Technische Universität München

Fakultät für Mathematik

Lehrstuhl für Numerische Mathematik (Prof. Dr. Barbara Wohlmuth)

Energy-Corrected Finite Element Method for Parabolic and Optimal Control Problems

Piotr Świerczyński, M.Sc.

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender: Prof. Dr. Boris Vexler
Prüfer der Dissertation:
1. Prof. Dr. Barbara Wohlmuth
2. Prof. Dr. Thomas Apel

Die Dissertation wurde am 21.11.2018 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 18.03.2019 angenommen.

Acknowledgements

First and foremost I would like to express my sincere gratitude to my supervisor Prof. Dr Barbara Wohlmuth for her invaluable guidance and continuous support throughout my doctoral studies. She has always found time to share her mathematical knowledge and insight and I have greatly benefited from the research environment she created.

I am deeply grateful to Prof. Dr Thomas Apel for dedicating his valuable time to read and review this thesis, as well as for many helpful discussions over the last three years.

I also acknowledge the financial support of the German Research Foundation (DFG) through the International Research Training Group (IGDK) 1754 Munich-Graz "Optimization and Numerical Analysis for Partial Differential Equations with Nonsmooth Structures". IGDK 1754 created a very stimulating environment, and I want to especially thank Dr Johannes Pfefferer and Dr Sergejs Rogovs for many fruitful discussions.

As a part of the IGDK 1754 program, I had the pleasure to spend several months visiting the Technical University of Graz. I want to thank colleagues there for a warm welcome and support.

Special thanks to my present and former colleagues at the Chair of Numerical Mathematics (M2) at the Technical University of Munich for the friendly and inspiring working atmosphere. I have particularly enjoyed collaborating with Dr Lorenz John and I am also very grateful to Dr Linus Wunderlich for proofreading this thesis. Moreover, I would like to thank them, as well as Ettore Vidotto and Markus Huber, for their friendliness and encouragement in developing my German language skills.

I want to thank my parents Liliana and Wojciech, as well as my brother Michał, for their continuous support. Without them, I would not be where I am today.

Finally, I want to thank my loving wife Anna for being the source of happiness, inspiration and purpose in my life.

Publications included in the thesis

- L. John, P. Swierczynski, and B. Wohlmuth, Energy corrected FEM for optimal Dirichlet boundary control problems, *Numerische Mathematik*, 139(4):913–938, 2018.
- P. Swierczynski and B. Wohlmuth, Energy-corrected FEM and explicit time-stepping for parabolic problems, (Submitted), 2018.
- P. Swierczynski and B. Wohlmuth, Maximum norm estimates for energy-corrected finite element method, *Numerical Mathematics and Advanced Applications ENUMATH 2017*, 126, Springer International Publishing, *doi:10.1007/978-3-319-96415-7*, 2018.

Publications not included in the thesis

- U. Khristenko, L. Scarabosio, P. Swierczynski, E. Ullmann, and B. Wohlmuth, Analysis of boundary effects on PDE-based sampling of Whittle-Matérn random fields, SIAM/ASA Journal on Uncertainty Quantification, (Accepted for publication), 2019.
- S. Dohr, Ch. Kahle, S. Rogovs, and P. Swierczynski, A FEM for an optimal control problem subject to the fractional Laplace equation, (Submitted), 2018.

Abstract

In this thesis, we study an energy-corrected finite element method mitigating the pollution effect on polygonal domains. We show that the general higher-order discretisation yields error of optimal order when measured in the maximum norm. We propose an energy-corrected discretisation of parabolic problems based on an explicit time-stepping. Finally, we show that the application of the energy-correction can improve the finite element approximation of optimal control problems.

Zusammenfassung

In dieser Arbeit wird die energiekorrigierte Finite-Elemente-Methode untersucht um den Verschmutzungseffekt auf polygonalen Gebieten einzudämmen. Wir zeigen, dass die allgemein höhere Ordnung Diskretisierung einen Fehler mit optimaler Ordnung liefert, wenn dieser in der Maximumsnorm gemessen wird. Wir führen eine energiekorrigierte Diskretisierung, die auf ein explizites Zeitschrittverfahren basiert, für parabolische Probleme ein. Schließlich zeigen wir, dass Energiekorrektur die Finite Element Approximation von Optimalsteuerungsprobleme verbessern kann.

Contents

List of Figures

Chapter 1

Introduction

Mathematical models play a vital role in natural, social and medical sciences, as well as in engineering applications. Time-dependent phenomena, such as radioactive decay, population growth or velocity of a falling body, can be described using equations involving derivatives in the temporal dimension. On top of that, especially in the area of continuum mechanics, models can also include derivatives in the spatial directions, resulting in partial differential equations. This mathematical formalism has proven very useful in describing a wide range of phenomena, from fluid flow and heat distribution to elasticity and growth of bacteria population depending on a nutrient consumption. Despite their importance, it is, in general, very difficult to solve partial differential equations analytically.

Due to this inherent difficulty and the importance of partial differential equations, many numerical methods for approximating the solutions have been developed. Among the most popular ones are finite difference, finite volume, boundary element and finite element methods, all of which approximate the partial differential equations with systems of algebraic equations. The last method of the mentioned here has gained considerable popularity due to its flexibility in representing complex computational domains and treatment of nonhomogeneous, local material properties. It is based on a discretisation of the continuous problem in some finite dimensional space and subdividing the domain of interest into small parts.

Among domains of particular interest in applied sciences are polygons, potentially including non-convex corners. The presence of corners in the computational domains negatively influences the regularity of the solutions of elliptic partial differential equations, by introducing a certain known type of singular functions into the solution. The appearance of singular functions in the solutions on polygonal domains diminishes the accuracy of the standard finite element method. Furthermore, the approximation is

suboptimal not only in a vicinity of the corner but also when considered in some positive distance from it - the behaviour known as *the pollution effect*.

Many approaches for improving the global convergence of the finite element solution exist already, such as adaptivity, graded mesh algorithms or enrichment of the finite element space by singular functions. These methods improve the approximation quality regaining the optimal convergence properties. However, they require a modification of the standard finite element solution in a neighbourhood of the corners and need changing the structure of the mesh or the function space, in which the solution is approximated.

Recently, an alternative method for eliminating the pollution effect in the finite element discretisation on polygonal domains has been developed - *the energy-correction method*. It is based on the observation that the approximation quality of the finite element solution, when measured in the $L^2(\Omega)$-norm, is limited by the approximation of the energy. The numerical algorithm modifies the bilinear form governing the system of equations so that the optimal approximation of the energy is regained. Moreover, this modification can be created in a way that modifies the problem only in some small area around the corner. It corresponds to the scaling of a fixed number of entries in the stiffness matrix governing the system of algebraic equations arising upon the discretisation. Using the energy-correction method, optimal convergence concerning the interpolation error is regained. As opposed to the schemes involving adaptivity or the mesh grading, the energy-correction works on uniformly refined meshes.

The energy-corrected finite element discretisation was studied in the setting of scalar elliptic problems for the piecewise linear [69, 85, 137] and for the piecewise polynomial case [86]. It was also further extended to the Stokes problem in fluid dynamics [91]. The energy-correction method serves as a common denominator for all the problems analysed in this thesis. We aim at answering several interesting questions regarding this type of discretisation and at showing some potential applications in the realm of time-dependent problems and optimisation.

We begin in Chapter 2 by introducing the notation and standard, as well as weighted, function spaces, which are used later in the thesis. We compare the known regularity results for the solutions of elliptic problems on polygons with the ones defined on domains with smooth boundaries. Then, we introduce the standard finite element method and discuss the pollution effect lowering the convergence order of the finite element discretisation in the whole considered area. For completeness, we also briefly discuss the gradual mesh refinement around the domains' corners. We give a

complete presentation of the known results for the energy-corrected discretisations of elliptic problems. The energy-correction method mitigates the pollution effect and yields optimal convergence in weighted Lebesgue and Sobolev norms. Additional post-processing step results in a further improvement of the approximation when the standard norms of the error are concerned. We conclude by showing some numerical experiments confirming the theoretical findings. The goal of this chapter is to present a relevant mathematical background and known results, on which we shall later build our analysis.

Having provided a summary of the existing error estimates for the energy-corrected finite element in weighted Lebesgue and Sobolev norms, in Chapter 3 we focus our interest on pointwise error estimates. The main result of our investigations states that the energy-correction yields optimal convergence order, compared to the interpolation error, in weighted maximum norm. This, in particular, means that the pollution effect, which can also be observed in errors in the maximum norm, is eliminated. The energy-corrected scheme yields the optimal approximation when considered far from the corner. The proof of this result is based on the dyadic decomposition of the computational domain around the corner.We complete the chapter presenting numerical experiments confirming the theoretical findings.

The pollution effect introduced for the elliptic partial differential equations can also be observed in the case of parabolic problems on polygonal domains. In Chapter 4, we show that the piecewise linear energy-corrected finite element discretisation of parabolic problems improves the convergence properties of the standard schemes and results in the optimal convergence, also when considered in the weighted Sobolev norms. As opposed to the methods involving mesh refinement around the singular corner, the energy-correction requires only uniform triangulation of the computational domain. This results in a non-restrictive stability requirement for the explicit time-stepping schemes. We propose a fully discrete scheme for the parabolic equations based on the piecewise-linear and enriched cubic energy-corrected finite element combined with the explicit time discretisation. These schemes are completed with the mass lumping strategy yielding fast solvers in time, and an additional post-processing step. We present numerical experiments confirming the theoretical results.

In Section 4.4, we provide extensive numerical studies and propose several extensions of the scheme. We show that the proposed discretisation can be successfully applied to problems with multiple singular corners, to problems including mild convective terms and when various quantities of interest are considered. Furthermore, we also provide a comparison of the energy-corrected scheme with other commonly used

discretisations. Finally, we show a potential application of the method to a problem of heat distribution in a geometry of a graphite nuclear power plant moderator brick. This problem poses an exciting challenge from our perspective, as it involves 16 non-convex corners in a three-dimensional setting.

In Chapter 5, we consider an optimisation problem subject to the Poisson equation with the Dirichlet boundary conditions being controlled. Moreover, we equip the problem with the regularisation in the energy space and with box-constraints on the control boundary. Similar problems arise, among other places, in the area of computational fluid dynamics and have recently achieved considerable attention. Error estimates for the finite element discretisation of such problems have been studied lately in the context of convex domains. We show that the presence of non-convexities in the polygonal domains diminishes the convergence properties of the standard finite element method, and the phenomenon analogous to the pollution effect can be observed. As a remedy, we propose an energy-corrected finite element approximation of the problem. Furthermore, we prove the optimal convergence properties of the method, regarding the known interpolation error estimates. This is done under the assumption that the control is placed in some positive distance from the re-entrant corners. Finally, we show a numerical implementation based on the primal-dual active set strategy and present the numerical experiments confirming the theoretical findings.

We conclude the thesis in Chapter 6 with the outlook and a summary of possible interesting further steps in the development of the energy-corrected finite element.

Chapters 3–5 present our novel results and are extended discussions of the analysis presented in [154], [153] and [92] respectively. When discussing the theory already existing in the literature, we provide only a relevant reference, whereas proofs always support the novel results.

Chapter 2

Mathematical Background

In this chapter, we introduce the mathematical tools necessary for derivation and understanding of the analysis presented in this work. Our goal is to construct a theoretical framework, in which all the investigations will be conducted. Moreover, we wish to familiarise the reader with the notation. Finally, we also want to describe the current state of the knowledge in the field of finite element approximations of elliptic problems on polygonal domains with the particular interest in the energy-correction method.

We begin with the classical functional analysis results, defining spaces on bounded domains in Euclidean spaces and discussing their elementary properties. Further, we discuss also weighted spaces, which are useful for describing the features of the problems defined on domains with corners. Then, we introduce elliptic problems and investigate their fundamental properties. We show, how these properties depend on the smoothness of the considered domain. Next, we introduce the finite element methods providing the numerical toolset for discretisation of the partial differential equations considered in this thesis. Finally, we show an exhaustive summary of the existing results in the area of the energy-corrected finite element methods.

2.1 Function spaces

In this section, we introduce the function spaces, which are further used as the basis for the analysis presented in this thesis. For the sake of completeness and in order to introduce the notation, we begin with defining the standard Hölder, Lebesgue and Sobolev spaces and summarize their basic properties stating known embedding and trace theorems. An exhaustive discussion of the above-mentioned function spaces can be found for example in [1, 71, 75, 113, 64, 160]. Following [100, 113], we extend given definitions to the case of weighted Sobolev spaces, which constitute a convenient framework for the analysis of boundary value problems on polygonal domains.

2.1.1 Standard function spaces

Here, we will be concerned with subsets Ω of the Euclidean space $\mathbb{R}^d$, where we denote by $d \in \mathbb{Z}_+$ the dimension of the space. We begin by stating the definition of the Banach spaces of continuously differentiable functions and functions with Hölder continuous derivatives.

Definition 2.1.1 (Spaces of continuously differentiable functions)**.** *Let k be a non-negative integer and $\Omega \subset \mathbb{R}^d$ be an open set. We denote by $C^k(\overline{\Omega})$ the space of all k-times differentiable functions $u : \Omega \to \mathbb{R}$, for which derivatives $D^{\mathbf{m}}v : \Omega \to \mathbb{R}$ are continuous and bounded for all multiindices $\mathbf{m}$ satisfying $|\mathbf{m}| \leq k$. We equip this space with the following norm*

$$\|v\|_{C^k(\overline{\Omega})} = \sum_{|\mathbf{m}|\leq k} \sup_{\Omega} |D^{\mathbf{m}}v|.$$

We also denote by $C_0^k(\Omega)$ a subset of $C^k(\overline{\Omega})$ consisting of functions compactly supported in Ω. Furthermore, when $k = 0$, we write $C^0(\overline{\Omega}) = C(\overline{\Omega})$ and $C_0^0(\Omega) = C_0(\Omega)$.

Definition 2.1.2 (Spaces of Hölder continuous functions)**.** *Let $\Omega \subset \mathbb{R}^d$, be an open set and $\sigma \in (0, 1)$. The Hölder space $C^{k,\sigma}(\Omega)$ consists of all functions $v \in C^k(\overline{\Omega})$, for which*

$$\sum_{|\mathbf{m}|=k} \sup_{\mathbf{x},\mathbf{y}\in\Omega,\, \mathbf{x}\neq\mathbf{y}} \frac{|D^{\mathbf{m}}v(\mathbf{x}) - D^{\mathbf{m}}v(\mathbf{y})|}{|\mathbf{x}-\mathbf{y}|^{\sigma}} < \infty.$$

The norm in this space is defined by

$$\|v\|_{C^{k,\sigma}(\Omega)} = \|v\|_{C^k(\overline{\Omega})} + \sum_{|\mathbf{m}|=k} \sup_{\mathbf{x},\mathbf{y}\in\Omega,\, \mathbf{x}\neq\mathbf{y}} \frac{|D^{\mathbf{m}}v(\mathbf{x}) - D^{\mathbf{m}}v(\mathbf{y})|}{|\mathbf{x}-\mathbf{y}|^{\sigma}}.$$

Definition 2.1.3 ($L^p(\Omega)$ spaces)**.** *Let $1 \leq p \leq \infty$ and let $\Omega \subset \mathbb{R}^d$, be a measurable set. The Lebesgue space $L^p(\Omega)$ consists of all measurable functions $v : \Omega \to \mathbb{R}$, for which the corresponding norm is finite*

$$\|v\|_{L^p(U)} = \Big(\int_\Omega |v|^p \Big)^{1/p}, \quad \text{for } 1 \leq p < \infty,$$
$$\|v\|_{L^\infty(U)} = \operatorname*{ess\,sup}_\Omega |v|.$$

Moreover, we identify functions, which differ only on a set of measure 0*.*

Remark 2.1.4. *$L^2(\Omega)$ space is a Hilbert space with an inner product defined by*

$$\big\langle u, v \big\rangle_{L^2(\Omega)} = \int_\Omega uv.$$

For the sake of notational simplicity, we skip the subscript in the inner product, when the space is clear from the context. Furthermore, we denote the norm in the $L^2(\Omega)$ space by

$$\|v\|_0 = \|v\|_{L^2(\Omega)}.$$

Now, we assume that $\Omega \subset \mathbb{R}^d$, is a bounded, open domain. We define Sobolev spaces, which constitute a standard framework for investigating the partial differential equations [71].

Definition 2.1.5 (Sobolev spaces)**.** *Let $k \in \mathbb{Z}_+ \cup \{0\}$ and $1 \leq p \leq \infty$. The Sobolev space $W^{k,p}(\Omega)$ consists of all measurable functions $v : \Omega \to \mathbb{R}$, for which*

$$D^{\mathbf{m}} v \in L^p(\Omega), \quad \text{for all } |\mathbf{m}| \leq k,$$

where $\mathbf{m}$ is a multiindex and $D^{\mathbf{m}} v$ is a $\mathbf{m}$-weak derivative of the function v.
We equip the Sobolev space with the following norm

$$\|v\|_{W^{k,p}(\Omega)} = \Big(\sum_{|\mathbf{m}| \leq k} \|D^{\mathbf{m}} v\|^p_{L^p(\Omega)} \Big)^{1/p}, \qquad \text{for } 1 \leq p < \infty,$$
$$\|v\|_{W^{k,\infty}(\Omega)} = \sum_{|\mathbf{m}| \leq k} \|D^{\mathbf{m}} v\|_{L^\infty(\Omega)}.$$

Moreover, we denote by $W^{k,p}_0(\Omega)$ the closure of $C^k_0(\Omega)$ in $W^{k,p}(\Omega)$.
We also define the following seminorms

$$|v|_{W^{k,p}(\Omega)} = \Big(\sum_{|\mathbf{m}| = k} \|D^{\mathbf{m}} v\|^p_{L^p(\Omega)} \Big)^{1/p}, \qquad \text{for } 1 \leq p < \infty,$$
$$|v|_{W^{k,\infty}(\Omega)} = \sum_{|\mathbf{m}| = k} \|D^{\mathbf{m}} v\|_{L^\infty(\Omega)}.$$

Note that $W^{0,p}(\Omega) = L^p(\Omega)$.

Remark 2.1.6 (Sobolev–Hilbert spaces)**.** *When $p = 2$, we use the notation $H^k(\Omega) = W^{k,2}(\Omega)$. Similarly as in the case of Lebesgue spaces, when $p = 2$, Sobolev spaces are also Hilbert spaces with the inner product*

$$\left\langle u, v \right\rangle_{H^k(\Omega)} = \sum_{|\mathbf{m}| \leq k} \left\langle D^{\mathbf{m}} u, D^{\mathbf{m}} v \right\rangle_{L^2(\Omega)}.$$

Consequently, we also use the standard notation $H_0^k(\Omega) = W_0^{k,2}(\Omega)$. Moreover, we define

$$H^{-1}(\Omega) = \left(H_0^1(\Omega) \right)^*.$$

So far we have only defined Sobolev spaces $W^{k,p}(\Omega)$ for positive integer exponents k. We would like to extend this notion by incorporating also fractional exponents by the introduction of Sobolev–Slobodeckij spaces. Such spaces will appear in the analysis of problems with reduced regularity due to the presence of non-convexities in the computational domain. Furthermore, fractional Sobolev spaces also play a significant role in prescribing the values on the boundary of the domain to functions from some Sobolev spaces. This will be clear later once the notion of traces is introduced in Section 2.1.3.

Definition 2.1.7 (Sobolev–Slobodeckij spaces)**.** *Let $s > 0$ be a non-integer and $1 \leq p \leq \infty$. Let also $\vartheta = s - \lfloor s \rfloor \in (0, 1)$. We define the Sobolev–Slobodeckij seminorm of a measurable function $v : \Omega \to \mathbb{R}$ by*

$$|v|_{W^{s,p}(\Omega)} = \sum_{|\mathbf{m}| = \lfloor s \rfloor} \left(\int_\Omega \int_\Omega \frac{|D^{\mathbf{m}} v(\mathbf{x}) - D^{\mathbf{m}} v(\mathbf{y})|^p}{|\mathbf{x} - \mathbf{y}|^{\vartheta p + d}} \right)^{1/p}, \qquad \text{for } 1 \leq p < \infty,$$

$$|v|_{W^{s,\infty}(\Omega)} = \max_{|\mathbf{m}| = \lfloor s \rfloor} \operatorname*{ess\,sup}_{(\mathbf{x},\mathbf{y}) \in \Omega \times \Omega} \frac{|D^{\mathbf{m}} v(\mathbf{x}) - D^{\mathbf{m}} v(\mathbf{y})|}{|\mathbf{x} - \mathbf{y}|^{\vartheta}}.$$

We define the space $W^{s,p}(\Omega)$ as a space of all functions $v \in W^{\lfloor s \rfloor, p}(\Omega)$, for which

$$|v|_{W^{s,p}(\Omega)} < \infty$$

and equip it with the norm

$$\|v\|_{W^{s,p}(\Omega)} = \|v\|_{W^{\lfloor s \rfloor, p}(\Omega)} + |v|_{W^{s,p}(\Omega)}.$$

Again, when $p = 2$, we also write $H^s(\Omega) = W^{s,p}(\Omega)$.

As desired, the standard Sobolev spaces and fractional Sobolev–Slobodeckij spaces form a natural hierarchy.

Remark 2.1.8 (Hierarchy of Sobolev spaces)**.** *Let $k \leq s \leq s' \leq k+1$ for some $k \in \mathbb{Z}_+ \cup \{0\}$, and let Ω be a Lipschitz domain. Then for all $1 \leq p \leq \infty$*

$$W^{k+1,p}(\Omega) \hookrightarrow W^{s',p}(\Omega) \hookrightarrow W^{s,p}(\Omega) \hookrightarrow W^{k,p}(\Omega).$$

Having defined all standard spaces, we are now in a position to recall some well-known results in the theory of Sobolev spaces.

Theorem 2.1.9 (Sobolev embedding theorem)**.** *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain, $s \in (0,1]$ and $1 \leq p < \infty$. Consider two cases:*

- *if $sp < d$, then $W^{s,p}(\Omega) \hookrightarrow L^q(\Omega)$ for and $q \leq p^* = \frac{dp}{d-sp}$.*
- *if $sp > d$, then $W^{s,p}(\Omega) \hookrightarrow C^{0,\sigma}(\Omega)$ for $\sigma < s - d/p$.*

Remark 2.1.10. *In the limit case $sp = d$ in Theorem 2.1.9, the mentioned embedding does not extend naturally to bounded functions. In particular $W^{1,d}(\Omega) \not\hookrightarrow L^\infty(\Omega)$. For a treatment of this case in the framework of Orlicz spaces, we would like to refer the reader to the Trudinger–Moser inequality [116, 162] and [1, Chapter 8].*

The proof of the theorem for $s = 1$ can be found in [71, Section 5.6] and [113, Section 4.4, Section 5.1]. A comprehensive proof of the Sobolev Embedding Theorem in the case of fractional spaces can be found in [64, Theorem 6.7, Theorem 8.2]. We also refer the reader to [39, 163].

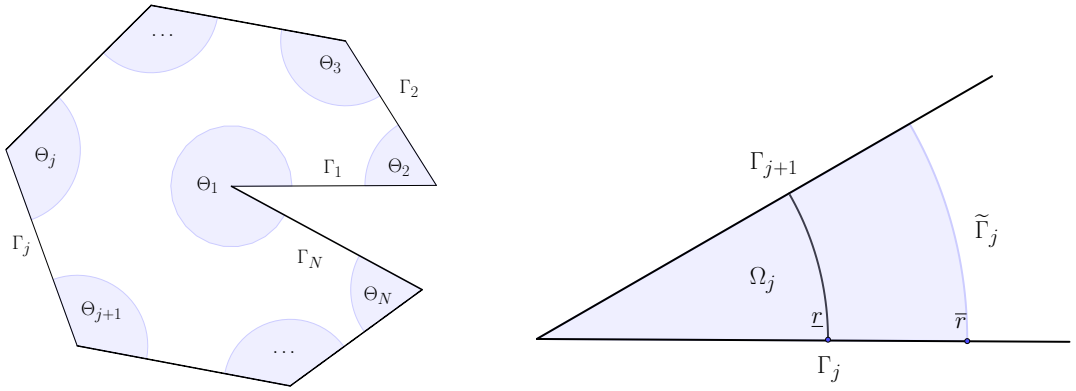
2.1.2 Weighted function spaces

As we shall see in Section 2.2, weighted Sobolev spaces constitute a very efficient framework for describing the regularity of solutions of elliptic problems on domains with corners. Now, we would like to focus our attention on two-dimensional polygonal domains.

Definition 2.1.11 (Polygonal domain)**.** *Let $N \in \mathbb{Z}_+$ and $\mathcal{S} = \{\mathbf{x}_j\}_{j=1}^N \subset \mathbb{R}^2$. Let also Γ_j, $j = 1, \ldots, N$, be a series of pairwise disjoint open intervals such that $\mathbf{x}_j = \overline{\Gamma}_j \cap \overline{\Gamma}_{j+1}$ with $\Gamma_{N+1} = \Gamma_1$. We call an open, bounded, and simply-connected set $\Omega \subset \mathbb{R}^2$ a polygonal domain if $\partial\Omega = \bigcup_{j=1}^N \overline{\Gamma}_j$.*
We call points $\mathbf{x}_i$ vertices or corners and intervals Γ_j the edges. Furthermore, we denote the angles between $\overline{\Gamma}_j$ and $\overline{\Gamma}_{j+1}$ by Θ_j.

We define (r_j, ϕ_j), $j = 1, \ldots, N$ *to be polar coordinates around respective corners. Here* $r_j(\mathbf{x})$ *denotes the distance of the point* $\mathbf{x} \in \mathbb{R}^2$ *from the* j*-th corner and* ϕ_j *denotes the angle, which is chosen so that* $\phi_j = 0$ *aligns with* Γ_j*.*

Further, we will use the notation introduced in the definition above. Notice that this definition allows only for angles of size $0 < \Theta_j < 2\pi$ excluding an important case of a so-called slit-domain. A visual representation of the notation is shown in Figure 2.1a. This notation will be used further throughout the work unless explicitly stated otherwise. From now on, we assume that $\Omega \subset \mathbb{R}^2$ is a two-dimensional polygonal domain with N corners of sizes $\Theta_1, \ldots \Theta_N$.



(a) Model of a polygon with N edges (b) Sector defined at one of the corners

Figure 2.1: Model polygonal domain (left) and a sector around one of the corners (right).

For a given vector $\vec{\beta} = (\beta_1, \ldots, \beta_N) \in \mathbb{R}^N$ we define the weights as

$$\mathbf{r}^{\vec{\beta}} = \prod_{i=1}^{N} r_i^{\beta_i}.$$

For the sake of notational simplicity, we use the convention $\vec{\beta}+\delta = (\beta_1+\delta, \ldots, \beta_N+\delta)$, where $\delta \in \mathbb{R}$.

As in the case of standard function spaces, we begin the disussion of weighted spaces with spaces of continuous and differentiable functions.

Definition 2.1.12 (Weighted Hölder spaces)**.** *Let* $k \in \mathbb{Z}_+ \cup \{0\}$ *and* $\sigma \in (0, 1)$*. We define space of weighted Hölder continuous functions* $N^{k,\sigma}_{\vec{\beta}}(\Omega)$ *consisting of all* k*-times continuously differentiable functions, for which*

$$\sum_{|\mathbf{m}|=k} \sup_{\mathbf{x},\mathbf{y}\in\Omega,\, \mathbf{x}\neq\mathbf{y}} \frac{|\mathbf{r}(\mathbf{x})^{\vec{\beta}} D^{\mathbf{m}} v(\mathbf{x}) - \mathbf{r}(\mathbf{y})^{\vec{\beta}} D^{\mathbf{m}} v(\mathbf{y})|}{|\mathbf{x} - \mathbf{y}|^{\sigma}} < \infty.$$

We consider this space with the following norm.

$$\|v\|_{N^{k,\sigma}_{\vec{\beta}}(\Omega)} = \sum_{|\mathbf{m}|\le k} \|\mathbf{r}^{\vec{\beta}-\sigma-k+|\mathbf{m}|} D^{\mathbf{m}} v\|_{C^0(\overline{\Omega})} + \sum_{|\mathbf{m}|=k} \sup_{\mathbf{x},\mathbf{y}\in\Omega,\, \mathbf{x}\neq\mathbf{y}} \frac{|\mathbf{r}(\mathbf{x})^{\vec{\beta}} D^{\mathbf{m}} v(\mathbf{x}) - \mathbf{r}(\mathbf{y})^{\vec{\beta}} D^{\mathbf{m}} v(\mathbf{y})|}{|\mathbf{x}-\mathbf{y}|^{\sigma}}.$$

Definition 2.1.13 (Weighted $L^p(\Omega)$ spaces)**.** *Let $\vec{\beta} \in \mathbb{R}^N$, $N \in \mathbb{Z}_+$. We define the weighted Lebesgue space $L^p_{\vec{\beta}}(\Omega)$, $1 \le p \le \infty$, as the set of all measurable functions $v : \Omega \to \mathbb{R}$, such that*

$$\|v\|_{L^p_{\vec{\beta}}(\Omega)} = \|v\|_{\vec{\beta}} = \left\|\mathbf{r}^{\vec{\beta}} v\right\|_{L^p(\Omega)} < \infty,$$

where $\mathbf{r}^{\vec{\beta}}$ is called 'the power weight', or simply 'the weight'.

We now move to the definition of two different types of weighted Sobolev spaces with weights depending on the distance r from a chosen point in the computational domain. An exhaustive discussion of more general types of weighted Sobolev spaces involving also different kinds of weights can be found in [100].

Definition 2.1.14 (Weighted Sobolev spaces $W^{k,p}_{\vec{\beta}}(\Omega)$ and $V^{k,p}_{\vec{\beta}}(\Omega)$)**.** *Suppose that $k \in \mathbb{Z}_+ \cup \{0\}$ and $1 \le p \le \infty$. We define the weighted Sobolev spaces as follows*

$$W^{k,p}_{\vec{\beta}}(\Omega) = \left\{ v - \textit{measurable} : \mathbf{r}^{\vec{\beta}} D^{\mathbf{m}} v \in L^p(\Omega), |\mathbf{m}| \le k \right\},$$
$$V^{k,p}_{\vec{\beta}}(\Omega) = \left\{ v - \textit{measurable} : \mathbf{r}^{\vec{\beta}-k+|\mathbf{m}|} D^{\mathbf{m}} v \in L^p(\Omega), |\mathbf{m}| \le k \right\}.$$

When $1 \le p < \infty$, we equip these spaces with the following norms

$$\|v\|_{W^{k,p}_{\vec{\beta}}(\Omega)} = \left(\sum_{|\mathbf{m}|\le k} \left\|\mathbf{r}^{\vec{\beta}} D^{\mathbf{m}} v\right\|^p_{L^p(\Omega)} \right)^{1/p},$$
$$\|v\|_{V^{k,p}_{\vec{\beta}}(\Omega)} = \left(\sum_{|\mathbf{m}|\le k} \left\|\mathbf{r}^{\vec{\beta}-k+|\mathbf{m}|} D^{\mathbf{m}} v\right\|^p_{L^p(\Omega)} \right)^{1/p},$$

and in the remaining case of $p = \infty$

$$\|v\|_{W^{k,\infty}_{\vec{\beta}}(\Omega)} = \sum_{|\mathbf{m}|\le k} \left\|\mathbf{r}^{\vec{\beta}} D^{\mathbf{m}} v\right\|_{L^\infty(\Omega)},$$
$$\|v\|_{V^{k,\infty}_{\vec{\beta}}(\Omega)} = \sum_{|\mathbf{m}|\le k} \left\|\mathbf{r}^{\vec{\beta}-k+|\mathbf{m}|} D^{\mathbf{m}} v\right\|_{L^\infty(\Omega)}.$$

Finally, we also define the seminorms

$$|v|_{W^{k,p}_{\vec{\beta}}(\Omega)} = |v|_{V^{k,p}_{\vec{\beta}}(\Omega)} = \Big(\sum_{|\mathbf{m}|=k} \big\| \mathbf{r}^{\vec{\beta}} D^{\mathbf{m}} v \big\|^p_{L^p(\Omega)} \Big)^{1/p}, \quad \text{for } 1 \le p < \infty,$$

$$|v|_{W^{k,\infty}_{\vec{\beta}}(\Omega)} = |v|_{V^{k,\infty}_{\vec{\beta}}(\Omega)} = \sum_{|\mathbf{m}|=k} \big\| \mathbf{r}^{\vec{\beta}} D^{\mathbf{m}} v \big\|_{L^\infty(\Omega)}.$$

When $p = 2$, we also denote $H^k_{\vec{\beta}}(\Omega) = V^{k,2}_{\vec{\beta}}(\Omega)$ and use the notation

$$\|v\|_{H^k_{\vec{\beta}}(\Omega} = \|v\|_{k,\vec{\beta}} = \Big(\sum_{|\mathbf{m}|\le k} \big\| \mathbf{r}^{\vec{\beta}-k+|\mathbf{m}|} D^{\mathbf{m}} v \big\|^2_{L^2(\Omega)} \Big)^{1/2}.$$

Note that $L^2_{\vec{\beta}}(\Omega) = H^0_{\vec{\beta}}(\Omega)$ and due to the special form of the weight $\mathbf{r}^{2\vec{\beta}}$, $H^k_{\vec{\beta}}(\Omega)$ are Hilbert spaces.

Let $\Gamma \subset \partial\Omega$. We define the space $H^k_{0,\vec{\beta}}(\Omega, \Gamma)$ as the closure of $\{v \in C^\infty(\Omega) : v|_\Gamma = 0\}$ in $H^k_{\vec{\beta}}(\Omega)$ norm. When $\Gamma = \partial\Omega$, we also write $H^k_{0,\vec{\beta}}(\Omega)$ instead.

The following embeddings describe the relationships between the weighted spaces introduced above

Theorem 2.1.15. *Let $\vec{\beta} = (\beta_1, \dots, \beta_N) \in \mathbb{R}^N$, $k \in \mathbb{Z}_+$ and $1 \le p \le \infty$. Then*

- *for all $l \in \mathbb{Z}_+$ we have $V^{k+l,p}_{\vec{\beta}+l}(\Omega) \hookrightarrow V^{k,p}_{\vec{\beta}}(\Omega)$*
- *if all $\beta_i \ne \frac{1}{2}(jp-1)$, where $j \in \mathbb{Z}_+$, then $W^{k,p}_{0,\vec{\beta}}(\Omega) \hookrightarrow V^{k,p}_{\vec{\beta}}(\Omega) \hookrightarrow W^{k,p}_{\vec{\beta}}(\Omega)$,*
- *for $\sigma > 0$ we have $N^{k,\sigma}_{\vec{\beta}} \hookrightarrow V^{k,p}_{\vec{\beta}}(\Omega)$.*

The proof of the first assertion can be found in [100, Theorem 8.11]. Note that it is an extension of Remark 2.1.8 to the framework of weighted Sobolev spaces. The second assertion is a consequence of [100, Theorem 9.4, Theorem 9.6]. The last one follows directly from the definitions of the considered spaces.

2.1.3 Traces

For sufficiently smooth functions and domains, the values on the boundary of the domain can be defined by a straightforward extension of the values inside the domain. However, pointwise values of functions from Sobolev spaces and their weak derivatives can, in general, be not well-defined. Spaces $H^s_0(\Omega)$, $s > \frac{1}{2}$, as closures of functions with compact support in Ω, can be intuitively understood as consisting of functions with homogeneous values on the boundary. In order to attach non-homogenous values on the boundary to functions in Sobolev spaces $H^s(\Omega)$, it is necessary to introduce the notion of traces.

Theorem 2.1.16 (Trace theorem on smooth domains)**.** *Let $\Omega \subset \mathbb{R}^d$ be domain with a C^{k+1}-boundary $\partial\Omega$ for some $k \in \mathbb{Z}_+ \cup \{0\}$. Let also $s \in \mathbb{R}$ and $j \in \mathbb{Z}_+ \cup \{0\}$ satisfy $j + 1/2 < s \leq k + 1$. The operator*

$$\mathrm{tr}_j u = \frac{\partial^j u}{\partial \mathbf{n}^j}\Big|_{\partial\Omega} : C^\infty(\overline{\Omega}) \to H^{s-j-1/2}(\partial\Omega)$$

has a unique continuous and surjective extension $\mathrm{tr}_j : H^s(\Omega) \to H^{s-j-1/2}(\partial\Omega)$.

We call tr_0 and tr_1 the Dirichlet and Neumann trace respectively. Moreover, we identify the values of functions from the space $H^s(\Omega)$ and values of their derivative in the direction orthogonal to the boundary $\mathbf{n}$, with the values given by the application of the corresponding trace operator and thus write also $\mathrm{tr}_0 u = u|_{\partial\Omega}$, $\mathrm{tr}_1 u = \partial_{\mathbf{n}} u|_{\partial\Omega}$. A more exhaustive discussion of traces of Sobolev functions can be found in [71, Section 5.5], [75, Chapter 18], [87, Section 4.2].

In Theorem 2.1.16, we introduced the notion of the trace defining the values of functions from Sobolev spaces on the boundary. So far, this was done only on domains with certain smoothness properties of the boundary, which are not satisfied by polygons. We will close this gap using standard Sobolev spaces. For the discussion regarding the traces of functions in weighted Sobolev spaces, we refer the reader to [100, Theorem 9.15] and [132, Section 2.2].

Suppose that $\Omega \subset \mathbb{R}^2$ is a polygonal domain in the sense of Definition 2.1.11. Let $\Gamma_i \subset \partial\Omega$, $i = 1, \ldots, J$, be some pairwise disjoint smooth parts of the boundary, for example edges of the polygon, and let $\Gamma = \bigcup_{i=1}^J \overline{\Gamma}_i$. The following theorem was proved in [78, Theorem 1.4.2].

Theorem 2.1.17. *Let $s \in \mathbb{R}$ and $j \in \mathbb{Z}_+ \cup \{0\}$ such that $j < s - 1/2$. Then the mapping*

$$\mathrm{tr}_j u = \frac{\partial^j u}{\partial \mathbf{n}^j}\Big|_{\Gamma_i} : C^\infty(\overline{\Omega}) \to H^{s-j-1/2}(\Gamma_i)$$

has a unique continuous and surjective extension

$$\mathrm{tr}_j : H^s(\Omega) \to H^{s-j-1/2}(\Gamma_i).$$

Due to the presence of corners in the polygonal domains, which violate the regularity assumptions imposed on the computational domain Ω in Theorem 2.1.16, the formulation presented there cannot be directly extended to parts of the boundary of a polygon consisting of several edges.

We restrict our attention only to two kinds of Sobolev spaces on boundaries of polygonal domains, since only these two spaces are necessary for our further considerations. For a complete treatment of such spaces, together with a formulation of more elaborate *compatibility conditions* imposed in higher order spaces, we refer to [78, Section 1.4].

Definition 2.1.18. *Let $\sigma \in \{1/2, 3/2\}$. We define spaces of piecewise Sobolev regularity by*

$$H^{\sigma}_{\mathrm{pw}}(\Gamma) = \big\{ v \in L^2(\Gamma) : \ v|_{\Gamma_i} \in H^{\sigma}(\Gamma_i) \big\}.$$

Moreover, when $\sigma = 3/2$, we impose an additional compatibility condition, namely, we require that all functions $v \in H^{3/2}(\Gamma)$ are continuous on Γ. For simplicity, we also write $H^{\sigma}_{\mathrm{pw}}(\partial\Omega) = H^{\sigma}(\partial\Omega)$.
For $\Gamma \subset \partial\Omega$, we define the subspace of functions with homogenous extension by

$$H^{\sigma}_{00}(\Gamma) = \Big\{ v = \tilde{v}\Big|_{\Gamma} : \ \tilde{v} \in H^{\sigma}(\partial\Omega), \ \mathrm{supp}\ \tilde{v} \subset \overline{\Gamma} \Big\}.$$

We denote the corresponding dual space by $H^{-1/2}(\Gamma) = \big(H^{1/2}_{00}(\Gamma)\big)^*$.

We can now formulate the final theorem of this section.

Theorem 2.1.19 (Trace theorem on polygons)**.** *Let $\sigma \in \{1/2, 3/2\}$ and $j \in \mathbb{Z}_+$ such that $j < \sigma$. Then the mapping*

$$\mathrm{tr}_j u = \frac{\partial^j u}{\partial \mathbf{n}^j}|_{\Gamma} : C^{\infty}(\overline{\Omega}) \to H^{\sigma-j}_{\mathrm{pw}}(\Gamma)$$

has a unique continuous and surjective extension

$$\mathrm{tr}_j : H^{\sigma+1/2}(\Omega) \to H^{\sigma-j}_{\mathrm{pw}}(\Gamma).$$

This theorem was proven in [77, Theorem 1.5.2.8] and in [78, Theorem 1.4.3].

2.2 Elliptic problems

In this section, we review some known regularity results for elliptic partial differential equations. We begin by briefly recalling the standard regularity theory for problems on domains with smooth boundaries. Further, we contrast these results with the regularity properties of solutions of elliptic problems on polygonal domains. As we shall see, the presence of corners in the domain introduces, in general, a certain type of singular functions to the solution. This, in turn, limits the Sobolev regularity, which can further influence the approximation properties of the finite element schemes discussed in Section 2.3.

Consider a model elliptic Poisson problem

$$-\Delta u = f \text{ in } \Omega, \quad u = 0 \text{ on } \partial\Omega, \tag{2.1}$$

where $\Omega \subset \mathbb{R}^2$ is an open, bounded and simply connected two-dimensional set.

Definition 2.2.1 (Weak solution)**.** *Let $f \in H^{-1}(\Omega)$. We say that $u \in H_0^1(\Omega)$ is the weak solution of the Poisson problem* (2.1) *if for all $v \in H_0^1(\Omega)$*

$$a(u,v) \coloneqq \int_\Omega \nabla u \cdot \nabla v = \left\langle f, v \right\rangle_{-1,1},$$

where $\left\langle \cdot, \cdot \right\rangle_{-1,1}$ is the duality pairing between H^{-1} and H_0^1.

Theorem 2.2.2 (Existence and uniqueness of the solution)**.** *The Poisson problem from Definition 2.2.1 has a unique solution $u \in H_0^1(\Omega)$.*

The proof of this theorem is based on the Riesz Representation Theorem or, alternatively, the more general Lax–Milgram Lemma. The application of them is justified by the fact that the bilinear form $a(\cdot,\cdot) : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ is bounded and coercive. A complete discussion of the theorem can be found in [71, Section 6.2, Theorem 3]. For the proof of the Lax–Milgram Lemma we send the reader to [71, Section 6.2, Theorem 1] and [131].

2.2.1 Elliptic problems on smooth domains

So far we only know that there exists a unique solution and we would now like to investigate its regularity properties. As we shall see, this depends on two factors: the regularity of the right-hand side forcing term f and the smoothness of the boundary $\partial\Omega$.

Theorem 2.2.3 (Boundary regularity)**.** *Let $m \in \mathbb{Z}_+$ and $\Omega \subset \mathbb{R}^2$ be a bounded domain with a C^{m+1} boundary $\partial\Omega$. Then, for each $f \in H^{m-1}(\Omega)$, there exists a unique weak solution of the boundary value problem* (2.2.1) *$u \in H_0^1(\Omega) \cap H^{m+1}(\Omega)$, which satisfies*

$$\|u\|_{H^{m+1}(\Omega)} \leq c\|f\|_{H^{m-1}(\Omega)}$$

for some constant $c > 0$ independent of f.

The proof of this theorem can be found in [71, Section 6.3, Theorem 5]. It is often called *the shift theorem*, since, upon the assumption of sufficient regularity of the domain, the regularity of the right-hand side f in Hilbert–Sobolev spaces translates directly to the regularity of the solution u in such spaces of the order higher by 2. Moreover, due to [71, Section 6.3, Theorem 6], the following also holds.

Remark 2.2.4. *Suppose that $f \in C^\infty(\Omega)$ for a bounded domain $\Omega \subset \mathbb{R}^2$ with a C^∞ boundary $\partial\Omega$. Then, the Poisson problem* (2.1) *has a unique smooth solution $u \in C^\infty(\overline{\Omega})$.*

2.2.2 Elliptic problems on polygonal domains

So far we have shown that on sufficiently regular domains the regularity of the solution of the Poisson problem (2.1) can be guaranteed by the smoothness assumptions on the forcing term f. However, as we shall see, this is not true anymore in the case of polygonal domains, which have only piecewise smooth boundaries.

Let $\Omega \subset \mathbb{R}^2$ be a bounded polygonal domain with N corners as introduced in Definition 2.1.11. In the presence of the corners in the computational domain a certain type of singular functions appear in the solution of the Poisson problem (2.1), see [77, Chapter 4], [99, Section 2.1.1] and [108], namely let for $i \in \mathbb{Z}_+$

$$s_{i,j}(r_j, \phi_j) = \eta(r_j) r_j^{\lambda_{i,j}} \sin(\lambda_{i,j}\phi_j), \quad \lambda_{i,j} = i\pi/\Theta_j. \tag{2.2}$$

Here, Θ_j is the angle of the j-th corner and (r_j, ϕ_j) are polar coordinates defined around it. Also, $\eta(r_j)$ is a smooth cut-off function equal to 1 for $r_j < \underline{r}$ and equal to 0, when $r_j > \overline{r}$, for some positive constants $\underline{r}, \overline{r}$ such that $\underline{r} < \overline{r}$. Let $\Omega_j = \Omega \cap B(\overline{r})$, where $B(\overline{r})$ denotes the ball of radius $\overline{r}$ centered at the corner, see Figure 2.1b. Note that $\overline{r}$ can be chosen so that the cut-off functions around different corners have disjoint supports Ω_j.

Later, we shall skip the index j, when it is clear from the context, which corner of the domain is considered or when it does not matter.

We would like now to summarize few important facts about the singular functions defined above and show, what kind of regularity properties they exhibit. We have for all $i \in \mathbb{Z}_+$

$$-\Delta s_i = 0, \text{ in } \Omega \cap \{r < \underline{r}\} \quad \text{and } s_i = 0 \text{ on } \partial\Omega.$$

Furthermore, $s_i \in H^s(\Omega)$ for $0 \leq s < 1 + \lambda_i$. If $\Theta > i\pi/k$ for a nonnegative integer k, then $s_i \notin H^{k+1}(\Omega)$. Consider, in particular, the case $\Theta > \pi$, so when the polygonal domain is not convex anymore and contains a re-entrant corner. Then we have $s_1 \notin H^2(\Omega)$.

Nevertheless, the regularity of the singular functions (2.2) can be precisely described using the framework of weighted Sobolev spaces introduced in Section 2.1.2.

Theorem 2.2.5 (Regularity of the singular functions in $H^{k+1}_\alpha(\Omega)$)**.** *Let $\alpha > k + \lambda_i$. Then the singular functions* (2.2) *satisfy*

$$s_i \in H^{k+1}_\alpha(\Omega).$$

The regularity description in terms of weighted Sobolev spaces is more precise, as it shows the locality of the singular behaviour. When considered in some fixed positive distance from the corner, functions s_i are in fact smooth and $s_i \in C^\infty(U)$, where $U \subset \Omega \setminus B(\epsilon)$ for some $\epsilon > 0$.

Now, we would like to recall some important results describing the regularity of the solutions of the Poisson problem (2.1) on polygonal domains.

Theorem 2.2.6 ($V^{k+1,p}_\beta(\Omega)$ regularity)**.** *Let k be a positive integer, $1 \leq p < \infty$ and let us for all $j = 1, \ldots, N$ take the weight coefficients β_j such that $k + 1 - \beta_j - 2/p \in \big(l_j\pi/\Theta_j, (l_j+1)\pi/\Theta_j\big)$ for some positive integers l_j. Let also $f \in H^{-1}(\Omega) \cap V^{k-1,p}_{\vec{\beta}}(\Omega)$. Then, there exists a unique solution $u \in H^1_0(\Omega)$. Moreover, the solution admits the following expansion*

$$u = U + \sum_{j=1}^{N} \sum_{i=1}^{l_j} k_{i,j} s_{i,j},$$

where $k_{i,j} \in \mathbb{R}$. The regular part of the solution satisfies $U \in V^{k+1,p}_{\vec{\beta}}(\Omega)$. Furthermore, for some $c > 0$ independent of f, the following a priori estimates hold

$$\sum_{j=1}^{N} \sum_{i=1}^{l_j} |k_{i,j}| + \|U\|_{V^{k+1,p}_{\vec{\beta}}(\Omega)} \leq c \|f\|_{V^{k-1,p}_{\vec{\beta}}(\Omega)}.$$

First investigations of the kind presented in the above theorem, together with the splitting of the solution into a singular and regular part and the use of weighted Sobolev spaces, can be found in [97]. The precise formulation and the proof of the above theorem in a general setting of $V_{\alpha}^{k,p}(\Omega)$ spaces comes from [99, Theorem 1.4.4, Theorem 2.6.1] and an extensive discussion in [99, Section 2.6.1]. The bounds on the norms can be found in [99, Theorem 1.4.1].

We would like to formulate two important implications of the aforementioned theorem, which will be used later in this work.

Theorem 2.2.7 ($H_{\alpha}^{k+1}(\Omega)$ regularity)**.** *Let us take some small $\epsilon > 0$ and let k be a positive integer. Furthermore, let us for all $j = 1, \ldots, N$ take the weight coefficients $\max(0, k - \lambda_{1,j} + \epsilon) \leq \alpha_j < k$ and $\tilde{\alpha}_j = \alpha_j - k + 1$. Let also $f \in H_{-\vec{\tilde{\alpha}}}^{k-1}(\Omega)$. Then, there exists a unique solution $u \in H_0^1(\Omega) \cap H_{\vec{\alpha}}^{k+1}(\Omega)$. Moreover, the solution admits the following expansion*

$$u = U + \sum_{j=1}^{N} \sum_{i:\lambda_{i,j}<k+\alpha_j} k_{i,j} s_{i,j},$$

where $k_{i,j} \in \mathbb{R}$. The regular part of the solution satisfies $U \in H_{-\vec{\tilde{\alpha}}}^{k+1}(\Omega)$. Furthermore, for some $c > 0$ independent of f, the following a priori estimates hold

$$\|u\|_{k+1,\vec{\alpha}} \leq c\|f\|_{k-1,-\vec{\tilde{\alpha}}}, \quad \text{and} \quad \sum_{j=1}^{N} \sum_{i:\lambda_{i,j}<k+\alpha_j} |k_{i,j}| + \|U\|_{k+1,-\vec{\tilde{\alpha}}} \leq c\|f\|_{k-1,-\vec{\tilde{\alpha}}}.$$

The formulation presented in Theorem 2.2.7 involves more restrictive assumption because the higher regularity of the regular part of the expansion is needed in the further investigations of the finite element methods. It is clear that due to the presence of the corners in the domain the Shift Theorem of the type Theorem 2.2.3 does not hold in standard Sobolev spaces. However, it can be shown that for the solution u to be in the space $H_{\vec{\alpha}}^{k+1}(\Omega)$ it is enough to assume that $f \in H_{\vec{\alpha}}^{k-1}(\Omega)$. This can be regarded as a version of the Shift Theorem in the weighted Sobolev spaces, where a shift in the order of spaces appears analogously as in Theorem 2.2.3. Similar results of this type, with the right-hand side f in certain fractional Sobolev space, can be found in [17] and [94].

Remark 2.2.8. *In the case of $k = 1$, the singular functions in the expansion above appear only around the corners of angle $\Theta_j > \pi$ and then, in general, $u \in H^2(\Omega_j)$ cannot be obtained. However, in the case of convex domains a Shift Theorem of the kind presented in Theorem 2.2.3 is obtained upon the choice of $m = 1$ regardless of*

the non-smoothness of the boundary and around corners of size $\Theta_j \leq \pi$ *we can expect* $u \in H^2(\Omega_j)$.

Apart from summarizing the regularity properties of the solutions of the Poisson problem on domains with corners, the above theorems provide additionally a splitting formula of the solution into regular and singular parts. This splitting will play an important role in the further investigations of finite element approximations on polygonal domains. The coefficients $k_{i,j}$ in these expansions are the so-called *stress-intensity factors* and the formulas for computing them can be found in [77, Section 8.4].

Let us define the dual singular functions of the form similar to (2.2)

$$s_{-i,j}(r_j, \phi_j) = -\eta(r_j) r_j^{-\lambda_{i,j}} \sin(\lambda_{i,j}\phi_j). \tag{2.3}$$

The stress-intensity factors can be then computed using the following representation

$$k_{i,j} = -\frac{1}{i\pi} \int_\Omega f s_{-i,j} + u \Delta s_{-i,j}. \tag{2.4}$$

2.3 Finite element methods

In this section, we discuss the finite element method for approximating the solutions of elliptic problems. We focus our attention on the model Poisson problem (2.1) defined on polygonal domains, see Definition 2.1.11. In Section 2.3.1, we present an abstract framework for the finite element and define the k-th order Lagrange elements used further in this work. Then, we introduce the standard finite element method in Section 2.3.2 and show that the presence of corners in the computational domain diminishes the convergence properties of the method not only in their vicinity but also in an arbitrary positive distance from them. This phenomenon, known as the pollution effect, is the motivation for introducing the energy-corrected finite element method, which is thoroughly reviewed in Section 2.3.3.

Section 2.3.2 is supposed to serve as a brief introduction to the theory of finite element methods. Together with the main results, we also introduce the notation used further in this work. It is worth noting that we do not limit our interest to the standard piecewise linear discretization, but present more general piecewise polynomial methods. Here, we are not aiming at providing a thorough study of finite element methods in general, as numerous great works in this field already exist. For more exhaustive discussions of the finite element methods for elliptic equations, we would like to refer the reader to [16, 23, 38, 56, 151].

After the introduction of the standard finite element, in Section 2.3.3 we review the known results in the theory of energy-correction schemes, which were developed to improve the convergence properties of the finite element for elliptic problems on polygonal domains.

2.3.1 Introduction to the finite element methods

We begin by defining the finite elements in a general, abstract setting following [38, 55, 56]. The introduction of the abstract finite element framework will later allow us to treat also the enriched spaces in Chapter 4 and the finite element discretization on the one-dimensional boundary of the domain in Chapter 5.

Definition 2.3.1 (Finite element)**.** *Let*

- $K \subset \mathbb{R}^d$ *be a bounded, closed set with a piecewise smooth boundary, which is star-shaped with respect to some ball* $B \subset K$*,*
- $\mathcal{P}$ *be a finite dimensional space of functions on* K*,*

- $\mathcal{N} = \{N_1, \ldots, N_J\}$ *be a basis of* $\mathcal{P}^*$.

We call the triplet $(K, \mathcal{P}, \mathcal{N})$ *a finite element. Furthermore, we call* K *the element or element domain, elements of* $\mathcal{P}$ *are called shape functions and* $\mathcal{N}$ *is the set of nodal variables.*

The dual basis $\mathcal{N}$ defines also the so called *nodal basis* $\mathcal{B} = \{\widehat{\phi}_l\}_{j=1}^J$ of $\mathcal{P}$ of functions satisfying $N_j(\widehat{\phi}_l) = 1$ if $j = l$ and $N_j(\widehat{\phi}_l) = 0$ otherwise. This allows us to define the interpolation operator on the element K.

Definition 2.3.2 (Interpolant)**.** *Let* $\{\widehat{\phi}_l\}_{j=1}^J$ *be the nodal basis of* $\mathcal{P}$*. If the actions of* N_i *on function* v *are well-defined, then we define the local interpolant on the element* K *as*

$$I_K v = \sum_{j=1}^{J} N_j(v)\widehat{\phi}_j.$$

Let now $P_k(K)$ denote the space of polynomials of order not larger than k on K. We now state a classical result of Bramble and Hilbert [34], which plays a crucial role in proving the interpolation error estimates.

Lemma 2.3.3 (Bramble–Hilbert Lemma)**.** *Let* $K \subset \mathbb{R}^d$ *be as in Definition 2.3.1 and* $1 \leq p \leq \infty$*. There exists a constant* $c > 0$ *independent of* p *and* u *such that for any* $u \in W^{m,p}(K)$ *there exists a polynomial* $p_{l-1} \in P_{l-1}(K)$ *such that for all* $m \in \{0, \ldots, l\}$

$$|u - p_{l-1}|_{W^{m,p}(K)} \leq c|u|_{W^{l,p}(K)}.$$

The first formulation of this lemma can be found in [34], where the domain K is assumed to satisfy a strong cone property. More general versions, together with a constructive proof and relaxed assumptions, can be found in [66]. For a complete discussion of the result we refer to [3, 112].

Having introduced the Bramble–Hilbert lemma, we can state the interpolation error estimate on the reference element $K \subset \mathbb{R}^d$

Theorem 2.3.4 (Interpolation on the reference element)**.** *Suppose that for some* $l \in \mathbb{Z}_+$ *we have* $P_{l-1}(K) \subset \mathcal{P}$ *and* $\mathcal{N} \subset \big(C^r(K)\big)^*$*. Let* $1 \leq p, q \leq \infty$ *and* $m \in \mathbb{Z}_+ \cup \{0\}$ *be such that* $l > r + d/p$*,* $m < l$ *and*

$$W^{l,p}(K) \hookrightarrow W^{m,q}(K).$$

Then, there exists a constant $c = c(p, q, K) > 0$ *such that for all* $v \in W^{l,p}(K)$

$$|v - I_K v|_{W^{m,q}(K)} \leq c|K|^{1/q - 1/p}|v|_{W^{l,p}(K)}.$$

The proof of the interpolation error estimate can be found in [38, Theorem 4.4.4], when $p = q$, and, following a constructive approach together with the study of the constants in the estimates, in [66, 67]. These proofs are based on the Deny–Lions Lemma, see [66]. For an exhaustive study of interpolation error estimates we refer the reader to [3]. Similar result, incorporating also fractional Sobolev spaces, was proven in [139] under stronger assumptions of convexity of the element K and $\mathcal{P} = P_{l-1}(K)$. Then the seminorm $|v|_{W^{l,p}(K)}$ can be replaced with a Sobolev–Slobodeckij seminorm introduced in Definition 2.1.7.

Now, we would like to move our considerations to a more particular case of k-th order Lagrange elements on triangles, where $k \in \mathbb{Z}_+$. In order to do this, we first introduce the reference triangle

$$\widehat{T} = \big\{\widehat{x}_1 \geq 0\big\} \cap \big\{\widehat{x}_2 \geq 0\big\} \cap \big\{1 - \widehat{x}_1 - \widehat{x}_2 \geq 0\big\},$$

where $\widehat{\mathbf{x}} = \big(\widehat{x}_1, \widehat{x}_2\big)^T \in \mathbb{R}^2$.

In the reference triangle $\widehat{T}$ we choose $(k+1)(k+2)/2$ nodal points for $i, j = 1, \ldots, k+1$ and $1 \leq i + j \leq k+2$,

$$\mathbb{X}^k = \big\{\widehat{\mathbf{x}}_{i,j} = ((i-1)/k, (j-1)/k)^T\big\}_{1 \leq i+j \leq k+2}.$$

In particular, all the points are pairwise distinct, 3 of them are placed in the vertices of $\widehat{T}$ and there are $k-1$ in the interior of each of the triangle's edges.

Notice that $\big(C(\widehat{T})\big)^* \subset \big(P_k(\widehat{T})\big)^*$. We define the nodal viariables $N^{\mathcal{L}}_{i,j} \in \big(C(\widehat{T})\big)^*$ by

$$N^{\mathcal{L}}_{i,j}(v) = v(\widehat{\mathbf{x}}_{i,j}), \quad \text{for all } v \in C(\widehat{T}),$$

where $\widehat{\mathbf{x}}_{i,j} \in \mathbb{X}^k$. We also denote $\mathcal{N}^{\mathcal{L}}_k = \big\{N^{\mathcal{L}}_{i,j}\big\}_{1 \leq i+j \leq k+2}$.

Definition 2.3.5 (Lagrange Finite Element)**.** *We call the triplet* $\big(\widehat{T}, P_k(\widehat{T}), \mathcal{N}^{\mathcal{L}}_k\big)$ *the Lagrange Finite Element.*

A discussion of Triangular Lagrange Elements can be found in [38, Section 3.2]. The following inverse inequality on the reference triangle can be found in [32, (6.12)].

Lemma 2.3.6 (Inverse inequality)**.** *Let* $0 \leq l \leq m \leq k$*. For the Lagrange finite element, the following inverse inequality holds for some* $c > 0$

$$|v|_{H^m(\widehat{T})} \leq c|v|_{H^l(\widehat{T})}, \quad \textit{for all } v \in P_k(\widehat{T}).$$

The choice of the nodal variables of the Lagrange Finite Element naturally defines the nodal basis of the $P_k(\widehat{T})$ space. We denote the elements of this basis by $\widehat{\phi}_i^{\widehat{T}}$.

Having defined the finite element on the reference triangle, we are now in the position to introduce the discretization of an arbitrary computational domain of interest. Let $\Omega \subset \mathbb{R}^2$ be a polygonal domain as in Definition 2.1.11 and let $\mathcal{T}_H$ be its triangulation, so $\overline{\Omega} = \bigcup_{T\in\mathcal{T}_H} T$. Triangles $T \subset \mathbb{R}^2$ are closed sets with pairwise disjoint interiors and without corners lying on the edges of any other triangle. Assume also that $\{\mathcal{T}_h\}_{h>0}$ is a series of conforming meshes constructed by a uniform refinement of $\mathcal{T}_H$. For each of the triangles T and mesh $\mathcal{T}_h$ we define

$$h_T = \sup_{\mathbf{x},\mathbf{y}\in T} |\mathbf{x}-\mathbf{y}|, \quad \text{and} \qquad h = \max_{T\in\mathcal{T}_h} h_T.$$

We call h a mesh size, which is the largest of the diameters of the elements in the triangulation. Notice that due to the assumption of the uniform refinement

$$ch \leq h_T \leq h, \quad \text{for all } T \in \mathcal{T}_h.$$

for some $c > 0$ independent of the refinemenet level.

The computational meshes defined above are quasi-uniform, which means that the minimal angle of each triangulation is uniformly bounded from below and that they satisfy the following property

$$\frac{\max_{T\in\mathcal{T}_h} |T|}{\min_{T\in\mathcal{T}_h} |T|} \leq c \tag{2.5}$$

for some $c > 0$ independent of the refinement level. In this thesis, unless stated otherwise, $c > 0$ will denote a generic constant independent of the mesh size h. Furthermore, in many proofs its value might change between the steps.

Each polygon can be exactly represented by a properly chosen set of triangles, which is not the case for domains with curved boundaries. The imperfections in the representation of the domains of the latter kind by a chosen triangulation or the use of curved elements can be a source of additional error, and, in general, needs to be thoroughly studied. Investigations of finite element methods on curved domains can be found, among others, in [38, Chapter 10].

For each element $T \in \mathcal{T}_h$, there exists an affine bijection $F_T : \widehat{T} \to T$. The basis functions $\widehat{\phi}_i^{\widehat{T}}$ of $P_k(\widehat{T})$ defined on the reference triangle, together with the afine transformations on all $T \in \mathcal{T}_h$, define the following functions on the domain Ω

$$\phi_i(\mathbf{x}) = \widehat{\phi}_l^{\widehat{T}}(\widehat{\mathbf{x}}), \quad \text{where } T \ni \mathbf{x} = F_T\big(\widehat{\mathbf{x}}\big), \text{ and } \widehat{\mathbf{x}} \in \widehat{T}.$$

The aforementioned affine transformations F_T identify each nodal point in the triangles $T \in \mathcal{T}_h$ with the corresponding nodal point in the reference triangle $\widehat{T}$. Let us denote all nodal points in the domain $\overline{\Omega}$ as

$$\mathcal{X} = \left\{ \mathbf{x} \in \overline{\Omega} : \ \mathbf{x} = F_T(\widehat{\mathbf{x}}_j) \text{ for some } T \in \mathcal{T}_h \text{ and a nodal point } \widehat{\mathbf{x}}_j \in \widehat{T} \right\}. \tag{2.6}$$

We number these points so that for some $K, L \in \mathbb{Z}_+$, points $\{\mathbf{x}_i\}_{i=1}^{K}$ belong to the interior of Ω, and the remaining $\{\mathbf{x}_i\}_{i=K+1}^{K+L}$ points lie on the boundary $\partial\Omega$. Note that for each point $\mathbf{x}_i \in \mathcal{X}$ there exists the corresponding basis function ϕ_i.

We define the discrete space $S_h^k(\Omega)$ introduced as the space of globally continuous, piecewise polynomial finite elements by

$$S_h^k(\Omega) = \left\{ v \in C(\overline{\Omega}) : \ v|_T \in P_k(T), \text{ for all } T \in \mathcal{T}_h \right\}. \tag{2.7}$$

Furthermore, let V_h^k be the subspace of finite element functions disappearing on the boundary

$$V_h^k(\Omega) = S_h^k(\Omega) \cap H_0^1(\Omega) \subset H_0^1(\Omega). \tag{2.8}$$

In the cases, when the choice of the domain Ω is clear from the context, we also write simply S_h^k and V_h^k. When $k = 1$, so in the case of piecewise linear finite element functions, we use the notation $S_h^1 = S_h$ and $V_h^1 = V_h$. The introduced discretisation spaces are indeed linear spaces spanned by the nodal basis functions associated with the nodal points.

$$V_h^k = \operatorname{span}\{\phi_i\}_{i=1}^{K}, \quad S_h^k = \operatorname{span}\{\phi_i\}_{i=1}^{K+L}. \tag{2.9}$$

The following inverse inequality is a fundamental property of the discrete spaces and follows from the equivalence of norms in the finite dimentional spaces, see [32, Section 6.8] for more details.

Lemma 2.3.7 (Global inverse inequality)**.** *Let $0 \le l \le m \le k$. There exists a constant $c > 0$ such that*

$$|v_h|_{H^m(\Omega)} \le c_i h_{\min}^{l-m} |v_h|_{H^l(\Omega)}, \quad \textit{for all } v_h \in S_h,$$

where $h_{\min} = \min_{T \in \mathcal{T}_h} h_T$. In particular, for uniformly refined meshes, we have

$$|v_h|_{H^m(\Omega)} \le c_i h^{l-m} |v_h|_{H^l(\Omega)}, \quad \textit{for all } v_h \in S_h.$$

Since functions ϕ_i are defined using the nodal basis functions, we should also refer to them as the nodal basis.

Definition 2.3.8 (Nodal interpolation)**.** *Let $\omega \subset \Omega$ be an open set. For any $k \in \mathbb{Z}_+$ we define the nodal interpolation $I_h^k : C(\overline{\omega}) \to S_h^k$ by*

$$I_h^k v = \sum_{i:\, \mathbf{x}_i \in \mathcal{X} \cap \overline{\omega}} v(\mathbf{x}_i)\phi_i.$$

Now, we would like to summarize the approximation properties of the nodal interpolation operator. These will depend not only on the order of the polynomials in the finite element basis but also on the regularity of the interpolated function. We begin by stating the result on one of the elements in the triangulation $\mathcal{T}_h$ only. This estimate is of a big importance, as all the interpolation error estimates used later follow from this one.

Lemma 2.3.9 (Local interpolation error estimates)**.** *Let $m, k \in \mathbb{Z}_+$, $m < s < \infty$, and $1 \leq p, q \leq \infty$. Assume also that $W^{s,p}(T) \hookrightarrow W^{m,q}(\Omega)$. Then, for some $c > 0$ and the nodal interpolation $I_h^k : C(\overline{T}) \to P_k(T)$ the following estimate holds for all $v \in W^{s,p}(T)$*

$$|v - I_h^k v|_{W^{m,p}(T)} \leq c h^{2/q-2/p+s-m} |v|_{W^{s,p}(T)}.$$

In the case $s \in \mathbb{Z}_+$, this result is a simple consequence of Theorem 2.3.4 upon observation that $|T| \leq ch^2$ for all $T \in \mathcal{T}_h$. The scaling $h^{2/q-2/p}$ comes from the Jacobian of the affine transformation between T and $\widehat{T}$. This is a classical result and can be found in [3] and [38, Theorem 4.4.20]. The proof in a more general case of fractional Sobolev spaces of order s follows an analogous path and can be found in [139].

Consequently, the application of the local interpolation error estimates to each of the elements in the triangulation $T \in \mathcal{T}_h$ leads to the following global interpolation error estimate.

Theorem 2.3.10. *Let $k \in \mathbb{Z}_+$ and $v \in H^s(\Omega)$ for some $1 < s \leq k+1$. Then the following interpolation error estimates hold for some $c > 0$ and the nodal interpolant $I_h^k : C(\overline{\Omega}) \to S_h^k$*

$$\|v - I_h^k v\|_0 \leq c h^s |v|_{H^s(\Omega)}, \text{ and } \quad \|\nabla(v - I_h^k v)\|_0 \leq c h^{s-1} |v|_{H^s(\Omega)}.$$

Note that due to Theorem 2.1.9, for $s > 1$ the functions $v \in H^s(\Omega)$ are continuous and hence, the interpolation $I_h^k v$ is well-defined.

As stated in Section 2.2, the regularity properties of the solution of the elliptic problem (2.1) can be efficiently described in the framework of weighted Sobolev

spaces. As these spaces allow for incorporating the knowledge about the loss of regularity in the neighbourhood of corners in the polygonal domains, we would like to use them also to describe the approximation properties of the finite element spaces. The following result can be found in [86, Lemma 11].

Theorem 2.3.11. *Let $k \in \mathbb{Z}_+$ and $v \in H^{k+1}_\alpha(\Omega)$ for some $\alpha < k$. Furthermore, let $\alpha - k \leq \beta \leq \alpha$. Then for the nodal interpolant $I^k_h : C(\overline{\Omega}) \to S^k_h$ the following error estimates hold for some $c > 0$*

$$\|v - I^k_h v\|_\beta \leq c h^{k+1+\beta-\alpha} |v|_{k+1,\alpha}, \text{ and } \quad \|\nabla(v - I^k_h v)\|_\beta \leq c h^{k+\beta-\alpha} |v|_{k+1,\alpha}.$$

2.3.2 Standard finite element method for elliptic problems

So far we have described the properties of the finite element spaces, which are independent of the potential equations, whose solutions we would like to approximate. Let us now consider the Poisson problem (2.1) in the variational formulation from Definition 2.2.1.

2.3.2.1 Finite element for elliptic problems

Definition 2.3.12. *We call $u_h \in V_h^k$ the finite element approximation of* (2.1) *if*

$$a(u_h, v_h) = \langle f, v_h \rangle_{-1,1}, \quad \textit{for all } v_h \in V_h^k. \tag{2.10}$$

From now on we will assume that $f \in L^2(\Omega)$, so the dual pairing $\langle f, v_h \rangle_{-1,1}$ is equivalent to the $L^2(\Omega)$ inner product of f and v_h.

Let $u \in H_0^1(\Omega)$ be the solution of (2.1), then the equation (2.10) can be reformulated as

$$a(u_h, v_h) = a(u, v_h), \quad \text{for all } v_h \in V_h^k. \tag{2.11}$$

This defines a linear operator $R_h^k : H_0^1(\Omega) \to V_h^k$ by $R_h^k u = u_h$, which we call *the Ritz projection*. In the case of the piecewise linear approximation, we also skip the superscript and write $R_h^1 u = R_h u$.

One of the fundamental properties of the finite element discretisation (2.10) is the so-called Galerkin orthogonality.

Theorem 2.3.13 (Galerkin orthogonality)**.** *Let $u_h \in V_h^k$ be the finite elmenent approximation defined in* (2.10) *of the solution $u \in H_0^1(\Omega)$ of the elliptic problem* (2.1)*. Then the following identity holds*

$$a(u - u_h, v_h) = 0, \quad \textit{for all } v_h \in V_h^k.$$

This means that the finite element approximation (2.10) is the orthogonal projection of the solution u onto the discrete space V_h^k with the inner product induced by the bilinear form $a(\cdot, \cdot)$.

Problem (2.10) can be equivalently rewritten as a linear system of equations, since V_h^k is a K-dimensional linear space. Let u_h be a linear combination of the basis functions in the space V_h^k

$$u_h = \sum_{i=1}^{K} u_i \phi_i.$$

Let also

$$\mathbf{M} = \Big[\big\langle \phi_i, \phi_j \big\rangle\Big]_{i,j=1}^{K}, \quad \mathbf{S} = \Big[a(\phi_i, \phi_j)\Big]_{i,j=1}^{K}, \quad \mathbf{F} = \Big(\big\langle f, \phi_j \big\rangle\Big)_{j=1}^{K}. \tag{2.12}$$

We call $\mathbf{M}$ and $\mathbf{S}$ the *mass* and *stiffness* matrices respectively. The problem introduced in Definition 2.3.12 can be then understood as finding $\mathbf{U} = \big(u_i\big)_{i=1}^{K} \in \mathbb{R}^K$ such that $\mathbf{SU} = \mathbf{F}$. Note that both mass and stiffness matrices are not only symmetric and positive definite but, due to the fact that the basis functions ϕ_i of the space V_h^k have local supports, they are also sparse.

We now move to summarizing the main approximation results of the finite element method defined in Definition 2.3.12.

Theorem 2.3.14 (Best approximation property)**.** *Let $u \in H_0^1(\Omega)$ be the weak solution of* (2.1) *with $f \in L^2(\Omega)$ and $u_h \in V_h^k$ be the solution of* (2.10)*. Then, the following best approximation property holds for some $c > 0$*

$$\|\nabla(u - u_h)\|_0 \leq c \inf_{v_h \in V_h^k} \|\nabla(u - v_h)\|_0.$$

In particular we also have

$$\|\nabla(u - u_h)\|_0 \leq c\|\nabla(u - I_h^k u)\|_0.$$

This means that, up to a multiplicative constant, the finite element method defined above yields the best possible approximation among all the possible discrete functions in V_h^k. Moreover, it approximates the real solution at least as well as the nodal interpolation, when measured in the $H_0^1(\Omega)$ seminorm. Note that the nodal interpolant $I_h^k u$ is well-defined, since $u \in H^s(\Omega)$ for some $s > 1$, see Theorem 2.2.7 and $H^s(\Omega) \hookrightarrow C(\overline{\Omega})$ due to the Sobolev Embedding Theorem 2.1.9.

The convergence order of the finite element scheme depends on the regularity of the solution, as well as on the order of the discretisation space.

Theorem 2.3.15. *Let k be a positive integer, $0 < s \leq k$, and $u \in H_0^1(\Omega) \cap H^{s+1}(\Omega)$ be the solution of* (2.1)*. Furthermore, let $u_h \in V_h^k$ be the finite element approximation* (2.10)*. Then the following error estimate is true*

$$\|\nabla(u - u_h)\|_0 \leq ch^s \|u\|_{H^{s+1}(\Omega)}.$$

The estimate in the $H_0^1(\Omega)$ semi-norm follows directly from the best approximation property in Theorem 2.3.14 and the interpolation error from Theorem 2.3.10.

2.3.2.2 Finite element on domains with corners

From now on, we would like to focus our attention on the influence that the presence of corners in the computational domain has on the finite element approximation of the solution of (2.1). To simplyfy the notation, we consider the singularites appearing around only one of the domain's corners. This is equivalent to the following technical assumption.

Assumption 2.3.16. *Let $k \in \mathbb{Z}_+$ be the chosen order of polynomials in the finite element space. At least one of the following holds for every corner except the first one in the computational domain, namely Θ_j, $j = 2, \ldots, N$*

- $\Theta_j \leq \pi/k$.
- *The singular functions in the expansion given by Theorem 2.2.7 disappear, namely $k_{i,j} = 0$ for all $i \in \mathbb{Z}_+$ such that $\frac{i\pi}{\Theta_j} \leq k$.*

This assumption means that the singular functions defined only around one of the corners in the computational domain Ω influence the regularity of the solution in the sense of the Theorem 2.2.6. It is important to notice that all the results presented later naturally extend to the more general case of multiple relevant corners in the domain and that the assumption above is introduced only to simplify the notation. From now on, by Θ be mean the size of the relevant corner and we set $\lambda_i = i\pi/\Theta$. Furthermore, we assume that the considered corner is located at the origin of the Euclidean space.

To understand precisely the approximation properties of the finite element method for the Poisson problem on polygonal domains, we need to first investigate the behaviour of the nodal interpolant of the singular functions (2.2).

Lemma 2.3.17 (Interpolation of the singular functions)**.** *Let s_i be the singular function as defined in* (2.2)*. Let us also take $\epsilon > 0$, and $\alpha = \max(0, k - \lambda_i + \epsilon)$. Then the following interpolation error estimates hold*

$$\|s_i - I_h^k s_i\|_0 \leq c h^{\min(k,\lambda_i)+1}, \quad \|\nabla(s_i - I_h^k s_i)\|_0 \leq c h^{\min(k,\lambda_i)},$$

Moreover,

$$\|s_i - I_h^k s_i\|_\alpha \leq c h^{k+1}, \quad \|\nabla(s_i - I_h^k s_i)\|_\alpha \leq c h^k.$$

We know that $s_i \in H^{1+\lambda_i-\epsilon}(\Omega)$ and $s_i \in H^{k+1}_\alpha(\Omega)$ for $\epsilon > 0$ and $\alpha = k - \lambda_i + \epsilon$. The direct application of the interpolation error estimates stated Theorem 2.3.10 and Theorem 2.3.11 yields the suboptimal error estimates

$$\|s_i - I_h^k s_i\|_0 \leq ch^{\min(k,\lambda_i)+1-\epsilon}, \quad \|\nabla(s_i - I_h^k s_i)\|_0 \leq ch^{\min(k,\lambda_i)-\epsilon}.$$

The ϵ-dependence can be eliminated by a direct computation, as done in [28]. Moreover, the interpolation error estimates in Lemma 2.3.17 are optimal in the order of h.

The estimates in the weighted norms follow from Theorem 2.3.11. The constants in Lemma 2.3.17 depend only on the regularity of the considered singular function, so their order i and the size of the corner, around which it is defined.

Lemma 2.3.17 guarantees better approximation order of the interpolation operator, when applied to the singular functions, than predicted by Theorem 2.3.11. The following is a direct consequence of the linearity of the Poisson problem (2.1).

Corollary 2.3.18. *Let the assumptions of Theorem 2.2.6 be satisfied. Let us also take $\epsilon > 0$, and $\alpha = \max(0, k - \lambda_i + \epsilon)$. The following interpolation error etimates for the solution u of the Poisson problem* (2.1) *hold*

$$\|u - I_h^k u\|_0 \leq ch^{\min(k,\lambda_i)+1}, \quad \|\nabla(u - I_h^k u)\|_0 \leq ch^{\min(k,\lambda_i)},$$

Moreover,

$$\|u - I_h^k u\|_\alpha \leq ch^{k+1}, \quad \|\nabla(u - I_h^k u)\|_\alpha \leq ch^k.$$

From these interpolation error estimates and the best approximation property, see Theorem 2.3.14, we immediately conclude the following error estimate for the Ritz projection (2.11) of the singular functions.

Corollary 2.3.19 (Approximation of the singular functions)**.** *Let $k \in \mathbb{Z}_+$ be the order of the finite element discretization and let s_i be the i-th singular function* (2.2) *around an angle of size Θ. Then for some $c > 0$*

$$\|\nabla(s_i - R_h^k s_i)\|_0 \leq ch^{\min(k,\lambda_i)}. \tag{2.13}$$

The proof of the sharpness of this result in the case $k = 1$ can be found in [128]. For higher-order elements, we verified the sharpness of this estimate numerically, see Table 2.1. The linearity of the Poisson problem (2.1) yields now the following finite element error estimate in the energy norm.

Theorem 2.3.20 (Error estimate in $H_0^1(\Omega)$)**.** *Let k be a positive integer and $u \in H_0^1(\Omega)$ be the solution of (2.1) with $f \in H^{k-1}(\Omega)$. Furthermore, let $u_h \in V_h^k$ be the finite element approximation (2.10). Then the following error estimate holds true*

$$\|\nabla(u - u_h)\|_0 \leq ch^{\min(k,\lambda_1)}\|f\|_{H^{k-1}(\Omega)}.$$

This is a direct consequence of the regularity result in Theorem 2.2.6, the best approximation property from Theorem 2.3.14 and of the approximation of the singular functions stated in Lemma 2.3.17.

It is worth noting that the increase of the order of polynomials in the finite element space does not, in general, improve the convergence order of the finite element method in the presence of corners in the computational domain. This is due to the influence, which the corners have on the regularity of the solution of the elliptic problems, see Theorem 2.3.15.

The loss of regularity due to the presence of the corners in the domain has a more severe effect on the finite element approximation, when measured in the $L^2(\Omega)$ norm, in which the best approximation property similar to the one stated in Theorem 2.3.14 does not hold. Let $u \in H_0^1(\Omega)$ and $u_h \in V_h^k$ be the solutions of the Poisson problem (2.1) and its finite element approximation (2.10) respectively. Then

$$\begin{aligned}\|\nabla(u - u_h)\|_0^2 &= |a(u - u_h, u - u_h)| \\ &= |a(u, u) - a(u_h, u_h)| \leq \|u - u_h\|_\alpha \|f\|_{-\alpha}. \end{aligned} \tag{2.14}$$

So the order of the finite element approximation in the weighted $L^2(\Omega)$ is, regardless of the weight, bounded from below by the approximation measured in the squared energy norm. This means that due to the reduced regularity of the solution of (2.1) also the convergence order of the finite element approximation (2.10) is not optimal in the sense of the interpolation error, even if measured in the weighted norm. This behavior is known as the so-called pollution effect, see [27, 43] and [38, Section 5.8.4].

Theorem 2.3.21 (Pollution effect)**.** *Let k be a positive integer. Let also $u \in H_0^1(\Omega) \cap H_\alpha^{k+1}(\Omega)$ be the solution of (2.1) with $f \in H_{-\alpha}^{k-1}$ for some $k - \lambda_1 < \alpha < k$. The finite element approximation $u_h \in V_h^k$ defined in (2.10) satisfies for some $c > 0$*

$$\|u - u_h\|_\alpha \geq c\|\nabla(u - u_h)\|_0^2.$$

Note that due to the sharpness of the estimate presented in Corollary 2.3.19, the finite element method yields lower convergence than the interpolation operator, see Corollary 2.3.18.

Remark 2.3.22. *Assume that the stress-intensity factor k_1 in the expansion given in Theorem 2.2.7 satisfies $k_1 \neq 0$. Then, due to the sharpness of the error estimates from Corollary 2.3.19, we expect the following convergence orders of the standard finite element method defined in* (2.10)

$$\|u - u_h\|_\alpha \geq ch^{2\lambda_1}, \quad \text{and} \quad \|u - u_h\|_0 \geq ch^{2\lambda_1}.$$

Furthermore, let $\widetilde{\Omega} \subset \Omega$ such that dist$(\tilde{\Omega}, \mathbf{0}) > \delta > 0$. Then

$$\|u - u_h\|_{L^2(\widetilde{\Omega})} \geq ch^{2\lambda_1}. \tag{2.15}$$

The pollution effect implies also that the convergence of the scheme measured in the $L^2(\Omega)$ norm also in a fixed distance away from the corner is suboptimal. The presence of the corners in the computational domain not only diminishes the regularity properties of the solutions of the elliptic problems but also negatively influences the convergence properties of their finite element approximation. As we see, regardless of the order of the finite element discretisation and the regularity of the right-hand side in (2.1), the order of approximation of the finite element scheme is significantly worse compared to the interpolation error stated in Theorem 2.3.10 in all subparts of the domain Ω.

2.3.2.3 Gradually refined triangulations

One of the most commonly used techniques for improving the accuracy of the finite element schemes on polygonal domains is the so-called mesh grading, see [5, 6, 9, 16, 37, 49, 144]. For the sake of completeness and comparison, we briefly introduce this method. In Section 2.3.1, we presented the triangulation of the polygonal domain Ω based on a uniform refinement. As opposed to the standard methods involving the uniform triangulations introduced there, the mesh grading requires the refinement of the computational mesh in the surrounding of the singular corner located at the origin. We define such meshes following [132, Section 3.2.1].

Definition 2.3.23 (Graded mesh)**.** *Let $0 < h < 1$ be the global mesh parameter and let $R > 0$. Furthermore, let*

$$r_T = \inf_{\mathbf{x} \in T} |\mathbf{x}|$$

be the distance of the triangulation's element from the considered corner. We call the mesh $\mathcal{T}_h$ a graded mesh with the grading parameter $\mu \in (0, 1]$ if there exist $c_1, c_2 > 0$ such that for each $T \in \mathcal{T}_h$

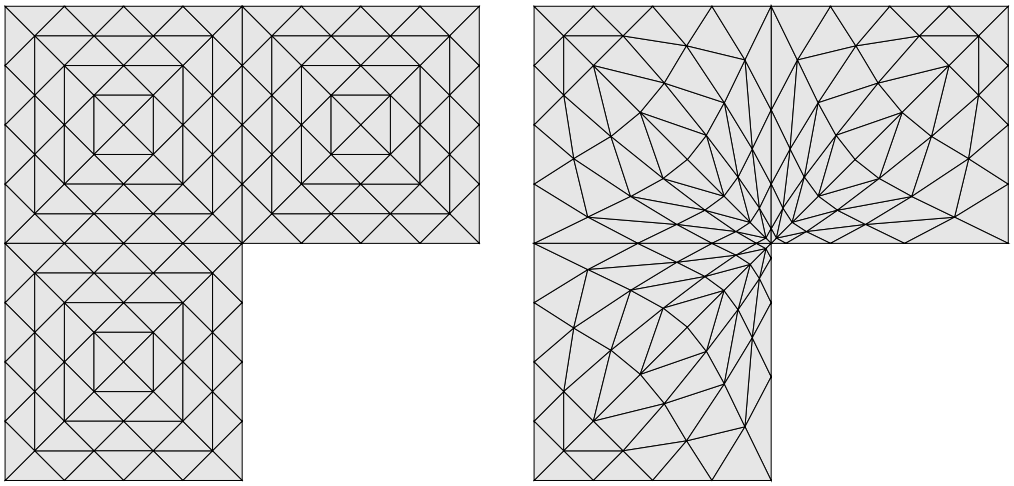


Figure 2.2: The graded mesh (right) with $\mu = 1/2$ and $R = 1$ constructed using the uniform triangulation of the L-shape domain (left)

$$\begin{aligned} c_1 h^{1/\mu} \leq h_T \leq c_2 h^{1/\mu}, &\quad \textit{if } r_T = 0, \\ c_1 h r_T^{1-\mu} \leq h_T \leq c_2 h r_T^{1-\mu}, &\quad \textit{if } 0 < r_T \leq R, \\ c_1 h \leq h_T \leq c_2 h, &\quad \textit{if } r_T > R, \end{aligned}$$

where $h_T = \operatorname{diam} T$.

The gradually refined meshes can be constructed using an initial coarse triangulation $\mathcal{T}_H$, which is uniformly refined until the desired mesh size $h > 0$ is obtained. Then, all the nodes of the triangulation within the chosen distance $R > 0$ from the corner are moved to get the triangulation sizes required in Definition 2.3.23. In Figure 2.2 we present the graded mesh obtained from a uniform triangulation of the L-shape domain. In the construction, we used parameters $\mu = 1/2$ and $R = 1$. Another mesh grading method involves the bisection of elements of the coarse mesh $\mathcal{T}_H$ until the conditions stated in Definition 2.3.23 are satisfied [50]. For the discussion of the construction of the graded meshes, we refer the reader to [132, Section 3.2.5].

Neither of the algorithms constructing the gradually refined triangulation significantly increases the number of nodes in the mesh, as it remains of order $\mathcal{O}(h^{-2})$. Although in the case of the first one this is obvious, a similar result for the latter algorithm is also known [9, Remark 3.1]. Moreover, the uniform meshes from Section 2.3.1 are obtained upon the choice $\mu = 1$.

The following theorem summarising the accuracy of the finite element scheme defined on the gradually refined triangulation can be found in [14] and [147, Theorem 4.3]. An exhaustive discussion of the interpolation error estimates on the graded meshes can be found in [132, Section 3.2.2].

Theorem 2.3.24 (Error estimates on graded meshes)**.** *Let $f \in L^2(\Omega)$, $u \in H_0^1(\Omega)$ be the unique solution of* (2.1) *and let $u_h \in V_h$ be its piecewise linear finite element approximation* (2.10) *on a triangulation satisfying requirements of Definition 2.3.23 with $\mu < \lambda_1$. Then*

$$\|u - I_h u\|_0 \leq ch^2 \|f\|_0, \quad \text{and} \quad \|\nabla(u - I_h u)\|_0 \leq ch\|f\|_0.$$

Moreover, the following a priori error estimates hold for some $c > 0$ independent of f

$$\|u - u_h\|_0 \leq ch^2 \|f\|_0, \quad \text{and} \quad \|\nabla(u - u_h)\|_0 \leq ch\|f\|_0.$$

This result shows that the finite element method on the mesh with a sufficiently strong grading towards the singular corner is capable of regaining the optimal convergence of the finite element scheme compared to the interpolation error. Moreover, the interpolation error is also improved compared to the one defined on the uniform meshes, see Lemma 2.3.17. Even though this requires the refinement of the computational mesh in the vicinity of the corner, since the number of nodes is asymptotically equivalent to the quasi-uniform case, it does not increase the complexity of the resulting finite element scheme.

2.3.2.4 Numerical examples

In order to illustrate the influence that the presence of corners can have on the finite element approximation of the solution of (2.1), we study the convergence of finite element methods using P_1, P_2 and P_3 Lagrange elements. We perform the computations on the L-shape domain $(-1, 1)^2 \setminus [0, 1] \times [-1, 0]$ with a known exact solution $u = s_1 + s_2 + s_3$ being a linear combination of the first three singular functions. We use the computational mesh presented in Figure 2.3. We measure the $L^2(\Omega)$ norm and $H_0^1(\Omega)$ seminorm, both standard and weighted, and summarise the results of the simulations in Table 2.1. For the weighted spaces we choose weights depending on the order of the finite element space used and set $\alpha = k - \lambda_1 + \epsilon$, $\epsilon = 10^{-4}$, so

$u \in H^{k+1}_{\alpha}(\Omega)$. This satisfies the assumptions of Theorem 2.3.11 and Corollary 2.3.18, and yields the following interpolation error estimates

$$\|u - I^k_h u\|_0 = \mathcal{O}\big(h^{1+\lambda_1}\big), \quad \text{and} \quad \|u - I^k_h u\|_\alpha = \mathcal{O}\big(h^{k+1}\big), \tag{2.16}$$

$$\|\nabla(u - I^k_h u)\|_0 = \mathcal{O}\big(h^{\lambda_1}\big), \quad \text{and} \quad \|\nabla(u - I^k_h u)\|_\alpha = \mathcal{O}\big(h^{k}\big), \tag{2.17}$$

and $\lambda_1 = 2/3$.

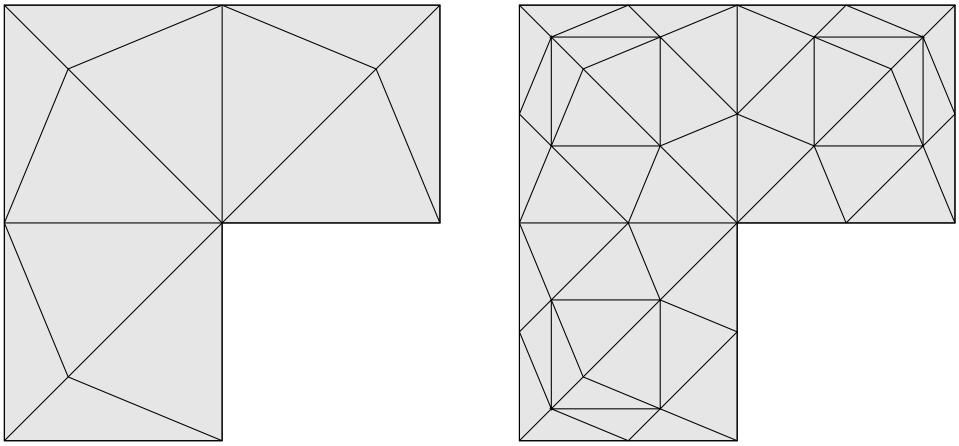


Figure 2.3: Initial triangulation of the L-shape domain (left) with a once refined mesh (right).

As predicted in Theorem 2.3.21, the presence of the corner in the domain leads to the suboptimal convergence of the finite element approximation compared to the interpolation error, when measured in standard and weighted $L^2(\Omega)$ norms. Moreover, the application of higher-order discretisation does not improve the convergence order of the scheme.

The convergence order of the piecewise linear discretisation is optimal in the sense of the interpolation error, when measured in the standard and weighted $H^1_0(\Omega)$ seminorm. However, the optimal convergence order in the weighted seminorm is no longer observed for the piecewise quadratic and piecewise cubic discretisations.

P_1 **elements**								
L	$\|u-u_h\|_0$	eoc	$\|u-u_h\|_\alpha$	eoc	$\|\nabla(u-u_h)\|_0$	eoc	$\|\nabla(u-u_h)\|_\alpha$	eoc
1	1.7834e-1	-	1.4343e-1	-	1.3521e-0	-	1.1553e-0	-
2	5.2930e-2	1.75	4.0224e-2	1.83	6.9847e-1	0.95	5.9203e-1	0.96
3	1.7248e-3	1.62	1.1918e-2	1.75	3.6515e-1	0.94	2.9959e-1	0.98
4	6.0922e-3	1.50	3.8509e-3	1.63	1.9416e-1	0.91	1.5116e-1	0.99
5	2.2675e-3	1.43	1.3516e-3	1.51	1.0550e-1	0.88	7.6206e-2	0.99
6	8.6841e-4	1.38	5.0253e-4	1.43	5.8788e-2	0.84	3.8408e-2	0.99
7	3.3754e-4	1.36	1.9300e-4	1.38	3.3624e-2	0.81	1.9355e-2	0.99
Expected		1.33		1.33		0.67		1.00
P_2 **elements**								
L	$\|u-u_h\|_0$	eoc	$\|u-u_h\|_\alpha$	eoc	$\|\nabla(u-u_h)\|_0$	eoc	$\|\nabla(u-u_h)\|_\alpha$	eoc
1	2.6468e-2	-	6.4823e-3	-	1.9930e-1	-	7.1927e-2	-
2	8.8458e-3	1.58	1.5853e-3	2.03	1.1978e-1	0.73	1.9817e-2	1.86
3	3.1144e-3	1.51	5.9664e-4	1.41	7.3789e-2	0.70	5.5919e-3	1.83
4	1.1390e-3	1.45	2.3609e-4	1.34	4.6038e-2	0.68	1.6696e-3	1.74
5	4.2751e-4	1.41	9.3683e-5	1.33	2.8888e-2	0.67	5.3813e-4	1.63
6	1.6341e-4	1.39	3.7176e-5	1.33	1.8170e-2	0.67	1.8766e-4	1.52
7	6.3257e-5	1.37	1.4753e-5	1.33	1.1439e-2	0.67	6.9498e-5	1.43
Expected		1.33		1.33		0.67		1.33
P_3 **elements**								
L	$\|u-u_h\|_0$	eoc	$\|u-u_h\|_\alpha$	eoc	$\|\nabla(u-u_h)\|_0$	eoc	$\|\nabla(u-u_h)\|_\alpha$	eoc
1	1.0128e-2	-	2.4257e-3	-	1.1720e-1	-	2.7053e-2	-
2	3.4783e-3	1.54	4.1394e-4	2.55	7.2282e-2	0.70	3.8455e-3	2.81
3	1.2389e-3	1.49	1.4982e-4	1.47	4.5144e-2	0.68	8.8484e-4	2.12
4	4.5456e-4	1.45	5.9242e-5	1.34	2.8341e-2	0.67	3.1463e-4	1.49
5	1.7067e-4	1.41	2.3504e-5	1.33	1.7829e-2	0.67	1.2328e-4	1.35
6	6.5204e-5	1.39	9.3270e-6	1.33	1.1225e-2	0.67	4.8861e-5	1.34
7	2.5226e-5	1.37	3.7013e-6	1.33	7.0700e-3	0.67	1.9387e-5	1.33
Expected		1.33		1.33		0.67		1.33

Table 2.1: Summary of convergence rates on the L-shape domain obtained using Lagrange finite elements P_1, P_2, P_3.

2.3.3 Energy-corrected finite element method

As we showed in the previous section, the standard finite element method on polygonal domains yields significantly lower convergence rates compared to the interpolation error in the standard and weighted $L^2(\Omega)$ norms. The goal of this section is to introduce an efficient method for mitigating the pollution effect summarised in Theorem 2.3.21.

In the past, several approaches have been proposed to improve the global convergence of the numerical solutions of elliptic problems on polygonal domains, such as graded mesh algorithms [5, 6, 9, 16, 37, 49], mesh adaptivity [13, 18, 19, 144] and h-p adaptivity [15, 62, 63, 70, 126, 134], in which the mesh refinement is combined with the increase in the order of locally used finite elements. Other types of methods for eliminating the pollution effect in the discrete approximation include the enrichment of the finite element space by singular functions, see, e.g., [28, 36, 41, 72, 111], post-processing based on the evaluation of the stress-intensity factors [96, 124] or extrapolation-based techniques [29, 30]. These methods improve the approximation quality but require a modification of the standard finite element solution in a $\mathcal{O}(1)$ neighbourhood of the re-entrant corner and require changing the structure of the mesh or the function space, in which the solution is approximated.

We follow a different approach, based on the so-called energy-corrected finite element method. The idea was originally proposed for finite difference schemes in [136, 138, 168]. Recently, it has been extended in the context of finite element methods, for the Poisson equation in [69, 88, 137] for the piecewise linear discretisation and in [86], as well as in [84, Chapter 4], using higher-order finite element spaces. In this section, we follow the ideas proposed there.

The energy-correction finite element was also developed for the Stokes problem in [91].

2.3.3.1 Definition of the energy-correction

Following Assumption 2.3.16, we restrict our considerations to only one corner in the computational domain.

As observed in Equation (2.14), the performance of the finite element approximation u_h of the solution u of the Poisson problem (2.1), when measured in the weighted $L^2(\Omega)$ norms, is limited in the following way

$$\|u - u_h\|_\alpha \gtrsim |a(u,u) - a(u_h,u_h)| \tag{2.18}$$

for any $\alpha \in \mathbb{R}$ and sufficiently regular forcing term f. The expression on the right-hand side can be understood as a difference in energies between the continuous problem (2.1) and its discretisation (2.10). Moreover, we know that it is bounded from below by an expression depending on the size of the corner in the domain (2.14).

The necessity of improving the approximation of the energy serves as a motivation for the energy-corrected finite element method described in this section. This method is based on a local modification of the bilinear form $a(\cdot,\cdot)$ defining the discrete problem (2.10). Hence, if we want to obtain the optimal approximation order in the sense of the best-approximation property stated in Theorem 2.3.14, the necessary condition

$$|a(u,u) - a_h(u_h,u_h)| \leq ch^{k+1}. \tag{2.19}$$

needs to be fulfilled. Naturally, this contradicts (2.14) in the case of the standard finite element approximation introduced in Section 2.3.2. Here, we will construct a modified bilinear form $a_h(\cdot,\cdot)$ on the discrete space so that (2.19) is satisfied. This modification will also prove sufficient for obtaining optimal convergence orders of the finite element scheme measured in weighted norms.

Let $k \in \mathbb{Z}_+$ be the order of the used finite element space. Similarly to (2.10), the modified finite element approximation of (2.1) reads: find $u_h^m \in V_h^k$ such that

$$a_h(u_h^m, v_h) = \langle f, v_h \rangle_\Omega \quad \text{for all } v_h \in V_h^k, \tag{2.20}$$

where the bilinear form is defined as

$$a_h(u,v) = a(u,v) - c_h(u,v). \tag{2.21}$$

The energy-corrected scheme defines a modified Ritz projection $R_h^m u = u_h^m$. We assume that $a_h(\cdot,\cdot) : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ is bilinear, continuous and elliptic, and that $c_h(\cdot,\cdot) : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ is symmetric. Furthermore, we assume that the support of the bilinear form $c_h(\cdot,\cdot)$ is contained a subset $\mathcal{S}_h$ consisting of few element patches around the considered corner. This means that

$$\text{If } u|_{\mathcal{S}_h} = 0 \text{ or } v|_{\mathcal{S}_h} = 0, \text{ then } c_h(u,v) = 0.$$

Finally, we assume that

$$|c_h(u,v)| \leq c\|u\|_{H^1(\mathcal{S}_h)}\|v\|_{H^1(\mathcal{S}_h)},$$

so that the correcting bilinear form is bounded in $H^1(\mathcal{S}_h)$.

Due to the small size of the support of the modification, the stiffness matrix $\mathbf{S}^m = \Big[a_h(\phi_i, \phi_j)\Big]_{i,j=1}^K$ associated with the energy-corrected discretisation (2.20) has only a fixed, independent of h, number of different entries compared with the standard stiffness matrix (2.12). Note also that for $c_h(u, v) \equiv 0$ we recover the standard finite element discretisation (2.10).

Similarly as in the case of the standard finite element discretisation (2.10), the orthogonality result resembling the Galerkin orthogonality introduced in Theorem 2.3.13 holds also for the energy-corrected discretisation.

Theorem 2.3.25 (Modified Galerkin orthogonality)**.** *Let $u_h^m \in V_h^k$ be the finite elmenent approximation defined in* (2.20) *of the solution $u \in H_0^1(\Omega)$ of the elliptic problem* (2.1)*. Then the following identity holds*

$$a(u - u_h^m, v_h) + c_h(u_h^m, v_h) = 0, \quad \textit{for all } v_h \in V_h^k.$$

Motivated by the necessary condition (2.19) for obtaining the optimal convergence of the energy-corrected scheme, we define the *energy defect function* as

$$g_h(v) = a(v, v) - a_h(R_h^m v, R_h^m v). \tag{2.22}$$

The analysis of the properties of the energy defect function will prove vital in the construction of the optimally accurate energy-corrected finite element method. Due to the modified Galerkin orthogonality introduced in Theorem 2.3.25, we can rewrite the energy defect function in a commonly used form

$$g_h(v) = a(v - R_h^m v, v - R_h^m v) - c_h(R_h^m v, R_h^m v). \tag{2.23}$$

We also define a more general entity

$$\widetilde{g}_h(v, w) = a(v - R_h^m v, w - R_h^m w) - c_h(R_h^m v, R_h^m w).$$

In particular, we have

$$\widetilde{g}_h(v, v) \equiv g_h(v).$$

Before moving to the investigations of the approximation properties of the energy-corrected scheme, we need to impose an important constraint on the computational mesh in the vicinity of the corner. We consider three types of patches, which are summarized in Figure 2.4, namely unstructured (G1), locally symmetric (G2) and consisting of identical isosceles triangles (G3). Choice of the right local mesh around

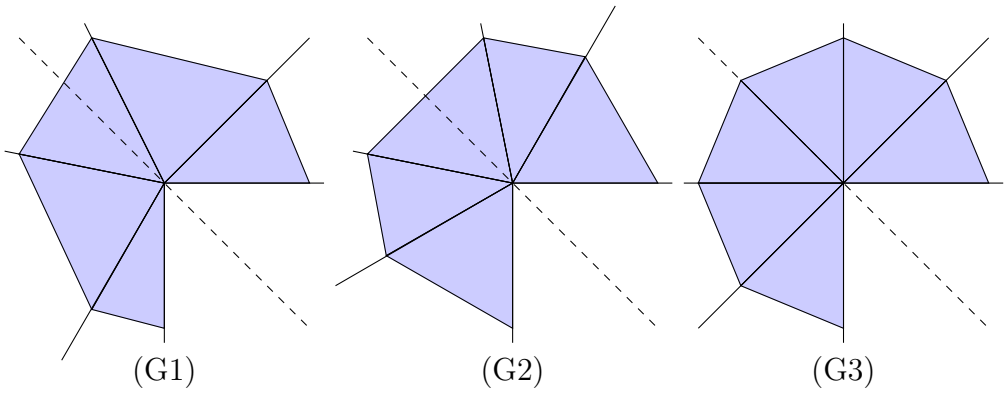


Figure 2.4: Constraints on the local mesh around the corner. Here, the dashed line is a bisection of the angle at the corner. In the first case (G1) we see an arbitrary patch. In (G2) we choose the mesh so that it is locally symmetric with respect to the bisection axis. In (G3) we see a grid consisting of identical isosceles triangles.

the considered corner depends on the order of polynomials in the finite element space and the size of the corner and is summarized below.

$$\begin{cases} \text{(G1)} & \text{if } k = 1 \text{ and } \Theta < \frac{3}{2}\pi \\ \text{(G2)} & \text{if } k = 1 \text{ and } \Theta \geq \frac{3}{2}\pi \text{ or } k = 2 \text{ and } \Theta < \frac{4}{3}\pi \\ \text{(G3)} & \text{if } k = 2 \text{ and } \Theta \geq \frac{4}{3}\pi \text{ or } k \geq 3 \end{cases} \tag{U}$$

From now on, whenever the energy-corrected finite element is considered, we will assume that the following holds.

Assumption 2.3.26. *We assume that for different singular functions s_i, s_j, for which $\lambda_i + \lambda_j < k + 1$, we have*

$$\int_{\mathcal{S}_h} \nabla s_i \cdot \nabla I_h^k s_j = \int_{\mathcal{S}_h} \nabla s_i \cdot \nabla s_j = 0.$$

This assumption was numerically studied and verified for the meshes satisfying condition (U) in [86, Section 6.1.3].

Finally, let us define the correction order as

$$N_k = \begin{cases} \left\lfloor (k+1)\frac{\Theta}{2\pi} \right\rfloor & \text{if } (k+1)\frac{\lambda_1}{2} \notin \mathbb{Z}_+, \\ (k+1)\frac{\Theta}{2\pi} - 1 & \text{otherwise.} \end{cases} \tag{2.24}$$

2.3.3.2 Convergence of the energy-corrected scheme

Now, we can state the main result describing the approximation property of the energy-corrected finite element discretisation.

Theorem 2.3.27 (Accuracy of the energy-correction)**.** *Let $k \in \mathbb{Z}_+$ be the order of the finite element space used in the discretisation* (2.20)*. Let also* $\max\big(0, k-\lambda_1\big) < \alpha < k$ *and* $\widetilde{\alpha} = \alpha - k + 1$*. Suppose that* $f \in H^{k-1}_{-\widetilde{\alpha}}(\Omega)$ *and* $u \in H^1_0(\Omega)$ *is the unique solution of* (2.1)*. If*

$$g_h(s_i) = \mathcal{O}\big(h^{k+1}\big), \quad \textit{for all } i \leq N_k, \tag{2.25}$$

then for some $c > 0$ *independent of* f *the following estimates hold*

$$\|u - u_h\|_\alpha \leq c h^{k+1} \|f\|_{k-1,-\widetilde{\alpha}}, \quad \textit{and} \quad \|\nabla(u - u_h)\|_\alpha \leq c h^k \|f\|_{k-1,-\widetilde{\alpha}}.$$

This result was first proven in [69, Theorem 2.4] in the case of a piecewise-linear discretisation. It was further extended to the form presented here in [86, Theorem 2.2].

Due to Theorem 2.2.7, the regularity assumption on the forcing term f implies that $u \in H^{k+1}_\alpha(\Omega)$. Hence, the error estimates are optimal in the sense of the interpolation error stated in Theorem 2.3.11. The correction order is the number of relevant singular functions (2.2) in the expansion provided by the Theorem 2.2.7 that need to be taken care of by the energy-corrected scheme so that the energy defect function converges optimally (2.25). The mixed term $\widetilde{g}_h(s_i, s_j)$ for $i \neq j$ disappear due to the Assumption 2.3.26. The presence of the positive weight α in the norms is necessary to compensate the lower approximation order in the vicinity of the corner stemming from the presence of the singular functions in the solution. However, the approximation exhibits the optimal order of convergence, when measured in some positive distance from the considered corner.

Corollary 2.3.28 (Approximation far from the corner)**.** *Suppose that the assumption of Theorem 2.3.27 are satisfied and let* $\widetilde{\Omega} \subset \Omega$ *be such that dist*$(\widetilde{\Omega}, 0) > \delta > 0$*. Then*

$$\|u - u_h\|_{L^2(\widetilde{\Omega})} \leq c h^{k+1} \|f\|_{k-1,-\widetilde{\alpha}}, \quad \textit{and} \quad \|\nabla(u - u_h)\|_{L^2(\widetilde{\Omega})} \leq c h^k \|f\|_{k-1,-\widetilde{\alpha}}.$$

So the energy-corrected finite element scheme recovers the optimal convergence properties in the sense of the interpolation error summarised in Theorem 2.3.10 when measured far from the corner. Hence, if we can assure that the condition (2.25) is satisfied, we will know that the pollution effect stated in Theorem 2.3.21 is eliminated from the approximation.

Naturally, we would also like to know, how the energy-corrected scheme performs in more regular cases when the singular functions are absent in the expansion given in Theorem (2.2.7). The following result was proven in [86, Lemma 3.1].

Lemma 2.3.29. *Let k be a positive integer and $u \in H^{k+1}_{-\widetilde{\alpha}}(\Omega) \cap H^1_0(\Omega)$ for some $1 - \lambda_1 < \widetilde{\alpha} < 1$. Then the following estimates hold for some $c > 0$*

$$\|u - u_h^m\|_{-\widetilde{\alpha}} \leq ch^{k+1}\|u\|_{k+1,-\widetilde{\alpha}}, \quad \text{and} \quad \|\nabla(u - u_h^m)\|_{-\widetilde{\alpha}} \leq ch^k\|u\|_{k+1,-\widetilde{\alpha}}. \tag{2.26}$$

Remark 2.3.30. *Since $L^2_{-\widetilde{\alpha}} \hookrightarrow L^2(\Omega)$, the following estimates hold under the assumptions of Lemma 2.3.29 are*

$$\|u - u_h^m\|_0 \leq ch^{k+1}\|u\|_{k+1,-\widetilde{\alpha}}, \quad \text{and} \quad \|\nabla(u - u_h^m)\|_0 \leq ch^k\|u\|_{k+1,-\widetilde{\alpha}}. \tag{2.27}$$

However, it is important to note that the regularity $u \in H^{k+1}_{-\widetilde{\alpha}}(\Omega) \cap H^1_0(\Omega)$ in negatively weighted norms is also necessary for the estimate (2.27) *to hold. This cannot be improved due to the low regularity of the dual problem around the corner.*

We would like to obtain similar estimates to the ones given in Theorem 2.3.27 in more commonly-used standard $L^2(\Omega)$ and $H^1_0(\Omega)$ norms. In order to do this, we will exploit the known structure of the solution in the post-processing step.

Suppose that for some positive integer k we have $\max(0, k - \lambda_1) < \alpha < k$ and $\widetilde{\alpha} = \alpha - k + 1$. Furthermore, assume that $f \in H^{k-1}_{-\widetilde{\alpha}}(\Omega)$. Thanks to Theorem 2.2.7 and (2.4), we know that

$$u = U + \sum_{i:\lambda_i < k+\alpha} k_i s_i. \tag{2.28}$$

Recall, the stress-intensity factors k_i in this expansion can be computed using formula (2.4). We define the approximate stress-intensity factors as

$$k_i^h = \frac{1}{i\pi} \int_\Omega f s_{-i} + u_h^m \Delta s_{-i}. \tag{2.29}$$

The following lemma was proposed in the case of the piecewise linear discretisation in [69, Theorem 2.6] and further extended to the piecewise polynomial case in [86].

Lemma 2.3.31. *Suppose that the assumptions of Theorem 2.3.27 are satisfied. Then for all $i \leq N_k$ and some $c > 0$*

$$|k_i - k_i^h| \leq ch^{k+1}\|f\|_{k-1,-\widetilde{\alpha}}.$$

Let $s_{i,h}^m = R_h^m s_i$. Due to the linearity of the modified Ritz projection and (2.28), we also have

$$u_h^m = U_h^m + \sum_{i:\lambda_i < k+\alpha} k_i s_{i,h}^m. \tag{2.30}$$

Notice that $U \in H^{k+1}_{-\widetilde{\alpha}}(\Omega)$, and hence, according to Lemma 2.3.30, it is regular enough for its finite element approximations to exhibit desired convergence. Now, having the accurate approximation of the stress-intensity factors through Lemma 2.3.31, we can replace the remaining approximations of the singular functions (2.30) by their continuous counterparts. Finally, we define the post-processed solution by

$$\widetilde{u}_h^m = u_h^m + \sum_{i:\lambda_i<k+\alpha} k_i^h\big(s_i - s_{i,h}^m\big). \tag{2.31}$$

This post-processing step is based on the enrichment of the discrete space by the singular functions after the solution of (2.20) is obtained. This results in an improved approximation property of the scheme, which is summarised in the following theorem.

Theorem 2.3.32. *Suppose that the assumptions of Theorem 2.3.27 hold. Then for some $c > 0$*

$$\|u - \widetilde{u}_h^m\|_0 \le ch^{k+1}\|f\|_{k-1,-\widetilde{\alpha}}, \quad \textit{and} \quad \|\nabla(u - \widetilde{u}_h^m)\|_0 \le ch^k\|f\|_{k-1,-\widetilde{\alpha}}.$$

For the proof of this theorem, we refer the reader to [86, Corollary 2.3]. This improved approximation of the scheme comes at a computational cost of solving N_k additional elliptic problems in the search of $s_{i,h}^m$ and N_k integral evaluations for the approximation of the stress-intensity factors (2.29).

2.3.3.3 Construction of the modification

We will now focus on showing that the assumptions on the energy defect function (2.25) can indeed be satisfied upon the right choice of the correcting bilinear form $c_h(\cdot,\cdot)$. Let us define

$$\mathcal{S}_h^1 = \{T \in \mathcal{T}_h \ : 0 \in \partial T\}, \quad \mathcal{S}_h^i = \{T \in \mathcal{T} \ : \partial T \cap \partial\mathcal{S}_h^{i-1} \neq \emptyset\}, i = 1, \dots, N_k.$$

So $\mathcal{S}_h^i$ denotes the i-th layer of elements counting from the corner.

We consider two types of modifications for the energy-corrected scheme

$$c_h^R(u,v) = \sum_{i=1}^{N_k} \gamma_i^R \int_{\mathcal{S}_h^i} \nabla u \cdot \nabla v, \quad c_h^F(u,v) = \sum_{i=1}^{N_k} \gamma_i^F \int_{\mathcal{S}_h^1} \hat{r}^{i-1} \nabla u \cdot \nabla v. \tag{2.32}$$

Here, $\hat{r}$ is a distance from the corner measured on the reference triangle and γ^R, γ^F are vectors of correction parameters, which will be specified later. The bilinear form $c_h^R(\cdot,\cdot)$ is supported in N_k element layers from the corner, which are visualised in Figure 2.5. Note that in the case of $k = 1$ and $\Theta > \pi$, we have $N_k = 1$ and both corrections reduce to the same type of modification

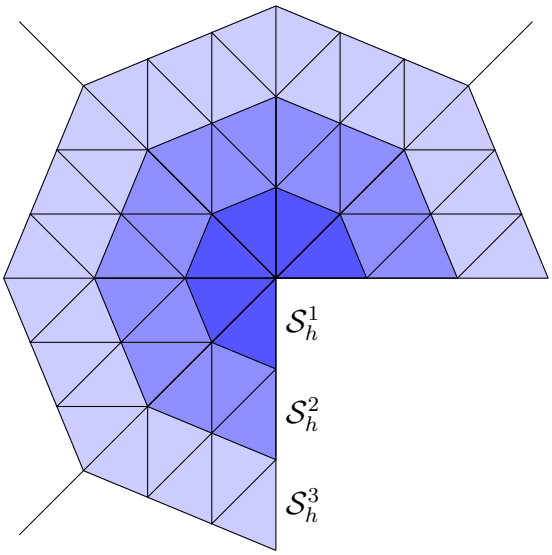


Figure 2.5: Local mesh around the re-entrant corner. In the higher-order energy-correction, modification $c_h^R(\cdot,\cdot)$ is supported in a layer of N_k elements around the re-entrant corner. In this case $N_k = 3$.

$$c_h(u,v) = \gamma \int_{\mathcal{S}_h^1} \nabla u \cdot \nabla v. \tag{2.33}$$

We also assume that $\|\gamma^R\|_{\ell^\infty} < 1$ and $\|\gamma^F\|_{\ell^\infty} < 1$, so that the coercivity of $a_h(\cdot,\cdot)$ is preserved. The modifications (2.32) were proposed in [86] and (2.33) was studied in [69].

We will now move to finding the right parameters for the modifications introduced in (2.32). For the better readability, we skip the superscripts and simply write γ instead of γ^R and γ^F. Let us define

$$\mathbf{g}_h(\gamma) = \big(g_h(s_i)\big)_{i=1}^{N_k}, \tag{2.34}$$

where $g_h(\cdot)$ is the energy defect function (2.22) with a suitable modification (2.32) and a given parameter vector $\gamma = \big(\gamma_{h,1}, \ldots, \gamma_{h,N_k}\big)$. For each refinement level with the mesh-size h we define a new vector of parameters as the solution of

$$\mathbf{g}_h(\gamma_h) = 0. \tag{2.35}$$

The following result shows that the assumptions of Theorem 2.3.27 with modifications (2.32) can indeed be satisfied.

Lemma 2.3.33. *Let* $\gamma^* = \lim_{h\to 0^+} \gamma_h$ *and suppose that*

$$|\gamma_i^* - \gamma_{h,i}| = \mathcal{O}\big(h^{k+1-\lambda_i}\big). \tag{2.36}$$

Then the assumption (2.25) *of Theorem 2.3.27 is satisfied with the modification* (2.32) *and* $\gamma = \gamma^*$.

This result was proven in [69, Lemma 5.3] for a piecewise linear discretisation and in [86, Lemma 5.1] for the general piecewise polynomial setting. Moreover, for piecewise linear finite element the solution of (2.35) exists and converges sufficiently fast to satisfy (2.36), see [69, Lemma 5.2, Remark 5.4]. However, for the higher-order discretisation, the assumption (2.36) was only confirmed numerically [86, Section 6.1.1].

Parameters γ_h defined as roots of the energy defect function (2.35) can be efficiently found using nested Newton algorithms proposed in [137] and standard Newton iteration studied also in [86, Section 6.1.1].

Throughout this work, we will assume that the optimal parameter γ^* provided by Lemma 2.3.33 is known. In all the numerical investigations we will use its approximation γ_h computed on a sufficiently many times refined mesh. It is worth noting that the parameter γ^* depends only on the size of the considered corner and on the one element patch of the triangulation around it. In particular, it is independent of the whole domain Ω and a particular choice of the right-hand side f in (2.1).

2.3.3.4 Numerical examples

To illustrate the approximation properties of the energy-corrected finite element scheme, we perform numerical examples in the same setting as in Section 2.3.2. We use the exact solution $u = s_1 + s_2 + s_3$ on the L-shape domain with the triangulation presented in Figure 2.3 and for the weighted norms we set $\alpha = k - \lambda_1 + \epsilon$, $\epsilon = 10^{-4}$. We summarise the convergence rates for the energy-corrected scheme in Table 2.2. The obtained errors are optimal in the sense of the interpolation error (2.16). Hence, the application of the energy-correction significantly improves the observed convergence properties of the finite element scheme yielding higher orders of convergence than the ones observed in Table 2.1 for the standard discretisation.

We would also like to point out that the use of $c^F(\cdot,\cdot)$ modification results in lower errors than the ones obtained with the help of $c^R(\cdot,\cdot)$. This was already observed in [86, Section 6.3] and can be attributed to the larger support of the latter modification.

Energy-corrected P_1 elements, $\gamma = 0.117531611518762$								
L	$\|u-u_h^m\|_0$	eoc	$\|u-u_h^m\|_\alpha$	eoc	$\|\nabla(u-u_h^m)\|_0$	eoc	$\|\nabla(u-u_h^m)\|_\alpha$	eoc
1	1.8706e-1	-	1.5053e-1	-	1.3609e-0	-	1.1511e-0	-
2	5.2642e-2	1.83	4.0585e-2	1.89	7.0341e-1	0.95	5.9341e-1	0.96
3	1.4130e-3	1.90	9.8875e-3	2.04	3.6794e-1	0.93	3.0048e-1	0.98
4	3.8881e-3	1.86	2.4112e-3	2.04	1.9603e-1	0.91	1.5161e-1	0.99
5	1.1092e-3	1.81	5.9519e-4	2.02	1.0682e-1	0.88	7.6424e-2	0.99
6	3.2650e-4	1.76	1.4834e-4	2.00	5.9717e-2	0.84	3.8514e-2	0.99
7	9.8405e-5	1.73	3.7233e-5	1.99	3.4265e-2	0.80	1.9406e-2	0.99
Expected		1.66		2.00		0.67		1.00
Energy-corrected P_2 elements with $c_h^R(\cdot,\cdot)$, $\gamma = (0.03152, -0.005533)$								
L	$\|u-u_h^m\|_0$	eoc	$\|u-u_h^m\|_\alpha$	eoc	$\|\nabla(u-u_h^m)\|_0$	eoc	$\|\nabla(u-u_h^m)\|_\alpha$	eoc
1	2.6630e-2	-	8.3594e-3	-	2.04155e-1	-	8.3001e-2	-
2	8.6672e-2	1.62	1.8170e-3	2.20	1.2252e-1	0.74	2.1620e-2	1.94
3	2.5959e-3	1.74	1.9483e-4	3.22	7.4899e-2	0.71	5.2438e-3	2.04
4	8.0270e-4	1.69	2.1625e-5	3.17	4.6629e-2	0.68	1.3438e-3	1.96
5	2.5114e-4	1.68	2.5898e-6	3.06	2.9238e-2	0.67	3.4983e-4	1.94
6	7.8905e-5	1.67	3.2494e-7	2.99	1.8385e-2	0.67	9.1232e-5	1.94
Expected		1.67		3.00		0.67		2.00
Energy-corrected P_2 elements with $c_h^F(\cdot,\cdot)$, $\gamma = (0.10241, -0.16802)$								
L	$\|u-u_h^m\|_0$	eoc	$\|u-u_h^m\|_\alpha$	eoc	$\|\nabla(u-u_h^m)\|_0$	eoc	$\|\nabla(u-u_h^m)\|_\alpha$	eoc
1	2.9144e-2	-	8.2033e-3	-	2.0737e-1	-	8.4710e-2	-
2	7.6256e-2	1.93	1.0192e-3	3.01	1.2226e-1	0.76	2.0511e-2	2.05
3	2.3590e-3	1.69	1.3273e-4	2.94	7.5078e-2	0.70	5.3884e-3	1.93
4	7.3770e-4	1.68	1.7356e-5	2.93	4.6798e-2	0.68	1.4122e-3	1.93
5	2.3170e-4	1.67	2.3238e-6	2.90	2.9355e-2	0.67	3.6887e-4	1.94
6	7.2901e-5	1.67	3.0767e-7	2.92	1.8461e-2	0.67	9.6027e-5	1.94
Expected		1.67		3.00		0.67		2.00
Energy-corrected P_3 elements with $c_h^R(\cdot,\cdot)$, $\gamma = (0.012891, -0.0023667)$								
L	$\|u-u_h^m\|_0$	eoc	$\|u-u_h^m\|_\alpha$	eoc	$\|\nabla(u-u_h^m)\|_0$	eoc	$\|\nabla(u-u_h^m)\|_\alpha$	eoc
1	1.0368e-2	-	3.2602e-3	-	1.1866e-1	-	3.2392e-2	-
2	3.3871e-3	1.61	4.3010e-4	2.92	7.3067e-2	0.70	4.5921e-3	2.82
3	1.0012e-3	1.76	3.0173e-5	3.82	4.5454e-2	0.68	5.0162e-4	3.19
4	3.0779e-4	1.70	1.8948e-6	3.99	2.8503e-2	0.67	5.8511e-5	3.10
5	9.6080e-5	1.68	1.1716e-7	4.02	1.7925e-2	0.67	7.1874e-6	3.03
6	3.0161e-5	1.67	7.2796e-9	4.01	1.1284e-2	0.67	9.0095e-7	3.00
Expected		1.67		4.00		0.67		3.00
Energy-corrected P_3 elements with $c_h^F(\cdot,\cdot)$, $\gamma = (0.042218, -0.07085)$								
L	$\|u-u_h^m\|_0$	eoc	$\|u-u_h^m\|_\alpha$	eoc	$\|\nabla(u-u_h^m)\|_0$	eoc	$\|\nabla(u-u_h^m)\|_\alpha$	eoc
1	1.1670e-2	-	3.2991e-3	-	1.2107e-1	-	3.5799e-2	-
2	3.1178e-3	1.90	1.5467e-4	4.41	7.3228e-2	0.73	3.3454e-3	3.42
3	9.7010e-3	1.68	9.5697e-6	4.01	4.5632e-2	0.68	4.0535e-3	3.04
4	3.0425e-4	1.67	5.9812e-7	4.00	2.8628e-2	0.67	5.0667e-5	3.00
5	9.5683e-5	1.67	3.8298e-8	3.96	1.8006e-2	0.67	6.4023e-6	2.98
6	3.0120e-5	1.67	2.7183e-9	3.82	1.1336e-2	0.67	8.1184e-7	2.98
Expected		1.67		4.00		0.67		3.00

Table 2.2: Summary of convergence rates on the L-shape domain obtained using energy-corrected finite elements P_1, P_2, P_3 with two different types of modification.

Chapter 3

Maximum norm error estimates

The study of maximum norm error estimates for the standard piecewise linear finite element approximation of elliptic problems on quasi-uniform meshes started with [119, 121, 122, 123], where suboptimal estimates were obtained. They were further improved to the optimal form of sharp estimates in [74, 146].

The first investigations on polygonal domains were conducted in [143]. There, it was shown that the presence of the corners in the computational domain significantly reduces the convergence order of the finite element method in the maximum norm, a phenomenon similar to the pollution effect introduced in Theorem 2.3.21. The standard approaches for mitigating the pollution effect in the finite element solution include refinement towards the corners in the domain. The $L^\infty(\Omega)$ norm estimates for the schemes based on mesh refinement were studied in [144] and [147, Chapter 4], and sharp estimates can be found in [135]. Similar techniques also lead to optimal estimates for the Neumann problem on polygonal domains, see [7].

The optimal error estimates for the higher-order finite element discretisation was studied in [140, 141].

In Section 2.3.3, we introduced the energy-correction method, which improves the convergence properties of the finite element methods on domains with corners, when measured in weighted $L^2(\Omega)$ norm and $H^1_0(\Omega)$ seminorm. We build on the analysis presented there and show that the energy-correction method converges optimally in the sense of the best approximation property also when the weighted $L^\infty(\Omega)$ norm is concerned.

We begin by stating the regularity results for the singular solutions of the Poisson problem on polygonal domains in the framework of weighted $W^{k,\infty}(\Omega)$ spaces. Next, we show some auxiliary results describing the behaviour of piecewise polynomial elements in weighted spaces. We prove an interpolation error estimate, which to our knowledge, cannot be found in the literature. Having all necessary tools at hand, in

Section 3.3.1, we move to proving the optimal convergence of the energy-corrected finite element, when measured in the weighted $L^\infty(\Omega)$ norm. The proof follows the ideas first introduced in [143, 144] and is based on a dyadic decomposition of the computational domain Ω around the considered corner. We conclude the chapter with numerical investigations illustrating and confirming the obtained results.

The optimal error estimates presented in this chapter mean that the presence of the corners in the computational domain does not have any influence on the approximation properties of the energy-corrected scheme in a positive distance from it. This stems from the fact that the energy-correction method successfully eliminates the pollution effect introduced in Theorem 2.3.21, the fact also apparent in the proof of the main result of this chapter.

The results presented in this chapter were partly published by the author and B. Wohlmuth in the paper entitled "Maximum norm estimates for energy- corrected finite element method" in Numerical Mathematics and Advanced Applications ENUMATH 2017 [154].

3.1 Regularity

In this chapter, we focus on the model Poisson problem (2.1) with the right-hand side f and the solution u, on polygonal domains Ω, see Definition 2.1.11. In Section 2.2.2, we discussed the regularity properties of the elliptic problems on domains with corners and introduced the singular functions (2.2). Furthermore, we investigated their regularity in standard and weighted Hilbert–Sobolev spaces $H^{k+1}(\Omega)$.

A similar analysis can be conducted in terms of the Sobolev space $W^{k+1,\infty}(\Omega)$. Similarly as before, we consider the behaviour around only one of the domain's corners, which we locate at the origin of the Euclidean space for simplicity. Note first that for positive integers k, if $\Theta > \frac{i\pi}{k+1}$, then $s_i \notin W^{k+1,\infty}(\Omega)$. This, in particular, means that around a corner of size $\Theta > \pi/2$, we have $s_1 \notin W^{2,\infty}(\Omega)$. However, similarly as in the aforementioned case, it is convenient to describe the regularity in weighted $V_\beta^{k+1,\infty}(\Omega)$ spaces.

Theorem 3.1.1 (Regularity of the singular functions in $V_\beta^{k+1,\infty}(\Omega)$)**.** *Let $\beta \geq k + 1 - \lambda_i$. Then the singular functions* (2.2) *satisfy*

$$s_i \in V_\beta^{k+1,\infty}(\Omega).$$

In Theorem 2.2.6, we established the shift of regularity between the forcing term f and the corresponding solution of the Poisson problem. However, the similar result

in the case $p = \infty$ cannot be expected and higher Hölder regularity assumption on f needs to be imposed. The following theorem is a consequence of [99, Theorem 1.4.5] and [99, Theorem 2.6.3].

Theorem 3.1.2 ($V_\beta^{k+1,\infty}(\Omega)$ regularity)**.** *Let k be a positive integer and $\sigma \in (0,1)$. Let us also take $\beta \in (k+1+\sigma-\lambda_1, k+1+\sigma+\lambda_1)$ and suppose that $f \in N_\beta^{k-1,\sigma}(\Omega)$. Then there exists a unique solution of the Poisson problem (2.1) $u \in N_\beta^{k+1,\sigma}(\Omega) \cap H_0^1(\Omega)$, which is continuous. Moreover, for some $c > 0$ independent of the right-hand side f*

$$\|u\|_{N_\beta^{k+1,\sigma}(\Omega)} \leq c\|f\|_{N_\beta^{k-1,\sigma}(\Omega)}$$

This, together with Theorem 2.2.7 and the embedding stated in Theorem 2.1.15, leads to the following result.

Corollary 3.1.3. *Suppose that the assumptions of Theorem 3.1.2 are satisfied. Furthermore, let $k - \lambda_1 < \alpha < k$ and $\tilde{\alpha} = \alpha - k + 1$. If $f \in N_\beta^{k-1,\sigma}(\Omega) \cap H_{-\tilde{\alpha}}^{k-1}(\Omega)$, then the solution $u \in V_\beta^{k+1,\infty}(\Omega) \cap H_\alpha^{k+1}(\Omega) \cap H_0^1(\Omega)$ of (2.1) is continuous. Moreover, it admits the following expansion*

$$u = U + \sum_{i:\, \lambda_i < k+\alpha} k_i s_i,$$

and the regular part of the solution satisfies $U \in V_\beta^{k+1,\infty}(\Omega) \cap H_{-\tilde{\alpha}}^{k+1}(\Omega) \cap H_0^1(\Omega)$ and is continuous. Furthermore, for some $c > 0$

$$\|u\|_{V_\beta^{k+1,\infty}(\Omega)} \leq c\|f\|_{N_\beta^{k-1,\sigma}(\Omega)}, \quad \text{and} \quad \|u\|_{H_\alpha^{k+1}(\Omega)} \leq c\|f\|_{H_{-\tilde{\alpha}}^{k-1}(\Omega)}.$$

3.2 Finite element error in maximum norm

Now, we move to the investigations of the quality of the finite element approximations of the solution u in the discrete spaces V_h^k of piecewise polynomial functions, which we introduced in (2.8). We first analyse the errors obtained using the nodal interpolation from Definition 2.3.8. These estimates will serve as a benchmark, with which we shall compare the quality of the finite element approximation. Moreover, the interpolation error estimates will be essential further in the proof of the optimal error estimates for the energy-corrected scheme.

The Bramble–Hilbert Lemma 2.3.3 in the classical form presented in Section 2.3.1 is insufficient to prove interpolation error estimates in weighted Sobolev norms. Therefore, we present its suitable extension, which was proposed in [12]. There, an even more general form than the one presented here can be found.

Lemma 3.2.1 (Bramble–Hilbert Lemma in weighted spaces)**.** *Let $K \subset \mathbb{R}^d$ be as in Definition 2.3.1 and $1 \leq p \leq \infty$. Let $m, l \in \mathbb{Z} \cup \{0\}$ and $\alpha, \beta \in \mathbb{R}$ be the weights satisfying*

$$\begin{aligned} 0 < l - \alpha - m, \quad \text{and } -d < \alpha p, \quad \text{when } 1 \leq p < \infty, \\ -d < \beta, \quad \text{and } \alpha < l, \quad \text{when } p = \infty, \end{aligned}$$

Moreover, if $\beta p + l - \alpha - m > 0$, then there exists a constant $c > 0$ independent of p and u such that for any $u \in W^{m,p}(\Omega)$ there exists a polynomial $p_{l-1} \in P_{l-1}(K)$ such that

$$|u - p_{l-1}|_{W^{m,p}_{\beta}(\Omega)} \leq c|u|_{W^{l,p}_{\alpha}(\Omega)}.$$

This result is essential for the proof of the interpolation error estimates.

Lemma 3.2.2 (Interpolation error in $L^{\infty}_{\beta}(\Omega)$)**.** *Let $I^k_h : C(\overline{\Omega}) \to V^k_h$ denote the standard nodal interpolation operator, see Definition 2.3.8. Then, for any function $u \in W^{k+1,\infty}_{\alpha}(\Omega) \cap C(\bar{\Omega})$, with $0 \leq \alpha < k+1$ and $\max(0, \alpha - k) \leq \beta \leq \alpha$ the following estimate holds*

$$\|u - I^k_h u\|_{L^{\infty}_{\beta}(\Omega)} \leq c h^{k+1+\beta-\alpha} |u|_{W^{k+1,\infty}_{\alpha}(\Omega)}$$

for some constant $c > 0$ independent of h and u.

Proof. Suppose that the value of the norm $\|u - I^k_h u\|_{L^{\infty}_{\beta}(\Omega)}$ is attained in the element $T^* \in \mathcal{T}_h$. Then we have

$$\|u - I^k_h u\|_{L^{\infty}_{\beta}(\Omega)} = \|u - I^k_h u\|_{L^{\infty}_{\beta}(T^*)}.$$

We consider two cases, namely when $\operatorname{dist}\big(0, T^*\big) \coloneqq d_{T^*} > 0$ and $\operatorname{dist}\big(0, T^*\big) = 0$. We begin our considerations with the former case.

Note that in this case there exists a constant $c_1 > 0$ such that

$$c_1 \sup_{\mathbf{x} \in T^*} |\mathbf{x}| \leq d_{T^*} = \inf_{\mathbf{x} \in T^*} |\mathbf{x}|. \tag{3.1}$$

Thus, we can write

$$\|u - I^k_h u\|_{L^{\infty}_{\beta}(T^*)} \leq c d^{\beta}_{T^*} \|u - I^k_h u\|_{L^{\infty}(T^*)} = c d^{\beta}_{T^*} \|\widehat{u} - I^k_{\widehat{T}} \widehat{u}\|_{L^{\infty}(\widehat{T})}, \tag{3.2}$$

where $\widehat{T}$ is the reference triangle and $I^k_{\widehat{T}}$ is the interpolation operator defined on it. For the estimate in the standard Sobolev norms on the reference element we can use Theorem 2.3.4, where we set $p = q = \infty$, $m = 0$ and $l = k + 1$. We arrive at

$$\begin{aligned}\|u - I^k_h u\|_{L^\infty_\beta(T^*)} &\le c d^\beta_{T^*} \|\widehat{u} - I^k_{\widehat{T}}\widehat{u}\|_{L^\infty(\widehat{T})} \\ &\le c d^\beta_{T^*} |\widehat{u}|_{W^{k+1,\infty}(\widehat{T})} \\ &\le c d^\beta_{T^*} h^{k+1} |u|_{W^{k+1,\infty}(T^*)} \\ &\le c d^{\beta-\alpha}_{T^*} h^{k+1} |u|_{W^{k+1,\infty}_\alpha(T^*)}\end{aligned}$$

The last inequality is the result of (3.1). The scaling h^{k+1} comes from the Jacobian of the transformation from the reference triangle to the element T^*. Since $d_{T^*} \ge h$ and $\beta \le \alpha$, we have $d^{\beta-\alpha}_{T^*} \le h^{\beta-\alpha}$. Hence, we finally obtain

$$\|u - I^k_h u\|_{L^\infty_\beta(T^*)} \le c h^{k+1+\beta-\alpha} |u|_{W^{k+1,\infty}_\alpha(T^*)}. \tag{3.3}$$

Suppose now that $\operatorname{dist}\big(0, T^*\big) = 0$. Since in T^* we have $r \le h$, we can write

$$\begin{aligned}\|u - I^k_{T^*} u\|_{L^\infty_\beta(T^*)} &= \|r^\beta (u - I^k_{T^*} u)\|_{L^\infty(T^*)} \\ &\le h^\beta \|u - I^k_{T^*} u\|_{L^\infty(T^*)} = h^\beta \|\widehat{u} - I^k_{\widehat{T}}\widehat{u}\|_{L^\infty(\widehat{T})}.\end{aligned} \tag{3.4}$$

Let now $p_k \in P_k(\widehat{T})$ be some polynomial of order k on $\widehat{T}$, which will be specified later. Since $\|I^k_h v\|_{L^\infty(\widehat{T})} \le c\|v\|_{L^\infty(\widehat{T})}$ holds for the nodal interpolation $I^k_{\widehat{T}}$, we have

$$\|\widehat{u} - I^k_{\widehat{T}}\widehat{u}\|_{L^\infty(\widehat{T})} \le \|\widehat{u} - p_k\|_{L^\infty(\widehat{T})} + \|I^k_{\widehat{T}}(p_k - \widehat{u})\|_{L^\infty(\widehat{T})} \le c\|\widehat{u} - p_k\|_{L^\infty(\widehat{T})}.$$

Thanks to Lemma 3.2.1 in which we take $\beta = 0, l = k + 1, p = \infty$ and $m = 0$, we obtain

$$\|\widehat{u} - I^k_{\widehat{T}}\widehat{u}\|_{L^\infty(\widehat{T})} \le c|\widehat{u}|_{W^{k+1,\infty}_\alpha(\widehat{T})}.$$

In combination with (3.4) this yields

$$\|u - I^k_{T^*} u\|_{L^\infty_\beta(T^*)} \le c h^\beta |\widehat{u}|_{W^{k+1,\infty}_\alpha(\widehat{T})}$$

However, we also have $\widehat{r} \sim hr$, where $\widehat{r}$ is the distance from the origin on the reference triangle. Hence, similarly as before, after transformation back to the physical triangle T^*, we get

$$\|u - I^k_h u\|_{L^\infty_\beta(T^*)} \le c h^{k+1+\beta-\alpha} |u|_{W^{k+1,\infty}_\alpha(T^*)}.$$

This, together with (3.3) completes the proof of the lemma.

□

Note that for a positive parameter α, we have $V^{k+1,\infty}_{\alpha}(\Omega) \hookrightarrow W^{k+1,\infty}_{\alpha}(\Omega)$ and also $|u|_{W^{k+1,\infty}_{\alpha}(\Omega)} = |u|_{V^{k+1,\infty}_{\alpha}(\Omega)}$. Hence, the assertion of Lemma 3.2.2 holds also when spaces $W^{k+1,\infty}_{\alpha}(\Omega)$ are replaced with $V^{k+1,\infty}_{\alpha}(\Omega)$.

Now, we would like to investigate the convergence of the standard finite element approximation $u_h \in V^k_h$ defined in Section 2.3.2, when measured in standard and weighted $L^\infty(\Omega)$ norms.

Let $u \in H^1_0(\Omega)$ be the solution of the Poisson problem (2.1). The following error estimates can be found in [143, Section 0].

Theorem 3.2.3 (Pollution effect in maximum norm)**.** *Let $k \in \mathbb{Z}_+$, and let the right-hand side $f \in N^{k-1,\sigma}_{\beta}(\Omega)$ for some $\beta \geq \max(0, k+1-\lambda_1)$ and $\sigma > 0$. Moreover, let $\widetilde{\Omega} \subset \Omega$ be separated from the corner, namely $\operatorname{dist}(\widetilde{\Omega}, 0) > \delta > 0$. Then for any $\epsilon > 0$ the standard finite element approximation defined in* (2.10) *satisfies for some $c > 0$ depending on u*

$$\|u - u_h\|_{L^\infty(\Omega)} \leq c h^{\min(k+1,\lambda_1)-\epsilon}, \quad \textit{and} \quad \|u - u_h\|_{L^\infty(\widetilde{\Omega})} \leq c h^{\min(k+1,2\lambda_1)-\epsilon}.$$

We will now present some numerical examples showing that the estimates provided in Theorem 3.2.3 are sharp, at least up to the choice of $\epsilon > 0$, and cannot be improved. This means that the standard finite element approximation yields suboptimal convergence rates compared to the nodal interpolation summarised in Lemma 3.2.2.

We consider the benchmark problem, which we already studied in Section 2.3.2.4 and Section 2.3.3.4. Let $\Omega \subset \mathbb{R}^2$ be the L-shape domain with the largest interior angle $\Theta = 3\pi/2$. Consider the exact solution $u = s_1 + s_2 + s_3$ and let $\beta = k + 1 - \lambda_1$. Then the exact solution satisfies $u \in V^{k+1,\infty}_{\beta}(\Omega)$.

We approximate the solution of the Poisson problem (2.1) with the standard piecewise polynomial finite element of order $k = 1, 2, 3$. The approximation errors obtained upon the discretisation on few consecutive refinement levels are summarised in Table 3.1. Note that the weights depend on the order of the discretisation. Furthermore, we take $\widetilde{\Omega} = \Omega \setminus \overline{B(0.5)}$.

The convergence rates are consistent with the findings of Theorem 3.2.3 implying the sharpness of the estimates. Moreover, the increase in the order of the polynomials used in the finite element method does not qualitatively improve the approximation properties. For the sake of completness and comparison, we include also errors measured in $L^\infty_\beta(\Omega)$-norm. Local application of the weight around the corner does not lead to a higher convergence order than observed for the errors measured far from it.

P_1 **elements**						
L	$\|u-u_h\|_{L^\infty(\Omega)}$	eoc	$\|u-u_h\|_{L^\infty_\beta(\Omega)}$	eoc	$\|u-u_h\|_{L^\infty(\widetilde{\Omega})}$	eoc
1	1.1320e-2	-	1.1320e-2	-	1.1320e-2	-
2	2.2810e-2	-1.01	9.0520e-3	0.32	2.2810e-2	-1.01
3	1.9100e-2	0.26	3.6329e-3	1.32	9.1494e-3	1.32
4	1.3227e-2	0.53	1.4699e-3	1.31	3.6697e-3	1.32
5	8.6316e-3	0.62	5.8631e-4	1.33	1.4714e-3	1.32
6	5.5122e-3	0.65	2.3372e-4	1.33	5.8787e-4	1.32
7	3.4911e-3	0.66	9.3089e-5	1.33	2.3430e-4	1.33
Expected		0.67		1.33		1.33
P_2 **elements**						
L	$\|u-u_h\|_{L^\infty(\Omega)}$	eoc	$\|u-u_h\|_{L^\infty_\beta(\Omega)}$	eoc	$\|u-u_h\|_{L^\infty(\widetilde{\Omega})}$	eoc
1	2.1860e-2	-	4.3375e-3	-	2.1860e-2	-
2	1.6048e-2	0.45	1.0654e-3	2.03	3.3171e-3	2.72
3	1.0679e-2	0.59	4.4250e-4	1.27	1.7702e-3	0.91
4	6.8697e-3	0.64	1.7485e-4	1.34	7.0756e-4	1.32
5	4.3632e-3	0.65	6.9152e-5	1.34	2.7985e-4	1.34
6	2.7575e-3	0.66	2.7408e-5	1.34	1.1089e-4	1.34
7	1.7393e-3	0.66	1.0877e-5	1.33	4.3985e-5	1.33
Expected		0.67		1.33		1.33
P_3 **elements**						
L	$\|u-u_h\|_{L^\infty(\Omega)}$	eoc	$\|u-u_h\|_{L^\infty_\beta(\Omega)}$	eoc	$\|u-u_h\|_{L^\infty(\widetilde{\Omega})}$	eoc
1	7.7357e-3	-	2.1568e-3	-	4.2138e-3	-
2	5.5680e-3	0.47	3.5099e-4	2.62	2.5791e-3	0.71
3	3.6846e-3	0.60	1.4046e-4	1.32	7.2086e-4	1.84
4	2.3655e-3	0.64	5.6261e-5	1.32	2.8052e-4	1.36
5	1.5012e-3	0.66	2.2319e-5	1.33	1.1121e-4	1.33
6	9.4847e-4	0.66	8.8568e-6	1.33	4.4130e-5	1.33
7	5.9819e-4	0.66	3.5148e-6	1.33	1.7513e-5	1.33
Expected		0.67		1.33		1.33

Table 3.1: Summary of convergence rates on the L-shape domain obtained using Lagrange finite elements P_1, P_2, P_3.

3.3 Maximum norm error estimates for the energy-correction

In this section, we investigate the maximum norm error estimates of the energy-corrected finite element scheme introduced in Section 2.3.3. We begin by stating and proving some auxiliary results regarding the properties of the finite element spaces. Then, we move to the formulation of the main result of this chapter, namely the finite element error estimates in the weighted maximum norm for the energy-corrected scheme.

Lemma 3.3.1 (Inverse inequality)**.** *For all $v_h \in V_h^k$ and all $\alpha \geq -1$, the following estimate holds. Moreover, when the triangulation is uniform, the constant is independent of the choice of the element.*

$$\|v_h\|_{L^\infty_\alpha(T)} \leq ch^{-1}\|v_h\|_{L^2_\alpha(T)} \quad \textit{for all } T \in \mathcal{T}.$$

Proof. The proof follows from the standard scaling argument and the equivalence of finite dimensional norms. Let us consider two separate cases, namely, when the triangle T lies in a positive distance d_T from the corner located at the origin, and when it is attached to the corner. We begin with the former. Notice that

$$c_1 \sup_{\mathbf{x}\in T} |\mathbf{x}| \leq d_T = \inf_{\mathbf{x}\in T} |\mathbf{x}|. \tag{3.5}$$

First, we apply these inequalities and transform the discrete function to the reference triangle $\widehat{T}$ to obtain

$$\|v_h\|_{L^\infty_\alpha(T)} \leq cd_T^\alpha \|v_h\|_{L^\infty(T)} = cd_T^\alpha \|\widehat{v}_h\|_{L^\infty(\widehat{T})}.$$

Due to the equivalence of norms in the finite dimensional spaces we arrive at

$$\|\widehat{v}_h\|_{L^\infty(\widehat{T})} \leq c\|\widehat{v}_h\|_{L^2(\widehat{T})} = ch^{-1}\|v_h\|_{L^2(T)}.$$

The scaling in the last inequality comes from the Jacobian of the transformation between the physical triangle T and the reference triangle $\widehat{T}$. Finally, gathering all the estimates together and applying (3.5) we obtain upon setting $c_3 = cc_1$

$$\|v_h\|_{L^\infty_\alpha(T)} \leq cd_T^\alpha h^{-1}\|v_h\|_{L^2(T)} \leq c_3 h^{-1}\|v_h\|_{L^2_\alpha(T)}. \tag{3.6}$$

Now, assume that the triangle T is attached to the corner. The distance from the corner on the reference triangle satisfies $\widehat{r} \sim hr$, so

$$\|v_h\|_{L^\infty_\alpha(T)} \leq h^\alpha \|\widehat{v}_h\|_{L^\infty_\alpha(\widehat{T})}.$$

Similarly as before, we use the equivalence of the norms on the finite dimensional spaces to obtain

$$\|\widehat{v}_h\|_{L^\infty_\alpha(\widehat{T})} \leq c\|\widehat{v}_h\|_{L^2_\alpha(\widehat{T})} \leq ch^{-1-\alpha}\|v_h\|_{L^2_\alpha(T)}.$$

In the last step we again applied the transformation between the reference and physical element. Combining all the estimates together we arrive at

$$\|v_h\|_{L^\infty_\alpha(T)} \leq c_2 h^{-1}\|v_h\|_{L^2_\alpha(T)}.$$

Choosing the bigger one of the constants c_2, c_3 and accounting for (3.6) completes the proof of the lemma. □

The result presented in the next lemma is a form of an inverse inequality between standard and weighted $L^2(\Omega)$ norms in the neighbourhood of the corner.

Lemma 3.3.2. *Let $\alpha > -1$ and $T^* \in \mathcal{T}$ be a single element of the triangulation that lies close to the corner located at the origin, namely $\max_{\mathbf{x}\in T^*} r(\mathbf{x}) < \tilde{c}h$ for some positive constant $\tilde{c}$. Then the following estimate holds for all $v_h \in V_h^k$*

$$h^\alpha\|v_h\|_{L^2(T^*)} \leq c\|v_h\|_{L^2_\alpha(T^*)}.$$

Proof. Let first $\alpha \geq 0$. If the element T^* is not attached to the corner, then it lies in the distance of at least h from it and hence the estimate is obvious.

Suppose then that the triangle T^* shares a vertex with the corner. We proceed similarly as before, exploiting the equivalence of norms in the finite dimensional spaces. Thus, we write

$$\begin{aligned} h^\alpha\|v_h\|_{L^2(T^*)} &= h^\alpha\|\widehat{v}_h\|_{L^2(\widehat{T})} \\ &\leq ch^\alpha\|\widehat{v}_h\|_{L^2_\alpha(\widehat{T})} = c\|\widehat{v}_h\|_{L^2_\alpha(T^*)}. \end{aligned}$$

Suppose now that $\alpha \in (-1, 0)$. Since $r \leq \tilde{c}h$, we have $h^\alpha \leq \frac{1}{\tilde{c}^\alpha} r^\alpha$ and

$$h^\alpha\|v_h\|_{L^2(T^*)} \leq c\|v_h\|_{L^2_\alpha(T^*)}$$

□

3.3.1 Main result

Following [7, 8, 142, 143], we introduce a dyadic decomposition around the corner of the domain Ω. This construction will be important in the proof of the global finite element estimates, where we will use local estimates on each part of the decomposition.

Definition 3.3.3 (Dyadic decomposition)**.** *Let $R > 0$. For $J = 0, \dots, I$ we set*

$$\Omega_J = \{x \in \Omega : \ d_{J+1} < |x| < d_J\},$$

where $d_J = 2^{-J}R$ and $d_I \leq c^* h$, $d_{I+1} = 0$.

Moreover, d_I is chosen so that the correction patch $\mathcal{S}_h$ of (2.21) is contained in Ω_I. We also define $\Omega_{-1} = \Omega \setminus \bigcup_{J=0}^{I} \Omega_J$ and $d_{-1} = \operatorname{diam}(\Omega)$.

Finally, we set

$$\Omega'_J = \Omega_{J+1} \cup \Omega_J \cup \Omega_{J-1}.$$

In the proof of the maximum norm estimates, we will exploit the known local estimates derived by Schatz and Wahlbin in [142, Theorem 5.1]. The form presented here can be found in [7, (3.11)] and is a special case of the estimate proven in [165, Theorem 10.1].

Lemma 3.3.4 (Interior maximum norm estimate)**.** *Let $J < I - 1$ and suppose that the following holds*

$$a(u - u_h, v_h) = 0 \quad \text{for all } v_h \in V_h^k(\Omega'_J).$$

Then, we have

$$\|u - u_h\|_{L^\infty(\Omega_J)} \leq c\Big(|\log h|^s \inf_{\chi \in V_h^k} \|u - \chi\|_{L^\infty(\Omega'_J)} + d_J^{-1} \|u - u_h\|_{L^2(\Omega'_J)}\Big),$$

where $s = 1$, when $k = 1$ and $s = 0$ otherwise.

Finally, we are in a position to state the main result of this chapter.

Theorem 3.3.5. *Let $\sigma \in (0, 1)$, $\beta \in \big(\max(0, k + 1 + \sigma - \lambda_1), k + 1 + \sigma + \lambda_1\big)$, and also $k - \lambda_1 < \alpha < k$, $\tilde{\alpha} = \alpha - k + 1$. Assume that $f \in N_\beta^{k-1,\sigma}(\Omega) \cap H_{-\tilde{\alpha}}^{k-1}(\Omega)$. Then the energy-corrected finite element approximation (2.20) of (2.1) admits the following convergence*

$$\|u - u_h^m\|_{L^\infty_\beta(\Omega)} \leq c h^{k+1} |\log h|^s (\|f\|_{N_\beta^{k-1,\sigma}(\Omega)} + \|f\|_{H_{-\tilde{\alpha}}^{k-1}(\Omega)}),$$

where $s = 1$, when $k = 1$ and $s = 0$ otherwise.

Proof. Without loss of generality, we assume that the corner lies at the origin. We also assume that $\Theta > \frac{\pi}{k+1}$ since otherwise standard methods provide the desired convergence order of the scheme, see Theorem 3.2.3.

In the proof we consider cases $J = I-1, I$ and the case $J < I-1$ separately.

For $J < I-1$ we can rely on Theorem 3.3.4. This result holds under the Galerkin orthogonality assumption $a(u-u_h^m, v_h) = 0$. Although this does not hold globally for the energy-corrected finite element scheme, it is satisfied for functions $v_h \in V_h^k(\Omega'_J)$ with support in Ω'_J. This is true due to the Galerkin orthogonality property for the energy-corrected scheme stated in Theorem 2.3.25, since $\Omega'_J \cap \mathcal{S}_h = \emptyset$.

As a consequence, we immediately obtain

$$\begin{aligned}\|u-u_h^m\|_{L^\infty_\beta(\Omega_J)} &\le d_J^\beta \|u-u_h^m\|_{L^\infty(\Omega_J)} \\ &\le c d_J^\beta \Big(|\log h|^s \inf_{\chi\in V_h^k} \|u-\chi\|_{L^\infty(\Omega'_J)} + d_J^{-1}\|u-u_h^m\|_{L^2(\Omega'_J)}\Big).\end{aligned}$$

Furthermore, since for some constant independent of J

$$d_J \le c \inf_{\mathbf{x}\in\Omega'_J} r(\mathbf{x}),$$

we can also write

$$\|u-u_h^m\|_{L^\infty_\beta(\Omega_J)} \le c\Big(|\log h|^s \inf_{\chi\in V_h^k} \|u-\chi\|_{L^\infty_\beta(\Omega'_J)} + \|u-u_h^m\|_{L^2_{\beta-1}(\Omega'_J)}\Big). \tag{3.7}$$

Now, we move our investigations to subregions, which are close to the corners of the domain, namely, we consider $J = I-1, I$. Let $T^* \in \mathcal{T}$ denote the element in the domain's triangulation, in which the maximum error of the scheme, when measured on Ω_J only, is attained. Note also that $T^* \subset \Omega'_J$. Then, for any $v_h \in V_h^k$

$$\|u-u_h^m\|_{L^\infty_\beta(\Omega_J)} \le \|u-u_h^m\|_{L^\infty_\beta(T^*)} \le \|u-\chi\|_{L^\infty_\beta(\Omega'_J)} + \|\chi-u_h^m\|_{L^\infty_\beta(T^*)} \tag{3.8}$$

We now focus our attention on the second term in this estimate. Applications of the inverse inequality between L^∞ and L^2 norms of discrete functions, as stated in Lemma 3.3.1 for $\alpha = 0$, and Lemma 3.3.2 give for some constants $c_J, c'_J > 0$

$$\begin{aligned}\|\chi-u_h^m\|_{L^\infty_\beta(T^*)} &\le c_J h^\beta \|\chi-u_h^m\|_{L^\infty(T^*)} \\ &\le c_J h^{\beta-1}\|\chi-u_h^m\|_{L^2(T^*)} \le c'_J \|\chi-u_h^m\|_{L^2_{\beta-1}(T^*)}.\end{aligned}$$

Let β' satisfy $\beta > \beta' > k+1+\sigma-\lambda_1$. Then $r^{\beta-\beta'-1} \in L^2(\Omega)$ and a simple application of the Hölder inequality leads to

$$\|u-\chi\|_{L^2_{\beta-1}(\Omega'_J)} \le c\|u-\chi\|_{L^\infty_{\beta'}(\Omega'_J)}.$$

Hence, for $c''_J > 0$ we have

$$\begin{aligned} c'_J\|\chi - u_h^m\|_{L^2_{\beta-1}(T^*)} &\leq c'_J\|u - u_h^m\|_{L^2_{\beta-1}(\Omega'_J)} + c'_J\|u - \chi\|_{L^2_{\beta-1}(\Omega'_J)} \\ &\leq c'_J\|u - u_h^m\|_{L^2_{\beta-1}(\Omega'_J)} + c''_J\|u - \chi\|_{L^\infty_{\beta'}(\Omega'_J)} \end{aligned}$$

Therefore, we obtain for all $\chi \in V_h^k$ using (3.8)

$$\|u - u_h^m\|_{L^\infty_\beta(\Omega_J)} \leq c\Big(\|u - \chi\|_{L^\infty_{\beta'}(\Omega'_J)} + \|u - u_h^m\|_{L^2_{\beta-1}(\Omega'_J)}\Big). \tag{3.9}$$

Combining (3.7) and (3.9) we see that for some $c > 0$ and for any $\beta > \beta' > k{+}1{+}\sigma{-}\lambda_1$

$$\|u - u_h^m\|_{L^\infty_\beta(\Omega)} \leq c\Big(|\log h|^s \inf_{\chi\in V_h^k} \|u - \chi\|_{L^\infty_{\beta'}(\Omega)} + \|u - u_h^m\|_{L^2_{\beta-1}(\Omega)}\Big).$$

We also use the embedding $L^\infty_{\beta'}(\Omega) \hookrightarrow L^\infty_\beta(\Omega)$, which allows us to replace the $L^\infty_\beta(\Omega)$-norm with the $L^\infty_{\beta'}(\Omega)$-norm in (3.7). Application of the interpolation error estimate from Lemma 3.2.2 and the energy-corrected finite element estimates in weighted $L^2(\Omega)$-norm stated in Theorem 2.3.27 completes the proof. Note that the weight $\alpha = \beta - 1$ is exactly the one required there.

When the considered angle satisfies $\frac{\pi}{k+1} < \Theta < \frac{\pi}{k}$, then $-1 < \beta - 1 < 0$ and the results from Lemma 2.3.29 need to be used instead. □

Remark 3.3.6. *In the case of general polygonal domains, the energy-correction needs to be applied to all corners, for which the singular functions* (2.2) *influence the regularity stated in Theorem 3.1.2. These are exactly the corners, for which* $\Theta > \frac{\pi}{k+1}$*, where* k *is the order of polynomials used in the finite element discretisation. A similar condition also appears in schemes employing mesh grading on domains with corners, see [144].*

Remark 3.3.7. *Using the same argumentation, the following error estimate in weighted maximum norm for the standard finite element approximation* $u_h \in V_h^k$ *defined in Section 2.3.2 can be shown*

$$\|u - u_h\|_{L^\infty_\alpha(\Omega)} \leq ch^{\min(k+1,2\lambda_1)-\epsilon},$$

where $\alpha > \lambda_1$*. This convergence order is significantly lower than the one resulting from the energy-corrected approximation. Its sharpness is confirmed in numerical experiments summarised in Table 3.1.*

The result of Theorem 3.3.5 means that the energy-corrected finite element exhibits, up to a logarithmic factor, the same approximation properties as the interpolation operator, see Lemma 3.2.2. The same correction bilinear form $c_h(\cdot,\cdot)$, see (2.21), is used as necessary for the optimal convergence of the scheme when measured in the weighted $L^2(\Omega)$ norm. This follows directly from the proof, as the $L^2_{\beta-1}(\Omega)$ error appears in (3.10).

The value of the weight β applied in the vicinity of the corner cannot be lowered. However, the weight does not have any effect on the convergence of the scheme considered in a positive distance from the corner. Hence, similar convergence order of the energy-corrected finite element scheme, when considered outside of some neighbourhood of the corner, follows.

Corollary 3.3.8. *Suppose that the assumptions of Theorem 3.3.5 are satisfied. Furthermore, let* $\widetilde{\Omega} \subset \Omega$ *satisfy* $\operatorname{dist}\big(\widetilde{\Omega}, 0\big) > 0$. *Then the energy-corrected finite element approximation yields*

$$\|u - u_h^m\|_{L^\infty(\widetilde{\Omega})} \leq c h^{k+1} |\log h|^s (\|f\|_{N^{k-1,\sigma}_\beta(\Omega)} + \|f\|_{H^{k-1}_{-\tilde{\alpha}}(\Omega)}),$$

where $s = 1$, *when* $k = 1$ *and* $s = 0$ *otherwise.*

Finally, we can say that the pollution effect in $L^\infty(\Omega)$ space summarised in Theorem 3.2.3 can be successfully eliminated using the energy-corrected scheme.

3.4 Numerical results

In this section, we present numerical experiments validating the analysis conducted above. For this purpose, we consider two examples with known analytical solutions.

First, we choose the L-shape domain $\Omega = (-1,1)^2 \setminus \big([0,1] \times [-1,0]\big)$ with the largest interior angle of size $\Theta = 3\pi/2$. In the second example, we set the domain $\Omega = (-1,1)^2 \setminus \{-x \leq y \leq 0\}$, which has the angle $\Theta = 7\pi/4$ located at the origin. We call the latter *the Pac-Man domain*[1]. The initial triangulations of both domains are presented in Figure 3.1.

We use a known exact solution $u = s_1 + s_2 + s_3$, where s_i are singular functions (2.2) corresponding to the considered angles. For the sake of presentation, it is chosen so that the energy-correction needs to be applied only around the origin, see Remark 3.3.6 in a more general case. In the experiments, we choose the weight

[1] We name the domain after a character from a computer game popular in the 1980s, whose figure included a re-entrant corner of size $7\pi/4$ and resembled the defined domain.

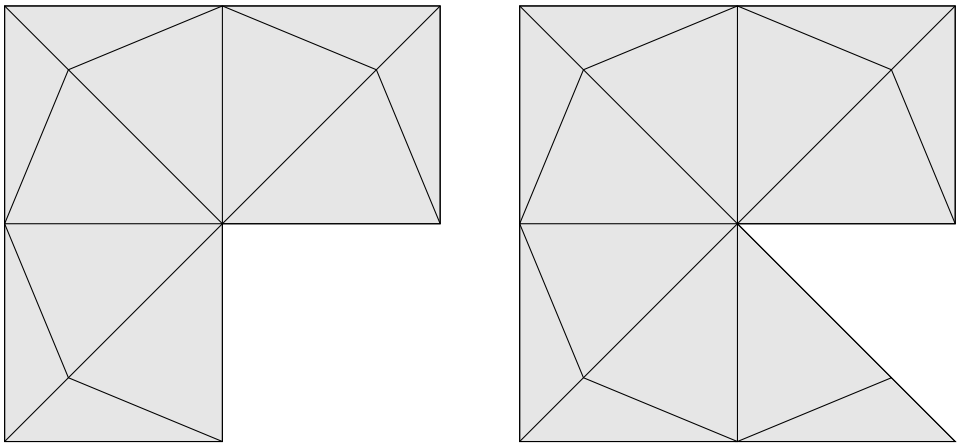


Figure 3.1: Initial triangulations of the L-shape domain (left) and Pac-Man domain (right).

$\beta = k + 1 - \lambda_1$, where $k = 1, 2, 3$ is the order of polynomials used in the discretizations and $\lambda_1 = \pi/\Theta$. This choice induces a slightly stronger norm than assumed in Theorem 3.3.5 but the optimal convergence order of the energy-corrected scheme can be observed regardless of this.

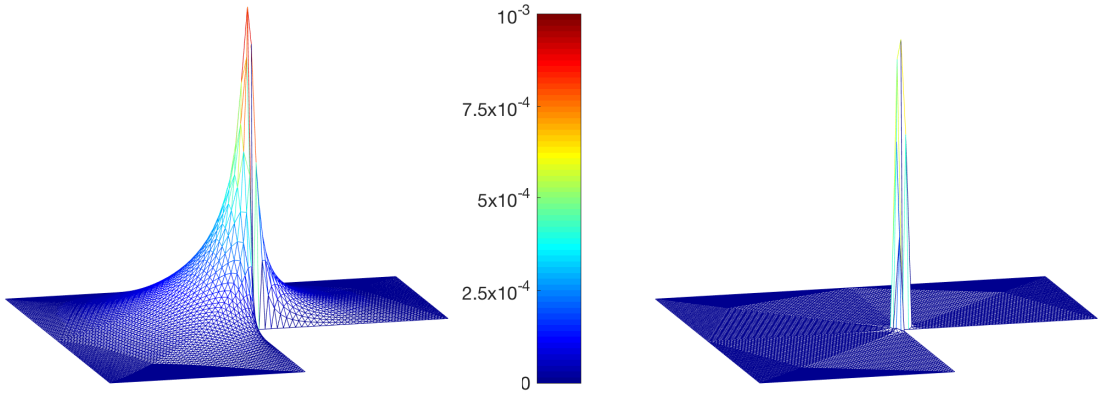


Figure 3.2: Comparison of errors of piecewise quadratic finite element schemes with (right) and without energy-correction (left) for $\Theta = 3\pi/2$.

The parameters γ in the modification (2.33) are computed using a version of the Newton algorithm for finding the roots of the energy defect function (2.22) described in [86, Section 6.1.1]. The precise values used in the experiments are included in the tables below. To fulfil the requirements (U) imposed on the mesh, we perform

computations with initial meshes including only identical isosceles triangles around the origin, 6 and 7 in the respective cases of L-shape and Pac-Man domains, see Figure 3.1.

In Table 3.2 and Table 3.3 we summarise the errors obtained using the energy-corrected finite element method in standard and weighted $L^\infty(\Omega)$ norms on the L-shape ($\Theta = 3\pi/2$) and Pac-Man ($\Theta = 7\pi/4$) domain respectively. In the case of the piecewise cubic energy-corrected approximation, we omit the first refinement level, since the initial coarse mesh does not allow for defining three necessary correction rings when the modification $c^R(\cdot,\cdot)$ is used. For comparison, we also include the errors obtained using the standard finite element method in the two leftmost parts of the tables. Finally, let $\widetilde{\Omega} = \Omega \setminus \overline{B(0.5)}$ be a part of the domain Ω containing all the points in a distance larger than 0.5 from the corner. In the rightmost parts of the tables we show the errors in the $L^\infty(\widetilde{\Omega})$ norm.

The application of the energy-correction does not, in general, improve the convergence order of the scheme, when measured in the standard $L^\infty(\Omega)$ norm, although astrong preasymptotic behaviour of the piecewise linear discretisation can be observed. This, however, can be expected, as the same approximation order is yielded by the interpolation operator in Lemma 3.2.2, where $\beta = 0$ is taken. Nevertheless, the numerical error of the energy-corrected scheme is in all studied cases quantitatively smaller than that of a standard finite element scheme. When the piecewise linear discretisation is considered, we observe a significant pre-asymptotic behaviour with a higher approximation order.

The numerically estimated convergence orders in the weighted $L^\infty_\beta(\Omega)$ norms confirm the predictions of Theorem 3.3.5. This is a significant improvement compared to the standard finite element discretisation, resulting in the order $k+1$ in place of $2\pi/\Theta$. Since the energy-corrected discretisation yields optimal convergence order, compared to the interpolation error, when measured in the weighted norm, see Lemma 3.2.2, it also converges optimally when considered in a fixed, positive distance from the corner, for instance when measured on $\widetilde{\Omega}$. The numerical investigations shown in the rightmost columns of the tables confirm these findings. Finally, we would like to note that the higher-order energy-correction methods using $c^F(\cdot,\cdot)$ give quantitatively better results than the methods involving the modification $c^R(\cdot,\cdot)$. This behaviour was already observed in [86] and we discussed it in Section 2.3.3.4.

In Figure 3.2, we plot the absolute values of the differences between the exact solutions and the piecewise quadratic finite element approximations. On the left-hand side, the error obtained using the standard discretisation is shown. The error is

substantial not only in the neighbourhood of the corner but propagates over the whole domain. On the other hand, the energy-corrected scheme successfully eliminates the pollution effect from Theorem 3.2.3. Therefore, the error of the energy-corrected scheme plotted on the right-hand side is large only in the vicinity of the corner and is significantly smaller in the remaining parts of the domain, when compared with the standard discretisation.

Energy-corrected P_1 elements, $\gamma = 0.117531611518762$										
L	$\|u-u_h\|_{L^\infty(\Omega)}$	eoc	$\|u-u_h\|_{L^\infty_\beta(\Omega)}$	eoc	$\|u-u_h^m\|_{L^\infty(\Omega)}$	eoc	$\|u-u_h^m\|_{L^\infty_\beta(\Omega)}$	eoc	$\|u-u_h^m\|_{L^\infty(\widetilde{\Omega})}$	eoc
1	1.1320e-2	-	1.1320e-2	-	5.7383e-2	-	5.7382e-2	-	5.7382e-2	-
2	2.2810e-2	-1.01	9.0520e-3	0.32	3.0394e-2	0.92	1.2062e-2	2.25	3.0394e-2	0.92
3	1.9100e-2	0.26	3.6329e-3	1.32	1.1326e-2	1.42	1.7838e-3	2.76	3.7011e-3	3.03
4	1.3227e-2	0.53	1.4699e-3	1.31	4.1396e-3	1.45	2.5873e-4	2.79	4.8867e-4	2.92
5	8.6316e-3	0.62	5.8631e-4	1.33	1.5643e-3	1.40	3.9999e-5	2.69	7.8657e-5	2.64
6	5.5122e-3	0.65	2.3372e-4	1.33	6.1449e-4	1.35	6.6710e-6	2.58	1.4737e-5	2.42
7	3.4911e-3	0.66	9.3089e-5	1.33	2.6763e-4	1.20	1.4622e-6	2.19	3.5095e-6	2.07
Expected		0.67		1.33		0.67		2.00		2.00
Energy-corrected P_2 elements with $c_h^R(\cdot,\cdot)$, $\gamma = (0.0315200, -0.005533)$										
L	$\|u-u_h\|_{L^\infty(\Omega)}$	eoc	$\|u-u_h\|_{L^\infty_\beta(\Omega)}$	eoc	$\|u-u_h^m\|_{L^\infty(\Omega)}$	eoc	$\|u-u_h^m\|_{L^\infty_\beta(\Omega)}$	eoc	$\|u-u_h^m\|_{L^\infty(\widetilde{\Omega})}$	eoc
1	2.1860e-2	-	4.3375e-3	-	2.2319e-2	-	1.5701e-2	-	2.2319e-2	-
2	1.6048e-2	0.45	1.0654e-3	2.03	1.2636e-2	0.45	1.9651e-3	3.00	9.9033e-3	1.17
3	1.0679e-2	0.59	4.4250e-4	1.27	7.4690e-3	0.59	1.5007e-4	3.71	4.2364e-4	4.55
4	6.8697e-3	0.64	1.7485e-4	1.34	4.6157e-3	0.64	1.5591e-5	3.27	2.3460e-5	4.17
5	4.3632e-3	0.65	6.9152e-5	1.34	2.8855e-3	0.65	1.6995e-6	3.20	1.8175e-6	3.69
6	2.7575e-3	0.66	2.7408e-5	1.34	1.8122e-3	0.66	1.9279e-6	3.14	1.5387e-7	3.56
Expected		0.67		1.33		0.67		3.00		3.00
Energy-corrected P_2 elements with $c_h^F(\cdot,\cdot)$, $\gamma = (0.1024101, -0.1680210)$										
L	$\|u-u_h\|_{L^\infty(\Omega)}$	eoc	$\|u-u_h\|_{L^\infty_\beta(\Omega)}$	eoc	$\|u-u_h^m\|_{L^\infty(\Omega)}$	eoc	$\|u-u_h^m\|_{L^\infty_\beta(\Omega)}$	eoc	$\|u-u_h^m\|_{L^\infty(\widetilde{\Omega})}$	eoc
1	2.1860e-2	-	4.3375e-3	-	2.5085e-2	-	1.8816e-2	-	2.3331e-2	-
2	1.6048e-2	0.45	1.0654e-3	2.03	1.0644e-2	1.24	4.1909e-4	5.49	1.6292e-3	3.84
3	1.0679e-2	0.59	4.4250e-4	1.27	5.7149e-3	0.90	4.4648e-5	3.23	8.1159e-5	4.33
4	6.8697e-3	0.64	1.7485e-4	1.34	3.2046e-3	0.83	4.9678e-6	3.17	1.2924e-5	2.65
5	4.3632e-3	0.65	6.9152e-5	1.34	1.8650e-3	0.78	6.4579e-7	2.94	1.7541e-6	2.88
6	2.7575e-3	0.66	2.7408e-5	1.34	1.1655e-3	0.68	8.1791e-8	2.98	2.2807e-7	2.94
Expected		0.67		1.33		0.67		3.00		3.00
Energy-corrected P_3 elements with $c_h^R(\cdot,\cdot)$, $\gamma = (0.012891, -0.0023667)$										
L	$\|u-u_h\|_{L^\infty(\Omega)}$	eoc	$\|u-u_h\|_{L^\infty_\beta(\Omega)}$	eoc	$\|u-u_h^m\|_{L^\infty(\Omega)}$	eoc	$\|u-u_h^m\|_{L^\infty_\beta(\Omega)}$	eoc	$\|u-u_h^m\|_{L^\infty(\widetilde{\Omega})}$	eoc
1	7.7357e-3	-	2.1568e-3	-	7.5928e-3	-	6.2087e-3	-	6.2087e-3	-
2	5.5680e-3	0.47	3.5099e-4	2.62	4.5109e-3	0.75	4.7663e-4	3.70	3.7451e-3	0.73
3	3.6846e-3	0.60	1.4046e-4	1.32	2.6220e-3	0.78	2.1275e-5	4.49	1.8711e-4	4.32
4	2.3655e-3	0.64	5.6261e-5	1.32	1.6339e-3	0.68	1.3379e-6	3.99	8.8010e-6	4.41
5	1.5012e-3	0.66	2.2319e-5	1.33	1.0248e-3	0.67	8.4020e-8	3.99	5.4275e-7	4.02
6	9.4847e-4	0.66	8.8568e-6	1.33	6.4449e-4	0.67	5.2556e-9	4.00	3.4767e-8	3.96
Expected		0.67		1.33		0.67		4.00		4.00
Energy-corrected P_3 elements with $c_h^F(\cdot,\cdot)$, $\gamma = (0.042218, -0.07085)$										
L	$\|u-u_h\|_{L^\infty(\Omega)}$	eoc	$\|u-u_h\|_{L^\infty_\beta(\Omega)}$	eoc	$\|u-u_h^m\|_{L^\infty(\Omega)}$	eoc	$\|u-u_h^m\|_{L^\infty_\beta(\Omega)}$	eoc	$\|u-u_h^m\|_{L^\infty(\widetilde{\Omega})}$	eoc
1	7.7357e-3	-	2.1568e-3	-	1.1025e-2	-	8.0471e-3	-	8.0471e-3	-
2	5.5680e-3	0.47	3.5099e-4	2.62	4.6299e-3	1.25	1.1971e-4	6.07	1.2066e-3	2.74
3	3.6846e-3	0.60	1.4046e-4	1.32	2.7603e-3	0.75	7.0327e-6	4.09	3.6505e-5	5.05
4	2.3655e-3	0.64	5.6261e-5	1.32	1.7723e-3	0.64	4.4799e-7	3.97	1.9441e-6	4.23
5	1.5012e-3	0.66	2.2319e-5	1.33	1.1321e-3	0.65	2.8391e-8	3.98	1.4697e-7	3.75
6	9.4847e-4	0.66	8.8568e-6	1.33	7.1711e-4	0.66	1.9533e-9	3.86	9.7768e-9	3.91
Expected		0.67		1.33		0.67		4.00		4.00

Table 3.2: Summary of convergence rates on the L-shape domain obtained using energy-corrected Lagrange finite elements P_1, P_2, P_3 with different types of correction bilinear forms.

Energy-corrected P_1 elements, $\gamma = 0.18617957$										
L	$\|u-u_h\|_{L^\infty(\Omega)}$	eoc	$\|u-u_h\|_{L^\infty_\beta(\Omega)}$	eoc	$\|u-u_h^m\|_{L^\infty(\Omega)}$	eoc	$\|u-u_h^m\|_{L^\infty_\beta(\Omega)}$	eoc	$\|u-u_h^m\|_{L^\infty(\tilde{\Omega})}$	eoc
1	1.5662e-2	-	1.5662e-2	-	7.2605e-2	-	7.2604e-2	-	7.2605e-2	-
2	3.7435e-2	-1.26	1.3907e-2	0.17	5.1037e-2	0.51	1.8960e-2	1.94	5.1037e-2	0.51
3	3.4296e-2	0.13	6.3299e-3	1.34	2.2443e-2	1.19	3.4897e-3	2.44	8.0102e-3	2.67
4	2.5699e-2	0.42	2.8704e-3	1.14	9.9787e-3	1.17	6.5552e-4	2.41	1.3809e-3	2.54
5	1.8064e-2	0.51	1.3122e-3	1.13	4.5199e-3	1.14	1.2830e-4	2.35	3.0774e-4	2.17
6	1.2382e-2	0.54	5.9458e-4	1.14	2.0894e-3	1.11	2.6127e-5	2.30	6.7657e-5	2.19
7	8.3975e-3	0.56	2.6944e-4	1.14	9.9080e-4	1.08	5.4307e-6	2.27	1.4072e-5	2.27
Expected		0.57		1.14		0.57		2.00		2.00
Energy-corrected P_2 elements with $c_h^R(\cdot,\cdot)$, $\gamma = (0.0736961, -0.0196752)$										
L	$\|u-u_h\|_{L^\infty(\Omega)}$	eoc	$\|u-u_h\|_{L^\infty_\beta(\Omega)}$	eoc	$\|u-u_h^m\|_{L^\infty(\Omega)}$	eoc	$\|u-u_h^m\|_{L^\infty_\beta(\Omega)}$	eoc	$\|u-u_h^m\|_{L^\infty(\tilde{\Omega})}$	eoc
1	3.6503e-2	-	6.7804e-3	-	3.7209e-2	-	3.3495e-2	-	3.3495e-2	-
2	2.9394e-2	0.31	2.2959e-3	1.56	2.8982e-2	0.36	5.0024e-3	2.74	2.6383e-2	0.34
3	2.1266e-2	0.47	1.0646e-3	1.11	1.4218e-2	1.03	4.4682e-4	3.48	1.6360e-3	4.01
4	1.4749e-2	0.53	4.7991e-4	1.15	8.9041e-3	0.68	4.3258e-5	3.37	1.1484e-4	3.83
5	1.0053e-2	0.55	2.1751e-4	1.14	5.7371e-3	0.63	4.4358e-6	3.29	1.1002e-5	3.38
6	6.8018e-3	0.56	9.8442e-5	1.14	3.7587e-3	0.61	4.7750e-7	3.22	1.0199e-6	3.43
Expected		0.57		1.14		0.57		3.00		3.00
Energy-corrected P_2 elements with $c_h^F(\cdot,\cdot)$, $\gamma = (0.2371897, -0.4160492)$										
L	$\|u-u_h\|_{L^\infty(\Omega)}$	eoc	$\|u-u_h\|_{L^\infty_\beta(\Omega)}$	eoc	$\|u-u_h^m\|_{L^\infty(\Omega)}$	eoc	$\|u-u_h^m\|_{L^\infty_\beta(\Omega)}$	eoc	$\|u-u_h^m\|_{L^\infty(\tilde{\Omega})}$	eoc
1	3.6503e-2	-	6.7804e-3	-	5.3662e-2	-	3.3748e-2	-	4.7699e-2	-
2	2.9394e-2	0.31	2.2959e-3	1.56	1.8689e-2	1.52	7.2941e-4	5.53	3.6019e-3	3.73
3	2.1266e-2	0.47	1.0646e-3	1.11	9.7133e-3	0.94	8.6554e-5	3.08	3.2576e-4	3.47
4	1.4749e-2	0.53	4.7991e-4	1.15	5.5317e-3	0.81	1.1240e-5	2.94	5.0920e-5	2.68
5	1.0053e-2	0.55	2.1751e-4	1.14	3.2878e-3	0.75	1.4445e-6	2.96	6.392e-6	3.01
6	6.8018e-3	0.56	9.8442e-5	1.14	2.0516e-3	0.68	1.8807e-7	2.94	8.391e-7	2.91
Expected		0.57		1.14		0.57		3.00		3.00
Energy-corrected P_3 elements with $c_h^R(\cdot,\cdot)$, $\gamma = (0.0406490, -0.0194762, 0.0045036)$										
L	$\|u-u_h\|_{L^\infty(\Omega)}$	eoc	$\|u-u_h\|_{L^\infty_\beta(\Omega)}$	eoc	$\|u-u_h^m\|_{L^\infty(\Omega)}$	eoc	$\|u-u_h^m\|_{L^\infty_\beta(\Omega)}$	eoc	$\|u-u_h^m\|_{L^\infty(\tilde{\Omega})}$	eoc
1	1.5888e-2	-	3.7222e-3	-	-	-	-	-	-	-
2	1.2811e-2	0.31	8.5171e-4	2.13	1.9448e-2	-	5.5003e-3	-	1.5520e-2	-
3	9.2634e-3	0.47	3.9812e-4	1.10	8.0764e-3	1.27	3.1353e-4	4.13	3.3758e-3	2.20
4	6.4263e-3	0.53	1.8010e-4	1.14	4.0592e-3	0.99	1.2057e-5	4.70	2.6139e-6	10.33
5	4.3819e-3	0.55	8.1681e-5	1.14	2.3892e-3	0.76	5.6577e-7	4.41	2.4501e-7	3.42
6	2.9657e-3	0.56	3.6977e-5	1.14	1.5798e-3	0.59	2.8453e-8	4.31	1.6478e-8	3.89
Expected		0.57		1.14		0.57		4.00		4.00
Energy-corrected P_3 elements with $c_h^F(\cdot,\cdot)$, $\gamma = (0.1227554, -0.3070823, 0.1419462)$										
L	$\|u-u_h\|_{L^\infty(\Omega)}$	eoc	$\|u-u_h\|_{L^\infty_\beta(\Omega)}$	eoc	$\|u-u_h^m\|_{L^\infty(\Omega)}$	eoc	$\|u-u_h^m\|_{L^\infty_\beta(\Omega)}$	eoc	$\|u-u_h^m\|_{L^\infty(\tilde{\Omega})}$	eoc
1	1.5888e-2	-	3.7222e-3	-	2.1093e-2	-	1.5679e-2	-	1.7765e-2	-
2	1.2811e-2	0.31	8.5171e-4	2.13	1.0037e-3	1.07	2.3385e-4	6.07	2.5179e-3	2.82
3	9.2634e-3	0.47	3.9812e-4	1.10	5.8585e-3	0.78	1.8147e-5	3.69	1.1997e-4	4.39
4	6.4263e-3	0.53	1.8010e-4	1.14	3.7843e-3	0.63	1.2072e-6	3.91	8.7340e-6	3.78
5	4.3819e-3	0.55	8.1681e-5	1.14	2.5085e-3	0.59	8.2371e-8	3.87	5.9223e-7	3.88
6	2.9657e-3	0.56	3.6977e-5	1.14	1.6668e-3	0.59	5.7123e-9	3.85	4.1840e-8	3.82
Expected		0.57		1.14		0.57		4.00		4.00

Table 3.3: Summary of convergence rates on the Pac-Man domain obtained using energy-corrected Lagrange finite elements P_1, P_2, P_3 with different types of correction bilinear forms.

Chapter 4

Energy-correction for parabolic problems

Numerical approximations of parabolic problems are of great interest due to the many technical applications, in which they appear, and have been extensively studied in multiple settings [45, 109, 133]. Standard numerical aapproximations involve finite difference methods [103, 156], but to allow for computations on more complicated domains finite volume [46] and finite element methods have been developed, among many others, in [33, 65, 73, 155, 166] and in [52, 53, 54, 60] for problems with restricted regularity. For a more exhaustive discussion of the standard finite element approximations of parabolic problems and an extensive list of references, we refer the reader to [157]. Recently, these investigations have been extended to the finite element error estimates in energy norm [68] and maximum norm in [104, 105, 106, 107, 130].

The presence of corners in the computational domain negatively influences regularity properties of the solutions of parabolic problems [21, 77], due to the appearance of certain known singular functions. For corners with angles $\Theta > \pi$, in general, H^2 regularity in space cannot be guaranteed. This introduces the pollution effect, diminishing the convergence order of the standard finite element schemes for parabolic problems [47]. Standard methods for improving the approximation properties of the finite element method in the elliptic setting include refinement [13, 16] and grading [6, 9, 144] of the mesh around the singular corner. These results can also be extended to parabolic problems on non-convex polygonal domains [47]. However, due to the very small mesh size in the vicinity of the re-entrant corner, the CFL condition [103], which guarantees the stability of an explicit time integration scheme, becomes very restrictive. This, in turn, means that explicit time-stepping schemes cannot be efficiently used, as they would require prohibitively small time-steps. Alternative approaches include finite volume element method, for which the

mesh refinement around the corner also improves the convergence properties of the scheme [48]. Also, in [51] the Fourier finite element method for efficient approximations of the solutions and stress-intensity factors was proposed.

In this chapter, we follow a different approach, based on the piecewise linear energy-corrected finite element method. As opposed to the methods involving mesh refinement in the vicinity of the corner, it can be successfully applied on quasi-uniform meshes, and hence, does not suffer from a too severe CFL condition. This permits the construction of fast explicit time-stepping schemes combined with the energy-corrected finite element in space.

This chapter is structured as follows: in Section 4.1 we discuss the regularity properties of the parabolic equations and in Section 4.2, we generalize the piecewise linear energy-corrected finite element to parabolic problems. We prove the error estimates for the semi-discretisation in space and for the fully discrete scheme with explicit time-stepping. The estimates are optimal in terms of the interpolation error. We illustrate the analysis with numerical investigations in Section 4.3. In Section 4.4, we present several potential extensions of the scheme. We show that it can be applied to the advection-diffusion problem with a moderate advection term. Furthermore, we introduce higher-order energy-corrected discretisation combined with mass-lumping technique and a post-processing approach for improving convergence properties of the scheme also in the vicinity of the singular corner. We complete the discretisation with an explicit Runge-Kutta time-stepping and show that the energy-corrected scheme exhibits superior performance compared with other commonly used discretisation methods. Finally, we present a potential application of the method, showing the flexibility of the energy-corrected finite element, which can be applied to problems with multiple re-entrant corners, and in three-dimensional settings.

The publication by the author and B. Wohlmuth entitled "Energy-corrected FEM and explicit time-stepping for parabolic problems" containing the results presented in this chapter is currently in preparation, [153].

4.1 Parabolic problem

Consider a standard heat equation defined on a domain $\Omega \subset \mathbb{R}^d$ in a time interval $[0, T]$ with $T > 0$

$$u_t - \Delta u = f \quad \text{in } \Omega \times (0, T), \tag{4.1}$$

$$u = 0 \quad \text{on } \partial\Omega \times [0, T], \tag{4.2}$$

$$u = u_0 \quad \text{in } \Omega \text{ at } t = 0. \tag{4.3}$$

Similarly as in the case of the elliptic equations (2.1), the classical twice continuously differentiable solution of the heat equation (4.1) might not exist. Therefore, we would like to define the solution in some weaker sense allowing for a less regular behaviour. To do this, we first need to introduce the Bochner spaces, which will prove useful for the time-dependent problems considered here. We follow the definitions presented in [71, Section 5.9.2]. For a comprehensive study of the Bochner spaces we refer the reader to [89].

We begin by defining the spaces of continuous functions with values in Banach spaces.

Definition 4.1.1. *Let X be a real Banach space, $T > 0$ and $\sigma \in (0, 1)$. The space $C^\sigma([0, T]; X)$ consists of all continuous functions $u : [0, T] \to X$, for which the following norm is finite*

$$\|u\|_{C^\sigma([0,T];X)} = \sup_{0 \le t \le T} \|u(t)\|_X + \sup_{0 \le t_1 < t_2 \le T} \frac{\|u(t_1) - u(t_2)\|_X}{|t_1 - t_2|^\sigma} < \infty.$$

Now, we extend the notion of measurability to the time-dependent functions.

Definition 4.1.2 (Strong measurability)**.** *Let X be a real Banach space and let $T > 0$.*

(i) We call a function $s : [0, T] \to X$ simple, if it can be written as

$$s(t) = \sum_{i=1}^{m} \chi_{E_i}(t) u_i,$$

where χ_B are indicator functions of a set B, $E_i \subset [0, T]$ are Lebesgue measurable sets and $u_i \in X$.

(ii) A function $u : [0, T] \to X$ is a strongly measurable function, if there exists a sequence of simple functions $s_m : [0, T] \to X$ such that

$$\lim_{m \to \infty} s_m(t) = u(t), \quad \text{for a.e. } t \in [0, T].$$

We can now introduce the Bochner spaces, which will allow us to define the solution of (4.1) in a weaker sense and investigate its regularity.

Definition 4.1.3 (Bochner spaces)**.** *Let $1 \leq p \leq \infty$, $T > 0$ and X be a real Banach space. The space $L^p(0,T;X)$ consists of all strongly measurable functions $u : [0,T] \to X$ such that*

$$\|u\|_{L^p(0,T;X)} = \left(\int_0^T \|u(t)\|_X^p \, \mathrm{d}t \right)^{1/p} < \infty, \quad \text{for } 1 \leq p < \infty$$
$$\|u\|_{L^\infty(0,T;X)} = \operatorname*{ess\,sup}_{t \in [0,T]} \|u(t)\|_X < \infty.$$

When X is a Sobolev space, then the Bochner spaces extend the definitions introduced in Section 2.1 to functions dependent on time. In particular, they allow for recognition of different natures of temporal and spatial dimensions.

Finally, we are in a position to define the weak solution of the heat equation (4.1).

Definition 4.1.4. *We define a weak solution of the heat equation* (4.1) *as a function $u \in L^2\big(0,T;H_0^1(\Omega)\big)$, $u_t \in L^2\big(0,T;H^{-1}(\Omega)\big)$ such that $u(0) = u_0$ a.e. and for all $v \in H_0^1(\Omega)$ and a.e. $0 < t \leq T$*

$$\langle u_t(t), v \rangle + a(u(t), v) = \langle f, v \rangle \tag{4.4}$$

where $a(u,v) = \big\langle \nabla u, \nabla v \big\rangle$.

Note that the bilinear form $a(\cdot,\cdot)$ is exactly the same as the one governing the Poisson equation from Definition 2.2.1. This will be important later in the construction of the finite element approximation, which will be directly inspired by the schemes considered in the elliptic setting before.

4.1.1 Regularity results

In order to show the impact of the corners in the computational domain on the solution of the problem (4.4), we first consider it in a smooth setting. The following result comes from [71, Chapter 7, Theorem 6] and can be found in a more general form in [22].

Theorem 4.1.5 (Parabolic regularity)**.** *Let $m \in \mathbb{Z}_+$ and let $\Omega \subset \mathbb{R}^2$ be an open domain with a C^{m+1} boundary. We assume that*

$$u_0 \in H^{2m+1}(\Omega), \quad \frac{\mathrm{d}^k f}{\mathrm{d}t^k} \in L^2(0,T;H^{2m-2k}(\Omega)), \ \text{for } k = 0, \ldots, m.$$

Suppose also that the m-th order compatibility condition holds, namely for

$$g_0 = u_0,\ g_1 = f(0) + \Delta u_0, \dots, g_m = \frac{\mathrm{d}^{m-1} f}{\mathrm{d}t^{m-1}}(0) + \Delta g_{m-1}$$

we have

$$g_k \in H_0^1(\Omega), \quad \textit{for all } k = 0, \dots, m.$$

Then, there exists a unique solution of (4.4)

$$\frac{\mathrm{d}^k u}{\mathrm{d}t^k} \in L^2(0, T; H^{2m+2-2k}(\Omega)), \quad \textit{for all } k = 0, \dots, m+1.$$

Furthermore, the following a priori estimate is true for some constant $c > 0$ depending only on m, Ω, T.

$$\sum_{k=0}^{m+1} \Big\| \frac{\mathrm{d}^k u}{\mathrm{d}t^k} \Big\|_{L^2(0,T;H^{2m+2-2k}(\Omega))} \leq c \Big(\sum_{k=0}^{m} \Big\| \frac{\mathrm{d}^k f}{\mathrm{d}t^k} \Big\|_{L^2(0,T;H^{2m-2k}(\Omega))} + \|u_0\|_{H^{2m+1}(\Omega)} \Big).$$

The author of [71] assumes that the boundary $\partial\Omega$ is smooth. This assumption can be relaxed to the ones included here.

This theorem is an equivalent of the elliptic Shift Theorem 2.2.3 and shows, how the regularity assumptions on the initial conditions u_0, the forcing term f and the smoothness of the boundary $\partial\Omega$ influence the regularity of the solution of the model parabolic problem (4.4).

An interesting consequence of Theorem 4.1.5, see [71, Chapter 7, Theorem 7], is the following result, which holds in a smooth setting.

Remark 4.1.6. *Let $\Omega \subset \mathbb{R}^2$ be a domain with a boundary $\partial\Omega$ of class C^∞. Assume also that*

$$u_0 \in C^\infty(\overline{\Omega}), \quad \textit{and } f \in C^\infty(\overline{\Omega} \times [0, T]).$$

Suppose also that the compatibility conditions stated in Theorem 4.1.5 hold true for all $m \in \mathbb{Z}_+ \cup \{0\}$. Then the parabolic problem from Definition 4.1.4 has a unique solution satisfying

$$u \in C^\infty(\overline{\Omega} \times [0, T]).$$

This means that for smooth domains, under sufficient smoothness assumptions on the initial conditions u_0 and the forcing term f, an arbitrary smoothness of the

solution of the heat equation (4.4) can be expected. However, this does not hold anymore, if polygonal domains are concerned.

From now on, unless stated otherwise, let $\Omega \subset \mathbb{R}^2$ be a bounded, non-convex, polygonal domain in the sense of Definition 2.1.11, that is a domain containing a re-entrant corner of size $\pi < \Theta < 2\pi$. As before, for the sake of simplicity we shall assume that this domain contains only one such corner located at the origin. Note however that the analysis presented here also applies in a more general setting of domains with an arbitrary number of re-entrant corners.

The following regularity result for the parabolic system (4.1) was proposed in [21].

Theorem 4.1.7. *Let $f \in C^{\sigma}\big([0,T], L^2(\Omega)\big)$, $\sigma > 0$. Then there exists a unique solution $u \in C([0,T], H^2_{\alpha}(\Omega))$ of (4.4), where $1 - \lambda_1 < \alpha < 1$. Moreover, there exists $U \in C\big([0,T]; H^2(\Omega)\big)$ and $k_1(t) \in C\big([0,T)\big) \cap C^1\big((0,T)\big)$ such that*

$$u(t, \mathbf{x}) = U(t, \mathbf{x}) + k_1(t) s_1(\mathbf{x})$$

where s_1 is a singular function defined in (2.2).

The solution of the parabolic problems on polygonal domains exhibits similar splitting into regular and singular parts as elliptic problems, see Theorem 2.2.6. Furthermore, we would like to stress out that the same singular functions arise in the solution of both elliptic and parabolic problems on non-convex polygonal domains. Slightly different, but equivalent form of the singular functions in the parabolic setting was discussed in [77, 78]. As opposed to the case of sufficiently smooth domains considered in Theorem 4.1.5 and Remark 4.1.6, the regularity of the parabolic problems on the non-convex polygonal domains cannot, in general, be improved by the choice of the more regular initial conditions u_0 or the forcing term f.

The regularity of parabolic equations on domains with conical points was further studied in [120, 164] with the results presented in the framework of weighted Sobolev spaces. The following theorem is a consequence of the analysis presented there.

Theorem 4.1.8. *Suppose that the assumptions of Theorem 4.1.7 are satisfied. Furthermore, let $f \in L^2\big(0,T; H^4_{-\alpha}(\Omega)\big)$, $\frac{\mathrm{d}f}{\mathrm{d}t} \in L^2\big(0,T; H^2_{-\alpha}(\Omega)\big)$, $\frac{\mathrm{d}^2 f}{\mathrm{d}t^2} \in L^2\big(0,T; L^2_{-\alpha}(\Omega)\big)$, and $u_0 \in H^3_{-\alpha}(\Omega)$. Assume also that the standard compatibility condition $f(0) + \Delta u_0 \in H^1(\Omega)$ is satisfied. Then,*

$$\max_{0 \le t \le T} \|\Delta u\|_{-\alpha} < \infty, \quad \text{and} \quad \int_0^T \|\Delta u_t\|^2_{-\alpha} < \infty, \quad \text{and} \quad \int_0^T \big\|u_{tt}\big\|^2_{\alpha} < \infty.$$

Note also that the results of Theorem 4.1.7 and Theorem 4.1.8 can be easily extended to multiple singular functions.

4.1.2 Finite element discretisation

In this section we introduce the standard finite element method for parabolic equations and discuss its properties. For now, we concentrate our attention on the piecewise linear approximation in the discrete spaces V_h introduced in (2.8). We will show that the impact, which the presence of the non-convex corners in the computational domain has on the approximation of the elliptic problems, is also directly inherited by the corresponding parabolic problems.

Definition 4.1.9. *We define the semi-discretisation of the problem* (4.4) *in space as finding a continuous function* $u_h : [0,T] \to V_h$ *such that,* $u_{h,t} \in L^2\big(0,T;V_h\big)$ *and*

$$\langle u_{h,t}(t), v_h\rangle + a(u_h(t), v_h) = \langle f, v_h\rangle \quad \text{for all } v_h \in V_h,\ t \in (0,T), \tag{4.5}$$

$$u_h(0) = \mathcal{P}_h u_0, \tag{4.6}$$

where $\mathcal{P}_h$ *is a linear operator on the space* V_h*.*

Definition 4.1.9 depends on the choice of the operator $\mathcal{P}_h$. For now, we assume that $\mathcal{P}_h = R_h$ is the Ritz projection (2.11).

We begin the error analysis of the finite element method for parabolic problems with summarising the known estimates for the standard semi-discrete approximation.

Theorem 4.1.10 (Error estimates for the standard FEM)**.** *Let* $\Omega \subset \mathbb{R}^2$ *be a polygonal domain with the largest interior angle of size* $\Theta > 0$*. Suppose that* $f \in L^2(0,T;H^2(\Omega))$*,* $\frac{\mathrm{d}f}{\mathrm{d}t} \in L^2(0,T;L^2(\Omega))$*,* $u_0 \in H_0^1(\Omega)$*,* $\Delta u_0 \in L^2(\Omega)$ *and* $f(0) + \Delta u_0 \in H_0^1(\Omega)$*. Then*

$$\max_{0\le t\le T} \|u(t) - u_h(t)\|_0 \le ch^\sigma \Big(\|\Delta u_0\|_0 + \|f(0)\|_0 + \int_0^T \big\| \frac{\mathrm{d}f}{\mathrm{d}t}(\tau) \big\|_0 \,\mathrm{d}\tau \Big),$$

where for any $\epsilon > 0$

$$\sigma = \begin{cases} 2, & \text{if } \Theta \le \pi \\ 2\lambda_1 - \epsilon, & \text{if } \Theta > \pi. \end{cases}$$

This means that the finite element approximation on convex polygonal domains exhibits optimal second-order convergence in $L^2(\Omega)$-norm compared to the interpolation error, see Theorem 2.3.10. The proof of this result can be found for example in [157, Theorem 19.2] for the convex case. Similar optimal convergence rates, when measured in the $H^1(\Omega)$-semi-norm and $L^\infty(\Omega)$-norm are also known [157, Theorem 19.6], the former under more restrictive smoothness assumptions on the domain.

As shown in [47, Theorem 3.3] and in [157, Theorem 19.4], the situation is significantly different when corners of angle $\Theta > \pi$ are present in the computational

domain. In the case of the non-convex polygonal domains the expected convergence order is not only lower than in the convex setting, but also considerably lower than the interpolation error estimate from Theorem 2.3.10. As we shall illustrate later using numerical examples, this cannot be improved by imposing more restrictive regularity assumptions on the initial conditions u_0 and the forcing term f. Moreover, the convergence properties are also worse than the ones obtained for the nodal interpolation operator in weighted norms in Theorem 2.3.11, and when measured in a positive distance from the re-entrant corner. This phenomenon is an equivalent of the pollution effect introduced in Theorem 2.3.21 in the case of elliptic equations.

These rates can be improved using suitable mesh-grading techniques, so that the optimal convergence in space is regained [157]. However, as we shall see in Section 4.3, the corresponding CFL condition for explicit time-stepping schemes gets very prohibitive and makes the use of explicit time-stepping schemes less attractive.

Theorem 4.1.11 (Mesh grading for parabolic problems)**.** *Let $\Omega \subset \mathbb{R}^2$ be a polygonal domain with the largest interior angle of size $\Theta > \pi$ and let $\delta \in (0, 1/2]$. Suppose that $f \in L^2(0,T;H^2(\Omega))$, $\frac{\mathrm{d}f}{\mathrm{d}t} \in L^2(0,T;H^\delta(\Omega))$, $u_0 \in H_0^1(\Omega)$, $\Delta u_0 \in L^2(\Omega)$ and $f(0) + \Delta u_0 \in H_0^1(\Omega)$. Then the finite element approximation $u_h(t)$ on a gradually refined mesh with $\mu < \lambda_1$, see Definition 2.3.23, exhibits the following convergence*

$$\max_{0\le t\le T} \|u(t) - u_h(t)\|_0 \le ch^2\Big(\|\Delta u_0\|_0 + \|f(0) + \Delta u_0\|_{H^\delta(\Omega)} + \int_0^T \big\|\frac{\mathrm{d}f}{\mathrm{d}t}(\tau)\big\|_{H^\delta(\Omega)}\,\mathrm{d}\tau\Big),$$

where the constant $c > 0$ depends only on the domain Ω, the final time $T > 0$ and the parameter δ.

We now move to the fully-discrete setting, where also the temporal domain is discretised. We discretise Equation (4.9) using the θ time-stepping scheme. Let us divide the time interval $[0, T]$ into $N \in \mathbb{Z}_+$ time steps of equal lengths Δt, so $t_n = n\Delta t$. The fully discrete approximation of the model problem (4.4) reads as follows.

Definition 4.1.12 (Fully discrete θ-scheme)**.** *Let $\theta \in [0, 1]$. We define the fully discrete solution $U_h^n \in V_h$ for $0 \le n \le N$ as the function satisfying*

$$\Big\langle \frac{U_h^{n+1} - U_h^n}{\Delta t}, v_h \Big\rangle + a\Big(\theta U_h^n + (1-\theta)U_h^{n+1}, v_h\Big) = \big\langle \theta f(t_n) + (1-\theta) f(t_{n+1}), v_h \big\rangle \tag{4.7}$$

for all $v_h \in V_h$. We also set $U_h^0 = R_h u_0$.

In particular, for $\theta = 0$ we recover the implicit (backward) Euler scheme and for $\theta = 1$ we have the explicit (forward) Euler scheme. Furthermore, when $\theta = 1/2$, the θ-scheme becomes the well-known Crank–Nicolson time discretisation.

In order to investigate the stability of the explicit Euler scheme with $\theta = 1$, we need to introduce the so-called CFL (Courant–Friedrichs–Lewy) condition.

Definition 4.1.13 (CFL condition)**.** *Consider the θ-scheme* (4.7) *with $\theta = 1$. We define the CFL condition as*

$$h_{\min}^{-2}\Delta t \leq c_s, \tag{4.8}$$

where $h_{\min} = \min_{T\in\mathcal{T}_h} h_T$ and c_s is a constant independent of the triangulation and the time-step Δt.

This condition was first introduced in [103] in the context of finite difference methods. The extension concerning the finite element methods for time-dependent problems can be found in [20, 35]. We send the reader to [157, Chapters 7,8] for the following result and an exhaustive analysis of the stability of the time discretisation methods with uniform refinement in space for parabolic problems. We include the proof of the stability on polygonal domains, to show, why it is important to include the smallest of the element sizes in Definition 4.1.13.

Theorem 4.1.14 (Stability of the fully discrete scheme)**.** *Let f and u_0 satisfy the assumptions of Theorem 4.1.7. Then the following stability estimate is true for the solution* (4.7) *in the case of $\theta = 0$ and $\theta = 1/2$*

$$\|U_h^n\|_0 \leq \|U_h^0\|_0 + 2\Delta t \sum_{k=0}^{n-1} \|f(t_n)\|_0.$$

Let $c_i > 0$ be the constant in the inverse inequality stated in Lemma 2.3.7 with $m = 1$ and $l = 0$. If, in addition, the CFL condition (4.8) *holds with $c_s = 2/c_i^2$, then the estimate is true also for the explicit Euler scheme, so when $\theta = 1$.*

Proof. The proofs in the cases of $\theta = 0$ and $\theta = 1/2$ can be found in [47, Theorem 5.1] and [47, Theorem 5.6] respectively. Therefore, we focus on the remaining case of $\theta = 1$.

Let us take $v_h = U_h^{n+1}$ in (4.7). We obtain

$$\langle U_h^{n+1} - U_h^n, U_h^{n+1}\rangle + \Delta t a\big(U_h^n, U_h^{n+1}\big) = \Delta t\langle f(t_n), U_h^{n+1}\rangle.$$

Note however that

$$\begin{aligned}\langle U_h^{n+1}, U_h^n\rangle &= \frac{1}{2}\Big(\|U_h^{n+1}\|_0^2 + \|U_h^n\|_0^2 - \|U_h^{n+1} - U_h^n\|_0^2\Big),\\ a\big(U_h^n, U_h^{n+1}\big) &= \frac{1}{4}\Big(a\big(U_h^{n+1} + U_h^n, U_h^{n+1} + U_h^n\big) - a\big(U_h^{n+1} - U_h^n, U_h^{n+1} - U_h^n\big)\Big).\end{aligned}$$

Hence

$$\begin{aligned}\frac{1}{2}\|U_h^{n+1}\|_0^2 - \frac{1}{2}\|U_h^n\|_0^2 + \frac{1}{2}\|U_h^{n+1} - U_h^n\|_0^2 + \frac{\Delta t}{4}a\big(U_h^{n+1} + U_h^n, U_h^{n+1} + U_h^n\big)\\ = \frac{\Delta t}{4}a\big(U_h^{n+1} - U_h^n, U_h^{n+1} - U_h^n\big) + \Delta t\big\langle f(t_n), U_h^{n+1}\big\rangle.\end{aligned}$$

Due to the assumed coercivity of the bilinear form we have

$$0 \le a\big(U_h^{n+1} + U_h^n, U_h^{n+1} + U_h^n\big).$$

Furthermore, due to the inverse inequality stated in Lemma 2.3.7, we get

$$a\big(U_h^{n+1} - U_h^n, U_h^{n+1} - U_h^n\big) = \|\nabla(U_h^{n+1} - U_h^n)\|_0^2 \le c_i^2 h_{\min}^{-2}\|U_h^{n+1} - U_h^n\|_0^2.$$

Thus we can write

$$\begin{aligned}\|U_h^{n+1}\|_0^2 - \|U_h^n\|_0^2 + \|U_h^{n+1} - U_h^n\|_0^2\\ \le \frac{c_i^2\Delta t h_{\min}^{-2}}{2}\|U_h^{n+1} - U_h^n\|_0^2 + 2\Delta t\big\langle f(t_n), U_h^{n+1}\big\rangle.\end{aligned}$$

Due to the assumed CFL condition, see Definition 4.1.13 we have

$$\|U_h^{n+1}\|_0^2 - \|U_h^n\|_0^2 \le 2\Delta t\big\langle f(t_n), U_h^{n+1}\big\rangle.$$

Applying the Cauchy–Schwarz inequality we arrive at

$$\|U_h^{n+1}\|_0^2 - \|U_h^n\|_0^2 \le 2\Delta t\|f(t_n)\|_0\big(\|U_h^{n+1}\|_0 + \|U_h^n\|_0\big).$$

Therefore also

$$\|U_h^{n+1}\|_0 - \|U_h^n\|_0 \le 2\Delta t\|f(t_n)\|_0$$

and by induction we finally obtain the desired stability result.

□

Theorem 4.1.14 provides the restriction on the growth of the solution in time in the $L^2(\Omega)$ norm. Although this restriction is guaranteed in the case of implicit schemes with $\theta = 0$ and $\theta = 1/2$, the stability in the fully explicit case of $\theta = 1$ is obtained only under the assumption of the CFL condition (4.8). This assumption cannot be dropped, see [157, Chapter 7]. Similar results exist also for other types of the time-stepping schemes, such as Runge–Kutta methods, hewever, they are beyond the scope of this work.

The stability estimate given in Theorem 4.1.14 is the basis for derivation of the error estimate for the finite element discretisation of parabolic equations.

Theorem 4.1.15 (Convergence of the fully discrete scheme)**.** *Suppose that the assumptions of Theorem 4.1.10 are satisfied. Consider the θ-scheme approximation introduced in Definition 4.1.12 and let for any $\epsilon > 0$*

$$\sigma = \begin{cases} 2, & \text{if } \Theta \leq \pi \\ 2\lambda_1 - \epsilon, & \text{if } \Theta > \pi. \end{cases}$$

If $\theta = 0$, or $\theta = 1$ and the CFL condition (4.8) is satisfied, then

$$\max_{0\leq n\leq N} \|u(t_n) - U_h^n\|_0 \leq c\big(h^\sigma + \Delta t\big)\Big(\|\Delta u_0\|_0 + \int_0^T \|\Delta u\|_0 \,\mathrm{d}\tau + \int_0^T \|\Delta u_t\|_0 \,\mathrm{d}\tau\Big).$$

Suppose now that $\theta = 1/2$ instead. If $\frac{\mathrm{d}^2 f}{\mathrm{d}t^2} \in L^1(0,T;L^2(\Omega))$, and $\frac{\mathrm{d}f}{\mathrm{d}t} f(0) + \Delta f(0) + \Delta^2 u_0 \in H^1(\Omega)$, then

$$\max_{0\leq n\leq N} \|u(t_n) - U_h^n\|_0 \leq c\big(h^\sigma + (\Delta t)^2\big)\Big(\|\Delta u_0\|_0 + \int_0^T \|\Delta u_t\|_0 \,\mathrm{d}\tau + \int_0^T \|u_{ttt}\|_0 \,\mathrm{d}\tau + \int_0^T \|\Delta u_{tt}\|_0 \,\mathrm{d}\tau\Big).$$

Finally, assuming that the finite element approximation U_h^n is obtained on a gradually refined triangulation with $\mu < \lambda_1$, and the assumptions of Theorem 4.1.11 are satisfied, we can set $\sigma = 2$.

This result in the case of uniform triangulation of the domain, can be found in [157, Theorem 1.5, Theorem 1.6] for $\theta = 0$ and $\theta = 1/2$, and in [157, Chapter 7,8] for $\theta = 1$. The variant of the scheme using graded meshes was considered in [47] for $\theta = 0$ and $\theta = 1/2$. The extension to the explicit Euler scheme, so when $\theta = 1$, is straightforward.

Theorem 4.1.15 summarises, upon certain regularity assumptions assuring that the right-hand sides in the bounds are finite, the error estimates for the fully discrete

scheme for parabolic equations. Application of the Crank-Nicolson time-stepping ($\theta = 1/2$) yields second order convergence in time, whereas implicit and explicit Euler schemes guarantee only first order convergence. Moreover, the CFL condition (4.8) needs to be satisfied in the case of the explicit time-stepping. The convergence order of the spatial discretisation is identical as in the semi-discrete setting, see Theorem 4.1.10 and Theorem 4.1.11, and is suboptimal compared to the interpolation error, when uniformly refined triangulation is concerned. Additional application of the mesh grading technique improves the convergence order of the method yielding the optimal second order of convergence.

In Theorem 4.1.11 we showed that the mesh grading technique is a feasible choice for the improvement of the spatial approximation order of the finite element approximation of parabolic equations. However, it requires a local refinement of the computational mesh around the singular corner.

For the CFL condition (4.8) to be satisfied in the case of graded meshes, the relationship $\Delta t \sim h^{2/\mu}$ needs to hold. For angles $\Theta > \pi$, this introduces a significantly more restrictive form of the CFL condition than the one necessary on the uniform meshes, so when $\mu = 1$ and $\Delta t \sim h^2$. In the following section, we will propose an energy-correction based method for recovering optimal convergence order of the finite element scheme on uniformly refined meshes.

4.2 Energy-corrected finite elements for parabolic problems

In this section, we investigate the energy-corrected finite element approximation of the parabolic problem (4.1). We begin by analysing the semi-discretisation, where the time variable is continuous. The reasoning is then further extended to a fully discrete case with the explicit Euler time-stepping.

4.2.1 Energy-corrected semi-discrete scheme

We define a modified semi-discrete finite element approximation $u_h^m : [0, T] \to V_h$ as

$$\left\langle u_{h,t}^m, v_h \right\rangle + a_h(u_h^m, v_h) = \left\langle f, v_h \right\rangle, \quad \text{for all } v_h \in V_h \tag{4.9}$$

$$u_h(0) = R_h^m u_0, \tag{4.10}$$

where $a_h(\cdot, \cdot)$ is an energy-corrected bilinear form introduced in Section 2.3.3 with the modification (2.33) and R_h^m is the corresponding modified Ritz projection. Moreover,

we also assume that the optimal modification parameter γ^* is known. Remember that its existence was guaranteed by Lemma 2.3.33. To impose the initial conditions we also use the corresponding modified Ritz projection R_h^m in place of the operator $\mathcal{P}_h$ in (4.5).

Let us denote by $c_c, c_b > 0$ the continuity and coercivity constants of the energy-corrected bilinear form $a_h(\cdot,\cdot)$ respectively, so

$$c_c\|\nabla v_h\|_0^2 \le a_h(v_h, v_h), \quad \text{and} \quad a_h(v_h, w_h) \le c_b\|\nabla v_h\|_0\|\nabla w_h\|_0, \quad \text{for all } v_h, w_h \in V_h.$$

Note that it is enough to set $c_c = 1 - \gamma^*$ and $c_b = 1 + \gamma^*$, where γ^* denotes the optimal parameter for the modification (2.33).

We denote by c_α the embedding constant from Theorem 2.1.15, so we have

$$\|v\|_{-\alpha} \le c_\alpha\|\nabla v\|_0, \quad \text{for all } v \in H_0^1(\Omega). \tag{4.11}$$

Before moving to the statement of the main convergence result of the scheme (4.9), we first show the stability of the scheme.

Lemma 4.2.1 (Stability of the discrete scheme)**.** *Let $f \in L^2\big(0,T;L^2_\alpha(\Omega)\big)$ for some $\alpha < 1$. The semi-discrete solution of problem* (4.9) *satisfies*

$$\|u_h^m\|_0^2 \le \|R_h^m u_0\|_0^2 + C^* \int_0^T \|f\|_\alpha^2 \,\mathrm{d}t.$$

Proof. By the coercivity of the bilinear form $a_h(\cdot,\cdot)$, we obtain

$$a_h(u_h^m, u_h^m) \ge c_1\|\nabla u_h^m\|_0^2.$$

Upon the choice $v_h = u_h^m$ in Eq. (4.9) we obtain

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u_h^m\|_0^2 + c_1\|\nabla u_h^m\|_0^2 \le \big(f, u_h^m\big).$$

A further application of the Cauchy-Schwarz and Young inequalities yields for any $\epsilon > 0$

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u_h^m\|_0^2 + c_c\|\nabla u_h^m\|_0^2 &\le \|f\|_\alpha\|u_h^m\|_{-\alpha} \\ &\le \frac{1}{2\epsilon}\|f\|_\alpha^2 + \frac{\epsilon}{2}\|u_h^m\|_{-\alpha}^2.\end{aligned}$$

Due to (4.11) we have

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u_h^m\|_0^2 + c_c\|\nabla u_h^m\|_0^2 \leq \frac{1}{2\epsilon}\|f\|_\alpha^2 + \frac{c_\alpha^2\epsilon}{2}\|\nabla u_h^m\|_0^2.$$

Choosing $\epsilon \leq 2c_c/c_\alpha^2$ we arrive at

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u_h^m\|_0^2 \leq C^*\|f\|_\alpha^2,$$

where $C^* \geq c_\alpha^2/4c_c$. Integrating both sides over the time interval $[0,T]$ completes the proof of the lemma. □

Remark 4.2.2. *In the case of the standard choice of the modified bilinear form* (2.33) *in the energy-corrected scheme we have* $c_c = 1-\gamma$ *and the stability constant* C^* *above can be reduced to* $C^* = c_\alpha^2/4$. *The embedding constant* c_α *is an equivalent of the Poincaré-Friedrichs constant in weighted spaces and depends only on the domain* Ω.

Theorem 4.2.3. *Suppose that functions* u_0 *and* f *satisfy the regularity requirements stated in Theorem 4.1.7 and Theorem 4.1.8 and let* $1-\lambda_1 < \alpha < 1$. *The energy-corrected semi-discretisation* (4.9) *of Problem* (4.1) *yields optimal convergence rate in weighted norm, namely for some* $c > 0$ *independent of* u

$$\max_{0\leq t\leq T}\|u-u_h^m\|_\alpha \leq ch^2\left(\max_{0\leq t\leq T}\|\Delta u(t)\|_{-\alpha}^2 + \int_0^T \|\Delta u_t(t)\|_{-\alpha}^2\,\mathrm{d}t\right)^{1/2}. \tag{4.12}$$

Proof. We proceed in a standard manner by splitting the discretisation error into two independent parts

$$u(t)-u_h^m(t) = \Big(u(t)-R_h^m u(t)\Big) + \Big(R_h^m u(t)-u_h^m(t)\Big) =: \rho+\eta, \tag{4.13}$$

where R_h^m denotes the energy-corrected Ritz projection defined in Section 2.3.3. Hence, due to Theorem 2.3.27

$$\|\rho\|_\alpha = \|u(t)-R_h^m u(t)\|_\alpha \leq ch^2\|\Delta u(t)\|_{-\alpha} \tag{4.14}$$

and

$$\|\rho_t\|_\alpha = \|u_t(t)-R_h^m u_t(t)\|_\alpha \leq ch^2\|\Delta u_t(t)\|_{-\alpha}. \tag{4.15}$$

Using the definition of the modified Ritz projection

$$a_h(R_h^m u(t), v_h) = a(u(t), v_h), \quad \text{for all } v_h \in V_h,$$

we obtain due to the definitions of the continuous solution (4.4) and the energy-corrected discretisation (4.9)

$$\langle \eta_t, v_h \rangle + a_h(\eta, v_h) = \langle -\rho_t, v_h \rangle.$$

Finally, due to Lemma 4.2.1, we get

$$\|R_h^m u(t) - u_h^m(t)\|_\alpha^2 = \|\eta(t)\|_\alpha^2 \leq \|\eta(0)\|_0^2 + C^* \int_0^t \|\rho_t\|_\alpha^2 \, \mathrm{d}t.$$

Note that due to Theorem 4.1.8, the right-hand side of the inequality above is well-defined. Moreover, the discrete initial conditions were chosen in a way that $\eta(0) = 0$.

Combining this with Equation (4.13) gives

$$\max_{0\leq t\leq T} \|u - u_h^m\|_\alpha^2 \leq \max_{0\leq t\leq T} \left(\|\rho\|_\alpha^2 + \|\eta\|_\alpha^2 \right) \leq \left(\max_{0\leq t\leq T} \|\rho\|_\alpha^2 + C^{*2} \int_0^T \|\rho_t\|_\alpha^2 \, \mathrm{d}t \right).$$

Finally, application of the results stated in (4.14) and (4.15) completes the proof. □

The right-hand side of Equation (4.12) is finite, see Theorem 4.1.8. The above theorem shows that the application of the energy-corrected finite element scheme to the parabolic equations results in the optimal accuracy of the scheme, when compared to the interpolation error.

4.2.2 Energy-corrected fully discrete scheme

Now, we move to the fully discrete setting, where also the temporal dimention is discretised. We consider only explicit Euler time-stepping, which later will serve as a foundation for building fast numerical schemes. The extension to a more general case of θ-scheme in time is straightforward.

Following the convention presented above for the fully discrete schemes, we denote the solution by capital letters. The fully discrete energy-corrected finite element approximation of the model problem (4.4) reads as follows: Find $U_h^{m,n} \in V_h$ for $0 \leq n \leq N$ such that

$$\left\langle \frac{U_h^{m,n+1} - U_h^{m,n}}{\Delta t}, v_h \right\rangle + a_h(U_h^{m,n}, v_h) = \langle f(t_n), v_h \rangle, \quad \text{for all } v_h \in V_h. \tag{4.16}$$

The initial condition, as before, is imposed using the modified Ritz projection

$$U_h^{m,0} = R_h^m u_0.$$

We begin the convergence analysis of the scheme by showing an auxiliary result bouding the finite difference in the formulation (4.16).

Let $c_i > 0$ denote the constant appearing in the inverse inequality from Lemma 2.3.7, when $l = 0$ and $m = 1$.

Lemma 4.2.4. *Suppose that $f \in C\big(0,T;L^2_\alpha(\Omega)\big)$ for some $0 \leq \alpha < 1$. Then for all $0 \leq n \leq N-1$*

$$\Big\|\frac{U_h^{m,n+1}-U_h^{m,n}}{\Delta t}\Big\|_0 \leq c_i h^{-1}\Big(c_\alpha\|f\|_\alpha + c_b\|\nabla U_h^{m,n}\|_0\Big).$$

Proof. Let us set $v_h = \frac{U_h^{m,n+1}-U_h^{m,n}}{\Delta t}$ in (4.16). Then, applying the Cauchy-Schwarz inequality and using the boundedness of the bilinear form $a_h(\cdot,\cdot)$, we get

$$\begin{aligned}\Big\|\frac{U_h^{m,n+1}-U_h^{m,n}}{\Delta t}\Big\|_0^2 &= \Big\langle f(t_n), \frac{U_h^{m,n+1}-U_h^{m,n}}{\Delta t}\Big\rangle - a_h\Big(U_h^{m,n}, \frac{U_h^{m,n+1}-U_h^{m,n}}{\Delta t}\Big)\\ &\leq \|f\|_\alpha\Big\|\frac{U_h^{m,n+1}-U_h^{m,n}}{\Delta t}\Big\|_{-\alpha} + c_b\|\nabla U_h^{m,n}\|_0\Big\|\nabla\frac{U_h^{m,n+1}-U_h^{m,n}}{\Delta t}\Big\|_0.\end{aligned}$$

Since $H_0^1(\Omega) \hookrightarrow L^2_{-\alpha}(\Omega)$, see Theorem 2.1.15, we obtain

$$\Big\|\frac{U_h^{m,n+1}-U_h^{m,n}}{\Delta t}\Big\|_0^2 \leq \Big(c_\alpha\|f\|_\alpha + c_b\|\nabla U_h^{m,n}\|_0\Big)\Big\|\nabla\frac{U_h^{m,n+1}-U_h^{m,n}}{\Delta t}\Big\|_0.$$

Finally, application of the inverse inequality from Lemma 2.3.7 yields the desired result. □

Now, we can state the stability result, which will prove crucial for showing the error estimates for the fully discrete scheme. The main difference between the following and the stablity estimate presented in Theorem 4.1.14 lies in the use of weighted spaces. Similarly as in the case of the standard norms, the explicit time-stepping scheme is stable only under an additional assumption that the CFL condition (4.8) is satisfied. We provide the precise value of the stability constant.

Theorem 4.2.5 (Stability of the fully discrete scheme)**.** *Suppose that for some $0 \leq \alpha < 1$ we have $f \in C\big(0,T;L^2_\alpha(\Omega)\big)$ and let $0 < \epsilon < \frac{1}{2}$, $0 < \delta < \frac{c_c}{c_\alpha^2}$. Suppose also that the CFL condition proposed in Definition 4.1.13 is satisfied with the constant $c_s = 2\frac{c_c - c_\alpha^2\delta}{c_i^2 c_b^2(1+\epsilon)}$. Then for some $c_{\epsilon,\delta} > 0$ independent of h and Δt, we have*

$$\|U_h^{m,n}\|_0^2 \leq \|U_h^{m,0}\|_0^2 + c_{\epsilon,\delta}\Delta t\sum_{k=0}^{n-1}\|f(t_n)\|_\alpha^2.$$

Proof. We set $v_h = U_h^{m,n}$ in (4.16). Notice that

$$U_h^{m,n} = \frac{U_h^{m,n+1} + U_h^{m,n}}{2} - \frac{\Delta t}{2}\frac{U_h^{m,n+1} - U_h^{m,n}}{\Delta t}.$$

Hence

$$\frac{\|U_h^{m,n+1}\|_0^2 - \|U_h^{m,n}\|_0^2}{2\Delta t} + a_h(U_h^{m,n}, U_h^{m,n}) = \big\langle f(t_n), U_h^{m,n}\big\rangle + \frac{\Delta t}{2}\Big\|\frac{U_h^{m,n+1} - U_h^{m,n}}{\Delta t}\Big\|_0^2. \tag{4.17}$$

Note that for any numbers $\epsilon, a, b > 0$ we have

$$(a+b)^2 \le \big(1 + \frac{1}{\epsilon}\big)a^2 + (1+\epsilon)b^2.$$

Therefore, Lemma 4.2.4 gives us

$$\Big\|\frac{U_h^{m,n+1} - U_h^{m,n}}{\Delta t}\Big\|_0^2 \le c_i^2 h^{-2} c_\alpha^2 \big(1 + \frac{1}{\epsilon}\big)\|f(t_n)\|_\alpha^2 + c_i^2 h^{-2} c_b^2 (1+\epsilon)\|\nabla U_h^{m,n}\|_0^2. \tag{4.18}$$

Furthermore, for any $\delta > 0$ we get due to the Cauchy-Schwarz inequality and the embedding from Theorem 2.1.15

$$\big\langle f(t_n), U_h^{m,n}\big\rangle \le \frac{1}{\delta}\|f(t_n)\|_\alpha^2 + c_\alpha^2 \delta \|\nabla U_h^{m,n}\|_0^2. \tag{4.19}$$

Using the coercivity of the bilinear form $a_h(\cdot,\cdot)$ and applying (4.17)–(4.19), we obtain

$$\begin{aligned}\frac{\|U_h^{m,n+1}\|_0^2 - \|U_h^{m,n}\|_0^2}{2\Delta t} + c_c\|\nabla U_h^{m,n}\|_0^2 \le{}& \Big(\frac{1}{\delta} + \frac{1}{2}\big(1+\frac{1}{\epsilon}\big)c_i^2 c_\alpha^2 h^{-2}\Delta t\Big)\|f(t_n)\|_\alpha^2 \\ &+ \Big(c_\alpha^2\delta + \frac{1}{2}c_i^2 c_b^2 (1+\epsilon)h^{-2}\Delta t\Big)\|\nabla U_h^{m,n}\|_0^2.\end{aligned}$$

The CFL condition (4.8) states that

$$h^{-2}\Delta t \le c_s = 2\frac{c_c - c_\alpha^2\delta}{c_i^2 c_b^2 (1+\epsilon)}$$

and therefore

$$\frac{\|U_h^{m,n+1}\|_0^2 - \|U_h^{m,n}\|_0^2}{2\Delta t} \le \Big(\frac{1}{\delta} + \frac{1}{2}\big(1+\frac{1}{\epsilon}\big)c_i^2 c_b^2 h^{-2}\Delta t\Big)\|f(t_n)\|_\alpha^2$$

Setting

$$c_{\epsilon,\delta} = 2\Big(\frac{1}{\delta} + \frac{1}{2}\big(1+\frac{1}{\epsilon}\big)c_i^2 c_b^2 c_s\Big)$$

and applying induction we finally obtain

$$\|U_h^{m,n}\|_0^2 \leq \|U_h^{m,0}\|_0^2 + c_{\epsilon,\delta}\Delta t \sum_{k=0}^{n-1} \|f(t_n)\|_\alpha^2.$$

□

Upon the right choice of the values ϵ, δ, we see that any $c_s < 2\frac{c_c}{c_i^2 c_b^2}$ is a feasible stability constant. Note however that when $\delta \to 0$ or $\epsilon \to 0$, then $c_{\epsilon,\delta} \to \infty$.

Finally, we are in a position to state the convergence result for the fully discrete scheme.

Theorem 4.2.6. *Suppose that functions u_0 and f satisfy the regularity requirements stated in Theorem 4.1.7 and Theorem 4.1.8 and let $1-\lambda_1 < \alpha < 1$. Suppose also that the CFL condition stated in Definition 4.1.13 holds with the constant $c_s = 2\frac{c_c - c_\alpha^2 \delta}{c_i^2 c_b^2 (1+\epsilon)}$ for some $0 < \epsilon < 1/2$, $0 < \delta < \frac{c_c}{c_\alpha^2}$. Then, the following error estimate for the energy-corrected discretisation $U_h^{m,n}$, see (4.16), of Problem (4.1) holds for some $c > 0$ independent of u*

$$\max_{0\leq n\leq N} \|u(t_n) - U_h^{m,n}\|_\alpha \tag{4.20}$$
$$\leq c(h^2 + \Delta t)\left(\max_{0\leq t\leq T} \|\Delta u(t)\|_{-\alpha}^2 + \int_0^T \|\Delta u_t(t)\|_{-\alpha}^2 \,\mathrm{d}t + \int_0^T \|u_{tt}\|_\alpha^2 \,\mathrm{d}t \right)^{1/2}.$$

Proof. Similarly as in the proof of Theorem 4.2.3, we begin the proof by splitting the error into two components.

$$u(t_n) - U_h^{m,n} = \Big(u(t_n) - R_h^m u(t_n)\Big) + \Big(R_h^m u(t_n) - U_h^{m,n}\Big) =: \rho^n + \eta^n, \tag{4.21}$$

where R_h^m denotes the energy-corrected Ritz projection defined in Section 2.3.3. Due to Theorem 2.3.27

$$\|\rho^n\|_\alpha = \|u(t_n) - R_h^m u(t_n)\|_\alpha \leq ch^2 \|\Delta u(t_n)\|_{-\alpha} \tag{4.22}$$

We focus now on estimating the remaining η^n component of the error. Due to the definition of the energy-corrected Ritz projection and the problem formulation (4.4), we have

$$\big\langle u_t(t_n), v_h \big\rangle + a_h(R_h^m u(t_n), v_h) = \big\langle f(t_n), v_h \big\rangle, \quad \text{for all } v_h \in V_h.$$

Therefore, equation (4.16) yields for all $v_h \in V_h$

$$
\begin{aligned}
a_h(\eta^n, v_h) &= \Big\langle \frac{U_h^{m,n+1} - U_h^{m,n}}{\Delta t} - u_t(t_n), v_h \Big\rangle \\
&= -\Big\langle \frac{R_h^m u(t_{n+1}) - R_h^m u(t_n)}{\Delta t} - \frac{U_h^{m,n+1} - U_h^{m,n}}{\Delta t}, v_h \Big\rangle \\
&\qquad - \Big\langle u_t(t_n) - \frac{R_h^m u(t_{n+1}) - R_h^m u(t_n)}{\Delta t}, v_h \Big\rangle \\
&= -\Big\langle \frac{\eta^{n+1} - \eta^n}{\Delta t}, v_h \Big\rangle + \Big\langle \frac{u(t_{n+1}) - u(t_n)}{\Delta t} - u_t(t_n), v_h \Big\rangle - \Big\langle \frac{\rho^{n+1} - \rho^n}{\Delta t}, v_h \Big\rangle.
\end{aligned}
$$

Thus, we can write

$$
\Big\langle \frac{\eta^{n+1} - \eta^n}{\Delta t}, v_h \Big\rangle + a_h(\eta^n, v_h) = \big\langle \psi_1^n + \psi_2^n, v_h \big\rangle, \quad \text{for all } v_h \in V_h,
$$

where

$$
\psi_1^n = \frac{u(t_{n+1}) - u(t_n)}{\Delta t} - u_t(t_n), \quad \text{and} \quad \psi_2^n = \frac{\rho^{n+1} - \rho^n}{\Delta t}.
$$

Thanks to the stability estimate stated in Theorem 4.2.5 we obtain

$$
\|\eta^n\|^2 \le \|\eta^0\| + 2\Delta t c_{\epsilon,\delta} \Big(\sum_{k=0}^{n-1} \|\psi_1^n\|_\alpha^2 + \sum_{k=0}^{n-1} \|\psi_2^n\|_\alpha^2 \Big). \tag{4.23}
$$

We now estimate ψ_1^n and ψ_2^n separately. Note that

$$
\psi_1^n = \frac{u(t_{n+1}) - u(t_n)}{\Delta t} - u_t(t_n) = -\frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} (t_{n+1} - t) u_{tt} \, \mathrm{d}t,
$$

and hence

$$
\|\psi_1^n\|_\alpha \le \sqrt{\Delta t} \Big(\int_{t^n}^{t^{n+1}} \|u_{tt}\|_\alpha^2 \, \mathrm{d}t \Big)^{1/2}. \tag{4.24}
$$

Further, due to the linearity of the modified Ritz projection, we have

$$
\psi_2^n = \frac{u(t_{n+1}) - u(t_n)}{\Delta t} - R_h^m \frac{u(t_{n+1}) - u(t_n)}{\Delta t},
$$

and thus, see Theorem 2.3.27, we also get

$$
\|\psi_2^n\|_\alpha \le c h^2 \Big\| \Delta \Big(\frac{u(t_{n+1}) - u(t_n)}{\Delta t} \Big) \Big\|_{-\alpha} = c h^2 \Big\| \frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} \Delta u_t \, \mathrm{d}t \Big\|_{-\alpha}.
$$

Further, applying Cauchy-Schwarz inequality, we obtain

$$
\|\psi_2^n\|_\alpha \le c \frac{h^2}{\sqrt{\Delta t}} \Big(\int_{t_n}^{t_{n+1}} \|\Delta u_t\|_{-\alpha}^2 \, \mathrm{d}t \Big)^{1/2}. \tag{4.25}
$$

Since the initial conditions in the discretisation are imposed using the modified Ritz projection, see (4.16), we automatically have $\eta^0 = 0$. Using this and combining (4.23) with (4.24) and (4.25), we arrive at

$$\|\eta^n\|^2 \leq c(\Delta t)^2 \int_0^T \|u_{tt}\|_\alpha^2 \,\mathrm{d}t + ch^4 \int_0^T \|\Delta u_t\|_{-\alpha}^2 \,\mathrm{d}t$$

Finally, combining this result with (4.22), and applying to the error splitting (4.21) we get

$$\max_{0\leq n\leq N} \|u(t_n) - U_h^{m,n}\|_\alpha^2 \leq c\big(h^4 + (\Delta t)^2\big)\Big(\max_{0\leq n\leq N} \|\Delta u(t_n)\|_{-\alpha}^2 + \int_0^T \|\Delta u_t\|_{-\alpha}^2 \,\mathrm{d}t + \int_0^T \|u_{tt}\|_\alpha^2 \,\mathrm{d}t\Big)$$

Hence, the proof is completed upon taking the square root of both sides of the inequality. The boundedness of the right-hand side is ensured by Theorem 4.1.8.

□

As opposed to the mesh grading strategy, the energy-correction works on uniform meshes with less restrictive stability conditions. We shall exploit this fact further in the next section, when creating fast time-stepping schemes.

4.3 Numerical results

In this section, we propose and numerically investigate a fast fully-discrete solver for parabolic problems based on energy-corrected finite element (4.16). We show that, as opposed to the algorithms involving mesh grading and adaptivity, explicit time-stepping schemes are a feasible choice in the proposed setting.

Let $\{\phi_i\}_{i=1}^K$ denote the nodal basis functions of the finite element space V_h and let $\mathbf{U}_n^m = \big(U_i^{m,n}\big)_{i=1}^K$ be the vector of coefficients of the finite element solution in this basis at the time step n. We also denote $\mathbf{F}^n = \Big[\big\langle f(t_n,\cdot),\phi_i\big\rangle\Big]_{i=1}^K$. Moreover, $\mathbf{M}$ and $\mathbf{S}^m$ are respectively the mass and modified stiffness matrices of the system introduced in (2.12).

The discrete systems (4.16) can be rewritten in a matrix-vextor formulation as

$$\mathbf{U}_{n+1}^m = \mathbf{U}_n^m + \Delta t \mathbf{M}^{-1}\Big[\mathbf{F}^n - \mathbf{S}^m \mathbf{U}_n^m\Big]. \tag{4.26}$$

The choice of a nodal, vertex-based quadrature rule for assembling the mass matrix leads to a lumped diagonal matrix $\widetilde{\mathbf{M}}$, which can be used in place of $\mathbf{M}$. The application of the nodal Gauss quadrature in the assembly process does not diminish the convergence of the finite element scheme, see [23] for more details. This results in a fast time-stepping scheme, where at each time step, multiplication by a diagonal matrix $\widetilde{\mathbf{M}}^{-1}$ and a sparse matrix $\mathbf{S}^m$ needs to be performed.

Stability of the explicit Euler scheme is guaranteed by the CFL condition (4.8), which is very prohibitive when mesh grading or adaptivity is concerned. However, this is not an issue in the case of the energy-corrected FEM, which works on uniform meshes. Then, balancing the error of order $\mathcal{O}(\Delta t)$ coming from the time-stepping discretisation with $\mathcal{O}(h^2)$ order of error measured in the weighted $L^2(\Omega)$-norm, see Theorem 4.2.6, exactly the same relationship, as the CFL stability requirement, needs to be kept.

In order to improve the convergence of the scheme at a fixed point in time $T > 0$, we complete the algorithm with a post-processing strategy introduced in (2.31). As stated in Equation (2.4), the stress-intensity factor for $i \in \mathbb{Z}_+ \cup \{0\}$, see Theorem 4.1.7, can be computed by

$$k_i(T) = -\frac{1}{i\pi} \int_\Omega \big(f(T) - u_t(T)\big) s_{-i} + u \Delta s_{-i}.$$

We define its discrete approximation using (2.29) as

$$k_i^h(T) = -\frac{1}{i\pi} \int_\Omega \left(f(T) - \frac{U_h^{m,N} - U_h^{m,N-1}}{\Delta t} \right) s_{-i} + U_h^m(T) \Delta s_{-i}.$$

For $1 - \lambda_1 < \alpha < 1$, this leads to the post-processed solution of the form

$$\widetilde{u}_h^m(T) = U_h^{m,N} + \sum_{i:\lambda_i < 1+\alpha} k_i^h(T) \big(s_i - s_{i,h}^m\big). \tag{4.27}$$

Note that the additional cost of performing the post-processing is equal to the cost of solving one or two additional elliptic equations and evaluating pu to two integrals.

In Table 4.1, we summarise the errors and the convergence rates of the proposed scheme. We choose the L-shape domain $\Omega = (-1,1)^2 \setminus \big([0,1] \times [-1,0]\big)$ with the largest interior angle of size $\Theta = 3\pi/2$ and the triangulation as previously shown in Figure 2.3. We also choose a known exact solution $u = \sin(t)s_1 + \sin(2t)s_2 - \sin(3t)s_3$, being a linear combination of singular functions (2.2) with smooth time-dependent coefficients.

The parameter γ in the modification (2.33) is computed using a version of the Newton algorithm described in [137] and in the numerical experiments we choose the

weight $\alpha = 1 - \lambda_1$. This choice of the weight induces a slightly stronger norm than assumed in Theorem 4.2.3 but the optimal convergence order of the energy-corrected scheme can be observed regardless of this.

We consider uniform refinement of the initial mesh and together with refining the mesh, we also divide the time-step Δt by 4, initially set to be equal to 0.1. For the purposes of comparison, in the first two columns of Table 4.1, we summarise the results obtained using the standard, uncorrected scheme. The suboptimal convergence rates in the sense of the interpolation error, when measured both in standard and weighted $L^2(\Omega)$-norms, are in line with the results in Theorem 4.1.10. For the energy-corrected scheme (4.16), we see that no pollution in the $L^2(\Omega)$-norm appears. Moreover, second-order convergence in the weighted norm means that the error is relatively large only in the vicinity of the re-entrant corner, so the pollution effect from Theorem 2.3.21 has been removed. Finally, the post-processing approach yields second-order convergence in the standard $L^2(\Omega)$-norm. Numerical tests confirm the theoretical results of Theorem 4.2.3.

In Figure 4.1 the evolution of the $L^2(\Omega)$ error in time is shown. The presented results were obtained on the 5th refinement level in space and time. The energy-corrected finite element consistently provides a smaller error than the standard finite element method. Moreover, the growth of the error in the latter is also faster than in the energy-corrected scheme.

Standard and energy-corrected P_1 elements, $\gamma = 0.18617957$										
L	$\|u - u_h\|_0$	rate	$\|u - u_h\|_\alpha$	rate	$\|u - u_h^m\|_0$	rate	$\|u - u_h^m\|_\alpha$	rate	$\|u - \tilde{u}_h^m\|_0$	rate
1	8.3556e-02		6.4626e-02		8.6186e-02		6.7405e-02		6.0499e-02	
2	2.9690e-02	1.49	2.0407e-02	1.66	2.7783e-02	1.63	1.9365e-02	1.80	1.9961e-02	1.60
3	1.1018e-02	1.43	6.8650e-03	1.57	7.9946e-03	1.79	4.7695e-03	2.02	5.0731e-03	1.98
4	4.1983e-03	1.39	2.4516e-03	1.49	2.2934e-03	1.80	1.1349e-03	2.07	1.1892e-03	2.09
5	1.6212e-03	1.37	9.1443e-04	1.42	6.7778e-04	1.75	2.7326e-04	2.06	2.7349e-04	2.12
6	6.3061e-04	1.36	3.5033e-04	1.38	2.0565e-04	1.72	6.6972e-05	2.03	6.3007e-05	2.12
Expected		1.33		1.33		1.66		2.00		2.00

Table 4.1: Summary of convergence rates obtained using two different approximations of the heat equation on the L-shape domain

4.4 Extensions

In this section, we present extensions of the methods introduced above. We show that the energy-corrected finite element can be applied to domains with multiple re-entrant corners, also in the presence of a moderate advection in the problem (4.1). Furthermore, we show a possible extension to the piecewise polynomial finite element

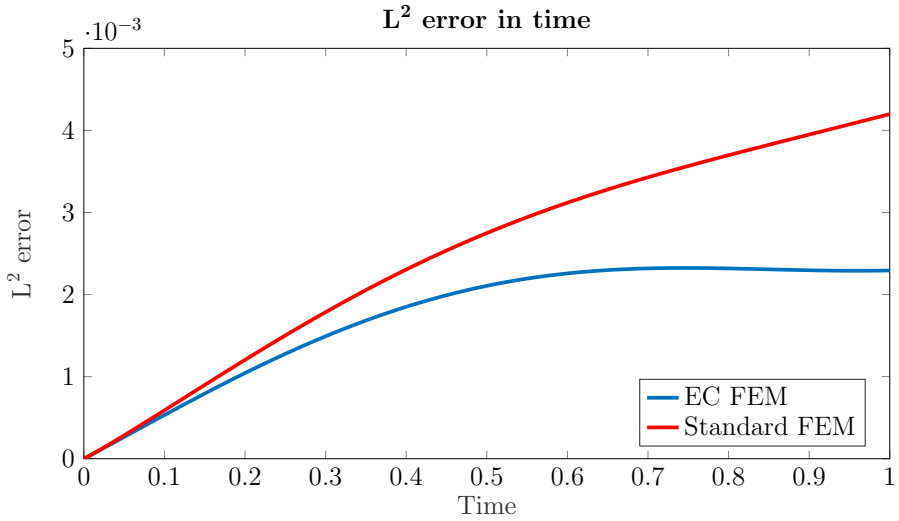


Figure 4.1: Development of the $L^2(\Omega)$ error in time.

and propose a fast explicit time-stepping scheme based on cubic elements combined with mass-lumping techniques. Finally, in order to show the flexibility of the energy-correction method, we present numerical experiments involving multiple re-entrant corners in three dimensions.

4.4.1 Advection-diffusion equation

We consider the following advection-diffusion problem

$$u_t + b \cdot \nabla u - \Delta u = f \quad \text{in } \Omega \times (0,T), \tag{4.28}$$

$$u = 0 \quad \text{on } \partial\Omega \times [0,T], \tag{4.29}$$

$$u = u_0 \quad \text{in } \Omega \text{ at } t = 0. \tag{4.30}$$

In the numerical example, we consider $T = 1$, $u_0 = 0$ and $b = (1,1)$, so the problem is equipped with moderate advection. The computational domain Ω, together with its triangulation, is presented in Figure 4.2 and consists of a rectangle $(0,4) \times (0,3)$ with a right, isosceles triangle cut out. There are three re-entrant corners in the domain Ω, two of sizes $7\pi/4$ and one of size $3\pi/2$. We use a computational grid with one-element patches around the singular corners consisting of the identical isosceles triangles. For the right-hand side we choose $f = \sin(\pi t)\Big((x-2)^2+(y-3/2)^2\Big)^{-1}$, which has a singularity in the middle of the cut-out triangle. For the time discretisation, we choose Explicit Euler time-stepping as described in Section 4.3 with an initial step-size $\Delta t = 0.02$, which is small enough to guarantee the stability of the scheme.

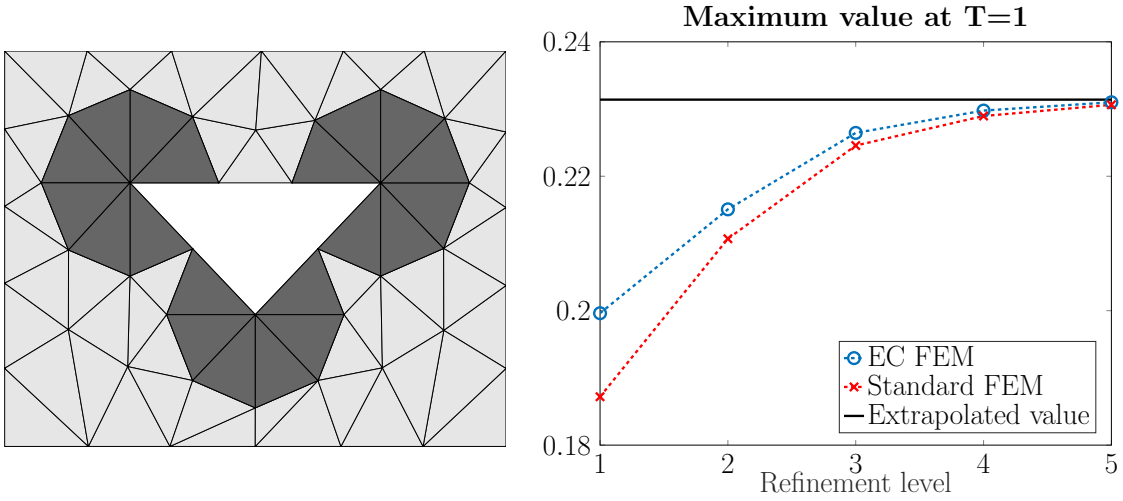


Figure 4.2: On the left-hand side a computational domain consisting of a rectangle $(0,4)\times(0,3)$ with a right-angled, isosceles triangle cut out of it is shown. The domain is triangulated so that one element patches around the re-entrant corners consist of identical isoscles triangles. On the right-hand side a convergence of a computed quantity of interest - maximum value in the domain, is shown.

We investigate the behaviour of a quantity of interest QoI $= \|u_h\|_{L^\infty(\Omega)}$ for standard finite element method and the energy-corrected finite element method on 5 consecutive refinement levels. The results of the simulations are summarised in the plot on the right-hand side of Figure 4.2. For completeness, we also include the extrapolated approximation $\|u\|_\infty^{ex}$ of the real value in the plot. The estimated order of convergence of $|\|u\|_\infty^{ex} - \|u_h\|_\infty|$ is equal to 1.85 and 1.55 in the case of the energy-corrected scheme and standard finite element respectively. The energy-corrected finite element can be successfully applied also in the cases of several different re-entrant corners in the domain and the presence of a moderate advection in the problem.

4.4.2 Higher-order FEM

So far we only analysed the piecewise linear discretisation in space. In this section, we discuss the feasibility of using higher-order discretisations. Moreover, we also present the mass lumping strategy for the higher-order elements based on the enrichment of the finite element spaces, the idea introduced in [57]. We apply the method to the parabolic problem (4.4) in a straightforward manner, suitably modifying the formulation (4.9) by using the more general form of the energy-correction (2.32).

We are aiming to construct a fast, explicit solver and this means that, in order to guarantee the stability of the method, the CFL condition $\Delta t \sim h^2$ needs to be

satisfied. The use of cubic finite element basis yields fourth-order convergence in weighted $L^2(\Omega)$-norm, and we would like to balance it with a second-order time-stepping scheme. Similarly to the piecewise linear case discussed in Section 4.3, optimal balancing of the errors stemming from the time and space discretisations means that the CFL condition is automatically satisfied and makes the use of explicit time-stepping scheme feasible. We use the second-order Runge-Kutta scheme, also known as the Heun's method, see [118, Section 2.4] and [152, Section 12.5]. The fully discrete scheme can be written as

$$\tilde{\mathbf{U}}^m_{h,n+1} = \mathbf{U}^m_{h,n} + \Delta t \mathbf{M}^{-1}\Big[\mathbf{F}^n - \mathbf{S}^m\mathbf{U}^m_{h,n}\Big] \tag{4.31}$$

$$\mathbf{U}^m_{h,n+1} = \mathbf{U}^m_{h,n} + \frac{1}{2}\Delta t \mathbf{M}^{-1}\Big[\big(\mathbf{F}^n - \mathbf{S}^m\mathbf{U}^m_{h,n}\big) + \big(\mathbf{F}^{n+1} - \mathbf{S}^m\tilde{\mathbf{U}}^m_{h,n+1}\big)\Big] \tag{4.32}$$

4.4.2.1 Mass-lumping for higher-order FEM

Note that the application of the mass-lumping strategy is not as straightforward as in the piecewise linear case. When the standard piecewise linear P_1 elements are considered, a diagonal matrix is obtained upon the choice of a suitable Gauss-Lobatto quadrature rule for computation of the matrix entries, where the quadrature points are located at the triangulation nodes. However, similar strategy in the case of higher-order finite elements cannot be applied, as it would result in negative or zero entries in the mass matrix. To overcome this, we follow the method proposed in [57] in the context of the wave equation. It is based on the enrichment of the finite element space with higher-order polynomial bubble functions, which are uniformly equal to 0 at the element's edges. Moreover, the mass matrix is assembled using a positive quadrature rule

$$\mathcal{I}_{\widehat{T}}(f) = |\widehat{T}| \sum_{\mathbf{x}_i \in \widehat{T}} \omega_i f(\mathbf{x}_i), \tag{4.33}$$

where ω_i are weights and $\mathbf{x}_i$ are quadrature points on the reference element $\widehat{T}$. Note that it suffices to define the quadrature rule on the reference triangle only. We will show the construction of the quadrature rule with quadrature points located in the nodal points of the enriched space. Such a construction yields a diagonal matrix $\widetilde{\mathbf{M}}$ used in place of $\mathbf{M}$ in scheme (4.31).

We now introduce mass-lumping techniques for locally enriched quadratic and cubic finite elements. Resulting quadrature rules have third and fourth order accuracy respectively and will contain only positive weights ensuring the stability of the scheme.

We shall also define a locally enriched finite element space with nodal points coinciding with the nodes of the quadrature rule. A more detailed discussion of the topic can be found in [57]. It was further extended to even higher-order approximation spaces [93, 117]. Moreover, the method is also known in the context of $\mathcal{Q}_k$ finite element spaces on quadrilateral meshes, see [58, 79].

In the construction of the quadrature rules and finite element spaces we demand two main assumptions to be satisfied, namely:

$$\text{The modified mass matrix is positive definite} \tag{A1}$$

$$\text{The weights in the quadrature formula are strictly positive.} \tag{A2}$$

Satisfying these two assumptions correspond to the invertability of the mass matrix and stability of the scheme. Since this cannot in principle be done for the standard $P_2(\widehat{T})$ and $P_3(\widehat{T})$ spaces, we shall consider their enrichments denoted by $P_k(\widehat{T}) \subset \widetilde{P}_k(\widehat{T}) \subset P_{k'}(\widehat{T})$ for some $k \leq k'$. We define a corresponding enriched finite element space as

$$\widetilde{S}_h^k = \{v_h \in H^1(\Omega) : \text{ for all } T \in \mathcal{T}_h \; v_h|_T \in \widetilde{P}_k(T)\}. \tag{4.34}$$

Moreover, to ensure sufficient accuracy of the integration formula, we add one more assumption

$$\text{The quadrature rule is exact in } P_{k+k'-2}(\widehat{T}). \tag{A3}$$

P_2 elements

For the sake of completeness, we begin by introducing the mass-lumping technique for piecewise quadratic finite element spaces. We define the enriched space of quadratic polynomials as $\widetilde{P}_2(\widehat{T}) = P_2(\widehat{T}) \oplus b$, where b is a bubble function, that is a third order polynomial disappearing on the edges of the triangle. We consider the nodal basis of the enriched discrete space. Note that in this way we introduce one extra degree of freedom in each triangle in the computational mesh $\mathcal{T}_h$. There are 7 nodal points defining the nodal basis, and we place 3 of them in the vertices (V), 3 in the midpoints (M) of the edges and one in the centre of the gravity (G) of the triangle. It is graphically summarised in Figure 4.3. We also set the values of the quadrature weights in the corresponding nodal points as

$$\omega_V = \frac{1}{20}, \quad \omega_M = \frac{2}{15}, \quad \omega_G = \frac{9}{20}.$$

The quadrature formula

$$\mathcal{I}_{\widehat{T}}(f) = |\widehat{T}|\Big(\omega_V \sum_i f(V_i) + \omega_M \sum_i f(M_i) + \omega_G \sum_i f(G_i)\Big)$$

is exact in the space P_3 and the mass matrix in the space $\widetilde{S}_h^2$ defined in Equation (4.34) is diagonal with strictly positive entries.

P_3 elements

The procedure in the case of $P_3(K)$ spaces is more complicated, as not only do we have to enrich the space, but also move the nodal points to different positions compared to the standard cubic Lagrange finite elements introduced in Section 2.3.1. To describe the construction accurately, we introduce the barycentric coordinates $\xi_1(\mathbf{x}), \xi_2(\mathbf{x}), \xi_3(\mathbf{x})$ being the distances of the point x from the corresponding vertices of the triangle.

We begin with the definition of the suitable polynomial space $P_3(\widehat{T}) \subset \widetilde{P}_3(\widehat{T}) = P_3(\widehat{T}) \oplus bP_1(\widehat{T}) \subset P_4(\widehat{T})$. This means that we enrich the qubic polynomial spaces with fourth order polynomials. We also need to define the nodal points so that the quadrature rule can be exact on $P_5(\widehat{T})$ polynomial functions, see Assumption A3. We divide these points into three groups: 3 vertices V, 6 edge points $M(\alpha)$ and 3 interior points $G(\beta)$, so in total 12 degrees of freedom depending on the parameters $0 < \alpha, \beta < 1/2$, $\beta \neq 1/3$. The edge points are chosen so that they are at a distance α and $1-\alpha$ from the corresponding vertices of the triangle and the interior points have barycentric coordinates $\xi_i = \beta$, $\xi_j = \frac{1-\beta}{2}$, $j \neq i$. This choice of nodal points is graphically summarised in the Figure 4.3. The basis functions in each triangle of the mesh can be divided into three groups:

- Functions associated with the interior points $G(\beta)$

$$\phi_i^G = \Big(\xi_i - \frac{1-\beta}{2}\Big)b.$$

- Functions associated with the interior points $M_{i,j}(\alpha)$

$$\phi_{i,j}^M = p_{i,j} - \frac{8b}{\beta(1-\beta)^2(3\beta-1)}\Big[A_1\Big(\xi_i - \frac{1-\beta}{2}\Big) + B_1\Big(\xi_j - \frac{1-\beta}{2}\Big) + C_1\Big(\xi_k - \frac{1-\beta}{2}\Big)\Big],$$

where

$$p_{i,j} = \frac{\xi_i \xi_j}{\alpha(1-\alpha)(2\alpha-1)}\big(\alpha\xi_i - (1-\alpha)\xi_j - (1-2\alpha)\xi_k\big),$$
$$A_1 = p_{i,j}(G_i), \quad B_1 = p_{i,j}(G_j), \quad C_1 = p_{i,j}(G_k).$$

Note that the values of the constants A_1, B_1, C_1 are independent of the choice of indices i, j.

- Functions associated with the vertices V_i

$$\phi_i^V = p_i - \frac{8b}{\beta(1-\beta)^2(3\beta-1)}\Big[A_2\Big(\xi_i - \frac{1-\beta}{2}\Big) + B_2 \sum_{l\neq i}\Big(\xi_l - \frac{1-\beta}{2}\Big)\Big],$$

where

$$p_i = \xi_i\Big(\sum_l \xi_l^2 - \frac{1-2\alpha+2\alpha^2}{\alpha(1-\alpha)}\xi_j\xi_k + \frac{2-7\alpha+7\alpha^2}{\alpha(1-\alpha)}\xi_i(\xi_j+\xi_k)\Big),$$
$$A_2 = p_i(G_i), \quad B_2 = p_i(G_j), j \neq i.$$

Moreover, constants A_2, B_2 are intependent of the choice of the indices i, j.

The basis functions constructed this way on the reference triangle $\widehat{T}$, together with the nodal evaluations, create the finite element in the sense of Definition 2.3.1.

There exists a unique quadrature formula of the form as in equation (4.33) satisfying Assumptions A1–A3 and it requires

$$\alpha = \frac{42 + 21\sqrt{7} - \sqrt{21(35+16\sqrt{7})}}{84+42\sqrt{7}}, \quad \beta = \frac{1}{3} + \frac{2}{21}\sqrt{7}.$$

It also defines uniquely strictly positive weights

$$\omega_V = 2\frac{919\sqrt{7}+2471}{124080\sqrt{7}+330960}, \quad \omega_M = 2\frac{\sqrt{7}(2+\sqrt{7})^4}{25280+9520\sqrt{7}}, \quad \omega_G = 2\frac{147+42\sqrt{7}}{400\sqrt{7}+1280}.$$

Finaly, the quadrature formula

$$\mathcal{I}_{\widehat{T}}(f) = |\widehat{T}|\Big(\omega_V \sum_{i:\mathbf{x}_i\in\widehat{T}} f(V_i) + \omega_M \sum_{i\neq j} f(M_{i,j}) + \omega_G \sum_i f(G_i)\Big)$$

is exact in the space $P_5(\widehat{T})$ and the mass matrix in the space $\widetilde{S}_h^3$ defined in equation (4.34) approximated with the use of the defined quadrature formula is diagonal and positive definite.

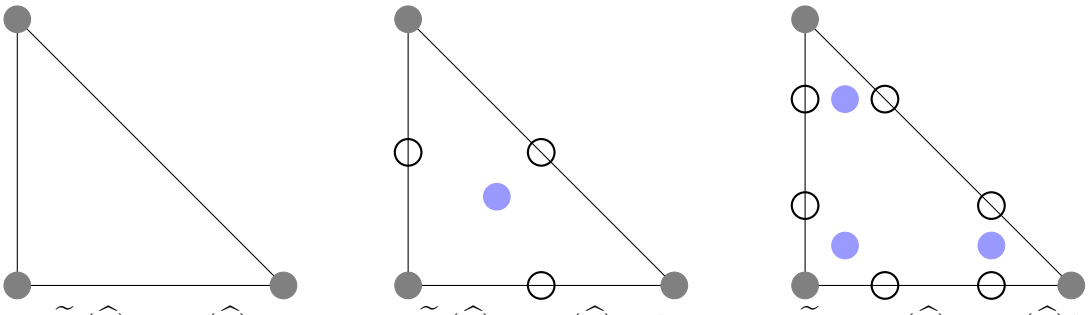


Figure 4.3: Graphical representation of the enriched finite element spaces. Grey circles represent nodal points in the vertices of the triangle, which always coincide with the standard choice for the piecewise polynomial finite element. Nodes on the edges have modified coordinates. Blue points correspond to the nodal points of the higher-order bubble functions enriching the finite element spaces.

4.4.2.2 Comparison of schemes

In the numerical study, we use a known exact solution $u = \sin(t)s_1 + \sin(2t)s_2 - \sin(3t)s_3$ on the L-shaped domain $\Omega = (-1,1)^2 \setminus \big([0,1] \times [-1,0]\big)$ with the largest interior angle of size $\Theta = 3\pi/2$ and the triangulation presented in Figure 2.3. We compare the $L^2(\Omega)$ and $L^2(\Omega')$ errors of the schemes at the last time step $T = 1$, where $\Omega' = \Omega \cap \{|x| > 0.25\}$.

In Figure 4.4, a standard $L^2(\Omega)$ error is shown. The standard piecewise linear finite element combined with an explicit Euler time-stepping and mass-lumping provides the worst results among the ones studied since its performance is limited by (4.1.10). Initially $\Delta t = 0.1$ is chosen and with each space refinement the time step is divided by 4.

Application of mesh grading improves the performance of the solver [47]. However, the use of explicit time-stepping schemes is infeasible due to the very restrictive CFL condition. To recover the optimal convergence order in the $L^2(\Omega)$-norm, it is necessary to grade the mesh towards the singular corner, introducing elements of the size $h^{1/\mu}$, where $\mu < \pi/\Theta$. This, in particular, means that in the case of the L-shape domain with $\Theta = 3\pi/2$, time-steps Δt smaller than $\mathcal{O}(h^3)$ need to be used. Therefore, we use the unconditionally stable Crank-Nicolson scheme in time allowing for coarser time discretisation. To keep the right balance between space and time discretisation errors, with each mesh refinement we divide the time-step by 2, beginning with $\Delta t = 0.1$.

The piecewise linear energy-corrected finite element scheme with explicit Euler time-stepping (4.26) yields significantly better results than the standard piecewise linear discretisations. It also gives comparable results with the mesh grading scheme

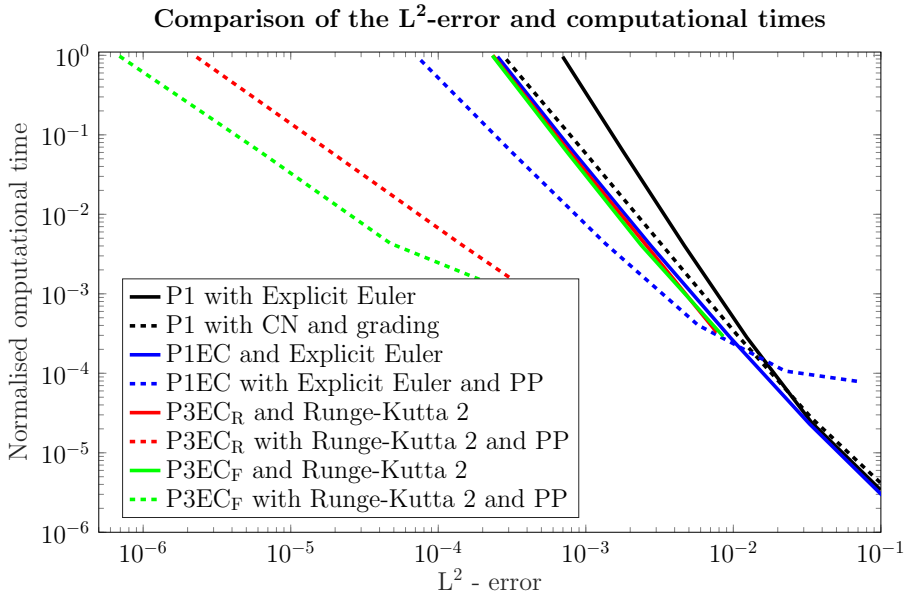


Figure 4.4: Comparison of the normalised computational time and accuracy of commonly used finite element discretisations of the parabolic problems with the proposed energy-corrected solvers on the L-shape domain. Schemes incorporating the post-processing (PP) are also included in the graph. We compare the $L^2(\Omega)$ error evaluated at the final time-step $T = 1$.

completed with Crank-Nicolson time-stepping. An application of the post-processing additionally improves the accuracy of the method resulting in a better error-to-time ratio than the mesh grading method equipped with Crank-Nicolson time-stepping. We use in-built MATLAB linear system solvers. Note that the application of fast iterative solvers, such as multigrid methods, could additionally improve the performance of the implicit methods.

Application of the cubic energy-corrected finite element scheme with second-order Runge-Kutta scheme in time gives similar results to the piecewise linear energy-corrected scheme. However, additional application of the post-processing yields the best results regarding the balance between the computational time and the accuracy of the scheme out of all tested methods. This can be attributed to the use of the scheme eliminating the pollution effect in the solution, completion with the post-processing strategy yielding optimal convergence in the standard norms, and the use of mass-lumping strategy. Note that the use of $c_h^F(\cdot,\cdot)$ modification gives quantitatively better results than $c_h^R(\cdot,\cdot)$. This phenomenon was previously observed in the case of eliptic equations in [86, Section 6.3] and in Section 2.3.3.4, and can be attributed to the smaller modification subregion in the computational domain.

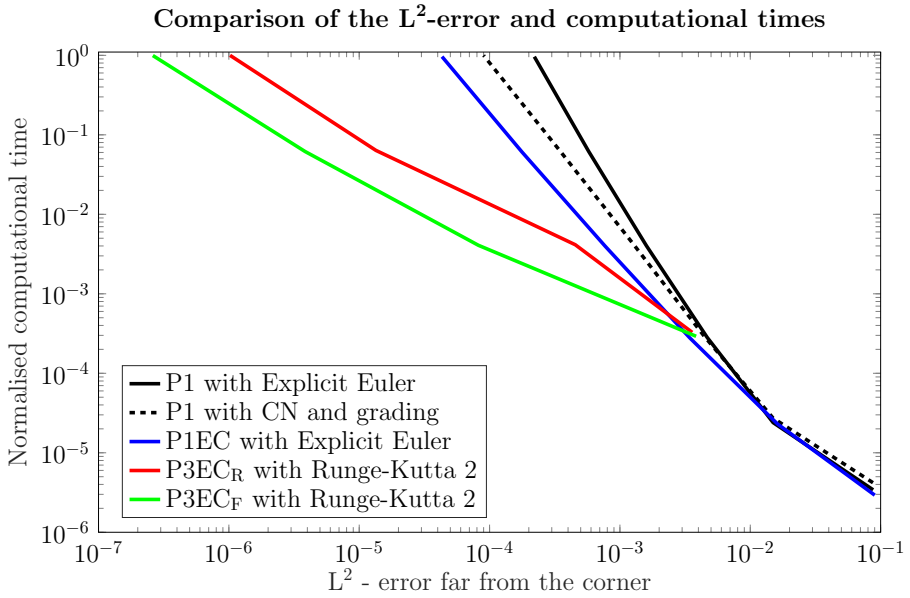
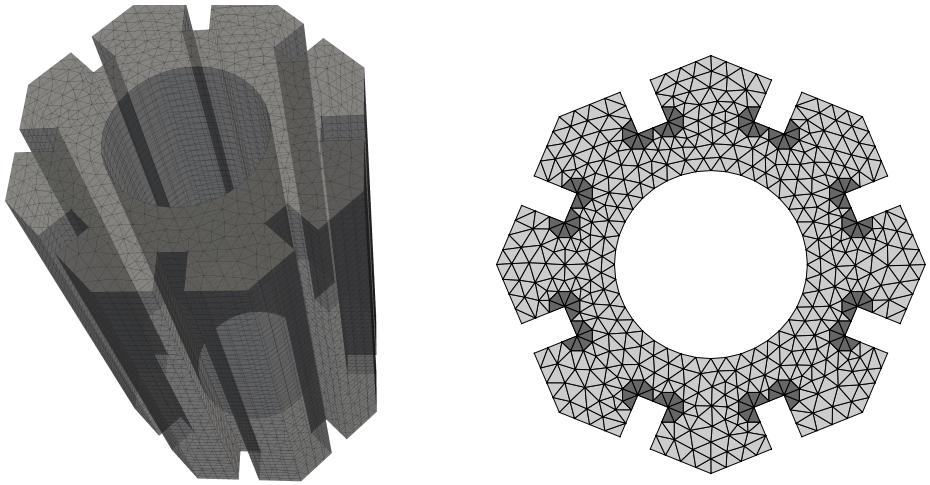


Figure 4.5: Comparison of the normalised computational time and accuracy of commonly used finite element discretisations of the parabolic problems with the proposed energy-corrected solvers on the L-shape domain. We compare the $L^2(\Omega')$ error evaluated at the final time-step $T = 1$.

As shown in Theorem 2.3.27, the energy-correction method gives optimal convergence rates in terms of the best-approximation property, however, in weighted norms. This, in particular, means that the method converges optimally when measured far from the re-entrant corner. Therefore, no additional post-processing needs to be applied, when one is interested in the solution far from the singular corner.

In Figure 4.5, a comparison of $L^2(\Omega')$ errors and normalised computational times of the previously described methods are shown. Again, due to the pollution effect, the standard finite element discretisation results in the worst error-to-time ratio. It can be improved by the application of the mesh grading together with Crank-Nicolson time-stepping, which yields only slightly worse results than the piecewise-linear energy-corrected scheme. The cubic energy-corrected finite element, together with Heun's time-stepping and mass-lumping strategy, results in by far the best method when the L^2-error far from the re-entrant corner is concerned. Small variations in the convergence rates in the cubic finite element scheme, when using $c_h^R(\cdot,\cdot)$ modification, appear because of insufficient initial resolution of the mesh. Again, modification $c_h^F(\cdot,\cdot)$ yields a better performance than $c_h^R(\cdot,\cdot)$.



(a) 3D geometry (b) Computational mesh of a cross-section

Figure 4.6: Geometry and a computational mesh of a graphite moderator brick of a nuclear power plant. In Figure 4.6a a complete 3D domain is shown.

4.4.3 Application

In this section, we apply the piecewise linear energy-corrected finite element method to a real 3D geometry of a graphite moderator brick of a nuclear power plant. Such a moderation type is commonly used in Advanced Gas-cooled Reactors (AGR) [125]. Efficient simulations of heat distribution in moderator bricks play an important role in the analysis of the material properties of the whole nuclear core, and accurate computations of temperature distribution can help determine the lifetime of nuclear materials, which often suffer from large temperature gradients and fast neutron fluxes [11].

The 3D geometry Ω of the graphite brick is presented in Figure 4.6a. The shape of the brick has a tensorial structure, with identical 2D horizontal cross-sections containing 16 equally-sized re-entrant corners with angle $\Theta = 3\pi/2$ at the external boundary. Heat transfer in a graphite moderator brick, in its simplest form, can be described by

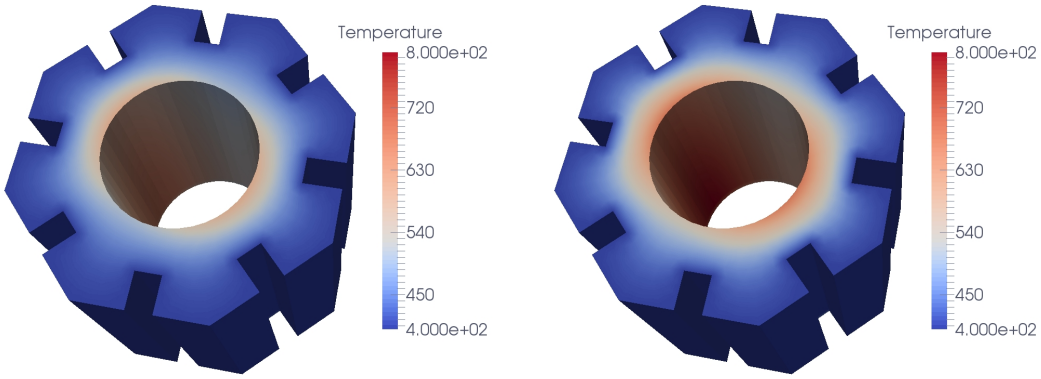
$$u_t - \Delta u = f \quad \text{in } \Omega \times (0,T), \tag{4.35}$$

$$u = g \quad \text{on } \partial\Omega_1 \times [0,T], \tag{4.36}$$

$$\partial_\nu u = 0 \quad \text{on } \partial\Omega_2 \times [0,T], \tag{4.37}$$

$$u = 0 \quad \text{on } \partial\Omega_3 \times [0,T], \tag{4.38}$$

$$u = u_0 \quad \text{in } \Omega \text{ at } t = 0. \tag{4.39}$$



(a) Temperature distribution at $t = 0.25$. (b) Temperature distribution at $t = 0.5$.

Figure 4.7: Temperature distribution in the geometry of a graphite moderator brick at two different time points. The temperature is shown in a cross-section of a domain at the height $L/2 = 1.5$. We would like to point out the nonuniformity of the solution and lack of rotational symmetry due to choice of the boundary conditions on the internal, cylindrical wall.

Here, $\partial\Omega_1$ is the interior boundary of the domain and g is the heating produced due to the nuclear reaction occurring in the fuel assembly. The system is thermally isolated from below on the $\partial\Omega_2$ part of the boundary, which is reflected by the uniform Neumann boundary conditions. Finally, the remaining part of the domain's boundary - $\partial\Omega_3$, is subject to a circulating coolant of a constant temperature. Note that the solution is rescaled, so that the temperature there is uniformly distributed.

Exploiting the tensorial structure of the domain Ω, we divide it into prismatic elements of equal length h_z in the vertical dimension. Moreover, each cross-section is triangulated as shown in Figure 4.6b. It is worth noting that around each of the 16 re-entrant corners in the cross-sections, we use identical one-element patches consisting of congruent isosceles triangles. This, together with the tensorial structure of the mesh, allows us to reuse the same parameter γ in the energy-corrected scheme (4.9) once computed in the two-dimensional setting, for all the corners. Following (4.16), we complete the finite element discretisation in space with the explicit Euler time-stepping scheme.

In the simulations, we choose a homogenuous initial temperature distribution $u_0 = 0$ and the heating on the internal boundary to be given by $g = 10\cos\big(\pi z/(2L)\big)\big(1 + \sin(\phi+4\pi t)\big)\sin(\pi t)$, where $L = 3$ is the height of the domain Ω and (r, ϕ, z), $z \in [0, L]$ are cylindrical coordinates associated with the domain. The simulation is performed using $N = 5000$ time steps with the final time at $T = 1$ and at the end of the

simulation the solution is rescaled to a physically meaningful value range by $U = 20u + 400$. The solution at two intermediate time-steps is presented in the vertical cross-sections of the domain taken in the middle at the height $z = 1.5$ at times $t = 0.25$ and $t = 0.5$.

To investigate the convergence of the scheme, we measure two different quantities of interest, namely the average temperature in the body at the final time-step and the average temperature in the whole space-time cylinder $\Omega \times [0, T]$

$$\mathrm{QoI}_1 = \frac{1}{|\Omega|} \int_\Omega u_h(T, x)\, \mathrm{d}x, \quad \mathrm{QoI}_2 = \frac{1}{T|\Omega|} \int_0^T \int_\Omega u_h(t, x)\, \mathrm{d}x\, \mathrm{d}t.$$

In order to investigate the convergence properties of the quantities of interest, we perform the computations on four different refinement levels. Initially, we use the mesh described above, which is then uniformly refined and the time-step size is divided by 4. The results of the computations are presented in Figure 4.8. We compare the accuracy of the standard finite element with the energy-corrected finite element method. Together with the quantities of interest QoI_1 and QoI_2, we include also the extrapolated values QoI_1^{ex} and QoI_2^{ex} in the plots. As presented in Figure 4.8, the application of energy-correction improves the approximation properties of the quantities of interest. Moreover, in the case of the standard finite element approximation the respective estimated orders of convergence of $|\mathrm{QoI}_1^{ex} - \mathrm{QoI}_1|$ and $|\mathrm{QoI}_2^{ex} - \mathrm{QoI}_2|$ are 1.37 and 1.23. The application of the energy-correction improves these orders and yields estimated values of 2.1 and 2.12, respectively.

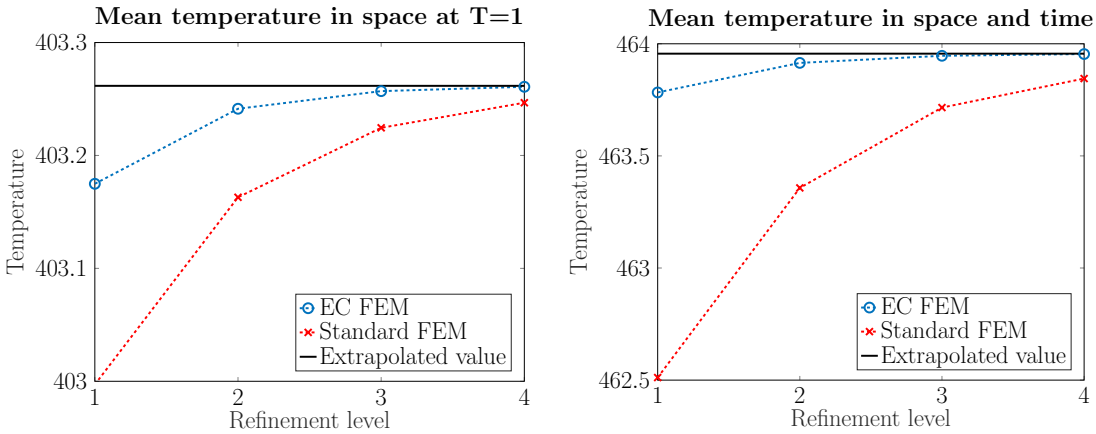


Figure 4.8: Comparison of the approximations of QoI_1 (left) and QoI_2 (right) obtained using standard and energy-corrected finite element scheme on four consecutive refinement levels in the space-time domain. Additionaly, extrapolated value of both quantities of interest are added to the plots for comparison.

Chapter 5

Energy-correction for optimal control problem

The results presented in this chapter were partly published by the author with L. John, and B. Wohlmuth in the paper entitled "Energy corrected FEM for optimal Dirichlet boundary control problems" in Numerische Mathematik, 139(4):913–938 in the year 2018, [92].

Here, we investigate the energy-corrected finite element approximation of a Dirichlet boundary optimal control problem on polygonal domains. We want to begin with a brief introduction to the field of optimal control by giving a general framework, in which our particular problem of interest is set. For this, we follow the presentation of [110].

Optimal control problems consist of several components and are concerned with the search for a control z, being an element of a set of admissible controls $\mathcal{U}_{\text{ad}}$. This set contains information about potential restrictions imposed on the control, such as its minimum or maximum pointwise value. Function z controls some process with an outcome u, which can be described as a solution of

$$F(u, z) = 0. \tag{5.1}$$

In problems of interest for us, the function $F(\cdot, \cdot)$ is a differential equation, potentially involving partial derivatives in various dimensions. The solution u can be observed only over some observation operator $\Phi(u)$. This operator, for instance, can be a pointwise evaluation of the function u, its average or maximum.

Finally, we equip the problem with the cost functional $\mathcal{J}(z)$, which describes, what we mean by the optimality. Often, the cost functional can be written in the form

$$\mathcal{J}(u, z) = \mathcal{C}\big(\Phi\big(u\big), z\big) + \mathcal{R}(z).$$

Functional $\mathcal{C}\big(\Phi\big(u(z)\big), z\big)$ measures, how close, for a given control z, the observed solution u is from the desired outcome. The second part $\mathcal{R}(z)$ is often interpreted as a regularisation term. It represents the cost yielded by the control and can be added for imposing additional regularity properties on the solution.

The optimal control problem can be written as

$$\arg\min_{z\in\mathcal{U}_{\mathrm{ad}}} \mathcal{J}(u,z), \quad \text{subject to } F(u,z)=0.$$

Optimal control problems constrained by partial differential equations occur in many technical applications, such as optimal heating and flow control, see, e.g., [61, 82, 90, 102] or optimal sensor placement [26, 40]. The precise form of the optimisation problem depends thus on the considered physical process. Equation (5.1) can be an elliptic [4, 7], parabolic [31, 44, 42], or hyperbolic partial differential equation [101]. Also, the considered spaces of admissible controls vary significantly, from subsets of L^2 spaces to measures [159]. Such models need to cover a wide range of computational domains, among them being polygonal domains with re-entrant corners [83].

The regularity theory of optimal control problems and finite element error estimates has been investigated in numerous articles, see, for instance, [4, 8, 43, 115]. As discussed in Section 2.3.2.2, in the presence of corners in the computational domain, the solutions of elliptic equations, in general, have singular components, even when smooth data are considered. Hence, the standard finite element approximations exhibit worse convergence properties than the nodal interpolation. Here, we follow the energy-correction based approach for improving the convergence properties of the finite element approximation of the optimal control problem. We apply this method to a certain class of optimal control problems, namely optimal Dirichlet boundary control in the energy space. The main result of this chapter shows the optimal convergence order of the energy-corrected finite element method for the approximation of this particular optimisation problem, see Theorem 5.2.11. It is worth noting that this methodology is not limited to this type of problem and can be extended to other types of optimal control, such as Neumann boundary control and distributed control. Also, different norms for the control might be considered, cf. [4, 8, 43, 61].

The remainder of the chapter is structured as follows: In Section 5.1, we introduce the optimal Dirichlet boundary control problem, recall the optimality system and present regularity results. Moreover, we present a numerical example illustrating the effect of a corner singularity on the accuracy and convergence of the finite element approximation. We derive optimal error estimates in weighted spaces for the optimal

control problem using the energy-corrected finite element method in Section 5.2. Finally, in Section 5.3, we present several numerical results, illustrating the presented theoretical results.

5.1 Dirichlet boundary control problem

Let $\Omega \subset \mathbb{R}^2$ be a bounded polygonal domain in the sense of Definition 2.1.11 with mutually disjoint boundary parts Γ_{D} and Γ_{C}, both of positive measure, satisfying $\Gamma := \partial\Omega = \overline{\Gamma}_{\mathrm{D}} \cup \overline{\Gamma}_{\mathrm{C}}$.

In this chapter, we consider as a model problem the optimal Dirichlet boundary control in the energy space. This type of optimal control problem was proposed in [110] and has been recently studied in [90, 127]. For a given desired state $\overline{u}$, right-hand side f and a cost-coefficient ρ, it is defined as a minimisation of the following tracking-type functional

$$\mathcal{J}(u,z) = \frac{1}{2}\|u-\overline{u}\|^2_{L^2(\Omega)} + \frac{\rho}{2}|z|^2_{H^{1/2}_{00}(\Gamma_{\mathrm{C}})}, \tag{5.2}$$

subject to the constraint

$$\begin{aligned} -\Delta u &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \Gamma_{\mathrm{D}}, \\ u &= z \quad \text{on } \Gamma_{\mathrm{C}}, \end{aligned} \tag{5.3}$$

and the control constraints

$$z_a \le z \le z_b \quad \text{a.e. on } \Gamma_{\mathrm{C}} \tag{5.4}$$

for some $z_a, z_b \in H^{1/2}_{00}(\Gamma_{\mathrm{C}})$.

In the notation of the framework introduced above, we set $F(u,z)$ to be the Poisson equation (5.3). Furthermore, $\mathcal{J}(u,z) = \mathcal{C}(u,z) + \mathcal{R}(z)$, where $\mathcal{C}(u,z)$ is the $L^2(\Omega)$ distance of the solution u from some desired state $\overline{u}$. The regularisation term $\mathcal{R}(z)$ is the energy norm of the control on the domain's boundary scaled by some positive parameter ρ. Similar problem with the regularisation involving $L^2(\Gamma)$ norm in place of the energy norm was studied in [4, 8, 43, 132, 135], where the control z in the Poisson problem (5.3) is applied on as the Dirichlet or Neumann boundary condition.

The use of the $H^{1/2}(\Gamma)$ norm on the boundary for the regularisation of the optimal control problem has several advantages compared to the $L^2(\Gamma_{\mathrm{C}})$ norm. First of all, it allows for the use of the weak formulation of (5.3) and the so-called ultra-weak form is not required. Moreover, the energy space $H^{1/2}(\Gamma)$ appears naturally as the trace of the Sobolev space $H^1(\Omega)$. Finally, it results in the higher regularity of the control than when the $L^2(\Gamma)$ regularisation is used. Hence, also better approximation order of the finite element method can be obtained, see [127, Section 4].

In particular, we are interested in the case of non-convex domains, i.e. Ω containing a re-entrant corner with interior angle $\pi < \Theta < 2\pi$. As before, we use the

notation $\lambda_i = i\pi/\Theta$. Moreover, the re-entrant corner is located on the Dirichlet boundary Γ_{D}, having a fixed positive distance from Γ_{C}, and we assume that the angle between Γ_{D} and Γ_{C} is not larger than $\pi/2$. This restricts the analysis to a certain subclass of optimal control problems, which nevertheless is not too restrictive and relevant in many application problems, see, e.g., [59, 102]. For simplicity, we consider a single re-entrant corner, located at the origin. The methodology and analysis presented here can be straightforwardly extended to a more general case of domains containing multiple re-entrant corners, see also Section 5.3 for numerical results.

The forthcoming analysis of the energy-correction scheme is carried out in weighted Lebesgue and weighted Sobolev spaces introduced in Section 2.1.2. We denote by $H^k_{0,\beta}(\Omega)$ the set of all functions from $H^k_\beta(\Omega)$ with zero trace and by $H^k_{0,\beta}(\Omega, \Gamma_{\mathrm{D}})$ all functions with zero trace on Γ_{D}. We skip the subscript β, when $\beta = 0$.

Recall, $H^{3/2}_{\mathrm{pw}}(\Gamma_{\mathrm{C}})$ is the Dirichlet trace space of $H^2_\beta(\Omega)$ with respect to Γ_{C}, see Section 2.1.3. We shall refer to [77, 87] for further details.

To introduce a convenient representation of the norm, we consider for a given $z \in H^{1/2}_{00}(\Gamma_{\mathrm{C}})$ the harmonic extension $u_z \in H^1(\Omega)$, which is defined as the unique weak solution of the problem

$$\begin{aligned} -\Delta u_z &= 0 \quad \text{in } \Omega, \\ u_z &= 0 \quad \text{on } \Gamma_{\mathrm{D}}, \\ u_z &= z \quad \text{on } \Gamma_{\mathrm{C}}\,. \end{aligned} \tag{5.5}$$

Definition 5.1.1 (Solution operator)**.** *For the harmonic extension, we introduce the solution operator* $\mathcal{H} : H^{1/2}_{00}(\Gamma_{\mathrm{C}}) \to H^1_0(\Omega, \Gamma_D)$ *assigning a weak solution of* (5.5) *to any given* $z \in H^{1/2}_{00}(\Gamma_{\mathrm{C}})$*, namely*

$$\mathcal{H}z = u_z \quad \in H^1_0(\Omega, \Gamma_D).$$

The following result is a special case of [98, Theorem 6.1.1].

Lemma 5.1.2. *Operator* $\mathcal{H} : H^{1/2}_{00}(\Gamma_{\mathrm{C}}) \to H^1_0(\Omega, \Gamma_D)$ *is bounded and coercive, namely for some constant* $c > 0$ *depending only on* Ω *and* Γ_C *such that for all* $z \in H^{1/2}_{00}(\Gamma_{\mathrm{C}})$ *we have*

$$\|\nabla \mathcal{H}z\|_0 \le c|z|_{H^{1/2}_{00}(\Gamma_{\mathrm{C}})}, \quad \textit{and} \quad \big\langle \mathcal{H}z, z \big\rangle \ge c|z|^2_{H^{1/2}_{00}(\Gamma_{\mathrm{C}})}.$$

Moreover, when $z \in H^{3/2}_{\mathrm{pw}}(\Gamma_C)$*, then* $\mathcal{H}z \in H^2_\alpha(\Omega)$ *for* $0 < 1 - \alpha < \lambda_1$.

From now on, we will assume that $0 \leq \alpha < 1$. Note that $H_0^1(\Omega, \Gamma_{\mathrm{D}}) \hookrightarrow L^2_{-\alpha}(\Omega)$, see Theorem 2.1.15.

For the realisation of the energy norm of the control we introduce the Steklov–Poincaré operator, also known as Dirichlet-to-Neumann map [158].

Definition 5.1.3. *We define the operator* $S : H_{00}^{1/2}(\Gamma_{\mathrm{C}}) \to H^{-1/2}(\Gamma_C)$ *by*

$$Sz = \partial_{\mathbf{n}} u_z \quad \text{in } H^{-1/2}(\Gamma_C), \tag{5.6}$$

where u_z *is the solution of* (5.5) *for a given* z.

Due to the Trace Theorem 2.1.19, the Steklov–Poincaré operator (5.6) is bounded and elliptic in $H_{00}^{1/2}(\Gamma_{\mathrm{C}})$. The seminorm in $H_{00}^{1/2}(\Gamma_{\mathrm{C}})$ can be equivalently written with the help of the duality product [127, Equation 2.5], namely

$$|z|^2_{H_{00}^{1/2}(\Gamma_{\mathrm{C}})} = \langle Sz, z \rangle_{\Gamma_{\mathrm{C}}} \quad \text{for all } z \in H_{00}^{1/2}(\Gamma_{\mathrm{C}}). \tag{5.7}$$

We assume that $f, \overline{u} \in L^2_{-\alpha}(\Omega)$ and $z_a, z_b \in H^{3/2}_{\mathrm{pw}}(\Gamma_{\mathrm{C}})$ with $z_a \leq z_b$ a.e. on Γ_{C}. These regularity assumptions can be relaxed for the existence and uniqueness of the solution of the optimal control problem. However, later we shall see that these more regular data are required to obtain optimal error convergence rates for the finite element approximation. Since we equip the problem with control constraints, we introduce the set of admissible controls by

$$\mathcal{U}_{\mathrm{ad}} = \big\{ z \in H_{00}^{1/2}(\Gamma_{\mathrm{C}}) : \ z_a \leq z \leq z_b \text{ a.e. on } \Gamma_{\mathrm{C}} \big\},$$

where z_a and z_b are assumed to be such that this set is not empty.

5.1.1 Primal and dual problems

In order to derive the dual problem, the solution of problem (5.3) is split into a homogeneous and an inhomogeneous part, namely $u = u_z + u_f$. Here, $u_f \in H_0^1(\Omega)$ is the solution of the Poisson problem with the right-hand side f and homogeneous Dirichlet boundary conditions: find $u_f \in H_0^1(\Omega)$ such that

$$a(u_f, v) = \langle f, v \rangle_\Omega, \quad \text{for all } v \in H_0^1(\Omega), \tag{5.8}$$

where the bilinear form is given by

$$a(u, v) = \langle \nabla u, \nabla v \rangle_\Omega \quad \text{for all } u, v \in H^1(\Omega).$$

Hence, the solution of (5.3) can be written as $u = \mathcal{H}z + u_f$.

For the optimality condition, we also need the adjoint solution operator. Therefore, for a given $\psi \in L^2_\alpha(\Omega)$, we consider the unique solution $p \in H^1_0(\Omega)$ of the adjoint problem

$$\begin{aligned} -\Delta p &= \psi \quad \text{in } \Omega, \\ p &= 0 \quad \text{on } \Gamma, \end{aligned} \tag{5.9}$$

with the variational formulation: find $p \in H^1_0(\Omega)$ such that

$$a(p, v) = \langle \psi, v \rangle_\Omega \quad \text{for all } v \in H^1_0(\Omega).$$

Then, the adjoint solution operator $\mathcal{H}^* : L^2_\alpha(\Omega) \to H^{-1/2}(\Gamma_{\mathrm{C}})$ for $\alpha \geq 0$, is given by

$$\mathcal{H}^*\psi = -\partial_{\mathbf{n}} p \quad \text{in } H^{-1/2}(\Gamma_{\mathrm{C}}).$$

Note, this simply follows from the fact that the adjoint operator defined in this way, satisfies

$$\langle \mathcal{H}^*\psi, w \rangle_{\Gamma_{\mathrm{C}}} = -a(p, \mathcal{H}w) + \langle \psi, \mathcal{H}w \rangle_\Omega = \langle \psi, \mathcal{H}w \rangle_\Omega \quad \text{for all } w \in H^{1/2}_{00}(\Gamma_{\mathrm{C}}).$$

5.1.2 Optimality conditions and regularity

Let us define $\overline{w} = \overline{u} - u_f$. Due to (5.7) and $u = \mathcal{H}z + u_f$ the cost functional in (5.2) can be rewritten as a reduced cost-functional

$$\widetilde{\mathcal{J}}(z) = \frac{1}{2}\|\mathcal{H}z - \overline{w}\|^2_{L^2(\Omega)} + \frac{\rho}{2}\langle Sz, z \rangle_{\Gamma_{\mathrm{C}}}.$$

The optimal control problem (5.2)–(5.4) is then equivalent to

$$\arg\min_{z \in \mathcal{U}_{\mathrm{ad}}} \widetilde{\mathcal{J}}(z). \tag{5.10}$$

The following result can be found in [161, Lemma 2.21].

Theorem 5.1.4. *The minimisation of the reduced cost functional* (5.10)*, and thus also the optimal control problem* (5.2)*, is equivalent to the variational inequality*

$$\langle \mathcal{H}^*\big(\mathcal{H}z - \overline{w}\big), w - z \rangle_{\Gamma_C} + \rho \, \langle Sz, w - z \rangle_{\Gamma_C} \geq 0 \quad \textit{for all } w \in \mathcal{U}_{\mathrm{ad}}.$$

Therefore, we introduce the operator $T_\rho : H^{1/2}_{00}(\Gamma_{\mathrm{C}}) \to H^{-1/2}(\Gamma_{\mathrm{C}})$ and right-hand side by

$$T_\rho = \mathcal{H}^*\mathcal{H} + \rho \, S, \quad g = \mathcal{H}^*(\overline{u} - u_f) \quad \text{in } H^{-1/2}(\Gamma_{\mathrm{C}}). \tag{5.11}$$

The optimal control problem (5.2)–(5.4) can then be equivalently formulated as finding $z \in \mathcal{U}_{\text{ad}}$ such that

$$\langle T_\rho z, w - z\rangle_{\Gamma_{\text{C}}} \geq \langle g, w - z\rangle_{\Gamma_{\text{C}}} \quad \text{for all } w \in \mathcal{U}_{\text{ad}}. \tag{5.12}$$

The following result describes some desirable properties of operator T_ρ.

Lemma 5.1.5. *The operator T_ρ is $H^{1/2}_{00}(\Gamma_{\text{C}})$-elliptic and bounded, namely for some constants $C_1^{T_\rho}, C_2^{T_\rho} > 0$ that depend only on the domain and the choice of the parameter ρ we have*

$$\langle T_\rho z, z\rangle_{\Gamma_{\text{C}}} \geq C_1^{T_\rho} |z|^2_{H^{1/2}_{00}(\Gamma_{\text{C}})}, \quad \|T_\rho z\|_{H^{-1/2}(\Gamma_C)} \leq C_2^{T_\rho} |z|_{H^{1/2}_{00}(\Gamma_{\text{C}})},$$

for all $z \in H^{1/2}_{00}(\Gamma_{\text{C}})$.

Proof. The proof of the properties mentioned above is based on the coercivity and boundedness of the Steklov–Poincaré operator S and the solution operator $\mathcal{H}$, see Lemma 5.1.2. We begin by showing the coercivity of T_ρ.

$$\begin{aligned}
\langle T_\rho z, z\rangle_{\Gamma_{\text{C}}} &= \langle \mathcal{H}^* \mathcal{H} z, z\rangle_{\Gamma_{\text{C}}} + \rho \langle Sz, z\rangle_{\Gamma_{\text{C}}} \\
&= \langle \mathcal{H} z, \mathcal{H} z\rangle_{\Gamma_{\text{C}}} + \rho \langle Sz, z\rangle_{\Gamma_{\text{C}}} \\
&\geq c_1 \|z\|^2_{H^{1/2}_{00}(\Gamma_{\text{C}})} + c_2 \rho \|z\|^2_{H^{1/2}_{00}(\Gamma_{\text{C}})} \geq c \min\{1, \rho\} \|z\|^2_{H^{1/2}_{00}(\Gamma_{\text{C}})}.
\end{aligned}$$

Similarly, we show the boundedness of the operator

$$\begin{aligned}
\|T_\rho z\|_{H^{-1/2}(\Gamma_{\text{C}})} &= \sup_{w \in H^{1/2}_{00}(\Gamma_{\text{C}})} \frac{\langle T_\rho z, w\rangle_{\Gamma_{\text{C}}}}{\|w\|_{H^{1/2}_{00}(\Gamma_{\text{C}})}} \\
&= \sup_{w \in H^{1/2}_{00}(\Gamma_{\text{C}})} \frac{\langle \mathcal{H} z, \mathcal{H} w\rangle_{\Gamma_{\text{C}}} + \rho \langle Sz, w\rangle_{\Gamma_{\text{C}}}}{\|w\|_{H^{1/2}_{00}(\Gamma_{\text{C}})}} \\
&\quad \sup_{w \in H^{1/2}_{00}(\Gamma_{\text{C}})} \frac{\|\nabla \mathcal{H} z\|_0 \|\nabla \mathcal{H} w\|_0 + \rho \|z\|_{H^{1/2}_{00}(\Gamma_{\text{C}})} \|w\|_{H^{1/2}_{00}(\Gamma_{\text{C}})}}{\|w\|_{H^{1/2}_{00}(\Gamma_{\text{C}})}} \\
&\leq c(1+\rho) \|z\|_{H^{1/2}_{00}(\Gamma_{\text{C}})}.
\end{aligned}$$

□

Note that the constants satisfy $C_1^{T_\rho} = \mathcal{O}(\min\{1, \rho\})$ and $C_2^{T_\rho} = \mathcal{O}(1 + \rho)$. Consequently, upon the direct application of [110, Theorem 1.2], we deduce that the elliptic variational inequality (5.12) of first kind has a unique solution.

Theorem 5.1.6 (Existence of the solution)**.** *Problem* (5.2)*, together with its equivalent formulation* (5.12)*, has a unique solution $z \in \mathcal{U}_{\text{ad}}$.*

Let us now introduce the Lagrange multiplier

$$\lambda = T_\rho z - g = \mathcal{H}^*(u - \overline{u}) + \rho\, Sz = -\partial_{\mathbf{n}} p + \rho\, \partial_{\mathbf{n}} u_z \in H^{-1/2}(\Gamma_{\mathrm{C}}), \tag{5.13}$$

where $p \in H^1_0(\Omega)$ is the solution of the problem (5.9) with $\psi = u - \overline{u}$ and $u_z = \mathcal{H}z$ is the harmonic extension defined in (5.5). The boundary Γ_{C} can be split into three parts, Γ_a, Γ_b called active zones and Γ_n called inactive zone, given by

$$\Gamma_a = \{\mathbf{x} \in \Gamma_{\mathrm{C}} :\ z(\mathbf{x}) = z_a(\mathbf{x})\}, \quad \Gamma_b = \{\mathbf{x} \in \Gamma_{\mathrm{C}} :\ z(\mathbf{x}) = z_b(\mathbf{x})\},$$
$$\Gamma_n = \Gamma_{\mathrm{C}} \setminus \big(\Gamma_a \cup \Gamma_b\big).$$

Following [90] we can reformulate the variational problem (5.12) equivalently as a Signiorini boundary value problem

$$\begin{aligned} -\Delta u_z = 0 \quad \text{in } \Omega, \qquad -\Delta p = u - \overline{u} \quad \text{in } \Omega, \\ u_z = 0 \quad \text{on } \Gamma_{\mathrm{D}}, \quad u_z = z_a \quad \text{on } \Gamma_a, \quad u_z = z_b \quad \text{on } \Gamma_b, \quad \rho\partial_{\mathbf{n}} u_z = \partial_{\mathbf{n}} p \quad \text{on } \Gamma_n \end{aligned} \tag{5.14}$$

equipped with the inequality constraints

$$\rho\partial_{\mathbf{n}} u_z \geq \partial_{\mathbf{n}} p \quad \text{on } \Gamma_a, \quad \rho\partial_{\mathbf{n}} u_z \leq \partial_{\mathbf{n}} p \quad \text{on } \Gamma_b, \quad z_a < z < z_b \quad \text{on } \Gamma_n.$$

These last constraints correspond to the Karush–Kuhn–Tucker conditions for the variational inequality (5.12). This is a standard primal-dual formulation of variational inequality (5.12), see [76, 95].

5.1.3 Weak formulation

In the following, we derive the weak formulation of the optimality system in a saddle-point form. For the adjoint problem (5.9) with $\psi = u - \overline{u}$, we obtain

$$a(p, v) = \langle u - \overline{u}, v\rangle_\Omega + \langle \partial_{\mathbf{n}} p, v\rangle_{\Gamma_{\mathrm{C}}} \quad \text{for all } v \in H^1_0(\Omega, \Gamma_{\mathrm{D}}). \tag{5.15}$$

Similarly, for the primal problem we have

$$a(u, v) - \langle \partial_{\mathbf{n}} u_z, v\rangle_{\Gamma_{\mathrm{C}}} = a(u_f, v) \quad \text{for all } v \in H^1_0(\Omega, \Gamma_{\mathrm{D}}),$$

and with the boundary condition $\rho\partial_{\mathbf{n}} u_z = \lambda + \partial_{\mathbf{n}} p$ on Γ_{C} we arrive at

$$a(u, v) = \frac{1}{\rho}\langle \partial_{\mathbf{n}} p + \lambda, v\rangle_{\Gamma_{\mathrm{C}}} + a(u_f, v) \quad \text{for all } v \in H^1_0(\Omega, \Gamma_{\mathrm{D}}). \tag{5.16}$$

Moreover, we demand the Lagrange multiplier λ to satisfy

$$\langle \lambda, w - z\rangle_{\Gamma_{\mathrm{C}}} \geq 0 \quad \text{for all } w \in \mathcal{U}_{\mathrm{ad}}.$$

Finally, combining (5.16) with (5.15) and the variational inequality leads for given $u_f \in H^1_0(\Omega)$ to the following variational formulation of the optimal control problem, see [90, Section 3.1.2] for more details: find $(u, p, \lambda) \in H^1_0(\Omega, \Gamma_{\mathrm{D}}) \times H^1_0(\Omega) \times H^{-1/2}(\Gamma_{\mathrm{C}})$ such that

$$\begin{aligned} a_\rho(u,v) - b(v,p) - \langle \lambda, v\rangle_{\Gamma_{\mathrm{C}}} &= \langle \overline{u}, v\rangle_\Omega + \rho\, a(u_f, v) && \text{for all } v \in H^1_0(\Omega, \Gamma_{\mathrm{D}}), \\ b(u,q) &= \langle f, q\rangle_\Omega && \text{for all } q \in H^1_0(\Omega), \\ \langle \lambda, w - u_{|\Gamma_{\mathrm{C}}}\rangle_{\Gamma_{\mathrm{C}}} &\geq 0 && \text{for all } w \in \mathcal{U}_{\mathrm{ad}}, \end{aligned} \tag{5.17}$$

where the bilinear forms are given by

$$a_\rho(u,v) = \langle u, v\rangle_\Omega + \rho\, \langle \nabla u, \nabla v\rangle_\Omega, \quad b(v,q) = \langle \nabla v, \nabla q\rangle_\Omega.$$

Remark 5.1.7. *In the absence of box-constraints, so when* $\mathcal{U}_{\mathrm{ad}} = H^{1/2}_{00}(\Gamma_{\mathrm{C}})$*, the variational inequality* (5.12) *turns into equality, and problem* (5.17) *turns into a simple 2-by-2 saddle-point problem of finding* $(u, p) \in H^1_0(\Omega, \Gamma_D) \times H^1_0(\Omega)$ *such that*

$$\begin{aligned} a_\rho(u,v) - b(v,p) &= \langle \overline{u}, v\rangle_\Omega + \rho\, a(u_f, v) && \textit{for all } v \in H^1_0(\Omega, \Gamma_D), \\ b(u,q) &= \langle f, q\rangle_\Omega && \textit{for all } q \in H^1_0(\Omega). \end{aligned} \tag{5.18}$$

From now on, we assume that the weight satisfies

$$1 - \lambda_1 < \alpha < 1.$$

For the solution, the following regularity result is valid.

Theorem 5.1.8. *Let* $\Omega \subset \mathbb{R}^2$ *be a bounded polygonal domain with one re-entrant corner with an interior angle* $\pi < \Theta < 2\pi$*. Furthermore, let* $f, \overline{u} \in L^2_{-\alpha}(\Omega)$*,* $z_a, z_b \in H^{3/2}_{\mathrm{pw}}(\Gamma_C)$*,* $z_a < z_b$*. Then the solution of Problem* (5.17)*, satisfies* $u \in H^1_0(\Omega, \Gamma_D) \cap H^2_\alpha(\Omega)$*,* $p \in H^1_0(\Omega) \cap H^2_\alpha(\Omega)$*,* $z \in H^{1/2}_{00}(\Gamma_{\mathrm{C}}) \cap H^{3/2}_{\mathrm{pw}}(\Gamma_C)$*, and the following splitting into regular and singular parts holds*

$$u = U + \sum_{\lambda_i < 1+\alpha} k_i s_i, \quad \textit{and} \quad p = P + \sum_{\lambda_i < 1+\alpha} l_i s_i.$$

Here $U, P \in H^2_{-\alpha}(\Omega)$ *and* s_i *are the singular functions* (2.2)*. Moreover, we obtain* $g \in H^{1/2}_{\mathrm{pw}}(\Gamma_C)$*, where* g *is defined in* (5.11)*.*

Proof. For the proof of this theorem in the case of convex Ω and $\alpha = 0$, we refer the reader to [127, Theorem 2.2]. Following the reasoning presented there, the proof can be generalised to polygonal domains containing re-entrant corners.

Due to Theorem 5.1.6, there exists a unique solution $z \in H^{1/2}_{00}(\Gamma_C)$. Furthermore, the primal state can be written as $u = \mathcal{H}z + u_f$, where u_f is the unique solution of (5.8). According to Theorem 2.2.6, we have $u_f \in H^2_\alpha(\Omega) \cap H^1_0(\Omega)$ and

$$u_f = U_f + \sum_{\lambda_i < 1+\alpha} k_i^f s_i \tag{5.19}$$

for some $U_f \in H^2_{-\alpha}(\Omega)$ and $k_i^f \in \mathbb{R}$.

We also know that $\mathcal{H}z \in H^1_0(\Omega, \Gamma_D)$, see Lemma 5.1.2, so $u \in H^1_0(\Omega, \Gamma_D)$.

Let η be a smooth cut-off function as defined in (2.2) with $\overline{r} > 0$ chosen so that $\overline{B_{\overline{r}}} \subset \Omega$, where B_R is a ball of radius R centered at the re-entrant corner. We can write $\mathcal{H}z = \eta\mathcal{H}z + (1-\eta)\mathcal{H}z$.

According to [127, Theorem 2.2], the Signiorini problem (5.17) does not exhibit any singular behaviour at the boundary and thus $(1-\eta)\mathcal{H}z \in H^2(\Omega)$. Finally, notice that $\eta\mathcal{H}z$ is the unique solution of the equation

$$-\Delta\tilde{u} = -\Delta\big(\eta\mathcal{H}z\big), \text{ in } \Omega, \quad \text{and} \quad \tilde{u} = 0, \text{ on } \Gamma.$$

Since in $B_{\underline{r}}$ we have $-\Delta\big(\eta\mathcal{H}z\big) = 0$, then $-\Delta\big(\eta\mathcal{H}z\big) \in L^2_{-\alpha}(\Omega)$. Theorem 2.2.6 guarantees $\tilde{u} \in H^2_\alpha(\Omega)$ and $\tilde{u}$ can be split into regular and singular part

$$\tilde{u} = \tilde{U} + \sum_{\lambda_i < 1+\alpha} \tilde{k}_i s_i, \tag{5.20}$$

where $\tilde{U} \in H^2_{-\alpha}(\Omega)$ and $\tilde{k}_i \in \mathbb{R}$.

Finally, taking into account (5.19), we get $u \in H^2_\alpha(\Omega)$ and the desired splitting into regular and singular parts

$$u = U + \sum_{\lambda_i < 1+\alpha} k_i s_i$$

holds true. Due to Trace Theorem 2.1.19, we also have $z \in H^{1/2}_{00}(\Gamma_C) \cap H^{3/2}_{\mathrm{pw}}(\Gamma_C)$.

Applying the results of Theorem 2.2.6, the adjoint state being the unique solution of (5.9) with $\phi = u - \overline{u}$ satisfies $p \in H^2_\alpha(\Omega)$ and for some $P \in H^2_{-\alpha}(\Omega)$ and $l_i \in \mathbb{R}$

$$p = P + \sum_{\lambda_i < 1+\alpha} l_i s_i.$$

□

The boundary control z and box constraints are located far from the re-entrant corner, so the regularity of the control variable z is not influenced by the non-convexity.

However, a singularity in the primal and adjoint states u and p being solutions of the Poisson equation, may occur in the vicinity of the re-entrant corner. Thus, we may not, in general, obtain desirable $H^2(\Omega)$ regularity of the solution and the use of weighted $H^2_\alpha(\Omega)$ space is necessary.

The angle between Γ_{C} and Γ_{D} is assumed to be at most of size $\pi/2$. Otherwise, another type of singular functions would appear in the solution around the corner, where the Neumann and Dirichlet boundaries in the Signiorini formulation (5.14) meet [108], and some additional numerical treatment of such singularities would need to be introduced. This, however, is beyond the scope of this work.

5.1.4 Finite element discretization

Now, we would like to show the negative influence that the presence of the re-entrant corner has on the convergence of the finite element approximation of the optimal control problem (5.17). To do this, we first introduce the finite element discretisation of the problem (5.18), where no box-constraints are imposed. Finally, we recall the known convergence result proposed in [127] in the convex setting.

Let $\mathcal{T}_h$ be a family of uniformly refined admissible and shape-regular triangulations of Ω, as introduced in Section 2.3. So far, we considered only problems with homogeneous conditions on the boundary. In this chapter, however, we are interested in problems with nonhomogeneous boundary values, and hence, we would like to redefine the discrete spaces as follows

$$V_h = S^1_h \cap H^1_0(\Omega, \Gamma_{\mathrm{D}}) \subset H^1_0(\Omega, \Gamma_{\mathrm{D}}), \qquad Q_h = S^1_h \cap H^1_0(\Omega) \subset H^1_0(\Omega).$$

Note that this is consistent with the notation used before, since when $\Gamma_{\mathrm{C}} = \emptyset$, spaces V_h and Q_h are equal.

From now on, we assume for the control constraints that $z_a, z_b \in H^{3/2}_{\mathrm{pw}}(\Gamma_{\mathrm{C}})$. Let us introduce the finite element space for the control as a trace operator space of V_h, namely

$$Z_h = V_h|_{\Gamma_{\mathrm{C}}} \cap H^{1/2}_{00}(\Gamma_{\mathrm{C}}) \subset H^{1/2}_{00}(\Gamma_{\mathrm{C}}).$$

The triangulation $\mathcal{T}_h$ defines a natural partition $\mathcal{Z}_h$ of the boundary Γ intro intervals of the length $\mathcal{O}(h)$. Also, the traces of the nodal basis functions $\{\phi_i\}_{i=1}^{K+L}$ associated with the space S_h, see (2.9), define the nodal basis $\{\phi_i|_\Gamma\}_{i=1}^{N}$ of the space Z_h on the partition $\mathcal{Z}_h$. Therefore, the nodal interpolation operator on the boundary of the domain Ω is well-defined.

Definition 5.1.9 (Interpolation on the boundary)**.** *Let $\mathcal{X}_{\Gamma_C}$ be the set of nodal points on the boundary Γ_C associated with the partition $\mathcal{Z}_h$. Let also $\Gamma_0 \subset \Gamma$ be an open set. We define the nodal interpolation operator $I_h : C(\overline{\Gamma_0}) \to Z_h$ by*

$$I_h v = \sum_{i:\, \mathbf{x}_i \in \mathcal{X}_{\Gamma_C} \cap \overline{\Gamma_0}} v(\mathbf{x}_i)\phi_i|_\Gamma.$$

Now, we state the result showing the accuracy of the approximation using the interpolation operator on the given partition $\mathcal{Z}_h$ of the boundary. It will serve as a benchmark, with which we shall compare the accuracy of the finite element schemes for the optimal control problem. A similar set of estimates to Theorem 2.3.10 also holds when the boundary is considered, see [149, Equation 10.11].

Theorem 5.1.10 (Interpolation on the boundary)**.** *Let $\Gamma_0 \subset \Gamma$ be an open set and let $v \in H^s(\Gamma_0)$ for some $1/2 < s \leq 2$. Then for any $0 \leq r \leq \min(1, s)$ the following interpolation error estimates hold for some $c > 0$ independent of v and h*

$$\|v - I_h v\|_{H^r(\Gamma_0)} \leq c h^{s-r} |v|_{H^s(\Gamma_0)}.$$

Note that the functions $v \in H^s(\Gamma_0)$, where $s > 1/2$, are continuous due to the embedding stated in Theorem 2.1.9. Hence, the interpolation operator is also well defined.

Even in the case of the convex polygonal domain, no more than $H^{1/2}_{00}(\Gamma_C)$ regularity for the control and $H^2(\Omega)$ regularity for the primal state u and the adjoint state p, in general, can be obtained, see Theorem 5.1.8. Hence, the application of the higher order finite element discretisation would not improve the convergence property of the finite element scheme. Therefore, we restrict our interest to the piecewise linear approximation only.

For now, we only state the finite element approximation of the optimal control problem in the absence of the box-constraints (5.4). This means that $\lambda = 0$ in (5.17) and the optimisation problem reduces to the saddle point formulation (5.18). We use this formulation of the problem for introducing the discretisation.

Definition 5.1.11. *The finite element discretisation of problem* (5.18) *reads: find $(u_h, p_h) \in V_h \times Q_h$ such that*

$$\begin{aligned} a_\rho(u_h, v_h) - b(v_h, p_h) &= \langle \overline{u}, v_h \rangle_\Omega + \rho\, a(u_{f,h}, v_h) && \text{for all } v_h \in V_h, \\ b(u_h, q_h) &= \langle f, q_h \rangle_\Omega && \text{for all } q_h \in Q_h. \end{aligned} \tag{5.21}$$

Here, $u_{f,h} \in Q_h$ is a standard finite element approximation of the Poisson (2.10) *with the right-hand side f. The discrete control is obtained applying the trace $z_h = u_h|_{\Gamma_C} \in Z_h$.*

The following convergence result was given in [127, Corollary 3.8] in a general setting including the box-constraints. However, for the sake of presentation, we show only its version restricted to the case relevant here.

Theorem 5.1.12 (Error estimates on convex polygonal domains)**.** *Let $\Omega \subset \mathbb{R}^2$ be a bounded, convex polygonal domain and let $f, \overline{u} \in L^2(\Omega)$. Then, the following estimates hold for the solution $(u_h, p_h, z_h) \in V_h \times Q_h \times Z_h$ of (5.21) and some constant $c(\overline{u}, f) > 0$ independent of h*

$$\|z - z_h\|_{H^{1/2}(\Gamma_C)} \leq c(\overline{u}, f)h.$$

Moreover,

$$\|u - u_h\|_0 \leq c(\overline{u}, f)h^2, \quad \text{and} \quad \|p - p_h\|_0 \leq c(\overline{u}, f)h^2.$$

Recently, these results have been improved and extended to a more general case of arbitrary polygonal domains in [167, Theorem 3].

Theorem 5.1.13 (Error estimates on polygonal domains)**.** *Let $\Omega \subset \mathbb{R}^2$ be a polygonal domain with the largest interior angle $\Theta > 0$ and suppose that $f, \overline{u} \in L^2(\Omega)$. Then, for any $\epsilon > 0$, the following estimates hold for the solution $(u_h, p_h, z_h) \in V_h \times Q_h \times Z_h$ of (5.21) and some constant $c(\overline{u}, f) > 0$ independent of h*

$$\|z - z_h\|_{H^{1/2}(\Gamma_C)} \leq c(\overline{u}, f)h^{\min(1, \lambda_1 - \epsilon)}.$$

Furthermore, if Ω is convex and $\overline{u} \in C^{0,\sigma}(\overline{\Omega})$ for some $\sigma > 0$, then there holds

$$\|z - z_h\|_{H^{1/2}(\Gamma_C)} \leq c(\overline{u}, f)h^{\min(3/2, \lambda_1 - \epsilon)}.$$

These convergence orders are optimal in the sense of the interpolation error on the boundary of the domain stated in Theorem 5.1.10. Also inside the domain Ω the primal and adjoint states u and p exhibit optimal convergence, see Theorem 2.3.10. Yet, the convergence on the boundary depends on the size of the largest interior angle in the domain. We will confirm this result numericaly and further, focus our attention on improving the convergence property of the finite elements approximation.

5.1.5 Numerical examples

In this section, we show the difference between the numerically obtained estimates of the convergence orders of the finite element schemes on convex and non-convex

domains. In the latter case, similarly as in the case of elliptic problems from Section 2.3.2.2 and parabolic problems discussed in Chapter 4, we observe significantly worse approximation properties than the ones yield by the interpolation operator.

For the sake of simplicity, we consider no box-constraints, so $\lambda = 0$ in formulation (5.17). Following (5.21), we denote the finite element solution of the discrete version of (5.17) by $(u_h, p_h) \in V_h \times Q_h$. As a computational domain we consider either the unit square $\Omega = (0,1)^2$ with $\Gamma_C = \Gamma$ or the non-convex domain $\Omega = (-1,2)^2 \backslash [0,1]^2$ with $\Gamma_C = \{-1,2\} \times [-1,2] \cup [-1,2] \times \{-1,2\}$ and $\Gamma_D = \Gamma \setminus \Gamma_C$, as illustrated in Figure 5.1.

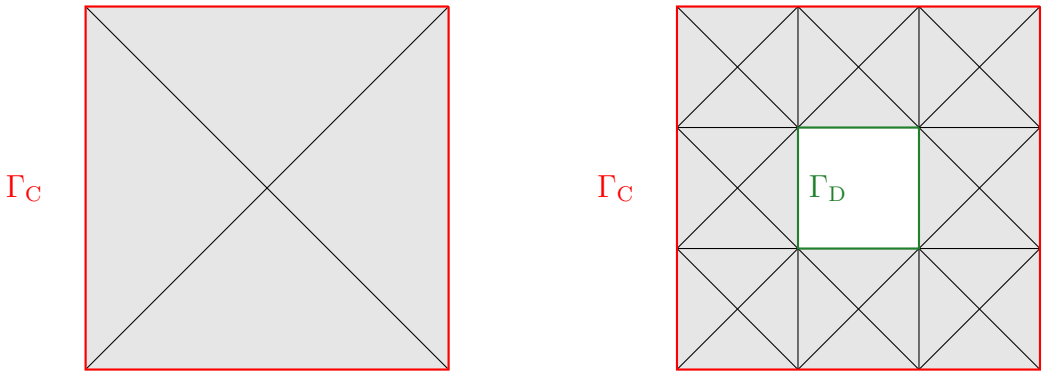


Figure 5.1: Computational domains with corresponding initial mesh, for the convex-case (left) and the non-convex-case (right).

In order to illustrate the influence of a re-entrant corner on the convergence of the finite element approximation we choose a smooth desired state $\overline{u} = ((x_1 - 1/2)^2 + (x_2 - 1/2)^2)^2$, right-hand side $f = 0$ and cost coefficient $\rho = 1$. Note, in the case of the non-convex domain we have four re-entrant corners, each with an interior angle of size $\Theta = 3/2\pi$.

For the numerical tests, we choose $\alpha = 1/3$. The weights are applied in the vicinity of every re-entrant corner in the domain. This choice of the weight parameters induces a slightly stronger norm than required by the theory presented in Section 2.3.3. This difference between the two norms cannot be observed numerically, and we make this particular choice of the weighting parameter α to keep the presentation consistent with the results presented in Section 5.3. Since the exact solution for these problems is unknown we take the solution of refinement level $L = 9$, denoted by $(u_{h_9}, p_{h_9}, z_{h_9})$, as a reference solution.

In Table 5.1, we present errors in the standard and weighted L^2-norms as well as estimated orders of convergence (eoc). The error on the boundary is also measured in the L^2-norm and we expect up to a half an order higher convergence order than

convex domain						
L	$\|u_{h_9} - u_h\|_0$	eoc	$\|u_{h_9} - u_h\|_\alpha$	eoc	$\|z_{h_9} - z_h\|_{L^2(\Gamma_C)}$	eoc
1	2.8098e–04	–	1.7264e–04	–	9.7098e–04	–
2	7.9573e–05	1.82	5.0858e–05	1.76	3.6516e–04	1.41
3	1.8388e–05	2.11	1.2022e–05	2.08	1.0333e–04	1.82
4	4.2616e–06	2.11	2.7865e–06	2.11	2.6673e–05	1.95
5	1.0259e–06	2.05	6.6849e–07	2.06	6.7037e–06	1.99
6	2.5248e–07	2.02	1.6422e–07	2.03	1.6628e–06	2.01
7	6.2377e–08	2.02	4.0545e–08	2.02	3.9999e–07	2.06
non-convex domain						
L	$\|u_{h_9} - u_h\|_0$	eoc	$\|u_{h_9} - u_h\|_\alpha$	eoc	$\|z_{h_9} - z_h\|_{L^2(\Gamma_C)}$	eoc
1	8.3320e–01	–	1.2354e+00	–	8.4197e–01	–
2	3.0218e–01	1.46	4.3269e–01	1.51	3.2203e–01	1.39
3	1.1299e–01	1.42	1.5131e–01	1.52	1.1319e–01	1.51
4	4.3259e–02	1.39	5.6016e–02	1.43	4.0961e–02	1.47
5	1.6621e–02	1.38	2.1209e–02	1.40	1.5269e–02	1.42
6	6.2772e–03	1.40	7.9527e–03	1.42	5.6907e–03	1.42
7	2.2261e–03	1.50	2.8085e–03	1.50	2.0069e–03	1.50

Table 5.1: Errors and estimated convergence orders for the optimal control problem in a convex and non-convex domain.

given in the estimates in energy-norm, see Theorem 5.1.12 and Theorem 5.1.13. In the numerical experiments, we obtain the predicted optimal convergence rates in the case of a convex domain for the approximation of the primal state. Furthermore, the convergence order for the approximation of the control is also optimal, when compared with the interpolation errors stated in Theorem 5.1.10. In the non-convex case, we observe a reduced convergence order, for the primal state in the standard $L^2(\Omega)$ and for the control in the $L^2(\Gamma_C)$-norm. Also, the consideration of a weighted norm does not increase the order of convergence.

This means that in the presence of a re-entrant corner a non-optimal convergence is obtained and the pollution effect, see Theorem 2.3.21, known in the case of elliptic and parabolic problems is present also in the optimal control problem. The need for eliminating the pollution motivates the application of an energy-corrected finite element approach.

5.2 Energy-corrected finite elements for optimal boundary control

In this section, we extend the energy-corrected finite element approach to the optimal Dirichlet boundary control problem introduced in Section 5.1. We will show that the suboptimal convergence results on non-convex polygonal domains discussed in Section 5.1.5 can be improved using the energy-corrected finite element method. The line of reasoning is similar to the one proposed in [127] in the case of the standard finite element method and convex domains. However, the use of the weighted Sobolev spaces is necessary.

Whenever we speak of the energy-corrected finite element method in this section, we assume that the optimal correction parameter γ^*, see (2.33) is known so that the assumption (2.25) is satisfied. Furthermore, we also impose the conditions (U) on the computational mesh in the vicinity of the re-entrant corner. The analysis presented here is based on the investigations of the approximative operators in the discrete spaces, defined with the help of the energy-corrected method. We will treat all the operators separately in Sections 5.2.1–5.2.3.

We define the set of discrete admissible controls as

$$\mathcal{U}_{\mathrm{ad},h} = \big\{ z_h \in Z_h : z_a(x_i) \leq z_h(x_i) \leq z_b(x_i) \quad \text{for all nodes } x_i \in \overline{\Gamma}_{\mathrm{C}} \big\},$$

which is a subset of $H^{1/2}_{00}(\Gamma_{\mathrm{C}})$. The corresponding finite element semi-discretisation of (5.12) reads then: find $z_h \in \mathcal{U}_{\mathrm{ad},h}$ such that

$$\langle T_\rho\, z_h, w_h - z_h \rangle_{\Gamma_{\mathrm{C}}} \geq \langle g, w_h - z_h \rangle_{\Gamma_{\mathrm{C}}} \quad \text{for all } w_h \in \mathcal{U}_{\mathrm{ad},h}. \tag{5.22}$$

Similarly as in the continuous setting presented in Theorem 5.1.6, due to the boundedness and coercivity of the operator T_ρ, see Lemma 5.1.5, problem (5.22) admits a unique solution [110, Theorem 1.2]. The following finite element error estimate can be found in [127, 148].

Theorem 5.2.1. *Let $z \in \mathcal{U}_{\mathrm{ad}}$ and $z_h \in \mathcal{U}_{\mathrm{ad},h}$ be the unique solutions of* (5.12) *and* (5.22) *respectively and let $z_a, z_b \in H^{3/2}_{\mathrm{pw}}(\Gamma_C)$, $g \in H^{1/2}_{\mathrm{pw}}(\Gamma_C)$. Then the following error estimates are satisfied for some $c > 0$*

$$\|z - z_h\|_{H^{1/2}_{00}(\Gamma_{\mathrm{C}})} \leq ch \|z\|_{H^{3/2}_{\mathrm{pw}}(\Gamma_C)}.$$

All the arguments presented there hold also in the presence of re-entrant corner in the domain as long as the control boundary has a fixed positive distance from the re-entrant corner.

5.2.1 Primal problem

Now we want to define a discrete counterpart of the solution operator $\mathcal{H}$. Therefore, let us first denote by $\mathcal{E}$ a continuous extension operator, vanishing in a fixed neighborhood of the re-entrant corner, which is both a mapping $\mathcal{E} : H^{1/2}_{00}(\Gamma_{\mathrm{C}}) \to H^1_0(\Omega, \Gamma_{\mathrm{D}})$ and $\mathcal{E} : H^{1/2}_{00}(\Gamma_{\mathrm{C}}) \cap H^{3/2}_{\mathrm{pw}}(\Gamma_{\mathrm{C}}) \to H^1_0(\Omega, \Gamma_{\mathrm{D}}) \cap H^2(\Omega)$.

Let η be a smooth cut-off function as defined in (2.2) with $\overline{r} > 0$ chose so that $\overline{B_{\overline{r}}} \subset \Omega$, where B_R is a ball of radius R centered at the re-entrant corner. Without the loss of generality, we can choose $\mathcal{E}z = (1 - \eta)\mathcal{H}$.

Furthermore, let $\Pi_h : H^1(\Omega) \to V_h$ be a Scott-Zhang quasi-interpolation operator, see [145] for details.

Definition 5.2.2 (Modified Ritz projection)**.** *For a given $\xi \in H^1_0(\Omega, \Gamma_D)$ we introduce the energy-corrected Ritz projection $R^m_h : H^1_0(\Omega, \Gamma_D) \to V_h$ as a solution of the following problem: find $R^m_h\xi \in V_h$ with $R^m_h\xi = \Pi_h\xi$ on Γ_C such that*

$$a_h(R^m_h\xi, v_h) = a(\xi, v_h) \quad \textit{for all } v_h \in V_h.$$

This definition of the modified Ritz projection is a straightforward extension of the one given in Section 2.3.3 for the problems with homogeneous Dirichlet boundary conditions.

Since $Q_h \subset V_h$, we then find $u^m_{0,h} \in Q_h$ as the solution of

$$a_h(u^m_{0,h}, q_h) = -a_h(R^m_h\mathcal{E}z, q_h) \quad \text{for all } q_h \in Q_h.$$

Now, we define the approximate solution operator $\widetilde{\mathcal{H}} : H^{1/2}_{00}(\Gamma_{\mathrm{C}}) \to L^2_{-\alpha}(\Omega)$ by

$$\widetilde{\mathcal{H}}z = u^m_{0,h} + R^m_h\mathcal{E}z. \tag{5.23}$$

Note, the continuous solution operator $\mathcal{H}z$ can be similarly expressed by $\mathcal{H}z = u_0 + \mathcal{E}z$. This definition is different than the one introduced in [127]. However, the advantages of our definition will be apparent once the discrete adjoint operator is considered, see Section 5.2.3.

Now, due to the results described in Section 2.3.3, the following error estimates are valid.

Lemma 5.2.3. *The approximate solution operator $\widetilde{\mathcal{H}} : H^{1/2}_{00}(\Gamma_{\mathrm{C}}) \to L^2_{-\alpha}(\Omega)$, defined in (5.23), is bounded. Moreover, if $z \in H^{1/2}_{00}(\Gamma_{\mathrm{C}}) \cap H^{3/2}_{\mathrm{pw}}(\Gamma_C)$, the following error estimates hold for some $c > 0$*

$$\|(\mathcal{H} - \widetilde{\mathcal{H}})z\|_\alpha \le ch^2\|z\|_{H^{3/2}_{\mathrm{pw}}(\Gamma_C)}, \quad \textit{and} \quad \|\nabla(\mathcal{H} - \widetilde{\mathcal{H}})z\|_\alpha \le ch\|z\|_{H^{3/2}_{\mathrm{pw}}(\Gamma_C)}.$$

Proof. Let $s \in \{0, 1\}$. From the definition of the solution operator $\mathcal{H}$ and approximate solution operator $\widetilde{\mathcal{H}}$ we see that

$$\|\nabla^s(\mathcal{H} - \widetilde{\mathcal{H}})z\|_\alpha \leq \|\nabla^s(u_0 - u_{0,h}^m)\|_\alpha + \|\nabla^s(\mathcal{E}z - R_h^m\mathcal{E}z)\|_\alpha.$$

Following equalities hold for all $q_h \in Q_h$

$$a_h(u_{0,h}^m, q_h) = -a_h(R_h^m\mathcal{E}z, q_h) = -a(\mathcal{E}z, q_h) = \langle \Delta\mathcal{E}z, q_h\rangle_\Omega.$$

Thus, due to Theorem 2.3.27, we have for some $c > 0$

$$\|\nabla^s(\mathcal{E}z - R_h^m\mathcal{E}z)\|_\alpha \leq ch^{2-s}\|\Delta\mathcal{E}z\|_{-\alpha} \leq ch^{2-s}\|\mathcal{E}z\|_{H^2(\Omega)} \leq ch^{2-s}\|z\|_{H^{3/2}_{\mathrm{pw}}(\Gamma_\mathrm{C})}.$$

Similarly, we obtain $\|\nabla^s(u_0 - u_{0,h}^m)\|_\alpha \leq ch^{2-s}\|z\|_{H^{3/2}_{\mathrm{pw}}(\Gamma_\mathrm{C})}$ and hence, the assertion follows. □

Let now $u_f \in H_0^1(\Omega)$ be the weak solution of problem (5.8) and $u_{f,h}^m \in Q_h$ its finite element approximation obtained using the energy-corrected scheme

$$a_h(u_{f,h}^m, q_h) = \langle f, q_h\rangle_\Omega \quad \text{for all } q_h \in Q_h.$$

Then, according to Theorem 2.3.27, for $f \in L^2_{-\alpha}(\Omega)$ we have for some $c > 0$ independent of f

$$\|u_f - u_{f,h}^m\|_\alpha \leq ch^2\|f\|_{-\alpha}. \tag{5.24}$$

5.2.2 Steklov–Poincaré operator

The Steklov–Poincaré operator S, defined in (5.6), can be expressed by

$$\langle Sz, w\rangle_{\Gamma_\mathrm{C}} = \langle \nabla\mathcal{H}z, \nabla\mathcal{E}w\rangle_\Omega, \quad \text{for all } w \in H_{00}^{1/2}(\Gamma_\mathrm{C}).$$

This, in turn, motivates the following definition of the approximate Steklov–Poincaré operator $\widetilde{S} : H_{00}^{1/2}(\Gamma_\mathrm{C}) \to H^{-1/2}(\Gamma_\mathrm{C})$ by

$$\langle \widetilde{S}z, w\rangle_{\Gamma_\mathrm{C}} = \langle \nabla\widetilde{\mathcal{H}}z, \nabla\mathcal{E}w\rangle_\Omega \quad \text{for all } w \in H_{00}^{1/2}(\Gamma_\mathrm{C}), \tag{5.25}$$

with the approximative solution operator $\widetilde{\mathcal{H}}$, see (5.23).

Lemma 5.2.4. *The approximative Steklov–Poincaré operator $\widetilde{S}$, defined in* (5.25)*, is bounded and $H_{00}^{1/2}(\Gamma_\mathrm{C})$-elliptic. For $z \in H_{00}^{1/2}(\Gamma_\mathrm{C}) \cap H^{3/2}_{\mathrm{pw}}(\Gamma_C)$ there holds the error estimate for some $c > 0$ independent of z*

$$\|(S - \widetilde{S})z\|_{H^{-1/2}(\Gamma_C)} \leq ch\|z\|_{H^{3/2}_{\mathrm{pw}}(\Gamma_C)}.$$

Proof. From the definition of $H^{-1/2}(\Gamma_C)$ norm and the application of the Cauchy–Schwarz inqeuality, we get

$$\begin{aligned}\|(S-\widetilde{S})z\|_{H^{-1/2}(\Gamma_C)} &= \sup_{0\neq w\in H_{00}^{1/2}(\Gamma_C)} \frac{\langle\nabla(\mathcal{H}-\widetilde{\mathcal{H}})z,\nabla\mathcal{E}w\rangle_\Omega}{\|w\|_{H_{00}^{1/2}(\Gamma_C)}}\\ &\leq \|\nabla(\mathcal{H}-\widetilde{\mathcal{H}})z\|_\alpha \sup_{0\neq w\in H_{00}^{1/2}(\Gamma_C)} \frac{\|\nabla\mathcal{E}w\|_{-\alpha}}{\|w\|_{H_{00}^{1/2}(\Gamma_C)}}.\end{aligned}$$

Function $\mathcal{E}w$ is equal to 0 in a fixed neighborhood of the re-entrant corner and thus

$$\|\nabla\mathcal{E}w\|_{-\alpha}\leq c\|\nabla\mathcal{E}w\|_0.$$

The assertion holds due to Lemma 5.2.3 and the boundedness of the extension operator $\mathcal{E}$. □

5.2.3 Adjoint problem

Now, we introduce the adjoint of the discrete solution operator. Consider the adjoint problem (5.9) and the operator $\mathcal{H}^*\psi=-\partial_{\mathbf{n}}p$ in $H^{-1/2}(\Gamma_C)$.

We define the energy-corrected finite element approximation of the adjoint problem as finding $p_h^m\in Q_h$ such that

$$a_h(p_h^m,q_h)=\langle\psi,q_h\rangle_\Omega\quad\text{for all } q_h\in Q_h.$$

For all $\psi\in L^2_\alpha(\Omega)$ we define the operator $\widetilde{\mathcal{H}}^*:L^2_\alpha(\Omega)\to H^{-1/2}(\Gamma_C)$ by

$$\langle\widetilde{\mathcal{H}}^*\psi,w\rangle_{\Gamma_C}=-a_h(p_h^m,R_h^m\mathcal{E}w)+\langle\psi,R_h^m\mathcal{E}w\rangle_\Omega\quad\text{for all } w\in H_{00}^{1/2}(\Gamma_C).\tag{5.26}$$

In particular we obtain by the definitions above for all $\psi\in L^2_\alpha(\Omega)$, $z\in H_{00}^{1/2}(\Gamma_C)$

$$\begin{aligned}\langle\widetilde{\mathcal{H}}^*\psi,z\rangle_{\Gamma_C} &= -a_h(p_h^m,R_h^m\mathcal{E}z)+\langle\psi,R_h^m\mathcal{E}z\rangle_\Omega\\ &= a_h(u_{0,h}^m,p_h^m)+\langle\psi,R_h^m\mathcal{E}z\rangle_\Omega\\ &= \langle\psi,u_{0,h}^m\rangle_\Omega+\langle\psi,R_h^m\mathcal{E}w\rangle_\Omega\\ &= \langle\psi,\widetilde{\mathcal{H}}z\rangle_\Omega,\end{aligned}$$

and thus $\widetilde{\mathcal{H}}^*$ is indeed the adjoint discrete solution operator.

According to Theorem 2.3.27 the following lemma holds.

Lemma 5.2.5. *Let $\psi\in L^2_{-\alpha}(\Omega)$, and let $p\in H_0^1(\Omega)$ and $p_h^m\in Q_h$ be solutions of the adjoint problem and its energy-corrected finite element approximation, respectively. Then, the following error estimates hold for some $c>0$ independent of ψ and h*

$$\|p-p_h^m\|_\alpha\leq ch^2\|\psi\|_{-\alpha},\quad\textit{and}\quad\|\nabla(p-p_h^m)\|_\alpha\leq ch\|\psi\|_{-\alpha}.$$

We begin the analysis of the discrete adjoint operator by showing an approximation result for the energy-corrected finite element in the limited regularity setting. In Theorem 2.3.27, we assumed that the solution satisfies $u \in H^2_\alpha(\Omega)$ for the piecewise linear finite element. Now, we will show what happens, if only $H^1_{-\alpha}(\Omega)$ regularity is assumed.

Lemma 5.2.6. *Let $w \in H^1_{-\alpha}(\Omega)$. Then the following error estimate holds for some $c > 0$ independent of w and h*

$$\|w - R_h^m w\|_\alpha \leq ch\|w\|_{1,-\alpha}.$$

Proof. Consider an adjoint problem

$$-\Delta\xi = r^{2\alpha}\big(w - R_h^m w\big) \text{ in } \Omega, \quad \text{and} \quad \xi = 0 \text{ on } \Gamma.$$

Let $\xi_h^m \in Q_h$ be the energy-corrected finite element approximation of ξ. Using the modified Galerkin orthogonality, see Theorem 2.3.25, we obtain

$$\begin{aligned}\|w - R_h^m w\|_\alpha^2 &= a(w - R_h^m w, \xi)\\ &= a(w - R_h^m w, \xi - \xi_h^m) - c_h(R_h^m w, \xi_h^m)\\ &= a(w, \xi - \xi_h^m).\end{aligned}$$

Now, applying the Cauchy–Schwarz inquality and the results of Theorem 2.3.27, we see that

$$\begin{aligned}\|w - R_h^m w\|_\alpha^2 &\leq \|w\|_{1,-\alpha}\|\xi - \xi_h^m\|_{1,\alpha}\\ &\leq ch\|w\|_{1,-\alpha}\big\|r^{2\alpha}\big(w - R_h^m w\big)\big\|_{-\alpha}\\ &= h\|w\|_{1,-\alpha}\big\|w - R_h^m w\big\|_\alpha\end{aligned}$$

and hence, the assertion follows. □

Finally, we can state the result describing the approximation accuracy of the adjoint discrete operator.

Lemma 5.2.7. *The approximate adjoint solution operator $\widetilde{\mathcal{H}}^*$, defined in* (5.26)*, is bounded. Moreover, for $\psi \in L^2_{-\alpha}(\Omega)$ the following error estimate holds for some $c > 0$ independent of ψ and h*

$$\|(\widetilde{\mathcal{H}}^* - \mathcal{H}^*)\psi\|_{H^{-1/2}(\Gamma_C)} \leq ch\|\psi\|_{-\alpha}.$$

Proof. Since $a_h(p_h^m, R_h^m \mathcal{E} w) = a(p_h^m, \mathcal{E} w)$, we obtain by the definition of the norm in $H^{-1/2}(\Gamma_C)$ and the adjoint (discrete) solution operator the estimate

$$\begin{aligned}
\|(\widetilde{\mathcal{H}}^* - \mathcal{H}^*)\psi\|_{H^{-1/2}(\Gamma_C)} &= \sup_{0 \neq w \in H_{00}^{1/2}(\Gamma_C)} \frac{\langle (\widetilde{\mathcal{H}}^* - \mathcal{H}^*)\psi, w \rangle_{\Gamma_C}}{\|w\|_{H_{00}^{1/2}(\Gamma_C)}} \\
&= \sup_{0 \neq w \in H_{00}^{1/2}(\Gamma_C)} \frac{a(p - p_h^m, \mathcal{E} w) + \langle \psi, R_h^m \mathcal{E} w - \mathcal{E} w \rangle_\Omega}{\|w\|_{H_{00}^{1/2}(\Gamma_C)}} \\
&\leq \|\nabla (p - p_h^m)\|_\alpha \sup_{0 \neq w \in H_{00}^{1/2}(\Gamma_C)} \frac{\|\nabla \mathcal{E} w\|_{-\alpha}}{\|w\|_{H_{00}^{1/2}(\Gamma_C)}} \\
&\quad + \|\psi\|_{-\alpha} \sup_{0 \neq w \in H_{00}^{1/2}(\Gamma_C)} \frac{\|\mathcal{E} w - R_h^m \mathcal{E} w\|_\alpha}{\|w\|_{H_{00}^{1/2}(\Gamma_C)}}.
\end{aligned}$$

From Lemma 5.2.6, we conclude $\|\mathcal{E} w - R_h^m \mathcal{E} w\|_\alpha \leq ch \|\mathcal{E} w\|_1$. Applying Lemma 5.2.5, together with the boundedness of the extension operator $\mathcal{E}$, completes the proof. □

Notice that $H_\alpha^2(\Omega) \hookrightarrow L_{\alpha-2}^2(\Omega) \hookrightarrow L_{-\alpha}^2(\Omega)$, see Theorem 2.1.15, and we thus can set $\psi = u - \overline{u} \in L_{-\alpha}^2(\Omega)$.

5.2.4 Operator estimates

The operator T_ρ and g in (5.22) cannot be represented exactly in the implementation as it would require a precise implementation of the operators $\mathcal{H}, \mathcal{H}^*$ and the Steklov–Poincaré operator S. Consequently, we introduce suitable approximations. Let us denote by $\widetilde{T}_\rho$ and $\tilde{g}$ approximations of T_ρ and g given in (5.11) defined as

$$\widetilde{T}_\rho = \widetilde{\mathcal{H}}^* \widetilde{\mathcal{H}} + \rho \widetilde{S}, \quad \tilde{g} = \widetilde{\mathcal{H}}^* (\overline{u} - u_{f,h}^m) \quad \text{in } H^{-1/2}(\Gamma_C).$$

Notice that the energy-correction method is included in the definitions of $\widetilde{T}_\rho$ and $\tilde{g}$. We now move to the definition of the fully-discrete solution of the optimal control problem.

Definition 5.2.8 (Discrete optimal control)**.** *We define the finite element approximation of the optimal control problem* (5.12) *as finding* $\tilde{z}_h \in \mathcal{U}_{\mathrm{ad},h}$ *such that*

$$\langle \widetilde{T}_\rho \tilde{z}_h, w_h - \tilde{z}_h \rangle_{\Gamma_C} \geq \langle \tilde{g}, w_h - \tilde{z}_h \rangle_{\Gamma_C} \quad \textit{for all } w_h \in \mathcal{U}_{\mathrm{ad},h}. \tag{5.27}$$

We first study the properties of the discrete operator $\widetilde{T}_\rho$. We begin by showing that it is elliptic and bounded. Similar properties also hold for the continuous counterpart of the operator, see Lemma 5.1.5.

Lemma 5.2.9. *The operator $\widetilde{T}_\rho$ is Z_h-elliptic and bounded, namely, for some $c > 0$ independent of h we have*

$$\langle \widetilde{T}_\rho z_h, z_h \rangle_{\Gamma_C} \geq c|z_h|^2_{H^{1/2}_{00}(\Gamma_C)}, \quad \|\widetilde{T}_\rho z_h\|_{H^{-1/2}(\Gamma_C)} \leq c|z_h|_{H^{1/2}_{00}(\Gamma_C)},$$

for all $z_h \in Z_h$.

Proof. Boundedness follows straight from the boundedness of the discrete operators $\widetilde{\mathcal{H}}$, $\widetilde{\mathcal{H}}^*$ and $\widetilde{S}$.

Recall, $\langle \widetilde{\mathcal{H}}^* \psi, z_h \rangle_{\Gamma_C} = \langle \psi, \widetilde{\mathcal{H}} z_h \rangle_\Omega$, thus we obtain

$$\langle \widetilde{T}_\rho z_h, z_h \rangle_{\Gamma_C} = \|\widetilde{\mathcal{H}} z_h\|^2_0 + \rho \langle \widetilde{S} z_h, z_h \rangle_{\Gamma_C} \quad \text{for all } z_h \in Z_h.$$

The assertion follows then from the ellipticity of the approximative Steklov-Poincaré operator and ellipticity of the operator $\widetilde{\mathcal{H}}$. □

Existence of a unique solution $\tilde{z}_h \in \mathcal{U}_{\mathrm{ad},h}$ of (5.27) follows from [110, Theorem 1.2]. The following error estimate is valid.

Lemma 5.2.10. *Let $z \in H^{1/2}_{00}(\Gamma_C) \cap H^{3/2}_{\mathrm{pw}}(\Gamma_C)$. Furthermore, let $z_h, \tilde{z}_h \in \mathcal{U}_{\mathrm{ad},h}$ be the solutions of (5.22) and (5.27) respectively. Then the following error estimate holds for some $c > 0$ independent of z and z_h*

$$\|z - \tilde{z}_h\|_{H^{1/2}_{00}(\Gamma_C)} \leq c\|z - z_h\|_{H^{1/2}_{00}(\Gamma_C)} + ch\Big(\|z\|_{H^{3/2}_{\mathrm{pw}}(\Gamma_C)} + \|\overline{u}\|_{-\alpha} + \|f\|_{-\alpha}\Big).$$

Proof. Due to the ellipticity and boundedness of the operator $\widetilde{T}_\rho$ and the definitions of variational problems (5.12) and (5.27), we obtain

$$\begin{aligned}
\|z_h - \tilde{z}_h\|^2_{H^{1/2}_{00}(\Gamma_C)} &\leq c\langle \widetilde{T}_\rho(z_h - \tilde{z}_h), z_h - \tilde{z}_h \rangle_{\Gamma_C} \\
&= c\Big(\langle \widetilde{T}_\rho z_h, z_h - \tilde{z}_h \rangle_{\Gamma_C} - \langle \widetilde{T}_\rho \tilde{z}_h, z_h - \tilde{z}_h \rangle_{\Gamma_C} \\
&\qquad + \langle T_\rho z_h, \tilde{z}_h - z_h \rangle_{\Gamma_C} - \langle T_\rho z_h, \tilde{z}_h - z_h \rangle_{\Gamma_C}\Big) \\
&\leq c\Big(\langle \widetilde{T}_\rho z_h, z_h - \tilde{z}_h \rangle_{\Gamma_C} + \langle g - \tilde{g}, z_h - \tilde{z}_h \rangle_{\Gamma_C} + \langle T_\rho z_h, \tilde{z}_h - z_h \rangle_{\Gamma_C}\Big) \\
&\leq c\Big(\|(T_\rho - \widetilde{T}_\rho) z_h\|_{H^{-1/2}(\Gamma_C)} + \|g - \tilde{g}\|_{H^{-1/2}(\Gamma_C)}\Big)\|z_h - \tilde{z}_h\|_{H^{1/2}_{00}(\Gamma_C)}.
\end{aligned}$$

Application of the triangle inequality yields

$$\begin{aligned}
\|z - \tilde{z}_h\|_{H^{1/2}_{00}(\Gamma_C)} \leq c\Big(&\|z - z_h\|_{H^{1/2}_{00}(\Gamma_C)} + \big\|\big(T_\rho - \widetilde{T}_\rho\big) z\big\|_{H^{-1/2}(\Gamma_C)} \\
&+ \|g - \tilde{g}\|_{H^{-1/2}(\Gamma_C)}\Big).
\end{aligned} \tag{5.28}$$

Now, using the triangle inequality and boundedness of the operator $\mathcal{H}^* : L^2_\alpha(\Omega) \to H^{-1/2}(\Gamma_{\mathrm{C}})$, we get

$$\begin{aligned}\|(T_\rho - \widetilde{T}_\rho)z\|_{H^{-1/2}(\Gamma_{\mathrm{C}})} &= \|(\mathcal{H}^*\mathcal{H} - \widetilde{\mathcal{H}}^*\widetilde{\mathcal{H}})z + \rho(S - \widetilde{S})z\|_{H^{-1/2}(\Gamma_{\mathrm{C}})} \\ &\leq \|(\mathcal{H}^* - \widetilde{\mathcal{H}}^*)\mathcal{H}z\|_{H^{-1/2}(\Gamma_{\mathrm{C}})} + \|\widetilde{H}^*(\mathcal{H} - \widetilde{\mathcal{H}})z\|_{H^{-1/2}(\Gamma_{\mathrm{C}})} \\ &\quad + \rho\|(S - \widetilde{S})z\|_{H^{-1/2}(\Gamma_{\mathrm{C}})} \\ &\leq \|(\mathcal{H}^* - \widetilde{\mathcal{H}}^*)\mathcal{H}z\|_{H^{-1/2}(\Gamma_{\mathrm{C}})} + c\|(\mathcal{H} - \widetilde{\mathcal{H}})z\|_\alpha \\ &\qquad + \rho\|(S - \widetilde{S})z\|_{H^{-1/2}(\Gamma_{\mathrm{C}})},\end{aligned}$$

where we used the fact that $u_z = \mathcal{H}z \in L^2_{-\alpha}(\Omega)$. Applying Lemma 5.2.3 together with Lemma 5.2.4 and Lemma 5.2.7, we have

$$\|(T_\rho - \widetilde{T}_\rho)z\|_{H^{-1/2}(\Gamma_{\mathrm{C}})} \leq ch\|z\|_{H^{3/2}_{\mathrm{pw}}(\Gamma_{\mathrm{C}})}. \tag{5.29}$$

Using a similar argument, we also see that

$$\begin{aligned}\|g - \tilde{g}\|_{H^{-1/2}(\Gamma_{\mathrm{C}})} &= \|\mathcal{H}^*(\overline{u} - u_f) - \widetilde{\mathcal{H}}^*(\overline{u} - u^m_{f,h})\|_{H^{-1/2}(\Gamma_{\mathrm{C}})} \\ &\leq \|(\mathcal{H}^* - \widetilde{\mathcal{H}}^*)\overline{u}\|_{H^{-1/2}(\Gamma_{\mathrm{C}})} + \|\widetilde{\mathcal{H}}^*(u_f - u^m_{f,h})\|_{H^{-1/2}(\Gamma_{\mathrm{C}})} \\ &\quad + \|(\mathcal{H}^* - \widetilde{\mathcal{H}}^*)u_f\|_{H^{-1/2}(\Gamma_{\mathrm{C}})} \\ &\leq \|(\mathcal{H}^* - \widetilde{\mathcal{H}}^*)\overline{u}\|_{H^{-1/2}(\Gamma_{\mathrm{C}})} + c\|u_f - u^m_{f,h}\|_\alpha \\ &\qquad + \|(\mathcal{H}^* - \widetilde{\mathcal{H}}^*)u_f\|_{H^{-1/2}(\Gamma_{\mathrm{C}})}.\end{aligned}$$

Application of (5.24) and Lemma 5.2.7 yields

$$\|g - \tilde{g}\|_{H^{-1/2}(\Gamma_{\mathrm{C}})} \leq c\Big(h^2\|f\|_{-\alpha} + h\big(\|\overline{u}\|_{-\alpha} + \|f\|_{-\alpha}\big)\Big). \tag{5.30}$$

Finally, combining the results of (5.29) and (5.30) with (5.28) completes the proof. □

5.2.5 Finite element approximation accuracy

After introducing technical approximation results in previous sections, we are finally in the position to focus on the main result of this chapter.

Theorem 5.2.11. *Let $z_a, z_b \in H^{3/2}_{\mathrm{pw}}(\Gamma_C)$, $z_a < z_b$ and $f, \overline{u} \in L^2_{-\alpha}(\Omega)$ be given. Furthermore, let $z \in H^{1/2}_{00}(\Gamma_{\mathrm{C}}) \cap H^{3/2}_{\mathrm{pw}}(\Gamma_C)$ and $\tilde{z}_h \in \mathcal{U}_{\mathrm{ad},h}$ be the unique solutions of (5.12) and (5.27), respectively. Then the following error estimates hold for some $c(\overline{u}, f, z_a, z_b) > 0$ independent of h*

$$\|z - \tilde{z}_h\|_{H^{1/2}_{00}(\Gamma_{\mathrm{C}})} \leq c(\overline{u}, f, z_a, z_b)h. \tag{5.31}$$

Moreover, let $(u,p) \in (H_0^1(\Omega,\Gamma_D)\cap H_\alpha^2(\Omega))\times(H_0^1(\Omega)\cap H_\alpha^2(\Omega))$ *be the respective solutions of primal problem* (5.3) *and dual problem* (5.9)*. Then for the energy-corrected finite element approximations also* $(u_h^m, p_h^m) \in V_h \times Q_h$ *the following error estimates hold*

$$\|u - u_h^m\|_\alpha \leq c(\overline{u}, f, z_a, z_b)h^2, \tag{5.32}$$

and

$$\|p - p_h^m\|_\alpha \leq c(\overline{u}, f, z_a, z_b)h^2. \tag{5.33}$$

Proof. The error estimate (5.31) in the energy norm $H_{00}^{1/2}(\Gamma_C)$ is a consequence of Lemma 5.2.10 together with Theorem 5.2.1.

The estimate for the primal state in (5.32) is a consequences of inequality (5.24) and Lemma 5.2.3. Finally, the estimate for the adjoint state in (5.33) was already presented in Lemma 5.2.5. □

Theorem 5.2.11 shows that the convergence of the energy-corrected finite element discretisation of the Dirichlet control problem yields optimal convergence in terms of the interpolation error summarised in Theorem 5.1.10. Moreover, the application of the energy-correction results in better convergence properties of the finite element method on non-convex domains, when measured in $H^{1/2}(\Gamma_C)$ norm, see Theorem 5.1.13 for comparison. Similarly, the state variable u and the adjoint state p are also approximated with the optimal order, however, in weighted norms.

Remark 5.2.12. *An additional application of the post-processing approach described in Section 2.3.3.2 further improves the approximation of the primal and dual states. In particular, it yields second-order convergence in the standard* $L^2(\Omega)$ *norm and the first order convergence in the standard* $H^1(\Omega)$ *norm.*

5.2.6 Primal-dual active set strategy

We shortly recall the use of the primal-dual active set strategy for the solution of the discrete variational inequality, see for instance [80, 81]. Let $n_C = \dim Z_h$, then there exist canonical isomorphisms $\boldsymbol{z} \in \mathbb{R}^{n_C} \leftrightarrow z_h \in Z_h$ and $\widetilde{\boldsymbol{g}} \in \mathbb{R}^{n_C} \leftrightarrow \widetilde{g} \in Z_h$, where $\boldsymbol{z}$ and $\widetilde{\boldsymbol{g}}$ denote the coefficient vectors of the control and the right-hand side $\tilde{g}$ respectively, when written in the nodal basis introduced in Section 5.1.4. We introduce the system matrix corresponding to the system (5.17)

$$\begin{pmatrix} \mathbf{M}_{II} + \varrho\mathbf{S}^m_{II} & \mathbf{M}_{IC} + \varrho\mathbf{S}^m_{IC} & \mathbf{S}^m_{II} & \\ \mathbf{M}_{CI} + \varrho\mathbf{S}^m_{CI} & \mathbf{M}_{CC} + \varrho\mathbf{S}^m_{CC} & \mathbf{S}^m_{CI} & -\mathbf{I}_{\mathrm{CC}} \\ \mathbf{S}^m_{II} & \mathbf{S}^m_{IC} & & \\ & \mathbf{I}_{\mathcal{A}^k} & & \mathbf{I}_{\mathcal{I}^k} \end{pmatrix}, \tag{5.34}$$

where we separate the degrees of freedom in the interior of the domain Ω and on the control boundary Γ_{C}, respectively indicated by I and C. Here,

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}_{II} & \mathbf{M}_{IC} \\ \mathbf{M}_{CI} & \mathbf{M}_{CC} \end{pmatrix} \quad \text{and} \quad \mathbf{S}^m = \begin{pmatrix} \mathbf{S}^m_{II} & \mathbf{S}^m_{IC} \\ \mathbf{S}^m_{CI} & \mathbf{S}^m_{CC} \end{pmatrix},$$

denote the finite element matrices which correspond to the mass matrix and the stiffness matrix introduced in (2.12), where for the latter the energy-correction scheme is applied. Furthermore, $\mathbf{I}_{\mathcal{A}^k}$, $\mathbf{I}_{\mathcal{I}^k}$ are diagonal matrices, with entry 1, if the corresponding degree of freedom belongs to one of the active sets, or with entry 0 if it belongs to the inactive set. Finally, $\boldsymbol{\chi}_{\mathcal{A}^k}(\mathbf{z}_a, \mathbf{z}_b)$ represents the components of $\mathbf{z}_a$ and $\mathbf{z}_b$ corresponding to the active sets $\mathcal{A}^k_a$ and $\mathcal{A}^k_b$, respectively.

The discrete variational inequality (5.27) in matrix-vector notation then reads

$$(\widetilde{\mathbf{T}}_\varrho \boldsymbol{z} - \tilde{\boldsymbol{g}}, \boldsymbol{w} - \boldsymbol{z}) \geq 0 \quad \text{for all } \boldsymbol{w} \in \mathbb{R}^{n_{\mathrm{C}}} \leftrightarrow w_h \in \mathcal{U}_{\mathrm{ad},h}.$$

Here, $\widetilde{\mathbf{T}}_\varrho \in \mathbb{R}^{n_C \times n_C}$ denotes the Schur complement matrix of the system matrix (5.34) with respect to the control boundary, see [127, Section 3.5]. For the solution of the variational inequality, we apply a standard semi-smooth Newton method, see [25, 80], which results in the primal-dual active set strategy. Therefore, we introduce the following discrete Lagrange multiplier

$$\boldsymbol{\lambda} = \widetilde{\mathbf{T}}_\varrho \boldsymbol{z} - \tilde{\boldsymbol{g}} \in \mathbb{R}^{n_{\mathrm{C}}}.$$

Let us denote the mesh nodes of $\mathcal{T}_h$ on the control boundary Γ_{C} by x_i, $i = 1, \ldots, n_{\mathrm{C}}$. Following [150] we introduce for the k-th iteration step of the primal-dual active set strategy and some positive constant $c \in \mathbb{R}_+$ (we choose $c = 1/10$) the inactive set

$$\mathcal{I}^k = \left\{ i \,:\, \lambda^k_i + c\,[z_a(x_i) - z^k_i] \leq 0 \right\} \cup \left\{ i \,:\, \lambda^k_i + c\,[z_b(x_i) - z^k_i] \geq 0 \right\},$$

and active sets

$$\mathcal{A}^k_a = \left\{ i \,:\, \lambda^k_i + c\,[z_a(x_i) - z^k_i] > 0 \right\}, \quad \mathcal{A}^k_b = \left\{ i \,:\, \lambda^k_i + c\,[z_b(x_i) - z^k_i] < 0 \right\}.$$

Algorithm 1 primal-dual active set strategy

1: Initialize $\mathbf{z}^0$, $\boldsymbol{\lambda}^0$
2: **for** $k \in \{0, \dots, k_{\max}\}$ **do**
3: Set active and inactive sets $\mathcal{A}_a^k$, $\mathcal{A}_b^k$ and $\mathcal{I}^k$
4: Solve

$$\widetilde{\mathbf{T}}_\varrho \boldsymbol{z}^{k+1} - \boldsymbol{\lambda}^{k+1} = \widetilde{\boldsymbol{g}}$$
$$\boldsymbol{z}^{k+1} = \boldsymbol{z}_j \ \text{ on } \mathcal{A}_j^k,\ j \in \{a, b\} \quad \text{and} \quad \boldsymbol{\lambda}^{k+1} = \mathbf{0} \ \text{ on } \mathcal{I}^k.$$

5: **if** $|\mathbf{z}^{k+1} - \mathbf{z}^k| + |\boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^k| < \epsilon$ **then**
6: Stop
7: **end if**
8: **end for**

Moreover, we denote the nodal interpolation vector of z_a and z_b by $\boldsymbol{z}_a$ and $\boldsymbol{z}_b \in \mathbb{R}^{n_C}$, respectively. For a given accuracy $\epsilon > 0$ and maximal iterations $k_{\max} \in \mathbb{N}$ the primal-dual active set strategy is summarized in Algorithm 1.

For the implementation, it is impractical to realize the Schur complement $\widetilde{T}_\varrho$ in step (4) of Algorithm 1. Thus, we replace it by the equivalent system matrix (5.34). The algebraic system reads then

$$\begin{pmatrix} \mathbf{M}_{II} + \varrho \mathbf{S}_{II}^m & \mathbf{M}_{IC} + \varrho \mathbf{S}_{IC}^m & \mathbf{S}_{II}^m & \\ \mathbf{M}_{CI} + \varrho \mathbf{S}_{CI}^m & \mathbf{M}_{CC} + \varrho \mathbf{S}_{CC}^m & \mathbf{S}_{CI}^m & -\mathbf{I}_{\mathrm{CC}} \\ \mathbf{S}_{II}^m & \mathbf{S}_{IC}^m & & \\ & \mathbf{I}_{\mathcal{A}^k} & & \mathbf{I}_{\mathcal{I}^k} \end{pmatrix} \begin{pmatrix} \boldsymbol{u}_I^{k+1} \\ \boldsymbol{z}^{k+1} \\ \boldsymbol{p}_I^{k+1} \\ \boldsymbol{\lambda}^{k+1} \end{pmatrix} = \begin{pmatrix} \overline{\boldsymbol{w}}_I \\ \overline{\boldsymbol{w}}_C \\ \mathbf{0} \\ \boldsymbol{\chi}_{\mathcal{A}^k}(\boldsymbol{z}_a, \boldsymbol{z}_b) \end{pmatrix},$$

where $\boldsymbol{u}_I^{k+1},\ \boldsymbol{p}_I^{k+1} \in \mathbb{R}^{n_I}$, $n_I = \dim Q_h$, denote the coefficient vectors of primal- and adjoint state, respectively.

5.3 Numerical results

In this section, we present numerical results illustrating the benefits of the energy-corrected finite element method. We consider the optimal control problem (5.2)–(5.4) in a square domain with a hole inside as presented in Section 5.1, see Figure 5.1, with maximal interior angles $\Theta = 3\pi/2$. Again, we choose $\alpha = 1/3$ for the weighted norms, with weights applied in the neighbourhood of each corner. This choice of α induces a slightly stronger norm than the one required in the analysis presented in Theorem 5.2.11. The desired state is given by

$$\overline{u} = \left(x_1^2 + x_2^2\right)^{-1/4} + \left((x_1+1)^2 + (x_2-1)^2\right)^{-1/4},$$

which satisfies $\overline{u} \in L^2_{-\vec{\alpha}}(\Omega)$. Furthermore, we choose $f = 0$ and $\rho = 1$ for simplicity. In this example, the desired state is less regular than in the previous one, presented in Section 5.1 since the singularities are placed at the re-entrant corner and on the control boundary Γ_C.

In the following, we consider two cases: numerical results with and without box-constraints, where we choose $z_a = -1$ and $z_b = 1$ in the latter. Initially, we consider a uniform mesh consisting of 32 triangles, locally symmetric around the re-entrant corners, which are then uniformly refined. Thus, after $L \in \mathbb{Z}_+$ refinements the mesh size is $h \sim 2^{-L}$. Since the exact solution is not known for the considered problem, we shall use the solution on refinement level $L = 9$ as the reference solution. Moreover, for the energy-corrected scheme, we use the asymptotic correction parameter γ^* obtained using semi-smooth Newton iterations discussed in Section 2.3.3.3. In the presence of the box-constraints, a standard primal-dual set strategy is used, see Section 5.2.6, with the accuracy $\epsilon = 10^{-10}$.

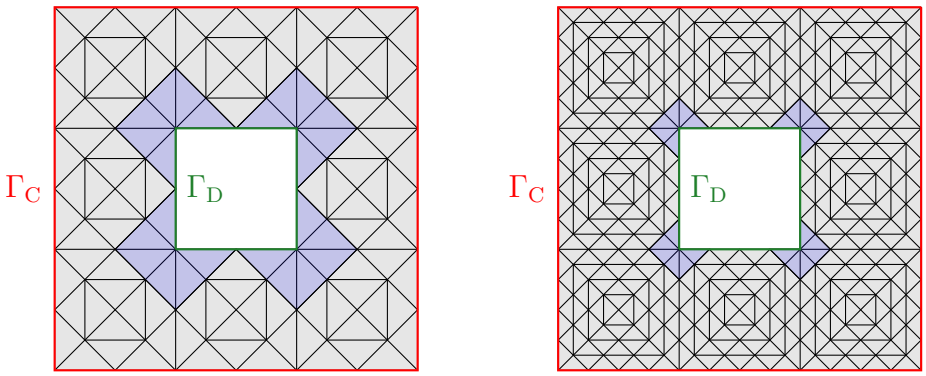


Figure 5.2: Computational domain with a once and twice refined initial mesh, illustration of the four correction patches (blue).

$\gamma = 0$						
L	$\|u_{h_9} - u_h\|_0$	eoc	$\|u_{h_9} - u_h\|_\alpha$	eoc	$\|z_{h_9} - z_h\|_{L^2(\Gamma_C)}$	eoc
1	3.2927e–01	–	5.1895e–01	–	4.0161e–01	–
2	1.2976e–01	1.34	1.9188e–01	1.44	1.5321e–01	1.39
3	5.0726e–02	1.36	7.2805e–02	1.40	5.8847e–02	1.38
4	1.9786e–02	1.36	2.8032e–02	1.38	2.2747e–02	1.37
5	7.6677e–03	1.37	1.0795e–02	1.38	8.7755e–03	1.37
6	2.9081e–03	1.40	4.0801e–03	1.40	3.3204e–03	1.40
7	1.0333e–03	1.49	1.4465e–03	1.50	1.1780e–03	1.50
Expected		1.33		1.33		1.33
$\gamma = 0.1382555154$						
L	$\|u^m_{h_9} - u^m_h\|_0$	eoc	$\|u^m_{h_9} - u^m_h\|_\alpha$	eoc	$\|z^m_{h_9} - z^m_h\|_{L^2(\Gamma_C)}$	eoc
1	2.1616e–01	–	3.2113e–01	–	2.3312e–01	–
2	5.4932e–02	1.98	6.5106e–02	2.30	5.5045e–02	2.08
3	1.6008e–02	1.78	1.4921e–02	2.13	1.3680e–02	2.01
4	4.8689e–03	1.72	3.5894e–03	2.06	3.4749e–03	1.98
5	1.5098e–03	1.69	8.9348e–04	2.01	9.0514e–04	1.94
6	4.7111e–04	1.68	2.2721e–04	1.98	2.3976e–04	1.92
7	1.4540e–04	1.70	5.7702e–05	1.98	6.2573e–05	1.94
Expected		1.66		2.00		1.50

Table 5.2: Errors and estimated convergence orders for the uncorrected and corrected method, without box-constraints.

$\gamma = 0$							
L	I	I^*	$\|u_{h_9} - u_h\|_0$	eoc	$\|u_{h_9} - u_h\|_\alpha$	eoc	$\|z_{h_9} - z_h\|_{L^2(\Gamma_C)}$	eoc
1	3	3	2.6854e–01	–	4.0506e–01	–	2.9459e–01	–
2	4	3	1.1077e–01	1.28	1.5584e–01	1.38	1.2061e–01	1.29
3	5	3	4.1695e–02	1.41	5.6803e–02	1.46	4.4153e–02	1.45
4	5	2	1.6337e–02	1.35	2.1981e–02	1.37	1.7271e–02	1.35
5	7	2	6.2750e–03	1.38	8.3790e–03	1.39	6.5879e–03	1.39
6	8	4	2.3797e–03	1.40	3.1665e–03	1.40	2.4990e–03	1.40
7	9	4	8.4602e–04	1.49	1.1233e–03	1.50	8.8878e–04	1.49
Expected				1.33		1.33		1.33
$\gamma = 0.1382555154$								
L	I	I^*	$\|u^m_{h_9} - u^m_h\|_0$	eoc	$\|u^m_{h_9} - u^m_h\|_\alpha$	eoc	$\|z^m_{h_9} - z^m_h\|_{L^2(\Gamma_C)}$	eoc
1	3	3	1.8026e–01	–	2.5346e–01	–	1.6928e–01	–
2	4	3	5.0371e–02	1.84	5.4724e–02	2.21	4.2410e–02	2.00
3	5	3	1.4646e–02	1.78	1.2164e–02	2.17	1.0091e–02	2.07
4	5	2	4.5100e–03	1.70	2.9774e–03	2.03	2.7256e–03	1.89
5	7	2	1.3952e–03	1.69	6.9192e–04	2.11	5.4485e–04	2.32
6	8	4	4.3743e–04	1.67	1.7295e–04	2.00	1.4431e–04	1.92
7	9	4	1.3548e–04	1.69	4.3587e–05	1.99	4.0949e–05	1.82
Expected				1.66		2.00		1.50

Table 5.3: Semi-smooth Newton iterations (I) and (I^*) for initial guess equal to 0 and equal to the approximation obtained at the previous refinement level, errors and estimated convergence orders for the uncorrected and corrected method, with box-constraints.

In Table 5.2 and Table 5.3, we present the errors and convergence rates of the finite element approximations with and without the application of the energy-correction.

Note that the optimal second order convergence for the primal state u is only recovered in weighted spaces, see Table 5.2. Also, the energy-corrected scheme gives better results than the uncorrected one. The pollution effect, as in Theorem 2.3.21, is eliminated, when the energy-corrected scheme is used. Thus, optimal approximation order compared to the best approximation in the L^2-norm is recovered. Also in the presence of box-constraints, the optimal convergence rates, compared to the interpolation error stated in Theorem 5.1.10 are observed for the energy-corrected finite element method.

In Table 5.3, we also include the number of semi-smooth Newton iterations, denoted by I, when the initial guess is set to 0 everywhere at the control boundary and I^*, when the initial guess for level L is set to be the solution obtained at level $L-1$. For the iteration numbers I we observe the typical logarithmic behaviour with respect to the number of degrees of freedom, while in the latter case (I^*) almost robust iteration numbers with respect to the problem size are observed, see [81]. Also, the iteration numbers do not differ for the standard and the energy-corrected finite element approximation.

Chapter 6

Conclusions and outlook

In this thesis, we discussed the applications of the energy-corrected finite element method for problems defined on polygonal domains. We began by introducing the necessary notation and mathematical background, including the regularity results for the elliptic equations in the framework of weighted Sobolev spaces. We also presented the pollution effect, being the suboptimal approximation of the finite element discretisation arising in the presence of corners in the computational domain. We then showed the existing theory of the energy-corrected finite element for the elliptic equations. This method is a common denominator of all the following parts of the thesis. In Chapters 3–5 we discussed the novel results concerning the energy-correction, supporting the theoretical findings with the relevant numerical experiments. The known auxiliary results are given without proofs but are always backed with the precise citations. The proofs always follow new results of our authorship.

We began by extending the known results to the estimates in the weighted $L^\infty(\Omega)$ norms. Using the dyadic decomposition of the computational domain, we showed that the energy-corrected finite element approximation exhibits the optimal pointwise convergence regarding the interpolation error in the whole computational domain. We did not restrict our considerations to the standard piecewise linear discretisation and treated the general case of the piecewise polynomial approximation spaces.

The pollution effect is known to diminish the approximation properties of the finite element methods also when parabolic problems are concerned. We showed that the energy-correction method yields the optimal discretisation error of such problems in the spatial dimensions. Moreover, since it allows for the use of uniform meshes, it is a feasible choice in combination with the explicit time-stepping schemes, as the CFL stability condition is naturally satisfied. We concluded this analysis with extensive numerical investigations, proposing the post-processing strategy and several possible extensions. Among them, we considered the convection-diffusion problem

with pointwise error estimates and enriched cubic finite element spaces with mass lumping techniques. We concluded the chapter proposing a potential application in the computation of the heat distribution within the three-dimensional geometry of a graphite nuclear moderator brick.

Finally, we investigated the discretisation error of the Dirichlet optimal control problem on non-convex polygonal domains using the problem formulation involving the regularisation in the energy $H_{00}^{1/2}(\Gamma_C)$ space. We proved that the energy-corrected approximation converges optimally in weighted Sobolev spaces. We also introduced an iterative solver based on the primal-dual active set strategy, convenient for the implementation.

The energy-corrected finite element is a relatively new method, first considered in [69], and a lot of questions regarding it remain open. So far, we restricted our interest to two-dimensional problems only. The derivation of an analogous method for the three-dimensional polyhedra is a necessary, however a challenging problem. The main difficulty posed in three dimensions is the presence of potentially two different types of singularities, introduced both by edges and corners. A known benchmark for this type of problems is the so-called Fichera corner, which is a domain created upon subtracting a small cube from a larger one, namely $\Omega = (-1,1)^3 \setminus [0,1]^3$. Mesh grading techniques for such problems were studied in [2, Section 4.4], and the analysis presented there could serve as the first step towards derivation of the corresponding estimates for the energy-corrected finite element.

Here, we began our investigations with the Poisson problem. An important extension would be a similar treatment of semi-linear and nonlinear elliptic equations. Also, different kinds of problems such as ones arising in continuum mechanics would be of interest. Optimal convergence results for the energy-corrected discretisation in weighted Sobolev spaces for the Stokes equation are already known [91]. Singular behaviour in the vicinity of corners in polygonal domains also arises in many other problems in fluid dynamis [24, 129]. We believe that the analysis of parabolic problems presented here, in combination with the energy-corrected discretisation for the Stokes equation, could serve as the first step towards the development of numerical solvers for the time-dependent flows.

Also, many interesting questions regarding the energy-corrected approximations of optimal control problems remain open. Here, we assumed that the control is located in some positive distance away from the re-entrant corners of the domain. It would be interesting to extend the analysis to a more general case of the control supported on the whole boundary of an arbitrary polygon. This, however, introduces

additional difficulties such as the necessity of using weighted Sobolev spaces on the boundary. Furthermore, the theory of the energy-corrected finite element would have to be extended to cover the problems with nonhomogeneous boundary conditions around the corners.

Furthermore, we observed that the energy-corrected scheme yields second order convergence on the boundary, which is half-an-order better than predicted by Theorem 5.2.11. This behaviour was already observed in [127] in the case of convex polygons, which we also confirmed numerically. The gap between the numerical and theoretical results can, in our opinion, be attributed to the higher regularity properties of the solution used in the experiments. We believe that the investigation of this phenomenon would benefit from the techniques and results presented in [114]. The first step in this direction was recently performed in [167].

Finally, different kinds of optimal control problems could be considered, such as having the distributed control inside the domain or controlling the Neumann boundary conditions of the state equation. These types of control could be combined with different kinds of regularisation, for example, $H^{-1/2}(\Gamma)$ energy regularisation for the Neumann control, as in [10]. Moreover, different types of state equations would be of interest among them being the Stokes equation describing the stationary flow with dominating viscous forces or the heat equation.

Bibliography

[1] R. A. Adams and J. J. F. Fournier. *Sobolev Spaces*. Academic Press, New York, London, 2003.

[2] T. Apel. *Anisotropic Finite Elements: Local Estimates and Applications*. Advances in Numerical Mathematics. Vieweg+Teubner Verlag, 1999.

[3] T. Apel. Interpolation in h-Version Finite Element Spaces. In *Encyclopedia of Computational Mechanics*, chapter 3. American Cancer Society, 2004.

[4] T. Apel, M. Mateos, J. Pfefferer, and A. Rösch. On the regularity of the solutions of Dirichlet optimal control problems in polygonal domains. *SIAM Journal on Control and Optimization*, 53(6):3620–3641, 2015.

[5] T. Apel and S. Nicaise. The finite element method with anisotropic mesh grading for elliptic problems in domains with corners and edges. *Mathematical Methods in the Applied Sciences*, 21(6):519–549, 1998.

[6] T. Apel, S. Nicaise, and J. Schöberl. A non-conforming finite element method with anisotropic mesh grading for the Stokes problem in domains with edges. *IMA Journal on Numerical Analysis*, 21(4):843–856, 2001.

[7] T. Apel, J. Pfefferer, S. Rogovs, and M. Winkler. L^∞-error estimates for Neumann boundary value problems on graded meshes. *ArXiv e-prints*, April 2018.

[8] T. Apel, J. Pfefferer, and A. Rösch. Finite element error estimates on the boundary with application to optimal control. *Mathematics of Computation*, 84(291):33–70, 2015.

[9] T. Apel, A.-M. Sändig, and J. R. Whiteman. Graded mesh refinement and error estimates for finite element solutions of elliptic boundary value problems in non-smooth domains. *Mathematical Methods in the Applied Sciences*, 19(1):63–85, 1996.

[10] T. Apel, O. Steinbach, and M. Winkler. Error estimates for Neumann boundary control problems with energy regularization. *Journal of Numerical Mathematics*, 24(4):207–233, 2016.

[11] J.D. Arregui Mena, L. Margetts, L. Evans, D.V. Griffiths, A. Shterenlikht, L. Cebamanos, and P.M. Mummery. The stochastic finite element method for nuclear applications. *ECCOMAS Congress 2016*, 2016.

[12] H. Atamni, M. El Hatri, and N. Popivanov. Polynomial approximation in weighted Sobolev space. *Comptes Rendus de l'Academie Bulgare des Sciences*, 54(3):3–25, 2001.

[13] I. Babuška. Finite element method for domains with corners. *Computing*, 6:264–273, 1970.

[14] I. Babuška, R. B. Kellogg, and J. Pitkäranta. Direct and inverse error estimates for finite elements with mesh refinements. *Numerische Mathematik*, 33(4):447–471, 1979.

[15] I. Babuška and M.B. Rosenzweig. A finite element scheme for domains with corners. *Numerische Mathematik*, 20(1):1–21, Feb 1972.

[16] I. Babuška and T. Strouboulis. *The finite element method and its reliability*. Numerical Mathematics and Scientific Computation. The Clarendon Press, Oxford University Press, New York, 2001.

[17] C. Bacuta, J.H. Bramble, and J. Xu. Regularity estimates for elliptic boundary value problems in Besov spaces. *Mathematics of Computation*, 72(244):1577–1595, 2002.

[18] C. Bacuta, V. Nistor, and L.T. Zikatanov. Improving the rate of convergence of high-order finite elements on polyhedra i: A priori estimates. *Numerical Functional Analysis and Optimization*, 26(6):613–639, 2005.

[19] C. Bacuta, V. Nistor, and L.T. Zikatanov. Improving the rate of convergence of high-order finite elements on polyhedra ii: Mesh refinements and interpolation. *Numerical Functional Analysis and Optimization*, 28(7-8):775–824, 2007.

[20] G.A. Baker, J.H. Bramble, and V. Thomée. Single step Galerkin approximations for parabolic problems. *Mathematics of Computation*, 31:818–847, 1977.

[21] J. Banasiak and G.F. Roach. On corner singularities of solutions to mixed boundary-value problems for second-order elliptic and parabolic equations. *Proceedings: Mathematical and Physical Sciences*, 433(1887):209–217, 1991.

[22] C. Bardos. A regularity theorem for parabolic equations. *Journal of Functional Analysis*, 7(2):311–322, 1971.

[23] S. Bartels. *Numerical Approximation of Partial Differential Equations*. Texts in Applied Mathematics. Springer International Publishing, 2016.

[24] M. Beneš and P. Kučera. Solutions to the Navier-Stokes Equations with Mixed Boundary Conditions in Two-Dimensional Bounded Domains. *ArXiv e-prints*, September 2014.

[25] M. Bergounioux, K. Ito, and K. Kunisch. Primal-dual strategy for constrained optimal control problems. *SIAM Journal on Control and Optimization*, 37(4):1176–1194 (electronic), 1999.

[26] A. Bermudez, P. Gamallo, and R. Rodriguez. Finite element methods in local active control of sound. *SIAM Journal on Control and Optimization*, 43(2):437–465, 2004.

[27] H. Blum. The influence of reentrant corners in the numerical approximation of viscous flow problems. In *Numerical treatment of the Navier-Stokes equations (Kiel, 1989)*, volume 30 of *Notes Numerical Fluid Mechanics*, pages 37–46. Vieweg, Braunschweig, 1990.

[28] H. Blum and M. Dobrowolski. On finite element methods for elliptic equations on domains with corners. *Computing*, 28(1):53–63, 1982.

[29] H. Blum and R. Rannacher. Extrapolation techniques for reducing the pollution effect of reentrant corners in the finite element method. *Numerische Mathematik*, 52(5):539–564, 1988.

[30] H. Blum and R. Rannacher. Finite element eigenvalue computation on domains with reentrant corners using Richardson extrapolation. *Journal of Computational Mathematics*, 8(4):321–332, 1990.

[31] L. Bonifacius and I. Neitzel. Second order optimality conditions for optimal control of quasilinear parabolic equations. *Mathematical Control and Related Fields*, 8:1, 2018.

[32] D. Braess. *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics*. Cambridge University Press, 3 edition, 2007.

[33] J. Bramble, A. Schatz, V. Thomée, and L. Wahlbin. Some convergence estimates for semidiscrete galerkin type approximations for parabolic equations. *SIAM Journal on Numerical Analysis*, 14(2):218–241, 1977.

[34] J.H. Bramble and S.R. Hilbert. Estimation of linear functionals on Sobolev spaces with application to Fourier transforms and spline interpolation. *SIAM Journal on Numerical Analysis*, 7(1):112–124, 1970.

[35] P. Brenner, M. Crouzeix, and V. Thomée. Single step methods for inhomogeneous linear differential equations in Banach space. *RAIRO. Analyse numérique*, 16(1):5–26, 1982.

[36] S. C. Brenner. Multigrid methods for the computation of singular solutions and stress intensity factors. I. Corner singularities. *Mathematics of Computation*, 68(226):559–583, 1999.

[37] S. C. Brenner, J. Cui, T. Gudi, and L.-Y. Sung. Multigrid algorithms for symmetric discontinuous Galerkin methods on graded meshes. *Numerische Mathematik*, 119(1):21–47, Sep 2011.

[38] S. C. Brenner and L. R. Scott. *The mathematical theory of finite element methods*, volume 15 of *Texts in Applied Mathematics*. Springer, New York, third edition, 2008.

[39] H. Brezis, P. Mironescu, and T. Kato. Gagliardo-Nirenberg, composition and products in fractional Sobolev spaces. *Journal of Evolution Equations*, 1(4):387–404, 2001.

[40] P. Brunner, C. Clason, M. Freiberger, and H. Scharfetter. A deterministic approach to the adapted optode placement for illumination of highly scattering tissue. *Biomedical Optics Express*, 3(7):1732–1743, 2012.

[41] Z. Cai and S. Kim. A finite element method using singular functions for the Poisson equation: corner singularities. *SIAM Journal on Numerical Analysis*, 39:286–299, 2001.

[42] E. Casas and K. Kunisch. Stabilization by sparse controls for a class of semilinear parabolic equations. *SIAM Journal on Control and Optimization*, 55(1):512–532, 2017.

[43] E. Casas, M. Mateos, and F. Tröltzsch. Error estimates for the numerical approximation of boundary semilinear elliptic control problems. *Computational Optimization and Applications. An International Journal*, 31(2):193–219, 2005.

[44] R. Casas and K. Kunisch. Parabolic control problems in space-time measure spaces. *ESAIM: COCV*, 22(2):355–370, 2016.

[45] Y.A. Cengel and A.J. Ghajar. *Heat and Mass Transfer: Fundamentals and Applications, 5th edition.* McGraw-Hill Education, 2014.

[46] P. Chatzipantelidis, R. D. Lazarov, and V. Thomée. Error estimates for a finite volume element method for parabolic equations in convex polygonal domains. *Numerical Methods for Partial Differential Equations*, 20(5):650–674, 2004.

[47] P. Chatzipantelidis, R. D. Lazarov, V. Thomée, and L. B. Wahlbin. Parabolic finite element equations in nonconvex polygonal domains. *BIT Numerical Mathematics*, 46(1):113–143, Nov 2006.

[48] P. Chatzipantelidis, R.D. Lazarov, and V. Thomée. Parabolic finite volume element equations in nonconvex polygonal domains. *Numerical Methods for Partial Differential Equations*, 25(3):507–525, 2009.

[49] L. Chen and H. Li. Superconvergence of gradient recovery schemes on graded meshes for corner singularities. *Journal of Computational Mathematics*, 28:11–31, 2010.

[50] L. Chen and C.-S. Zhang. AFEM@matlab: a Matlab package of adaptive finite element methods. Technical report, University of Maryland at College Park, 2006.

[51] H.J. Choi and J.R. Kweon. The Fourier finite element method for the corner singularity expansion of the Heat equation. *Computers and Mathematics with Applications*, 69(1):13–30, 2015.

[52] G. Choudury. Fully discrete Galerkin approximations of parabolic boundary-value problems with nonsmooth boundary data. *Numerische Mathematik*, 57(1):179–203, Dec 1990.

[53] G. Choudury. Fully discrete approximations of parabolic boundary-value problems with nonsmooth boundary data. *Applied Mathematics and Optimization*, 31(1):41–55, Jan 1995.

[54] K. Chrysafinos and L. Hou. Error estimates for semidiscrete finite element approximations of linear and semilinear parabolic equations under minimal regularity assumptions. *SIAM Journal on Numerical Analysis*, 40(1):282–306, 2002.

[55] P. G. Ciarlet and J. L. Lions. *Handbook of Numerical Analysis, vol. II. Finite Element Methods (Part 1).* North Holland, Amsterdam, New York, Oxford, 1991.

[56] P.G. Ciarlet. *Finite Element Method for Elliptic Problems.* Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2002.

[57] G. Cohen, P. Joly, J.E. Roberts, and N. Tordjman. Higher order triangular finite elements with mass lumping for the wave equation. *SIAM Journal on Numerical Analysis*, 38(6):2047–2078, 2001.

[58] G. Cohen, P. Joly, and N. Tordjman. Construction and analysis of higher order finite elements with mass lumping for the wave equation. *Proceedings of the Second International Conference on Mathematical and Numerical Aspects of Wave Propagation Phenomena*, pages 152–160, 1993.

[59] S. S. Collis, K. Ghayour, M. Heinkenschloss, M. Ulbrich, and S. Ulbrich. Optimal control of unsteady compressible viscous flows. *International Journal for Numerical Methods in Fluids*, 40(11):1401–1429, 2002.

[60] X. Da. Non-smooth initial data error estimates with the weight norms for the linear finite element method of parabolic partial differential equations. *Applied Mathematics and Computation*, 54(1):1–24, 1993.

[61] J. C. de los Reyes and K. Kunisch. A semi-smooth Newton method for control constrained boundary optimal control of the Navier-Stokes equations. *Nonlinear Analysis. Theory, Methods and Applications. An International Multidisciplinary Journal.*, 62(7):1289–1316, 2005.

[62] L. Demkowicz, J.T. Oden, W. Rachowicz, and O. Hardy. Toward a universal h-p adaptive finite element strategy, part 1. constrained approximation and data structure. *Computer Methods in Applied Mechanics and Engineering*, 77(1):79–112, 1989.

[63] L. Demkowicz, W. Rachowicz, and Ph. Devloo. A fully automatic hp-adaptivity. *Journal of Scientific Computing*, 17(1):117–142, Dec 2002.

[64] E. Di Nezza, G. Palatucci, and E. Valdinoci. Hitchhiker's guide to the fractional Sobolev spaces. *ArXiv e-prints*, April 2011.

[65] J. Douglas and T.F. Dupont. Galerkin methods for parabolic equations. *SIAM Journal on Numerical Analysis*, 7:575–626, 1970.

[66] T.F. Dupont and R. Scott. Polynomial Approximation of Functions in Sobolev Spaces. *Mathematics of Computation*, 34(150):441–463, 1980.

[67] R. Durán. On polynomial approximation in Sobolev spaces. *SIAM Journal on Numerical Analysis*, 20(5):985–988, 1983.

[68] H. Egger. Energy-norm error estimates for finite element discretization of parabolic problems. *ArXiv e-prints*, July 2015.

[69] H. Egger, U. Rüde, and B. Wohlmuth. Energy-corrected finite element methods for corner singularities. *SIAM Journal on Numerical Analysis*, 52(1):171–193, 2014.

[70] M. Elliotis, G. Georgiou, and Ch. Xenophontos. Solving Laplacian problems with boundary singularities: A comparison of a singular function boundary integral method with the p/hp version of the finite element method. *Applied Mathematics and Computation*, 169(1):485–499, October 2005.

[71] L.C. Evans. *Partial differential equations*. American Mathematical Society, Providence, R.I., 2010.

[72] G. Fix. Higher-order Rayleigh-Ritz approximations. *Journal of Mathematics and Mechanics*, 18:645–657, 1968/1969.

[73] G. Fix and N. Nassif. On finite element approximations in time dependent problems. *Numerische Mathematik*, 19:127–135, 1972.

[74] J. Frehse and R. Rannacher. Eine L^1-Fehlerabschätzung für diskrete Grundlösungen in der Methode der finiten Elemente. *Bonner Mathematische Schriften*, 89:92–114, 1976.

[75] L. Giovani. *A First Course in Sobolev Spaces: Second Edition*, volume 181 of *Graduate Studies in Mathematics*. American Mathematical Society, 2017.

[76] R. Glowinski, J. L. Lions, and R. Trémoliéres. *Numerical Analysis of Variational Inequalities*. Elsevier, 2011.

[77] P. Grisvard. *Elliptic problems in nonsmooth domains.* Pitman, Boston, 1985.

[78] P. Grisvard. *Singularities in Boundary Value Problems*. Recherches en mathématiques appliquées. Masson, 1992.

[79] J.P. Hennart. *Topics in Finite Element Discretization of Parabolic Evolution Problems*. Lecture Notes in Mathematics 909. Springer Verlag, Berlin, Heidelberg, 1982.

[80] M. Hintermüller, K. Ito, and K. Kunisch. The primal-dual active set strategy as a semismooth Newton method. *SIAM Journal on Optimization*, 13(3):865–888, August 2002.

[81] M. Hintermüller, V. A. Kovtunenko, and K. Kunisch. The primal-dual active set method for a crack problem with non-penetration. *IMA Journal of Applied Mathematics*, 69(1):1–26, 2004.

[82] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE constraints*, volume 23 of *Mathematical Modelling: Theory and Applications*. Springer, New York, 2009.

[83] D. Hömberg, C. Meyer, J. Rehberg, and W. Ring. Optimal control for the thermistor problem. *SIAM Journal on Control and Optimization*, 48(5):3449–3481, 2009/10.

[84] T. Horger. *Complexity Reduction for Finite Element Methods with Applications to Eigenvalue Problems*. PhD thesis, Technische Universität München, 2016.

[85] T. Horger, M. Huber, U. Rüde, Ch. Waluga, and B. Wohlmuth. Energy-corrected finite element methods for scalar elliptic problems. In Assyr Abdulle, Simone Deparis, Daniel Kressner, Fabio Nobile, and Marco Picasso, editors, *Numerical Mathematics and Advanced Applications - ENUMATH 2013*, pages 19–36, Cham, 2015. Springer International Publishing.

[86] T. Horger, P. Pustejovska, and Wohlmuth B. Higher order energy-corrected finite element methods. 2017. arXiv:1704.05638 [math.NA], `https://arxiv.org/abs/1704.05638`.

[87] G. C. Hsiao and W. L. Wendland. *Boundary integral equations.* Springer, Berlin, 2008.

[88] M. Huber, L. John, P. Pustejovska, U. Rüde, C. Waluga, and B. Wohlmuth. Solution techniques for the Stokes system: a priori and a posteriori modifications, resilient algorithms. In *Proceedings of the 8th International Congress on Industrial and Applied Mathematics*, pages 109–134. Higher Ed. Press, Beijing, 2015.

[89] T. Hytönen, J. van Neerven, M. Veraar, and L. Weis. *Bochner spaces*, pages 1–66. Springer International Publishing, Cham, 2016.

[90] L. John. *Optimal Boundary Control in Energy Spaces Preconditioning and Applications.* Monographic Series TU Graz, Computation in Engineering and Science, vol. 24, 2014.

[91] L. John, P. Pustejovska, B. Wohlmuth, and U. Rüde. Energy-corrected finite element methods for the Stokes system. *IMA Journal of Numerical Analysis*, 37(2):687–729, 2017.

[92] L. John, P. Swierczynski, and B. Wohlmuth. Energy corrected FEM for optimal Dirichlet boundary control problems. *Numerische Mathematik*, 139(4):913–938, 2018.

[93] S. Jund and S. Salmon. Arbitrary high-order finite element schemes and high-order mass lumping. *International Journal of Applied Mathematics and Computer Science*, 17(3):375–393, 2007.

[94] R.B. Kellogg. *Interpolation Between Subspaces of a Hilbert Space.* Technical note: Institute for Fluid Dynamics and Applied Mathematics. Institute for Fluid Dynamics and Applied Mathematics, 1971.

[95] N. Kikuchi and J. T. Oden. *Contact Problems in Elasticity: A Study of Variational Inequalities and Finite Element Methods.* Studies in Applied and Numerical Mathematics. Society for Industrial and Applied Mathematics, Philadelphia, 1988.

[96] S. Kim and H.-C. Lee. A finite element method for computing accurate solutions for Poisson equations with corner singularities using the stress intensity factor. *Computers and Mathematics with Applications*, 71(11):2330–2337, June 2016.

[97] V. A. Kondratiev. Boundary value problems for elliptic equations in domains with conical or angular points. *Transactions of Moscow Mathematical Society*, 16:227–313, 1967.

[98] V. A. Kozlov, V. G. Mazya, and J. Rossmann. *Elliptic boundary value problems in domains with point singularities*, volume 52 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1997.

[99] V.A. Kozlov, V.G. Mazya, and J. Rossmann. *Spectral Problems Associated with Corner Singularities of Solutions to Elliptic Equations*. Mathematical surveys and monographs. American Mathematical Society, 2001.

[100] A. Kufner. *Weighted Sobolev spaces*. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1985. Translated from Czech.

[101] K. Kunisch, P. Trautmann, and B. Vexler. Optimal control of the undamped linear wave equation with measure valued controls. *SIAM Journal on Control and Optimization*, 54(3):1212–1244, 2016.

[102] K. Kunisch and B. Vexler. Optimal vortex reduction for instationary flows based on translation invariant cost functionals. *SIAM Journal on Control and Optimization*, 46(4):1368–1397, 2007.

[103] H. Lewy, K. Friedrichs, and R. Courant. Über die partiellen Differenzengleichungen der mathematischen Physik. *Mathematische Annalen*, 100:32–74, 1928.

[104] D. Leykekhman and B. Vexler. Discrete maximal parabolic regularity for Galerkin finite element methods. *ArXiv e-prints*, May 2015.

[105] D. Leykekhman and B. Vexler. Pointwise best approximation results for Galerkin finite element solutions of parabolic problems. *SIAM Journal on Numerical Analysis*, 54(3):1365–1384, 2016.

[106] D. Leykekhman and B. Vexler. Global and interior pointwise best approximation results for the gradient of galerkin solutions for parabolic problems. *SIAM Journal on Numerical Analysis*, 55(4):2025–2049, 2017.

[107] B. Li. Analyticity, maximal regularity and maximum-norm stability of semidiscrete finite element solutions of parabolic equations in nonconvex polyhedra. *ArXiv e-prints*, February 2017.

[108] Z.C. Li and T.T. Lu. Singularities and treatments of elliptic boundary value problems. *Mathematical and Computer Modelling*, 31(8–9):97–145, 2000.

[109] J.H. Lienhard IV and J.H. Lienhard V. *A Heat Transfer Textbook*. Phlogiston Press, Cambridge, MA, 4th edition, 2017.

[110] J. L. Lions. *Optimal Control of Systems Governed by Partial Differential Equations*. Springer-Verlag,, Berlin, Heidelberg, 1971.

[111] J. M.-S. Lubuma and K. C. Patidar. Towards the implementation of the singular function method for singular perturbation problems. *Applied Mathematics and Computation*, 209:68–74, 2009.

[112] J. Mandel. The Bramble-Hilbert Lemma. *ArXiv e-prints*, October 2007.

[113] V. Maz'ya. *Sobolev Spaces with Applications to Elliptic Partial Differential Equations*. Grundlehren der Mathematischen Wissenschaften. Springer-Verlag Berlin Heidelberg, 2011. Originally published under Vladimir G. Maz'ya in Springer Series of Soviet Mathematics. 2nd revised and augmented edition.

[114] J. Melenk and B. Wohlmuth. Quasi-optimal approximation of surface based Lagrange multipliers in finite element methods. *SIAM Journal on Numerical Analysis*, 50(4):2064–2087, 2012.

[115] C. Meyer and A. Rösch. Superconvergence properties of optimal control problems. *SIAM Journal on Control and Optimization*, 43(3):970–985 (electronic), 2004.

[116] J. Moser. A sharp form of an inequality by N. Trudinger. *Indiana University Mathematics Journal*, 20:1077–1092, 1971.

[117] W. A. Mulder. New triangular mass-lumped finite elements of degree six for wave propagation. *Progress In Electromagnetics Research*, 141:671–692, 2013.

[118] C.-D. Munz and T. Westermann. *Numerische Behandlung Gewöhnlicher und Partieller Differenzialgleichungen*. Springer Berlin Heidelberg, Wiesbaden, 2012.

[119] F. Naterer. Über die punktweise Konvergenz Finiter Elemente. *Numerische Mathematik*, 25:67–77, 1975.

[120] M.H. Nguyen and T.A. Nguyen. Regularity of solutions of initialboundary value problems for parabolic equations in domains with conical points. *Journal of Differential Equations*, 245:1801–1818, 2008.

[121] J. A. Nitsche. Lineare Spline-Funktionen und die Methoden von Ritz für elliptische Randwertprobleme. *Archive for Rational Mechanics and Analysis*, 36:348–355, 1970.

[122] J. A. Nitsche. L^∞-convergence of finite element approximations. In *Mathematical Aspects of Finite Element Methods*, pages 261–274. Springer Berlin Heidelberg, 1977.

[123] J. A. Nitsche and A. H. Schatz. Interior estimates for Ritz-Galerkin methods. *Mathematics of Computation*, 28:937–958, 1974.

[124] B. Nkemzi and M. Jung. A Postprocessing Finite Element Strategy for Poisson's Equation in Polygonal Domains: Computing the Stress Intensity Factors. In T. Apel and O. Steinbach, editors, *Advanced Finite Element Methods and Applications*, pages 153–173. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[125] E. Nonbol. Description of the advanced gas cooled type of reactor (AGR). *Nordic Nuclear Safety Research*, 1996.

[126] J.T. Oden, L. Demkowicz, W. Rachowicz, and T.A. Westermann. Toward a universal h-p adaptive finite element strategy, part 2. a posteriori error estimation. *Computer Methods in Applied Mechanics and Engineering*, 77(1):113–180, 1989.

[127] G. Of, T. X. Phan, and O. Steinbach. An energy space finite element approach for elliptic Dirichlet boundary control problems. *Numerische Mathematik*, 129(4):723–748, 2015.

[128] L.A. Oganesyan and L.A. Rukhovets. *Variational-difference methods for the solution of elliptic equations*. Izd. Akad. Nauk Armyanskoi SSR, 1979 (In Russian).

[129] M. Ortl and A.-M. Sändig. Regularity of Viscous Navier–Stokes Flows in Nonsmooth Domains. In *Boundary Value Problems and Integral Equations in Nonsmooth Domains*, volume 167 of *Lecture Notes in Pure and Applied Mathematics*, pages 185–201. 1993.

[130] C. Palencia. Maximum norm analysis of completely discrete finite element methods for parabolic problems. *SIAM Journal on Numerical Analysis*, 33(4):1654–1668, 1996.

[131] W.V. Petryshyn. Constructional proof of Lax–Milgram Lemma and its applications to non-k-p.d. abstract and differential operator equation. *SIAM Journal on Numerical Analysis*, 3(2):404–420, 1965.

[132] J. Pfefferer. *Numerical analysis for elliptic Neumann boundary control problems on polygonal domains*. PhD thesis, Universität der Bundeswehr München, 2014.

[133] H. Price and R. Varga. Error bounds for semi-discrete galerkin approximations of parabolic problems with applications to petroleum reservoir mechanics. *Numerical Solution of Field Problems in Continuum Physics, American Mathematical Society, Providence, R.I*, pages 74–94, 1970.

[134] W. Rachowicz, J.T. Oden, and L. Demkowicz. Toward a universal h-p adaptive finite element strategy part 3. design of h-p meshes. *Computer Methods in Applied Mechanics and Engineering*, 77(1):181–212, 1989.

[135] S. Rogovs. *Pointwise Error Estimates for Boundary Control Problems on Polygonal Domains*. PhD thesis, Universität der Bundeswehr München, 2018.

[136] U. Rüde. Local corrections for eliminating the pollution effect of reentrant corners. Technical Report TUM-INFO-02-89-I01, Institut für Informatik, Technische Universtät München, 1989.

[137] U. Rüde, C. Waluga, and B. Wohlmuth. Nested Newton strategies for energy-corrected finite element methods. *SIAM Journal on Scientific Computing*, 36(4):A1359–A1383, 2014.

[138] U. Rüde and C. Zenger. On the treatment of singularities in the multigrid method. In Wolfgang Hackbusch and Ulrich Trottenberg, editors, *Multigrid Methods II*, volume 1228 of *Lecture Notes in Mathematics*, pages 261–271. Springer Berlin / Heidelberg, 1986. 10.1007/BFb0072651.

[139] A. M. Sanchez and R. Arcangeli. Estimations des erreurs de meilleure approximation polynomiale et d'interpolation de Lagrange dans les espaces de Sobolev d'ordre non entier. *Numerische Mathematik*, 45(2):301–321, 1984.

[140] A.H. Schatz. Pointwise error estimates and asymptotic error expansion inequalities for the finite element method on irregular grids: Part I. global estimates. *Mathematics of Computation*, 67(223):877–899, 1998.

[141] A.H. Schatz. Pointwise error estimates and asymptotic error expansion inequalities for the finite element method on irregular grids: Part II. interior estimates. *Mathematics of Computation*, 38:1269–1293, 2000.

[142] A.H. Schatz and L.B. Wahlbin. Interior maximum norm estimates for finite element methods. *Mathematics of Computation*, 31(138):414–442, 1977.

[143] A.H. Schatz and L.B. Wahlbin. Maximum norm estimates in the finite element method on plane polygonal domains. part 1. *Mathematics of Computation*, 32(141):73–109, 1978.

[144] A.H. Schatz and L.B. Wahlbin. Maximum norm estimates in the finite element method on plane polygonal domains. part 2, refinements,. *Mathematics of Computation*, 33(146):465–492, 1979.

[145] L.R. Scott and S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Mathematics of Computation*, 54(190):483–493, 1990.

[146] R. Scott. Optimal l^∞ estimates for the finite element method on irregular meshes. *Mathematics of Computation*, 30(136):681–697, 1976.

[147] D. Sirch. *Finite Element Error Analysis for PDE-constrained Optimal Control Problems: The Control Constrained Case Under Reduced Regularity*. PhD thesis, Technische Universität München, 2010.

[148] W. Spann. On the boundary element method for the Signorini problem of the Laplacian. *Numerische Mathematik*, 65:337–356, 1993.

[149] O. Steinbach. *Numerical Approximation Methods for Elliptic Boundary Value Problems*. Springer-Verlag New York, 2008.

[150] O. Steinbach. Boundary element methods for variational inequalities. *Numerische Mathematik*, 126(1):173–197, 2014.

[151] G. Strang and G.J. Fix. *An analysis of the finite element method*. Wellesley-Cambridge Press, 1988.

[152] E. Süli and D.F. Mayers. *An Introduction to Numerical Analysis*. Cambridge University Press, 2003.

[153] P. Swierczynski and B. Wohlmuth. Energy-corrected FEM and explicit time-stepping for parabolic problems. 2018. In preparation.

[154] P. Swierczynski and B. Wohlmuth. Maximum norm estimates for energy-corrected finite element method. *Numerical Mathematics and Advanced Applications ENUMATH 2017*, 126, 2018.

[155] V. Thomée. Some Convergence Results for Galerkin Methods for Parabolic Boundary Value Problems. In *Mathematical Aspects of Finite Elements in Partial Differential Equations*, pages 55–88. Elsevier, 1974.

[156] V. Thomée. *Finite difference methods for linear parabolic equations*. Handbook of Numerical Analysis vol I. Finite Difference Methods 1, P.G. Ciarlet and J.L. Lions, eds., North-Holland, Amsterdam, 1990.

[157] V. Thomée. *Galerkin Finite Element Methods for Parabolic Problems (Springer Series in Computational Mathematics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.

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

[159] P. Trautmann, B. Vexler, and A. Zlotnik. Finite element error analysis for measure-valued optimal control problems governed by a 1d wave equation with variable coefficients. *Mathematical Control and Related Fields*, 8(2):411–449, 2018.

[160] H. Triebel. *Interpolation Theory, Function Spaces, Differential Operators*. VEB Deutscher Verlag der Wissenschaften, Berlin, 1978.

[161] F. Tröltzsch. *Optimal control of partial differential equations. Theory, methods and applications.* American Mathematical Society, Providence, 2010.

[162] N.S. Trudinger. On imbeddings into Orlicz spaces and some applications. *Journal of Mathematics and Mechanics*, 17:473–483, 1967.

[163] S.V. Uspenskii. An imbedding theorem for S. L. Sobolevs classes $W^{p,r}$ of fractional order. *Soviet Mathematics Doklady*, 1:132–133, 1960.

[164] L. Vu trong and D. Loi. The first initial-boundary value problem for parabolic equations in a cone with edges. 2(60):394–404, 07 2015.

[165] L.B. Wahlbin. Local behavior in finite element methods. In *Finite Element Methods (Part 1)*, volume 2 of *Handbook of Numerical Analysis*, pages 353 – 522. Elsevier, 1991.

[166] M.F. Wheeler. A priori L2 error estimates for Galerkin approximations to parabolic partial differential equations. *SIAM Journal on Numerical Analysis*, 10:723–759, 1973.

[167] M. Winkler. Error estimates for variational normal derivatives and Dirichlet control problems with energy regularization. *ArXiv e-prints*, August 2018.

[168] C. Zenger and H. Gietl. Improved difference schemes for the Dirichlet problem of Poisson's equation in the neighbourhood of corners. *Numerische Mathematik*, 30(3):315–332, 1978.