L^2(\mathbb{R})$, and $A = A_1$ defined in (9.52) with $\theta = 1$. In this case the Lemma 9.2 becomes

$$\int_{\mathcal{S}'} e^{s\|\omega\|_{-\theta}^2} d\mu(\omega) = \begin{cases} \left(\frac{\sqrt{2}}{\pi\sqrt{-s}} \sinh\left(\frac{\pi\sqrt{-s}}{\sqrt{2}}\right)\right)^{-\frac{1}{2}}, & s < 0 \\ 1, & s = 0 \\ \left(\frac{\sqrt{2}}{\pi\sqrt{s}} \sin\left(\frac{\pi\sqrt{s}}{\sqrt{2}}\right)\right)^{-\frac{1}{2}}, & 0 < s < 2 \\ +\infty, & s \geq 2, \end{cases}$$

therefore, we can construct the general $\sigma(s)$-Fourier-Hermite polynomials for $s < 2$; see Theorem 9.10. We consider the modeling choice described in Section 9.7.2, i.e., $\phi_x(\cdot) = \phi(\cdot - x)$. To avoid numerical integration errors in the computation of the functions $a_j(x)$ in (9.35), we choose the function ϕ as

$$\phi(x) = e^{-\frac{1}{2}x^2}. \tag{9.58}$$

In order to compute the discretization errors, let $\hat{u}$ and f be given by

$$\hat{u}(x,\omega) = \frac{x(1-x)}{2} e^{-\langle \omega, \phi_x \rangle} \tag{9.59}$$

and

$$\begin{aligned} f(x,\omega) &= 1 - \frac{1-2x}{2}\langle \omega, \phi_x' \rangle + \frac{x(1-x)}{2}\langle \omega, \phi_x'' \rangle & (9.60)\\ &= 1 + \sum_{j=0}^{\infty} \left(\frac{x(1-x)}{2} a_j'(x) \right)' \langle \omega, \eta_j \rangle. & (9.61) \end{aligned}$$

It is easy to see that $\hat{u}$ in (9.59) is the exact solution of the problem (9.5) or (9.15) with right-hand side f given by (9.60).

It is easy to see by using the generating function identity (9.20) and direct calculations that the following results hold:

Lemma 9.31 *For ϕ, $\hat{u}$ and f defined in (9.58), (9.59) and (9.60), respectively, we have*

1. $\|\phi\|^2_{L^2(\mathbb{R})} = \sqrt{\pi}$ *and* $\|\phi'\|^2_{L^2(\mathbb{R})} = \frac{\sqrt{\pi}}{2}$.
2. $a_j(x) := (\phi_x, \eta_j) = \sqrt{\frac{\sqrt{\pi}}{2^{j-1}(j-1)!}} x^{j-1} e^{-\frac{1}{4}x^2}$.
3. *The process f belongs to $\mathcal{U}_s^{-1}$ for all $s < 2$.*
4. $\hat{u}(x,\omega) = \sum_{\boldsymbol{\alpha}\in\boldsymbol{\mathcal{J}}} \hat{u}_{\boldsymbol{\alpha}}(x) H_{\boldsymbol{\alpha}}(\omega) \in \widehat{\mathcal{U}}_s^1$ *for all $s \in \mathbb{R}$ where*

$$\hat{u}_{\boldsymbol{\alpha}}(x) = e^{\frac{1}{2}\|\phi\|^2_{L^2(\mathbb{R})}} \frac{x(1-x)a_{\boldsymbol{\alpha}}(x)}{2\boldsymbol{\alpha}!} \text{ with } a_{\boldsymbol{\alpha}}(x) = \prod_{j=1}^{d(\boldsymbol{\alpha})} a_j(x)^{\alpha_j}.$$

5. $|\hat{u}|^2_{\mathcal{U}_0^1} = \left(\frac{1}{12} + \frac{\|\phi'\|^2_{L^2(\mathbb{R})}}{120} \right) e^{2\|\phi\|^2_{L^2(\mathbb{R})}}$.

Throughtout this section, we solve the discrete problem (9.33) with $s = 0$ to obtain

$$\hat{u}^{N,K,h}(x,\omega) = \sum_{\boldsymbol{\alpha}\in\boldsymbol{\mathcal{J}}^{N,K}} \hat{u}^{N,K,h}_{\boldsymbol{\alpha}}(x) H_{\boldsymbol{\alpha}}(\omega).$$

In Figure 9.1 we show the approximation $\hat{u}^{N,K,h}_{\boldsymbol{\alpha}}(x) - \hat{u}_{\boldsymbol{\alpha}}(x)$ corresponding to $\boldsymbol{\alpha} = (0,0,\ldots)$ for $K = N = k$ with $k = 0,1,2,3,4$, $h = \frac{1}{4}$ and $\epsilon = 0, \frac{1}{2}$. For these small values of k (and $1/h$) we already observe that $\epsilon = \frac{1}{2}$ gives a slightly better approximation than $\epsilon = 0$. Here we can observe the fast convergence of the computed coefficient to the exact one. A similar behavior is also observed for $\boldsymbol{\alpha} = (1,0,\ldots)$; see Figure 9.2.

In Tables 9.1 and 9.2 we show the seminorm error $|\hat{u} - \hat{u}^{N,K,h}|_{\mathcal{U}_0^1}$ for $\epsilon = \frac{1}{2}$ and $\epsilon = 0$, respectively, and take $h = \frac{1}{32}$. We recall that this seminorm involves the computation of $|\hat{u}_{\boldsymbol{\alpha}} - \hat{u}^{N,K,h}_{\boldsymbol{\alpha}}|^2_{H^1(D)}$ for all $\boldsymbol{\alpha} \in \boldsymbol{\mathcal{J}}^{N,K}$, and the computation of $|\hat{u}_{\boldsymbol{\alpha}}|^2_{H^1(D)}$ for all $\boldsymbol{\alpha} \in \boldsymbol{\mathcal{J}} \backslash \boldsymbol{\mathcal{J}}^{N,K}$. In these tables we see clearly the decay of the errors with respect to N and K.

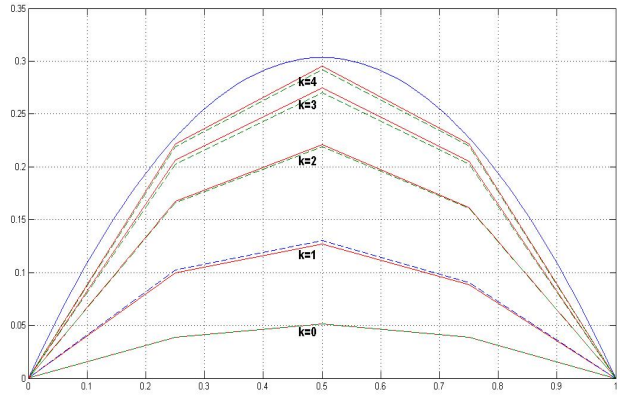

Figure 9.1: Approximation of the coefficient $u_{(0,0,0,\ldots)}$ for $K = N = k$, $h = \frac{1}{4}$ and $\epsilon = \frac{1}{2}$ (solid line) and $\epsilon = 0$ (dashed line).

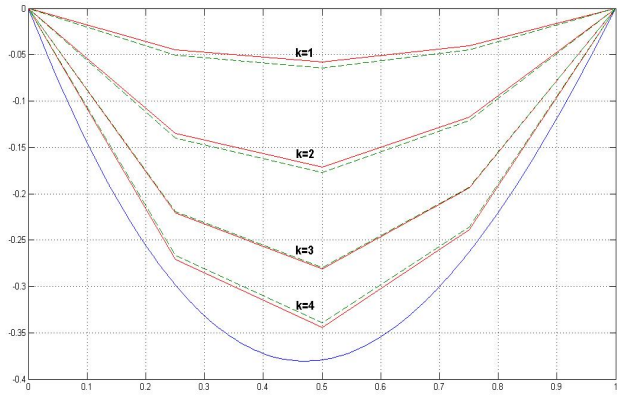

Figure 9.2: Approximation of the coefficient $u_{(1,0,0,\ldots)}$ for $K = N = k$, $h = \frac{1}{4}$ and $\epsilon = \frac{1}{2}$ (solid line) and $\epsilon = 0$ (dashed line).

We now analyze further the previous decay. Define the energy norm

$$|\hat{u} - \hat{u}^{N,K,h}|_a := a(\hat{u} - \hat{u}^{N,K,h}, \hat{u} - \hat{u}^{N,K,h})^{\frac{1}{2}}.$$

In Table 9.3 we present errors in the seminorm $|\cdot|_{\mathcal{U}_0^1}$ and in the energy norm $|\cdot|_a$. Here $h = \frac{1}{16}, \frac{1}{32}$, $\epsilon = \frac{1}{2}$ and $K = N = k$ for several values of k. The second row of Table 9.3 shows the number of Wiener-Chaos terms. The third row shows the $\mathcal{U}_0^1-$ interpolation error and, in parenthesis, the rate of convergence when we truncate the degree and the length of the polynomials. We can see a fast convergence rate when we increase k. The fourth row shows the error between the discrete and the exact solution in the $\mathcal{U}_0^1$ seminorm for $h = 1/16$ and $h = 1/32$; we also show, in parenthesis, the rate of convergence for $h = 1/16$. The decay follows the same behavior as in third row, however not so fast. This deterioration is not due to the mesh size h; see the errors associated to $h = 1/16$ and $h = 1/32$. The deterioration is due to the constant in the a priori error estimate in Theorem 9.24. The larger this constant is, the higher value of k is necessary for observing the asymptotic

decay behavior. In order to minimize the effect of this constant, we now measure the error in the energy norm. In the fifth row we observe a faster decay in the error measured in the energy norm. In Table 9.4 we have the same data as in the Table 9.3 except that now we take $\epsilon = 0$. This case shows clearly a slower decay in energy norm for higher values of k than for the case $\epsilon = \frac{1}{2}$. This is also observed in a less extent in the $|\cdot|_{\mathcal{U}_0^1}$-norm.

$K \downarrow N \rightarrow$	1	2	3	4	5	6	7
1	1.6153	1.4358	1.2512	1.1108	1.0293	0.9918	0.9777
2	1.5970	1.3575	1.0672	0.7948	0.5896	0.4670	0.4104
3	1.5942	1.3446	1.0340	0.7299	0.4814	0.3084	0.2081
4	1.5938	1.3429	1.0296	0.7214	0.4669	0.2849	
5	1.5938	1.3426	1.0291	0.7206	0.4659		

Table 9.1: Total error in the seminorm $|\cdot|_{\mathcal{U}_0^1}$. Here $h = 1/32$, $\epsilon = \frac{1}{2}$.

$K \downarrow N \rightarrow$	1	2	3	4	5	6	7
1	1.6067	1.4272	1.2466	1.1104	1.0307	0.9933	0.9786
2	1.5863	1.3439	1.0560	0.7890	0.5886	0.4683	0.4118
3	1.5832	1.3301	1.0213	0.7222	0.4786	0.3086	0.2093
4	1.5828	1.3282	1.0167	0.7133	0.4637	0.2848	*
5	1.5827	1.3280	1.0162	0.7125	0.4626	*	*

Table 9.2: Total error in the seminorm $|\cdot|_{\mathcal{U}_0^1}$. Here $h = 1/32$, $\epsilon = 0$.

k	0	1	2	3	4	5	6
$\binom{K+N}{K}$	1	2	6	20	70	252	924
$\lvert\hat{u} - Q^{N,K}\hat{u}\rvert_{\mathcal{U}_0^1}$	1.6284	1.3761	0.9767	0.6162	0.3570	0.1920	0.0964
		(1.18)	(1.41)	(1.59)	(1.73)	(1.86)	(1.99)
$\lvert\hat{u} - \hat{u}^{N,K,h}\rvert_{\mathcal{U}_0^1}$	1.7292	1.6157	1.3590	1.0375	0.7281	0.4626	
	1.7291	1.6153	1.3575	1.0340	0.7214	0.4659	
		(1.07)	(1.18)	(1.31)	(1.43)	(1.55)	
$\lvert\hat{u} - \hat{u}^{N,K,h}\rvert_a$	0.4319	0.3691	0.2598	0.1573	0.0836	0.0454	
	0.4318	0.3688	0.2589	0.1552	0.0790	0.0279	
		(1.17)	(1.42)	(1.67)	(1.96)	(2.83)	

Table 9.3: Errors for $K = N = k$, $h = 1/16, 1/32$ and $\epsilon = \frac{1}{2}$. For $h = 1/32$ we have added in parenthesis the reduction factor, when passing to next value of k, corresponding to the projection and finite element error in the seminorm $|\cdot|_{\mathcal{U}_0^1}$ and the finite element error in the energy norm.

9.9 Conclusions and Final Comments

We consider the white noise theory in a general setup to construct and characterize adequate spaces to prove the existence and uniqueness of the solution of the ordinary (rather than Wick) product stochastic pressure equation. We introduce the weak form of the stochastic pressure equation with different spaces for the solution and test functions and prove the continuous inf-sup condition.

We propose a generalization of the approximation proposed in Benth and Theting [2002] and Roman and Sarkis [2006]. By incorporating a weight to measure the exponential decay of the solution and the test functions in the white noise probability, we circumvent the ellipticity of the problem and establish the well-poseness of the problem and provide a priory error estimates for a wide class of

k	0	1	2	3	4	5
$\lvert\hat{u}-\hat{u}^{N,K,h}\rvert_{\mathcal{U}_0^1}$	1.7292	1.6072	1.3454	1.0249	0.7202	0.4746
	1.7291	1.6067	1.3439	1.0213	0.7133	0.4626
		(1.08)	(1.2)	(1.33)	(1.43)	(1.54)
$\lvert\hat{u}-\hat{u}^{N,K,h}\rvert_a$	0.4319	0.3661	0.2611	0.1659	0.0971	0.0533
	0.4318	0.3658	0.2602	0.1639	0.0932	0.0454
		(1.18)	(1.41)	(1.59)	(1.76)	(2.05)

Table 9.4: Errors for $K = N = k$, $h = 1/16, 1/32$ and $\epsilon = 0$. For $h = 1/32$ we have added in parenthesis the reduction factor, when passing to next value of k, corresponding to the projection and finite element error in the seminorm $|\cdot|_{\mathcal{U}_0^1}$ and the finite element error in the energy norm.

norms. The chosen approximation leads to the solution of a positive symmetric linear system. We choose a particular model to test numerically our results. The regularity result of the pressure equation for the types of norms considered in this paper is an object of study of a separated work; see Galvis and Sarkis [2008].

Bibliography

Babuška, I. and Chatzipantelidis, P. (2002). On solving elliptic stochastic partial differential equations. *Comput. Methods Appl. Mech. Engrg.*, 191(37-38):4093–4122.

Babuška, I., Nobile, F., and Tempone, R. (2007). A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM J. Numer. Anal.*, 45(3):1005–1034.

Babuška, I., Tempone, R., and Zouraris, G. E. (2004). Galerkin finite element approximations of stochastic elliptic partial differential equations. *SIAM J. Numer. Anal.*, 42(2):800–825 (electronic).

Benth, F. E. and Gjerde, J. (1998). Convergence rates for finite element approximations of stochastic partial differential equations. *Stochastics Stochastics Rep.*, 63(3-4):313–326.

Benth, F. E. and Theting, T. G. (2002). Some regularity results for the stochastic pressure equation of Wick-type. *Stochastic Anal. Appl.*, 20(6):1191–1223.

Berezanskiĭ, Y. M. (1986). *Selfadjoint operators in spaces of functions of infinitely many variables*, volume 63 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI. Translated from the Russian by H. H. McFaden, Translation edited by Ben Silver.

Bogachev, V. I. (1998). *Gaussian measures*, volume 62 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI.

Cao, Y. (2006). On convergence rate of Wiener-Ito expansion for generalized random variables. *Stochastics*, 78(3):179–187.

Cochran, W. G., Kuo, H.-H., and Sengupta, A. (1998). A new class of white noise generalized functions. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.*, 1(1):43–67.

Da Prato, G. (2006). *An introduction to infinite-dimensional analysis*. Universitext. Springer-Verlag, Berlin. Revised and extended from the 2001 original by Da Prato.

Da Prato, G. and Zabczyk, J. (1992). *Stochastic equations in infinite dimensions*, volume 44 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge.

Frauenfelder, P., Schwab, C., and Todor, R. A. (2005). Finite elements for elliptic problems with stochastic coefficients. *Comput. Methods Appl. Mech. Engrg.*, 194(2-5):205–228.

Furtado, F. and Pereira, F. (2003). Crossover from nonlinearity controlled to heterogeneity controlled mixing in two-phase porous media flows. *Comput. Geosci.*, 7(2):115–135.

Galvis, J. and Sarkis, M. (2008). Regularity results for the ordinary product stochastic pressure equation. In preparation.

Ghanem, R. (1999a). Ingredients for a general purpose stochastic finite elements implementation. *Comput. Methods Appl. Mech. Engrg.*, 168(1-4):19–34.

Ghanem, R. G. (1999b). Stochastic finite elements for heterogeneous media with multiple random non-gaussian properties. *ASCE J. Engrg. Mech.*, 125(1):24–40.

Ghanem, R. G. and R., K. (1999). Numerical solution of spectral stochastic finite element systems. *Comp. Methods Appl. Mech. Engrg.*, (129):289–303.

Ghanem, R. G. and Spanos, P. D. (1991). *Stochastic finite elements: a spectral approach*. Springer-Verlag, New York.

Hida, T. (1980). *Brownian motion*, volume 11 of *Applications of Mathematics*. Springer-Verlag, New York. Translated from the Japanese by the author and T. P. Speed.

Hida, T., Kuo, H.-H., Potthoff, J., and Streit, L. (1993). *White noise*, volume 253 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht. An infinite-dimensional calculus.

Holden, H., Øksendal, B., Ubøe, J., and Zhang, T. (1996). *Stochastic partial differential equations*. Probability and its Applications. Birkhäuser Boston Inc., Boston, MA. A modeling, white noise functional approach.

Jin, C., Cai, X.-C., and Li, C. (2007). Parallel domain decomposition methods for stochastic elliptic equations. *SIAM J. Sci. Comput.*, 29(5):2096–2114 (electronic).

Keese, A. (2003). A review of recent developments in the numerical solution of stochastic partial differential equations (stochastic finite elements). Technical Report 2003-06, Institute of Scientific Computing, , Technical University Braunschweig, Informatikbericht.

Kuo, H. H. (1975). *Gaussian measures in Banach spaces*. Springer-Verlag, Berlin. Lecture Notes in Mathematics, Vol. 463.

Kuo, H.-H. (1996). *White noise distribution theory*. Probability and Stochastics Series. CRC Press, Boca Raton, FL.

Malliavin, P. (1995). *Integration and probability*, volume 157 of *Graduate Texts in Mathematics*. Springer-Verlag, New York. With the collaboration of Hélène Airault, Leslie Kay and Gérard Letac, Edited and translated from the French by Kay, With a foreword by Mark Pinsky.

Matthies, H. G. and Keese, A. (2005). Galerkin methods for linear and nonlinear elliptic stochastic partial differential equations. *Comput. Methods Appl. Mech. Engrg.*, 194(12-16):1295–1331.

Nobile, F., Tempone, R., and Webster, C. (2007). An anisotropic sparse grid stochastic collocation method for elliptic partial differential equations with random input data. MOX-Report No. 04/2007.

Obata, N. (1994). *White noise calculus and Fock space*, volume 1577 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin.

Riesz, F. and Sz.-Nagy, B. (1990). *Functional analysis*. Dover Books on Advanced Mathematics. Dover Publications Inc., New York. Translated from the second French edition by Leo F. Boron, Reprint of the 1955 original.

Roman, L. J. and Sarkis, M. (2006). Stochastic Galerkin method for elliptic SPDEs: a white noise approach. *Discrete Contin. Dyn. Syst. Ser. B*, 6(4):941–955 (electronic).

Schwab, C. and Todor, R. A. (2003). Sparse finite elements for stochastic elliptic problems—higher order moments. *Computing*, 71(1):43–63.

Shigekawa, I. (2004). *Stochastic analysis*, volume 224 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI. Translated from the 1998 Japanese original by the author, Iwanami Series in Modern Mathematics.

Theting, T. G. (2000). Solving Wick-stochastic boundary value problems using a finite element method. *Stochastics Stochastics Rep.*, 70(3-4):241–270.