TECHNISCHE UNIVERSITÄT WIEN

DISSERTATION

Finite Element Analysis of the Heterogeneous Helmholtz Equation and Least Squares Methods

ausgeführt zum Zwecke der Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften unter der Leitung von

Univ.-Prof. Jens Markus Melenk, PhD
E101 – Institut für Analysis und Scientific Computing, TU Wien

eingereicht an der Technischen Universität Wien
Fakultät für Mathematik und Geoinformation

von

Dipl.-Ing. Maximilian Bernkopf
Matrikelnummer: 01125373
Hafnersteig 10/5
1010 Wien

Diese Dissertation haben begutachtet:

1. **Prof. Dr. Serge Nicaise**
   Laboratoire de Mathématiques et leurs Applications de Valenciennes,
   Université Polytechnique Hauts-de-France
2. **Prof. Dr. Stefan Sauter**
   Institut für Mathematik,
   Universität Zürich
3. **Prof. Jens Markus Melenk, PhD**
   Institut für Analysis und Scientific Computing,
   TU Wien

Wien, am 3. Juli 2021

Kurzfassung

Die vorliegende Arbeit befasst sich mit drei großen Themenblöcken. Zu Beginn der Arbeit betrachten wir eine kleinste Quadrate Methode zur numerischen Diskretisierung der homogenen Helmholtz Gleichung. Es wird eine Konvergenztheorie dieser Methode bewiesen, welche explizit in der Wellenzahl ist. Weiters betrachten wir eine kleinste Quadrate Methode zur Diskretisierung einer partiellen Differentialgleichung zweiter Ordnung, welche zuvor in ein System von Gleichungen erster Ordnung umformuliert wird. Für diese Methode wird unter minimalen Regularitätsannahmen an die Daten Optimalität bewiesen. Schließlich betrachten wir eine Klasse von zeitharmonischen Wellenphänomenen in stückweise glatten Medien. Für diese Klasse von Problemen wird eine Regularitätstheorie bewiesen, welche explizit in der Wellenzahl ist. Diese Regularitätstheorie wiederum erlaubt eine vollständige Konvergenzanalyse von Galerkin Verfahren für diese Problemklasse.

Abstract

The present thesis is concerned with three main topics. The first one being a least squares finite element approach for numerical discretizations of the homogeneous Helmholtz equation. We perform a wavenumber-explicit convergence theory for this method. Secondly, we prove optimality for a first order system least squares finite element method applied to a second order partial differential equation focusing on minimal regularity assumptions on the data. Finally, we consider a class of time-harmonic wave propagation problems in piecewise smooth media. For these problems, a wavenumber-explicit regularity theory is performed. This in turn allows for a complete and wavenumber-explicit convergence analysis of a Galerkin method applied to our model class.

Acknowledgement

First and foremost, I want to express my uttermost gratitude to my advisor Jens Markus Melenk. Without his mathematical guidance, patience, and encouragement over the past years the present thesis would not have been possible.

I want to thank Serge Nicaise and Stefan Sauter, for giving me the opportunity to visit both of them at their home universities. The research stays abroad have been very fruitful and widened my mathematical horizon. Furthermore, I want to express my gratitude to the numerous people I had the pleasure of meeting and working together during my time in Zürich and Valenciennes.

I also want to thank my colleagues from Vienna, especially at the Institute for Analysis and Scientific Computing at TU Wien. The office was not only a place for stimulating mathematical discussion but also fun and exciting.

On a personal note, I want to thank my parents for their support over the years as well as my girlfriend Melanie for being there for me in any circumstances.

Finally, I am very grateful for the financial support from the Austrian Science Fund (FWF) through the doctoral school *Dissipation and dispersion in nonlinear PDEs* (grant W1245).

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Dissertation selbstständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt bzw. die wörtlich oder sinngemäß entnommenen Stellen als solche kenntlich gemacht habe.

Wien, am 3. Juli 2021

Maximilian Bernkopf

Contents

1. Introduction and contributions

The present thesis is divided into four main chapters. In Chapter 3 we consider a first order system least squares (FOSLS) finite element method applied the homogeneous Helmholtz equation. Chapters 4 and 5 cover the optimality of a FOSLS method under minimal regularity assumptions on the data. In Chapter 4 we consider homogeneous boundary conditions, whereas in Chapter 5 the case of inhomogeneous boundary conditions is covered. Finally, in Chapter 6 we perform wavenumber-explicit regularity theory for a class of time-harmonic wave propagation problems in heterogeneous media and apply these results to derive a wavenumber-explicit convergence analysis for the Galerkin discretization of these problems.

1.1. Contributions of Chapter 3

The model problem of Chapter 3 is the following homogeneous Helmholtz problem:

$$\begin{aligned} -\Delta u - k^2 u &= f \quad \text{in } \Omega, \\ \partial_n u - iku &= g \quad \text{on } \Gamma, \end{aligned} \tag{1.1}$$

where the wavenumber $k \geq k_0 > 0$ is real. For large k, the numerical solution of (1.1) is challenging due to the requirement to resolve the oscillatory nature of the solution. A second challenge arises in classical, $H^1(\Omega)$-conforming discretizations of (1.1) from the fact that the Galerkin method is *not* an energy projection, and a meaningful approximation is only obtained under more stringent conditions on the mesh size h and the polynomial degree p than purely approximation theoretical considerations suggest. This shortcoming has been analyzed in the literature. In particular, as discussed in more detail in [MS11, EM12], the analyses [Ihl98, IB95, IB97, Ain04, MS10, MS11, EM12] show that high order methods are much better suited for the high-frequency case of large k than low order methods. Alternatives to the classical Galerkin methods which are still based on high order methods include stabilized methods [FW09, FX13a, FX13b, ZW13], hybridizable methods [CLX13], least squares type methods [CQ17, LMMR00] and Discontinuous Petrov Galerkin methods, [PD17, DGMZ12]. An attractive feature of least squares type methods is that the resulting linear system is always solvable and that they feature quasi-optimality, albeit in some nonstandard residual norms. Motivated by the results of [CQ17] we show for a first order system least squares method an *a priori* estimate in the more tractable $L^2(\Omega)$-norm under the scale resolution condition

$$\frac{kh}{p} \leq c_1 \qquad \text{and} \qquad p \geq c_2(\log k + 1).$$

For that, we closely follow [CQ17]. Our key refinement over [CQ17] is an improved regularity estimate for the solution of a suitable dual problem (cf. Lemma 3.3.1 vs. [CQ17, Lemma 5.1]) that allows us to establish the improved p-dependence in the $L^2(\Omega)$-error estimate (cf. Theorem 3.5.1 vs. [CQ17, Thm. 2.5]). As a tool, which is of independent interest, we develop approximation operators in Raviart-Thomas and Brezzi-Douglas-Marini spaces with optimal (in h and p) approximation rates simultaneously in $L^2(\Omega)$ and $\boldsymbol{H}(\Omega, \mathrm{div})$.

1.2. Contributions of Chapter 4

Motivated by the numerical findings of Chapter 3 we further investigate the optimality of a first order system least squares method applied to a second order elliptic model problem endowed with homogeneous boundary conditions in Chapter 4. Least Squares Finite Element Methods (LSFEM) are an important class of numerical methods for the solution of partial differential equations with a variety of applications. The main idea of the LSFEM is to reformulate the partial differential equation of interest as a minimization problem, for which a variety of tools is available. For example, even for nonsymmetric or indefinite problems, as showcased in Chapter 3, the discretization with the least squares approach leads to symmetric, positive definite systems, which can be solved with well-established numerical technologies. Furthermore, the least squares technique is naturally quasi-optimal, albeit in a problem-dependent norm. For second order PDEs the most common least squares approach is that of rewriting the equation as a first order least squares system that can be discretized with established finite element techniques. A benefit is that many quantities of interest are approximated directly without the need of postprocessing. We mention [BG09] as a classical monograph on the topic as well as the papers [Jes77, CLMM94, CMM97a, BG05]. Chapter 4 considers a Poisson-like second order model problem written as a system of first order equations. For the discretization, an $\boldsymbol{H}(\Omega, \mathrm{div}) \times H^1(\Omega)$-conforming least squares formulation is employed. Even though our model problem in its standard $H^1(\Omega)$ formulation is coercive our methods and lines of proof can most certainly be applied to other problems, see Chapter 3 as well as [BM19, CQ17] for an application to the Helmholtz equation. The LSFEM is typically quasi-optimal in some problem-dependent energy norm, which is, however, somewhat intractable; *a priori* error estimates in more familiar norms such as the $L^2(\Omega)$ norm of the scalar variable are thus desirable. Numerical examples in Chapter 3 suggested convergence rates in standard norms such as the $L^2(\Omega)$-norm which, to our best of knowledge, are not explained by the current theory. We develop such a convergence theory with minimal assumptions on the regularity of the right-hand side.

Our main contribution is an optimal $L^2(\Omega)$ based convergence result for the least squares approximation u_h to the scalar variable u. Furthermore, we derive hp error estimates for the gradient of the scalar variable u, which do not seem to be available in the current literature, as well as an hp error estimate for the vector variable $\boldsymbol{\varphi}$ in the $L^2(\Omega)$ norm, which is available in the literature for a pure h-version. These optimality results are new in the sense that we achieve optimal convergence rates under minimal regularity assumptions on the data. Here, we call a method optimal in a certain norm, if the norm of the error made by the method is of the same order as the best approximation of the employed space.

To highlight our contribution, we present an overview of the current results available in the literature:

In [Jes77] the author considered the classical model problem $-\Delta u = f$ with inhomogeneous Dirichlet boundary condition $u = g$ in some smoothly bounded domain Ω. Unlike the methodology of Chapter 4 the least squares formulation employs vector valued $H^1(\Omega)$ functions instead of $\boldsymbol{H}(\Omega, \text{div})$ for the vector variable. The corresponding finite element spaces are chosen such that they satisfy simultaneous approximation properties in $L^2(\Omega)$ and $H^1(\Omega)$ for both the scalar variable u and the vector variable $\boldsymbol{\varphi}$. Using a duality argument akin to the one used in this thesis the author arrived at the error estimate

$$\|u - u_h\|_{0,\Omega} \lesssim h \, \|(\boldsymbol{\varphi} - \boldsymbol{\varphi}_h, u - u_h)\|_b \, ,$$

see [Jes77, Thm. 4.1], where $\|(\cdot,\cdot)\|_b$ denotes the corresponding energy norm. At this point higher order convergence rates are just a question of approximation properties in $\|(\cdot,\cdot)\|_b$, see [Jes77, Lemma 3.1] for a precise statement. As stated after the proof of [Jes77, Thm. 4.1], one can extract optimal convergence rates for sufficiently smooth data f and g. The smoothness of the data is important as the following considerations show: For the case of a smooth boundary Γ and $f \in L^2(\Omega)$ and $g \in H^{3/2}(\Gamma)$, elliptic regularity gives $u \in H^2(\Omega)$. Therefore u can be approximated by globally continuous piecewise polynomials of degree greater or equal to one with an error $O(h^2)$ in the $L^2(\Omega)$ norm, which is achieved by classical FEM, due to the Aubin-Nitsche trick. In contrast, the above least squares estimate does not give the desired rate: The norm $\|(\boldsymbol{\varphi} - \boldsymbol{\varphi}_h, u - u_h)\|_b$ contains a term of the form

$$\|\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{\varphi}_h)\|_{0,\Omega} = \|f - \nabla \cdot \boldsymbol{\varphi}_h\|_{0,\Omega} \, ,$$

from which no further convergence rate can be extracted, since f is only in $L^2(\Omega)$.

In [CLMM94] (see also [CMM97a]) the problem $-\nabla \cdot (A \nabla u) + Xu = f$ with uniformly elliptic diffusion matrix A and X a linear differential operator of order at most one together with homogeneous mixed Dirichlet and Neumann boundary conditions was considered. The least squares formulation presented therein employs the same spaces as the present work. Apart from nontrivial norm equivalence results, see [CLMM94, Thm. 3.1], they also derived the following estimate of the least squares approximation

$$\|u - u_h\|_{1,\Omega} + \|\boldsymbol{\varphi} - \boldsymbol{\varphi}_h\|_{\boldsymbol{H}(\Omega,\text{div})} \lesssim h^s(\|u\|_{s+1,\Omega} + \|\boldsymbol{\varphi}\|_{s+1,\Omega}),$$

assuming $u \in H^{s+1}(\Omega)$ and $\boldsymbol{\varphi} \in \boldsymbol{H}^{s+1}(\Omega)$. This result is then optimal in the stated norm, however, the assumed regularity is somewhat unsatisfactory, in the sense that if the solution $u \in H^{s+1}(\Omega)$ then the relation $\nabla u + \boldsymbol{\varphi} = 0$ merely provides the regularity $\boldsymbol{\varphi} \in \boldsymbol{H}^s(\Omega)$ and not the assumed regularity $\boldsymbol{\varphi} \in \boldsymbol{H}^{s+1}(\Omega)$.

Finally, in [BG05] the same model problem, as well as the same least squares formulation, is considered. The main goal of [BG05] is to establish $L^2(\Omega)$ error estimates for u and $\boldsymbol{\varphi}$. In [BG05, Lemma 3.4] a result similar to [Jes77, Thm. 4.1] is obtained. This result, however, suffers from the same drawback as elaborated above. Furthermore, they prove optimality of the error of the vector variable $\boldsymbol{\varphi}$ in the $L^2(\Omega)$ norm, see [BG05, Cor. 3.7].

The main tools for *a priori* error estimates in more tractable norms such as $L^2(\Omega)$ instead of the energy norm in a least squares setting are, as it is done in the present thesis and the above literature, duality arguments, which lead to an estimate of the form

$$\|u - u_h\|_{0,\Omega} \lesssim h \, \|(\boldsymbol{\varphi} - \boldsymbol{\varphi}_h, u - u_h)\|_b \, .$$

As elaborated above it is not possible to extract the desired optimal rate from this estimate directly. In the proof of one of our main results (Theorem 4.3.12) we exploit the duality argument in a more delicate way, which allows us to lower the regularity requirements on $\boldsymbol{\varphi}$ to what could be expected from the regularity of the data f. Key components in the proof are the $\boldsymbol{H}(\Omega, \operatorname{div})$-conforming approximation operators $\boldsymbol{I}_h^0$ and $\boldsymbol{I}_h$ (cf. Lemmas 4.3.3, 4.3.6), which are also of independent interest.

1.3. Contributions of Chapter 5

Extending the results of Chapter 4 we consider inhomogeneous Robin boundary conditions in Chapter 5. These boundary conditions contribute to additional boundary terms in the bilinear form b. As discussed above, the main argument in deriving error estimates in more tractable norms, are duality arguments. The additional boundary terms arising due to the inhomogeneous Robin boundary conditions lead to a more delicate analysis. First off, the regularity of the dual solutions of Chapter 4 is further limited due to the boundary terms, cf. Theorem 4.2.1 vs. Theorem 5.2.1. Furthermore, an additional duality argument for the normal trace of the vector variable needs to be performed, see Theorem 5.2.4. Finally, the operator $\boldsymbol{I}_h$ needs to be adjusted in order to account for the additional boundary term, see Lemma 5.3.5 in comparison to Lemma 4.3.6.

1.4. Contributions of Chapter 6

In Chapter 6 we analyze the Galerkin discretization of a class of heterogeneous time-harmonic wave propagation problems in a high-frequency regime. The prototypical model problem is the time-harmonic homogeneous Helmholtz equation with wavenumber $k > 0$

$$-\Delta u - k^2 u = f. \tag{1.2}$$

The solution u to (1.2) inherits a highly oscillatory behavior. On the one hand, numerical schemes need to resolve this oscillatory nature and therefore require a large number of degrees of freedom. On the other hand, standard conforming Galerkin discretizations result in an indefinite formulation. To ensure stability on the discrete level more restrictive conditions than kh to be small need to be met. In the sequence of papers [MS10, MS11] the superiority of the *hp* Finite Element Method (*hp*-FEM) compared to a pure *h*-FEM was established. These results rely on a wavenumber-explicit regularity theory. Therein it is shown for a class of homogeneous Helmholtz problems, that the solution u admits a decomposition $u = u_F + u_A$ into a finite regularity part u_F and an analytic part u_A. The finite regularity part u_F features favorable k-explicit bounds. The analytic part u_A captures the oscillatory behavior of the solution. Apart from being of independent interest, this regularity theory enters the analysis of Galerkin discretizations when establishing quasi-optimality. Here, the approximability of an appropriate adjoint problem yields quasi-optimality. This adjoint problem is again of similar character to the homogeneous Helmholtz equation. Therefore, the aforementioned regularity theory applies. On a conceptual level, the superiority of the *hp*-FEM is due to the fact that unfavorable k-dependence of the analytic part u_A can be overcome, since the *hp* finite element space features exponential approximation properties for smooth functions. Assuming the solution operator is polynomially bounded in k, it is shown in [MS10, MS11], that under the scale resolution condition

$$\frac{kh}{p} \leq c_1 \qquad \text{and} \qquad p \geq c_2(\log k + 1) \tag{1.3}$$

quasi-optimality of the *hp*-FEM holds with wavenumber-independent constants.

The first results concerning a wavenumber-explicit splitting of the solution to Helmholtz problems can be found in [MS10, Lemma 3.5] and [MS11, Thm. 4.10 and 4.20], covering Dirichlet-to-Neumann boundary conditions on a sphere, interior Robin and exterior Dirichlet boundary conditions, respectively. Later, these results were generalized in [EM12] to polygonal domains and in [MPS13] to higher order Sobolev data, i.e., $f \in H^s(\Omega)$ and $g \in H^{s+1/2}(\Gamma)$. We also mention that similar splittings are available for the time-harmonic homogeneous Maxwell problem. See [MS21, Sec. 7.2] for the standard $\boldsymbol{H}(\Omega, \mathrm{curl})$ formulation as well as [NT20, Sec. 4.1.3] for an elliptic system formulation. We point out that previous splittings rely on a wavenumber-explicit analysis of the Newton potential. Our present approach circumvents this by relying solely on an operator S_k^+ which can be viewed as a parametrix of the Helmholtz solution operator S_k^- for high-frequency data. The recently published preprint [LSW20] derives similar results to the present work for the Dirichlet-to-Neumann map with a sphere as a coupling interface.

In Chapter 6 we consider an abstract class of heterogeneous time-harmonic wave propagation problems. These problems include heterogeneous Helmholtz problems with piecewise smooth coefficients. For these problems inhomogeneous Robin, Dirichlet-to-Neumann and second order absorbing boundary conditions are covered. Furthermore, perfectly matched layers and volume damping problems fit into our framework. We generalize the regularity theory developed in [MS10, MS11] and prove an analogous splitting of the solution u of our heterogeneous model class into a finite regularity part u_F and an analytic part u_A, see Theorem 6.3.10. The finite regularity part u_F is (piecewise) H^2 and features favorable wavenumber-explicit bounds. The analytic regularity part u_A is (piecewise) analytic with wavenumber-explicit bounds. This regularity theory allows for the wavenumber-explicit analysis of higher order Galerkin discretizations of the considered problems. We prove quasi-optimality under the scale resolution conditions (1.3) of the *hp*-FEM applied to this class of problems assuming polynomial well-posedness of the solution operator, see Theorem 6.6.3. Furthermore, we generalize the above splitting to higher order Sobolev data, in turn allowing for a complete convergence analysis of the method. Finally, we derive the following results which are of independent interest: In Lemma 6.5.4 we present a shift theorem for second order absorbing boundary conditions. In Lemma 6.5.12 a splitting of the Dirichlet-to-Neumann map for the exterior Helmholtz equation is derived. In fact the Helmholtz Dirichlet-to-Neumann map DtN_k can be written as $\mathrm{DtN}_k = \mathrm{DtN}_0 + kB + [\![\partial_n \tilde{A}]\!]$, where DtN_0 denotes the Dirichlet-to-Neumann map for the Laplacian, B is an operator of order zero featuring wavenumber-independent bounds and $\tilde{A}$ maps into a class of analytic functions.

2. Background and notation

Throughout this thesis, if not otherwise stated, the following notation applies. We introduce the usual Lebesgue and Sobolev spaces below, as a standard reference we mention [McL00]. In spatial dimension $d = 2, 3$ let $\Omega \subset \mathbb{R}^d$ be a bounded domain with smooth boundary $\Gamma \coloneqq \partial\Omega$. For $p \in [1, \infty]$ we denote by $L^p(\Omega)$ the usual Lebesgue spaces, by $\|\cdot\|_{L^p(\Omega)}$ the corresponding norm. For $s \geq 0$ and $p \in [1, \infty]$ we denote by $W^{s,p}(\Omega)$ the standard Sobolev spaces, with norm $\|\cdot\|_{W^{s,p}(\Omega)}$, with $W^{0,p}(\Omega) = L^p(\Omega)$. For the special case $p = 2$ we denote by $H^s(\Omega)$ the space $W^{s,2}(\Omega)$, and write $\|\cdot\|_{s,\Omega}$ for the corresponding norm. For $u, v \in L^2(\Omega)$ we denote by $(u, v)_\Omega$ the $L^2(\Omega)$ inner product. For $s \geq 0$ we denote by $\widetilde{H}^{-s}(\Omega)$ the dual of $H^s(\Omega)$. For $t \geq 0$ we denote by $H^t(\Gamma)$ the Sobolev space on the boundary Γ and write $\|\cdot\|_{t,\Gamma}$ for the corresponding norm. We denote by $H^{-t}(\Gamma)$ the dual space of $H^t(\Gamma)$. For $u, v \in L^2(\Gamma)$ we denote by $\langle u, v\rangle_\Gamma$ the $L^2(\Gamma)$ inner product. Furthermore, we write (u, v) for the duality pairing in the volume, and $\langle u, v\rangle$ for the duality pairing on the boundary. Furthermore, we introduce the spaces $\boldsymbol{H}(\Omega, \mathrm{div})$ and $\boldsymbol{H}(\Omega, \mathrm{curl})$ of square integrable vector fields, with square integrable weak divergence and rotation, respectively, see [Mon03, BBF13] for further details. We denote by $\boldsymbol{n}$ the outward unit normal vector on the boundary Γ. In Chapter 4 we consider different boundary conditions on parts of Γ. Therefore, let Γ consist of two disjoint parts Γ_D and Γ_N. We also consider subspaces of $H^1(\Omega)$, $\boldsymbol{H}(\Omega, \mathrm{div})$ and $\boldsymbol{H}(\Omega, \mathrm{curl})$ with additional boundary conditions. Summarizing, we will be working with the following spaces:

$$\begin{aligned} H^1(\Omega) &= \{u \in L^2(\Omega) \colon \nabla u \in \boldsymbol{L}^2(\Omega)\}, \\ H^1_D(\Omega) &= \{u \in H^1(\Omega) \colon u = 0 \text{ on } \Gamma_D\}, \\ H^1_0(\Omega) &= \{u \in H^1(\Omega) \colon u = 0 \text{ on } \Gamma\}, \end{aligned}$$

$$\begin{aligned} \boldsymbol{H}(\Omega, \mathrm{curl}) &= \{\boldsymbol{\varphi} \in \boldsymbol{L}^2(\Omega) \colon \nabla \times \boldsymbol{\varphi} \in \boldsymbol{L}^2(\Omega)\}, \\ \boldsymbol{H}_N(\Omega, \mathrm{curl}) &= \{\boldsymbol{\varphi} \in \boldsymbol{H}(\Omega, \mathrm{curl}) \colon \boldsymbol{n} \times \boldsymbol{\varphi} = 0 \text{ on } \Gamma_N\}, \\ \boldsymbol{H}_0(\Omega, \mathrm{curl}) &= \{\boldsymbol{\varphi} \in \boldsymbol{H}(\Omega, \mathrm{curl}) \colon \boldsymbol{n} \times \boldsymbol{\varphi} = 0 \text{ on } \Gamma\}, \end{aligned}$$

$$\begin{aligned} \boldsymbol{H}(\Omega, \mathrm{div}) &= \{\boldsymbol{\varphi} \in \boldsymbol{L}^2(\Omega) \colon \nabla \cdot \boldsymbol{\varphi} \in L^2(\Omega)\}, \\ \boldsymbol{H}_N(\Omega, \mathrm{div}) &= \{\boldsymbol{\varphi} \in \boldsymbol{H}(\Omega, \mathrm{div}) \colon \boldsymbol{\varphi} \cdot \boldsymbol{n} = 0 \text{ on } \Gamma_N\}, \\ \boldsymbol{H}_0(\Omega, \mathrm{div}) &= \{\boldsymbol{\varphi} \in \boldsymbol{H}(\Omega, \mathrm{div}) \colon \boldsymbol{\varphi} \cdot \boldsymbol{n} = 0 \text{ on } \Gamma\}. \end{aligned}$$

Additionally, for $s \geq 0$ set $\boldsymbol{H}^s(\Omega, \mathrm{div}) = \{\boldsymbol{\varphi} \in \boldsymbol{H}^s(\Omega) \colon \nabla \cdot \boldsymbol{\varphi} \in H^s(\Omega)\}$. Throughout this thesis $\mathcal{T}_h$ will denote a triangulation of the computational domain Ω and will consist of elements K. Since we are dealing with smooth boundaries we employ curved elements. We make the following assumptions on the triangulation.

Assumption 2.0.1 (quasi-uniform regular meshes)**.** Let $\widehat{K}$ be the reference simplex. Each element map $F_K\colon \widehat{K} \to K$ can be written as $F_K = R_K \circ A_K$, where A_K is an affine map and the maps R_K and A_K satisfy, for constants $C_{\text{affine}}, C_{\text{metric}}, \rho > 0$ independent of K:

$$\begin{aligned} &\|A'_K\|_{L^\infty(\widehat{K})} \leq C_{\text{affine}} h_K, &&\|(A'_K)^{-1}\|_{L^\infty(\widehat{K})} \leq C_{\text{affine}} h_K^{-1}, \\ &\|(R'_K)^{-1}\|_{L^\infty(\tilde{K})} \leq C_{\text{metric}}, &&\|\nabla^n R_K\|_{L^\infty(\tilde{K})} \leq C_{\text{metric}} \rho^n n! \qquad \forall n \in \mathbb{N}_0. \end{aligned}$$

Here, $\tilde{K} = A_K(\widehat{K})$ and $h_K > 0$ denotes the element diameter.

On the reference element $\widehat{K}$ we introduce the Raviart-Thomas and Brezzi-Douglas-Marini elements:

$$\begin{aligned} \mathcal{P}_p(\widehat{K}) &:= \operatorname{span}\{\boldsymbol{x}^{\boldsymbol{\alpha}} \colon |\boldsymbol{\alpha}| \leq p\}, \\ \mathbf{BDM}_p(\widehat{K}) &:= \mathcal{P}_p(\widehat{K})^d, \\ \mathbf{RT}_{p-1}(\widehat{K}) &:= \left\{\boldsymbol{p} + \boldsymbol{x} q \colon \boldsymbol{p} \in \mathcal{P}_{p-1}(\widehat{K})^d, q \in \mathcal{P}_{p-1}(\widehat{K})\right\}. \end{aligned}$$

Note that trivially $\mathbf{RT}_{p-1}(\widehat{K}) \subset \mathbf{BDM}_p(\widehat{K}) \subset \mathbf{RT}_p(\widehat{K})$. We also recall the classical Piola transformation, which is the appropriate change of variables for $\boldsymbol{H}(\Omega, \operatorname{div})$. For a function $\boldsymbol{\varphi} : K \to \mathbb{R}^d$ and the element map $F_K\colon \widehat{K} \to K$ its Piola transformation $\widehat{\boldsymbol{\varphi}} : \widehat{K} \to \mathbb{R}^d$ is given by

$$\widehat{\boldsymbol{\varphi}} = (\det F'_K)(F'_K)^{-1} \boldsymbol{\varphi} \circ F_K.$$

We consider the following global finite element spaces:

$$\begin{aligned} &S_{p_s}(\mathcal{T}_h) \subseteq H^1(\Omega), &&\mathbf{N}_{p_v}(\mathcal{T}_h) \subseteq \boldsymbol{H}(\Omega, \operatorname{curl}), &&\mathbf{RT}_{p_v-1}(\mathcal{T}_h) \subseteq \mathbf{BDM}_{p_v}(\mathcal{T}_h) \subseteq \boldsymbol{H}(\Omega, \operatorname{div}), \\ &S^D_{p_s}(\mathcal{T}_h) \subseteq H^1_D(\Omega), &&\mathbf{N}^N_{p_v}(\mathcal{T}_h) \subseteq \boldsymbol{H}_N(\Omega, \operatorname{curl}), &&\mathbf{RT}^N_{p_v-1}(\mathcal{T}_h) \subseteq \mathbf{BDM}^N_{p_v}(\mathcal{T}_h) \subseteq \boldsymbol{H}_N(\Omega, \operatorname{div}), \\ &S^0_{p_s}(\mathcal{T}_h) \subseteq H^1_0(\Omega), &&\mathbf{N}^0_{p_v}(\mathcal{T}_h) \subseteq \boldsymbol{H}_0(\Omega, \operatorname{curl}), &&\mathbf{RT}^0_{p_v-1}(\mathcal{T}_h) \subseteq \mathbf{BDM}^0_{p_v}(\mathcal{T}_h) \subseteq \boldsymbol{H}_0(\Omega, \operatorname{div}). \end{aligned}$$

The spaces $S_p(\mathcal{T}_h)$, $\mathbf{BDM}_p(\mathcal{T}_h)$, and $\mathbf{RT}_{p-1}(\mathcal{T}_h)$ are given by standard transformation and (contravariant) Piola transformation of functions on the reference element:

$$\begin{aligned} S_p(\mathcal{T}_h) &:= \left\{u \in H^1(\Omega) \colon \left.u\right|_K \circ F_K \in \mathcal{P}_p(\widehat{K}) \text{ for all } K \in \mathcal{T}_h\right\}, \\ \mathbf{BDM}_p(\mathcal{T}_h) &:= \left\{\boldsymbol{\varphi} \in \boldsymbol{H}(\operatorname{div}, \Omega) \colon (\det F'_K)(F'_K)^{-1} \left.\boldsymbol{\varphi}\right|_K \circ F_K \in \mathbf{BDM}_p(\widehat{K}) \text{ for all } K \in \mathcal{T}_h\right\}, \\ \mathbf{RT}_{p-1}(\mathcal{T}_h) &:= \left\{\boldsymbol{\varphi} \in \boldsymbol{H}(\operatorname{div}, \Omega) \colon (\det F'_K)(F'_K)^{-1} \left.\boldsymbol{\varphi}\right|_K \circ F_K \in \mathbf{RT}_{p-1}(\widehat{K}) \text{ for all } K \in \mathcal{T}_h\right\}, \end{aligned}$$

where the polynomial approximation of the scalar and vector variable is denoted by $p_s \geq 1$ and $p_v \geq 1$, respectively. For brevity we also denote by $\mathbf{V}_{p_v}(\mathcal{T}_h)$ either the space $\mathbf{RT}_{p_v-1}(\mathcal{T}_h)$ or $\mathbf{BDM}_{p_v}(\mathcal{T}_h)$. The spaces $\mathbf{V}^N_{p_v}(\mathcal{T}_h)$ and $\mathbf{V}^0_{p_v}(\mathcal{T}_h)$ are denoted analogously. Furthermore, the Nédélec space $\mathbf{N}_{p_v}(\mathcal{T}_h)$ is either of type one or two, depending on the choice of $\mathbf{V}_{p_v}(\mathcal{T}_h)$. The same convention applies to spaces with boundary conditions. See again [Mon03, BBF13] for further details. Further notational conventions will be:

- lower case Roman letters like u and v will be reserved for scalar valued functions;

- lower case boldface Greek letters like $\boldsymbol{\varphi}$ and $\boldsymbol{\psi}$ will be reserved for vector valued functions;
- K denotes the physical element and $\widehat{K}$ denotes the reference element;
- quantities without a $\widehat{\cdot}$ will be either global quantities or quantities defined on the physical element K, whereas quantities with a $\widehat{\cdot}$ are related to the reference element $\widehat{K}$;
- quantities like u_h and $\boldsymbol{\varphi}_h$ will be reserved for functions from the corresponding finite element space, again scalar and vector valued, respectively;
- if not stated otherwise discrete functions without a $\tilde{\cdot}$ will be in some sense fixed, e.g., resulting from a certain discretization scheme, whereas functions with a $\tilde{\cdot}$ will be arbitrary, e.g., when dealing with quasi-optimality results;
- generic constants will either be denoted by C or hidden inside a $\lesssim$ and will be independent of the wavenumber k, the mesh size h and the polynomial degree p, if not otherwise stated.

3. First order system least squares method for homogeneous Helmholtz problems

In the present chapter we analyze the *hp* version of a first order system least squares method applied to the homogeneous Helmholtz equation in a high-frequency regime. The homogeneous Helmholtz equation is reformulated as a minimization problem corresponding to a first order system of equations. The results of the current chapter are part of [BM19] motivated by the work [CQ17].
The outline of this chapter is as follows. In Section 3.1 we introduce a homogeneous Helmholtz model problem. In Section 3.2 we present the first order system least squares (FOSLS) method itself, followed by Section 3.3, where we prove a refined and wavenumber-explicit duality argument for the $L^2(\Omega)$ norm of the scalar variable (Lemma 3.3.1), which is later used to derive an *a priori* estimate (Theorem 3.5.1) of the method. Key ingredients are the results of [MPS13], where a frequency explicit splitting of the solution to our model problem (3.1) is performed when the data has higher order Sobolev regularity. Section 3.4 is concerned with the approximation properties of Raviart-Thomas and Brezzi-Douglas-Marini spaces. We follow the methodology of [MS10] in order to construct approximation operators, which are not only p-optimal and approximate simultaneously in $L^2(\Omega)$ and $H^1(\Omega)$, but also admit an elementwise construction. Section 3.5 is then devoted to the *a priori* estimate. Concluding, we give numerical examples which complement the theoretical findings and compare the method to the classical FEM in Section 3.6.

3.1. Model problem

In the present chapter we consider the following homogeneous Helmholtz problem:

$$\begin{aligned} -\Delta u - k^2 u &= f \quad \text{in } \Omega, \\ \partial_n u - iku &= g \quad \text{on } \Gamma, \end{aligned} \tag{3.1}$$

where $k \geq k_0 > 0$ is real. For a general discussion and presentation of current results concerning the numerical discretization of (3.1) we refer to Section 1.1. Throughout this chapter, if not otherwise stated, we assume the following:

Assumption 3.1.1. In spatial dimension $d = 2, 3$ the bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$ has an analytic boundary $\Gamma := \partial\Omega$. The wavenumber k satisfies $k \geq k_0 > 0$. Furthermore, $f \in L^2(\Omega)$ and $g \in L^2(\Gamma)$.

Remark 3.1.2. Under Assumption 3.1.1 we may apply [BSW16, Thm. 1.8] to conclude that the solution $u \in H^1(\Omega)$ satisfies the *a priori* bound

$$\|u\|_{1,\Omega} + k\|u\|_{0,\Omega} \lesssim \|f\|_{0,\Omega} + \|g\|_{0,\Gamma}, \tag{3.2}$$

with hidden constant independent of k. ∎

3.2. First order system least squares method and auxiliary results

In the present section we introduce the method of [CQ17] and list some auxiliary results for later reference.

3.2.1. First order system least squares formulation

Starting from the second order formulation (3.1) we introduce the additional variable $\boldsymbol{\varphi} = ik^{-1}\nabla u$ to formally arrive at the first order system of equations

$$\nabla \cdot \boldsymbol{\varphi} + iku = -ik^{-1}f \quad \text{in } \Omega, \tag{3.3a}$$
$$\nabla u + ik\boldsymbol{\varphi} = \mathbf{0} \quad \text{in } \Omega, \tag{3.3b}$$
$$k^{1/2}(\boldsymbol{\varphi} \cdot \boldsymbol{n} + u) = ik^{-1/2}g \quad \text{on } \Gamma. \tag{3.3c}$$

In the following we employ the complex Hilbert spaces

$$\boldsymbol{V} = \{\boldsymbol{\varphi} \in \boldsymbol{H}(\Omega, \operatorname{div})\colon \boldsymbol{\varphi} \cdot \boldsymbol{n} \in L^2(\Gamma)\} \quad \text{and} \quad W = H^1(\Omega),$$

where $\boldsymbol{V}$ is endowed with the usual graph norm and W with the classical $H^1(\Omega)$-norm. On $\boldsymbol{V} \times W$ we introduce the sesquilinear form b and the functional F by

$$\begin{aligned} b((\boldsymbol{\varphi}, u), (\boldsymbol{\psi}, v)) &:= (\nabla \cdot \boldsymbol{\varphi} + iku, \nabla \cdot \boldsymbol{\psi} + ikv)_\Omega + (\nabla u + ik\boldsymbol{\varphi}, \nabla v + ik\boldsymbol{\psi})_\Omega + \\ &\quad\; k\langle \boldsymbol{\varphi} \cdot \boldsymbol{n} + u, \boldsymbol{\psi} \cdot \boldsymbol{n} + v\rangle_\Gamma, \\ F((\boldsymbol{\psi}, v)) &:= (-ik^{-1}f, \nabla \cdot \boldsymbol{\psi} + ikv)_\Omega + \langle ig, \boldsymbol{\psi} \cdot \boldsymbol{n} + v\rangle_\Gamma. \end{aligned}$$

If $u \in H^1(\Omega)$ is the weak solution to (3.1) then the pair $(\boldsymbol{\varphi}, u)$ with $\boldsymbol{\varphi} = ik^{-1}\nabla u$ is in fact in $\boldsymbol{V} \times W$ due to the assumed regularity of the data and the domain and therefore satisfies

$$b((\boldsymbol{\varphi}, u), (\boldsymbol{\psi}, v)) = F((\boldsymbol{\psi}, v)) \quad \forall (\boldsymbol{\psi}, v) \in \boldsymbol{V} \times W. \tag{3.4}$$

For a given regular mesh $\mathcal{T}_h$ we consider the finite element spaces $\boldsymbol{V}_h = \mathbf{RT}_p(\mathcal{T}_h) \subset \boldsymbol{V}$ or $\boldsymbol{V}_h = \mathbf{BDM}_p(\mathcal{T}_h) \subset \boldsymbol{V}$ and $W_h = S_p(\mathcal{T}_h) \subset W$, where $\mathbf{RT}_p(\mathcal{T}_h)$ denotes the Raviart-Thomas space and $\mathbf{BDM}_p(\mathcal{T}_h)$ the Brezzi-Douglas-Marini space; see Chapter 2 for further detail and definition as well as Section 3.4 for further approximation theoretical results. The FOSLS method is to find $(\boldsymbol{\varphi}_h, u_h) \in \boldsymbol{V}_h \times W_h$ such that

$$b((\boldsymbol{\varphi}_h, u_h), (\boldsymbol{\psi}_h, v_h)) = F((\boldsymbol{\psi}_h, v_h)) \quad \forall (\boldsymbol{\psi}_h, v_h) \in \boldsymbol{V}_h \times W_h. \tag{3.5}$$

Remark 3.2.1. Based on the *a priori* estimate (3.2) reference [CQ17, Thm. 2.4] asserts the estimate

$$\|\boldsymbol{\varphi}\|_{0,\Omega}^2 + \|u\|_{0,\Omega}^2 + k\|\boldsymbol{\varphi} \cdot \boldsymbol{n} + u\|_{0,\Gamma}^2 \lesssim b((\boldsymbol{\varphi}, u), (\boldsymbol{\varphi}, u)) \quad \forall (\boldsymbol{\varphi}, u) \in \boldsymbol{V} \times W,$$

with hidden constant independent of k, which immediately gives uniqueness. Together with the fact that the pair $(\boldsymbol{\varphi}, u)$ with $\boldsymbol{\varphi} = ik^{-1}\nabla u$ is a solution, we have unique solvability of (3.4). ∎

3.2.2. Auxiliary results

Our refined duality argument in Lemma 3.3.1 hinges on the following decomposition result.

Proposition 3.2.2 ([MPS13, Thm. 4.5] combined with [BSW16, Thm. 1.8])**.** *Let $\Omega \subset \mathbb{R}^d, d \in \{2,3\}$, be a bounded Lipschitz domain with an analytic boundary Γ. Fix $s \in \mathbb{N}_0$. Then there exists a constant $\gamma > 0$ independent of k such that for every $f \in H^s(\Omega)$ and $g \in H^{s+1/2}(\Gamma)$ the solution u of* (3.1) *can be written as $u = u_A + u_{H^{s+2}}$, where, for all $n \in \mathbb{N}_0$, there holds*

$$\|u_A\|_{1,\Omega} + k\|u_A\|_{0,\Omega} \lesssim \|f\|_{0,\Omega} + \|g\|_{1/2,\Gamma}, \tag{3.6a}$$

$$\|\nabla^{n+2}u_A\|_{0,\Omega} \lesssim k^{-1}\gamma^n \max\{n,k\}^{n+2} (\|f\|_{0,\Omega} + \|g\|_{1/2,\Gamma}), \tag{3.6b}$$

$$\|u_{H^{s+2}}\|_{s+2,\Omega} + k^{s+2}\|u_{H^{s+2}}\|_{0,\Omega} \lesssim \|f\|_{s,\Omega} + \|g\|_{s+1/2,\Gamma}. \tag{3.6c}$$

Remark 3.2.3. Interpolation between $L^2(\Omega)$ and $H^{s+2}(\Omega)$ in Proposition 3.2.2 gives estimates for other Sobolev norms: Since we have for any $v \in H^m(\Omega)$

$$\|v\|_{j,\Omega} \lesssim \|v\|_{m,\Omega}^{\frac{j}{m}} \|v\|_{0,\Omega}^{\frac{m-j}{m}}, \qquad j \in \{0,\dots,m\},$$

Proposition 3.2.2 implies for $j \in \{0,\dots,s+2\}$

$$k^{s+2-j}\|u_{H^{s+2}}\|_{j,\Omega} \lesssim \|f\|_{s,\Omega} + \|g\|_{s+1/2,\Gamma}.$$

■

Furthermore, we often use the multiplicative trace inequality. We remind the reader of the general form, even though we only need it in the special case $s = 1$.

Proposition 3.2.4 ([Mel05, Thm. A.2])**.** *Let $\Omega \subset \mathbb{R}^d$ be a Lipschitz domain and $s \in (1/2, 1]$. Then there exists a constant $C > 0$ such that for all $u \in H^s(\Omega)$ there holds*

$$\|u\|_{0,\Gamma} \le C\|u\|_{0,\Omega}^{1-1/(2s)}\|u\|_{s,\Omega}^{1/(2s)},$$

where the left-hand side is understood in the trace sense.

3.3. Duality argument

We extend the results of [CQ17, Lemma 5.1]. To that end, we show that the function $\boldsymbol{\psi}_{H^2} \in \boldsymbol{H}^1(\mathrm{div},\Omega)$, constructed therein, can actually be modified to satisfy $\boldsymbol{\psi}_{H^2} \in \boldsymbol{H}^2(\Omega)$ and still allow for wavenumber-explicit higher order Sobolev norm estimates.

Lemma 3.3.1. *For any $(\boldsymbol{\varphi}, w) \in \boldsymbol{V} \times W$ there exists $(\boldsymbol{\psi}, v) \in \boldsymbol{V} \times W$ such that $\|w\|_{0,\Omega}^2 = b((\boldsymbol{\varphi}, w), (\boldsymbol{\psi}, v))$. The pair $(\boldsymbol{\psi}, v)$ admits a decomposition $\boldsymbol{\psi} = \boldsymbol{\psi}_A + \boldsymbol{\psi}_{H^2}$, $v = v_A + v_{H^2}$, where $\boldsymbol{\psi}_A$ and v_A are analytic in Ω, $\boldsymbol{\psi}_{H^2} \in \boldsymbol{H}^2(\Omega)$, and $v_{H^2} \in H^2(\Omega)$. Furthermore, there exists a constant $\gamma > 0$ independent of k such that for all $n \in \mathbb{N}_0$*

$$\|\boldsymbol{\psi}_A\|_{1,\Omega} + k\|\boldsymbol{\psi}_A\|_{0,\Omega} \lesssim k\|w\|_{0,\Omega}, \tag{3.7a}$$

$$\|v_A\|_{1,\Omega} + k\|v_A\|_{0,\Omega} \lesssim k\|w\|_{0,\Omega}, \tag{3.7b}$$

$$\|\nabla^{n+2}\boldsymbol{\psi}_A\|_{0,\Omega} + \|\nabla^{n+2}v_A\|_{0,\Omega} \lesssim \gamma^n \max\{n,k\}^{n+2} \|w\|_{0,\Omega}, \tag{3.7c}$$

$$\|\boldsymbol{\psi}_{H^2}\|_{2,\Omega} + k\|\boldsymbol{\psi}_{H^2}\|_{1,\Omega} + k^2\|\boldsymbol{\psi}_{H^2}\|_{0,\Omega} \lesssim \|w\|_{0,\Omega}, \tag{3.7d}$$

$$\|v_{H^2}\|_{2,\Omega} + k\|v_{H^2}\|_{1,\Omega} + k^2\|v_{H^2}\|_{0,\Omega} \lesssim \|w\|_{0,\Omega}. \tag{3.7e}$$

Proof. The proof follows the ideas of [CQ17, Lemma 5.1]; for the readers' convenience we recapitulate the important steps of the proof. The novelty over [CQ17] is the ability to choose $\boldsymbol{\psi}_{H^2} \in \boldsymbol{H}^2(\Omega)$ together with $\|\boldsymbol{\psi}_{H^2}\|_{2,\Omega} \lesssim \|w\|_{0,\Omega}$.

Consider the problem

$$\begin{aligned} -\Delta z - k^2 z &= w \qquad \text{in } \Omega, \\ \partial_n z + ikz &= 0 \qquad \text{on } \Gamma. \end{aligned}$$

For any $\boldsymbol{\varphi} \in \boldsymbol{V}$ we have, using the weak formulation and integrating by parts,

$$\begin{aligned} \|w\|_{0,\Omega}^2 &= (\nabla w, \nabla z)_\Omega - k^2 (w, z)_\Omega - ik\langle w, z\rangle_\Gamma \\ &= (ik\boldsymbol{\varphi} + \nabla w, \nabla z)_\Omega - (ik\boldsymbol{\varphi}, \nabla z)_\Omega - k^2(w, z)_\Omega - ik\langle w, z\rangle_\Gamma \\ &= (ik\boldsymbol{\varphi} + \nabla w, \nabla z)_\Omega + (\nabla \cdot \boldsymbol{\varphi} + ikw, -ikz)_\Omega + \langle \boldsymbol{\varphi} \cdot \boldsymbol{n} + w, ikz\rangle_\Gamma. \end{aligned}$$

Applying Proposition 3.2.2 together with Remark 3.2.3 we decompose z into $z = z_A + z_{H^2}$ with z_A analytic and $z_{H^2} \in H^2(\Omega)$. Furthermore, we have for all $n \in \mathbb{N}_0$,

$$\|z_A\|_{1,\Omega} + k\|z_A\|_{0,\Omega} \lesssim \|w\|_{0,\Omega}, \tag{3.8a}$$

$$\|\nabla^{n+2} z_A\|_{0,\Omega} \lesssim k^{-1}\gamma^n \max\{n, k\}^{n+2} \|w\|_{0,\Omega}, \tag{3.8b}$$

$$\|z_{H^2}\|_{2,\Omega} + k\|z_{H^2}\|_{1,\Omega} + k^2\|z_{H^2}\|_{0,\Omega} \lesssim \|w\|_{0,\Omega}. \tag{3.8c}$$

Let $(\boldsymbol{\psi}, v) \in \boldsymbol{V} \times W$ solve

$$\begin{aligned} \nabla \cdot \boldsymbol{\psi} + ikv &= -ikz \qquad \text{in } \Omega, \\ \nabla v + ik\boldsymbol{\psi} &= \nabla z \qquad \text{in } \Omega, \\ k^{1/2}(\boldsymbol{\psi} \cdot \boldsymbol{n} + v) &= ik^{1/2} z \qquad \text{on } \Gamma. \end{aligned}$$

Indeed, this system is uniquely solvable by Remark 3.2.1. This gives the desired representation such that $\|w\|_{0,\Omega}^2 = b((\boldsymbol{\varphi}, w), (\boldsymbol{\psi}, v))$. Using the decomposition $z = z_A + z_{H^2}$ we obtain $(\boldsymbol{\psi}, v) = (\tilde{\boldsymbol{\psi}}_A, \tilde{v}_A) + (\tilde{\boldsymbol{\psi}}_{H^2}, \tilde{v}_{H^2})$, where

$$\begin{aligned} \nabla \cdot \tilde{\boldsymbol{\psi}}_A + ik\tilde{v}_A &= -ikz_A \qquad \text{in } \Omega, \\ \nabla \tilde{v}_A + ik\tilde{\boldsymbol{\psi}}_A &= \nabla z_A \qquad \text{in } \Omega, \\ k^{1/2}(\tilde{\boldsymbol{\psi}}_A \cdot \boldsymbol{n} + \tilde{v}_A) &= ik^{1/2} z_A \qquad \text{on } \Gamma, \end{aligned}$$

$$\begin{aligned} \nabla \cdot \tilde{\boldsymbol{\psi}}_{H^2} + ik\tilde{v}_{H^2} &= -ikz_{H^2} \qquad \text{in } \Omega, \\ \nabla \tilde{v}_{H^2} + ik\tilde{\boldsymbol{\psi}}_{H^2} &= \nabla z_{H^2} \qquad \text{in } \Omega, \\ k^{1/2}(\tilde{\boldsymbol{\psi}}_{H^2} \cdot \boldsymbol{n} + \tilde{v}_{H^2}) &= ik^{1/2} z_{H^2} \qquad \text{on } \Gamma. \end{aligned}$$

One can immediately verify that

$$\begin{aligned} -\Delta(\tilde{v}_A - z_A) - k^2(\tilde{v}_A - z_A) &= 2k^2 z_A \qquad \text{in } \Omega, \\ \partial_n(\tilde{v}_A - z_A) - ik(\tilde{v}_A - z_A) &= (1+i)kz_A \qquad \text{on } \Gamma, \end{aligned} \tag{3.9}$$

as well as

$$\begin{aligned} -\Delta(\tilde{v}_{H^2} - z_{H^2}) - k^2(\tilde{v}_{H^2} - z_{H^2}) &= 2k^2 z_{H^2} \qquad \text{in } \Omega, \\ \partial_n(\tilde{v}_{H^2} - z_{H^2}) - ik(\tilde{v}_{H^2} - z_{H^2}) &= (1+i)kz_{H^2} \qquad \text{on } \Gamma. \end{aligned} \tag{3.10}$$

Note that the right-hand sides in equation (3.9) are analytic. This fact is used in [CQ17, Lemmas 4.4 and 5.1] to prove the following bounds for all $n \in \mathbb{N}_0$:

$$\|\nabla^{n+2}\tilde{v}_A\|_{0,\Omega} \lesssim \gamma^n \max\{n,k\}^{n+2}\|w\|_{0,\Omega}, \tag{3.11a}$$
$$\|\tilde{v}_A\|_{1,\Omega} + k\|\tilde{v}_A\|_{0,\Omega} \lesssim k\|w\|_{0,\Omega}, \tag{3.11b}$$
$$\|\nabla^{n+2}\tilde{\boldsymbol{\psi}}_A\|_{0,\Omega} \lesssim \gamma^n \max\{n,k\}^{n+2}\|w\|_{0,\Omega}, \tag{3.11c}$$
$$\|\tilde{\boldsymbol{\psi}}_A\|_{0,\Omega} + k\|\tilde{\boldsymbol{\psi}}_A\|_{0,\Omega} \lesssim k\|w\|_{0,\Omega}. \tag{3.11d}$$

Since $\tilde{v}_{H^2} - z_{H^2} = S_k^-(2k^2 z_{H^2}, (1+i)kz_{H^2})$, where S_k^- denotes the solution operator for (3.1), we can exploit the regularity of the right-hand sides in equation (3.10). Applying Proposition 3.2.2 with $s=1$ as well as Remark 3.2.3 we decompose $\tilde{v}_{H^2} - z_{H^2} = \hat{v}_A + \hat{v}_{H^3}$, where $\hat{v}_A$ is analytic and $\hat{v}_{H^3} \in H^3(\Omega)$. For every $j \in \{0,1,2,3\}$ we have

$$\begin{aligned} k^{3-j}\|\hat{v}_{H^3}\|_{j,\Omega} &\lesssim \|2k^2 z_{H^2}\|_{1,\Omega} + \|(1+i)kz_{H^2}\|_{3/2,\Gamma} \\ &\lesssim \underbrace{k^2\|z_{H^2}\|_{1,\Omega}}_{\overset{(3.8c)}{\lesssim} k\|w\|_{0,\Omega}} + \underbrace{k\|z_{H^2}\|_{3/2,\Gamma}}_{\lesssim k\|z_{H^2}\|_{2,\Omega} \overset{(3.8c)}{\lesssim} k\|w\|_{0,\Omega}} \\ &\lesssim k\|w\|_{0,\Omega}. \end{aligned}$$

Summarizing the above we have

$$k^{-1}\|\hat{v}_{H^3}\|_{3,\Omega} + \|\hat{v}_{H^3}\|_{2,\Omega} + k\|\hat{v}_{H^3}\|_{1,\Omega} + k^2\|\hat{v}_{H^3}\|_{0,\Omega} \lesssim \|w\|_{0,\Omega}. \tag{3.12}$$

In order to analyze the behavior of $\hat{v}_A$ we first estimate

$$\|2k^2 z_{H^2}\|_{0,\Omega} + \|(1+i)kz_{H^2}\|_{1/2,\Gamma} \overset{(3.8c)}{\lesssim} \|w\|_{0,\Omega}.$$

We therefore conclude, again with Proposition 3.2.2, that

$$\|\hat{v}_A\|_{1,\Omega} + k\|\hat{v}_A\|_{0,\Omega} \lesssim \|w\|_{0,\Omega}, \tag{3.13a}$$
$$\|\nabla^{n+2}\hat{v}_A\|_{0,\Omega} \lesssim k^{-1}\gamma^n \max\{n,k\}^{n+2}\|w\|_{0,\Omega}. \tag{3.13b}$$

We turn to the final decompositions with associated norm bounds.
Final decomposition of v:

$$v = \tilde{v}_A + \tilde{v}_{H^2} = \tilde{v}_A + \underbrace{\tilde{v}_{H^2} - z_{H^2}}_{=\hat{v}_A + \hat{v}_{H^3}} + z_{H^2} = \underbrace{\tilde{v}_A + \hat{v}_A}_{=:v_A} + \underbrace{\hat{v}_{H^3} + z_{H^2}}_{=:v_{H^2}}.$$

Verification of (3.7b):

$$\begin{aligned} \|v_A\|_{1,\Omega} + k\|v_A\|_{0,\Omega} &\le \underbrace{\|\tilde{v}_A\|_{1,\Omega} + k\|\tilde{v}_A\|_{0,\Omega}}_{\overset{(3.11b)}{\lesssim} k\|w\|_{0,\Omega}} + \underbrace{\|\hat{v}_A\|_{1,\Omega} + k\|\hat{v}_A\|_{0,\Omega}}_{\overset{(3.13a)}{\lesssim} \|w\|_{0,\Omega}} \\ &\lesssim k\|w\|_{0,\Omega}. \end{aligned}$$

Verification of (3.7e):

$$
\begin{aligned}
&\|v_{H^2}\|_{2,\Omega} + k\|v_{H^2}\|_{1,\Omega} + k^2\|v_{H^2}\|_{0,\Omega} \\
&\qquad \leq \underbrace{\|\hat{v}_{H^3}\|_{2,\Omega} + k\|\hat{v}_{H^3}\|_{1,\Omega} + k^2\|\hat{v}_{H^3}\|_{0,\Omega}}_{\overset{(3.12)}{\lesssim}\ \|w\|_{0,\Omega}} \\
&\qquad\quad + \underbrace{\|z_{H^2}\|_{2,\Omega} + k\|z_{H^2}\|_{1,\Omega} + k^2\|z_{H^2}\|_{0,\Omega}}_{\overset{(3.8c)}{\lesssim}\ \|w\|_{0,\Omega}} \\
&\qquad \lesssim \|w\|_{0,\Omega}.
\end{aligned}
$$

Final decomposition of $\boldsymbol{\psi}$: Since $-ik\tilde{\boldsymbol{\psi}}_{H^2} = \nabla(\tilde{v}_{H^2} - z_{H^2}) = \nabla\hat{v}_A + \nabla\hat{v}_{H^3}$, we decompose $\tilde{\boldsymbol{\psi}}_{H^2} = \hat{\boldsymbol{\psi}}_A + \hat{\boldsymbol{\psi}}_{H^2}$ accordingly such that $-ik\hat{\boldsymbol{\psi}}_A = \nabla\hat{v}_A$ and consequently $-ik\hat{\boldsymbol{\psi}}_{H^2} = \nabla\hat{v}_{H^3}$. The final decomposition takes the form

$$
\boldsymbol{\psi} = \tilde{\boldsymbol{\psi}}_A + \tilde{\boldsymbol{\psi}}_{H^2} = \underbrace{\tilde{\boldsymbol{\psi}}_A + \hat{\boldsymbol{\psi}}_A}_{=:\boldsymbol{\psi}_A} + \underbrace{\hat{\boldsymbol{\psi}}_{H^2}}_{=:\boldsymbol{\psi}_{H^2}}.
$$

Verification of (3.7a):

$$
\begin{aligned}
&\|\boldsymbol{\psi}_A\|_{1,\Omega} + k\|\boldsymbol{\psi}_A\|_{0,\Omega} \\
&\qquad \leq \underbrace{\|\tilde{\boldsymbol{\psi}}_A\|_{1,\Omega} + k\|\tilde{\boldsymbol{\psi}}_A\|_{0,\Omega}}_{\overset{(3.11d)}{\lesssim}\ k\|w\|_{0,\Omega}} + \|\hat{\boldsymbol{\psi}}_A\|_{1,\Omega} + k\|\hat{\boldsymbol{\psi}}_A\|_{0,\Omega} \\
&\qquad \lesssim k\|w\|_{0,\Omega} + k^{-1}\|\nabla\hat{v}_A\|_{1,\Omega} + \|\nabla\hat{v}_A\|_{0,\Omega} \\
&\qquad \lesssim k\|w\|_{0,\Omega} + k^{-1} \underbrace{\|\hat{v}_A\|_{1,\Omega}}_{\overset{(3.13a)}{\lesssim}\ \|w\|_{0,\Omega}} + k^{-1} \underbrace{\|\nabla^2\hat{v}_A\|_{0,\Omega}}_{\overset{(3.13b)}{\lesssim}\ k\|w\|_{0,\Omega}} + \underbrace{\|\hat{v}_A\|_{1,\Omega}}_{\overset{(3.13a)}{\lesssim}\ \|w\|_{0,\Omega}} \\
&\qquad \lesssim k\|w\|_{0,\Omega}.
\end{aligned}
$$

Verification of (3.7c): This is an immediate consequence of (3.11a), (3.11c), (3.13b), and the fact that $-ik\hat{\boldsymbol{\psi}}_A = \nabla\hat{v}_A$.

Verification of (3.7d): Since $-ik\hat{\boldsymbol{\psi}}_{H^2} = \nabla\hat{v}_{H^3}$ we estimate

$$
\begin{aligned}
&\|\boldsymbol{\psi}_{H^2}\|_{2,\Omega} + k\|\boldsymbol{\psi}_{H^2}\|_{1,\Omega} + k^2\|\boldsymbol{\psi}_{H^2}\|_{0,\Omega} \\
&\qquad = k^{-1}\|\nabla\hat{v}_{H^3}\|_{2,\Omega} + \|\nabla\hat{v}_{H^3}\|_{1,\Omega} + k\|\nabla\hat{v}_{H^3}\|_{0,\Omega} \\
&\qquad \leq \underbrace{k^{-1}\|\hat{v}_{H^3}\|_{3,\Omega} + \|\hat{v}_{H^3}\|_{2,\Omega} + k\|\hat{v}_{H^3}\|_{1,\Omega}}_{\overset{(3.12)}{\lesssim}\ \|w\|_{0,\Omega}} \\
&\qquad \lesssim \|w\|_{0,\Omega},
\end{aligned}
$$

which concludes the proof.

□

3.4. Approximation properties of RT and BDM spaces

In the present section we analyze the approximation properties of Raviart-Thomas and Brezzi-Douglas-Marini spaces. To that end, we will prove the existence of a polynomial approximation operator acting on functions defined on the reference element having certain desirable properties, as outlined below. This operator will then be used to construct a global polynomial approximation operator by means of the Piola transformation.

3.4.1. Preliminaries

For the remainder of this chapter we assume Assumption 2.0.1 to be satisfied. We recall the definition of the Sobolev space $H^{1/2}_{00}(\omega)$. If ω is an edge of a triangle or face of a tetrahedron, then the norm $\|\cdot\|_{H^{1/2}_{00}(\omega)}$ is given by

$$\|u\|^2_{H^{1/2}_{00}(\omega)} := \|u\|^2_{1/2,\omega} + \left\| \frac{u}{\sqrt{\operatorname{dist}(\cdot,\partial\omega)}} \right\|^2_{0,\omega},$$

and the space $H^{1/2}_{00}(\omega)$ is the completion of $C^\infty_0(\omega)$ under this norm. Since this norm is induced by a scalar product the space $H^{1/2}_{00}(\omega)$ is a Hilbert space.

3.4.2. Polynomial approximation on the reference element

We construct a polynomial approximation operator on the reference element $\widehat{K}$:

Definition 3.4.1. Let $\widehat{K}$ be the reference simplex in $\mathbb{R}^d$, $s > d/2$ and $p \in \mathbb{N}$. We define the operator $\widehat{\Pi}_p : H^s(\widehat{K}) \to \mathcal{P}_p(\widehat{K})$ by the following consecutive minimization steps:

(i) Fix $\widehat{\Pi}_p u$ in the vertices: $(\widehat{\Pi}_p u)(\widehat{V}) = u(\widehat{V})$ for all $d+1$ vertices $\widehat{V}$ of $\widehat{K}$.

(ii) Fix $\widehat{\Pi}_p u$ on the edges: for every edge $\hat{e}$ of $\widehat{K}$ the restriction $(\widehat{\Pi} u_p)\big|_{\hat{e}}$ is the unique minimizer of

$$\mathcal{P}_p(\hat{e}) \ni \pi \mapsto p\|u-\pi\|^2_{0,\hat{e}} + \|u-\pi\|^2_{H^{1/2}_{00}(\hat{e})}, \qquad \text{s.t. } \pi \text{ satisfies } (i). \tag{3.14}$$

(iii) Fix $\widehat{\Pi}_p u$ on the faces (only for $d=3$): for every face $\hat{f}$ of $\widehat{K}$ the restriction $(\widehat{\Pi} u_p)\big|_{\hat{f}}$ is the unique minimizer of

$$\mathcal{P}_p(\hat{f}) \ni \pi \mapsto p^2\|u-\pi\|^2_{0,\hat{f}} + \|u-\pi\|^2_{1,\hat{f}}, \qquad \text{s.t. } \pi \text{ satisfies } (i),\ (ii). \tag{3.15}$$

(iv) Fix $\widehat{\Pi}_p u$ in the volume: $\widehat{\Pi}_p u$ is the unique minimizer of

$$\mathcal{P}_p(\widehat{K}) \ni \pi \mapsto p^2\|u-\pi\|^2_{0,\widehat{K}} + \|u-\pi\|^2_{1,\widehat{K}}, \qquad \text{s.t. } \pi \text{ satisfies } (i),\ (ii),\ (iii). \tag{3.16}$$

It is convenient to construct an approximant Iu of a function u in an elementwise fashion. The drawback is that one has to check if the approximant is in fact in the finite element space. A useful property to achieve this is the following: The restriction of the approximant $Iu\big|_E$ to lower dimensional entities E of the mesh, i.e., edges, faces or vertices, is completely determined by the corresponding restriction of u. To put this rigorously, we employ the following concept:

Definition 3.4.2 (restriction property)**.** Let $\widehat{K}$ be the reference simplex in $\mathbb{R}^d$, $s > d/2$, and $p \in \mathbb{N}$. A polynomial $\pi \in \mathcal{P}_p(\widehat{K})$ is said to satisfy the *restriction property* of polynomial degree p for $u \in H^s(\widehat{K})$, if it satisfies (i), (ii), (iii) of Definition 3.4.1.

Remark 3.4.3. Note that the minimizations in the definition of the operator $\widehat{\Pi}_p$ are uniquely solvable. This is due to the fact that these minimizations are constrained minimizations of norms induced by Hilbert spaces. These constraints are given by an affine subspace $\mathcal{V}_p^u \subset \mathcal{P}_p(\widehat{K})$, the space of all polynomials satisfying the restriction property for u. Step (iv) is therefore the orthogonal projection onto the space $\mathcal{V}_p^u$ with respect to the scalar product inducing the norm

$$|||u|||^2 \coloneqq p^2\|u\|^2_{0,\widehat{K}} + \|u\|^2_{1,\widehat{K}}.$$

Furthermore, the affine space $\mathcal{V}_p^u$ can be written as $\mathcal{V}_p^u = \pi^u + \mathcal{P}_p^0$ for some $\pi^u \in \mathcal{V}_p^u$, where $\mathcal{P}_p^0(\widehat{K}) \subset \mathcal{P}_p(\widehat{K})$ is the space of polynomials vanishing on $\partial\widehat{K}$. The operator $\widehat{\Pi}_p$ can, apart from being the solution to a minimization problem, also be written as:

$$\widehat{\Pi}_p u = \operatorname{argmin}\{|||u-\pi||| : \pi \in \mathcal{V}_p^u\} = \pi^u + \widehat{\Pi}_{\mathcal{P}_p^0}(u-\pi^u), \tag{3.17}$$

where $\widehat{\Pi}_{\mathcal{P}_p^0}$ denotes the orthogonal projection onto the space $\mathcal{P}_p^0(\widehat{K})$, again with respect to the scalar product inducing $|||\cdot|||$. The operator $\widehat{\Pi}_p : H^s(\widehat{K}) \to \mathcal{P}_p(\widehat{K})$ is furthermore linear. This is easily seen when one explicitly constructs the Steps (i), (ii), (iii) in Definition 3.4.1: First, one picks polynomials $\pi_{\widehat{V}}$, which are one at the vertex $\widehat{V}$ and zero on all the others. Consider the mapping $\widehat{\Pi}_{\widehat{V}} : u \mapsto \sum_{\widehat{V}} u(\widehat{V})\pi_{\widehat{V}}$. This realizes Step (i). Next one considers the mapping $\tilde{\Pi}_{\hat{e}} : z \mapsto \operatorname{argmin}\{p\|u-\pi\|^2_{0,\hat{e}} + \|u-\pi\|^2_{H_{00}^{1/2}(\hat{e})} : z(\widehat{V}) = 0 \text{ for all vertices } \widehat{V}\}$ and extending it to the reference element. Step (ii) is then realized by the map $\widehat{\Pi}_{\hat{e}} : u \mapsto \widehat{\Pi}_{\widehat{V}}u + \tilde{\Pi}_{\hat{e}}(u - \widehat{\Pi}_{\widehat{V}}u)$. One can easily continue this procedure for Step (iii) and (iv). As a composition of linear operators $\widehat{\Pi}_p$ is therefore also linear. ∎

Remark 3.4.4. Definition 3.4.2 of the restriction property was introduced in [MS10, Def. 5.3] under the name *element-by-element construction*. This is due to the fact that when working in $S_p(\mathcal{T}_h) \subset H^1(\Omega)$, a polynomial, which is constructed in an elementwise fashion on the reference simplex $\widehat{K}$, satisfying the restriction property is already an element of the conforming element space $S_p(\mathcal{T}_h)$. However, when working in $\boldsymbol{H}(\Omega, \operatorname{div})$ or $\boldsymbol{H}(\Omega, \operatorname{curl})$ one only needs continuity of the inter element normal or tangential trace. Furthermore, it is necessary to use the Piola transformation to go back and forth between the reference element and the physical element to ensure that normal and tangential vectors are mapped appropriately. For the purpose of this chapter we therefore use the name restriction property, rather than element-by-element construction. ∎

In the Propositions 3.4.5, 3.4.7, and 3.4.8 we recall certain useful results concerning approximation properties of polynomials satisfying the restriction property. These results can be found in [MS10].

Proposition 3.4.5 ([MS10, Thm. B.4]). *Let $\widehat{K}$ be the reference triangle or reference tetrahedron. Let $s > d/2$. Then there exists $C > 0$ (depending only on s and d) and for every p a linear operator $\widehat{\Pi}_p^{\mathrm{MS}} \colon H^s(\widehat{K}) \to \mathcal{P}_p(\widehat{K})$, such that $\widehat{\Pi}_p^{\mathrm{MS}} u$ satisfies the restriction property of Definition 3.4.2 as well as*

$$p\|u - \widehat{\Pi}_p^{\mathrm{MS}} u\|_{0,\widehat{K}} + \|u - \widehat{\Pi}_p^{\mathrm{MS}} u\|_{1,\widehat{K}} \leq C p^{-(s-1)} |u|_{s,\widehat{K}} \qquad \forall p \geq s-1. \tag{3.18}$$

Remark 3.4.6. The operator $\widehat{\Pi}_p^{\mathrm{MS}}$ does in general not preserve polynomials $q \in \mathcal{P}_p(\widehat{K})$. See also [MR20] for operators with the projection property. ■

Proposition 3.4.7 ([MS10, Lemma C.2]). *Let $d \in \{1,2,3\}$, and let $\widehat{K} \subset \mathbb{R}^d$ be the reference simplex. Let $\gamma, \tilde{C} > 0$ be given. Then there exist constants $C, \sigma > 0$ that depend solely on γ and $\tilde{C}$ such that the following is true: For any function u that satisfies for some C_u, h, $R > 0$ and $\kappa > 1$ the conditions*

$$\|\nabla^n u\|_{0,\widehat{K}} \leq C_u (\gamma h)^n \max\{n/R, \kappa\}^n \qquad \forall n \in \mathbb{N}_{\geq 2},$$

and for any polynomial degree $p \in \mathbb{N}$ that satisfies

$$\frac{h}{R} + \frac{\kappa h}{p} \leq \tilde{C}$$

there holds

$$\inf_{\pi \in \mathcal{P}_p(\widehat{K})} \|u - \pi\|_{W^{2,\infty}(\widehat{K})} \leq C C_u \left[\left(\frac{h/R}{\sigma + h/R} \right)^{p+1} + \left(\frac{h\kappa}{\sigma p} \right)^{p+1} \right].$$

Proposition 3.4.8 ([MS10, Lemma C.3]). *Assume the hypotheses of Proposition 3.4.7. Then one can find a polynomial $\pi \in \mathcal{P}_p(\widehat{K})$ that satisfies*

$$\|u - \pi\|_{W^{1,\infty}(\widehat{K})} \leq C C_u \left[\left(\frac{h/R}{\sigma + h/R} \right)^{p+1} + \left(\frac{h\kappa}{\sigma p} \right)^{p+1} \right],$$

and additionally satisfies the restriction property of Definition 3.4.2.

It is not clear whether the polynomial $\widehat{\Pi}_p^{\mathrm{MS}} u$ has the same approximation properties as the polynomial given by Proposition 3.4.8. However, it is desirable to have both the simultaneous approximation properties in $L^2(\widehat{K})$ and $H^1(\widehat{K})$ as stated in Proposition 3.4.5 as well as the exponential approximation properties of an analytic function as stated in Proposition 3.4.8. In the following we will show that the operator $\widehat{\Pi}_p$ constructed in Definition 3.4.1 has these properties.

Theorem 3.4.9 (Properties of $\widehat{\Pi}_p$). *Let $\widehat{K}$ be the reference triangle or reference tetrahedron. Let $s > d/2$. Let $\widehat{\Pi}_p \colon H^s(\widehat{K}) \to \mathcal{P}_p(\widehat{K})$ be given by Definition 3.4.1. Then the following holds:*

(i) *The operator* $\widehat{\Pi}_p$ *is linear and satisfies the restriction property of Definition 3.4.2.*

(ii) *The operator* $\widehat{\Pi}_p$ *preserves* $\mathcal{P}_p(\widehat{K})$*, i.e.,* $\widehat{\Pi}_p q = q$ *for all* $q \in \mathcal{P}_p(\widehat{K})$*.*

(iii) *There exists* $C_s > 0$ *(depending only on* s *and* d*) such that*

$$p\|u - \widehat{\Pi}_p u\|_{0,\widehat{K}} + \|u - \widehat{\Pi}_p u\|_{1,\widehat{K}} \leq C_s p^{-(s-1)} |u|_{s,\widehat{K}} \qquad \forall p \geq s-1.$$

(iv) *For given* γ*,* $\tilde{C} > 0$*, there exist constants* C_A*,* $\sigma > 0$ *that depend solely on* γ *and* $\tilde{C}$ *such that the following is true: For any function* u *and polynomial degree* p *that satisfy the assumptions of Proposition 3.4.7 there holds*

$$\|u - \widehat{\Pi}_p u\|_{W^{1,\infty}(\widehat{K})} \leq C_A C_u \left[\left(\frac{h/R}{\sigma + h/R} \right)^{p+1} + \left(\frac{h\kappa}{\sigma p} \right)^{p+1} \right].$$

Idea: The crucial points of Theorem 3.4.9 are items (iii) and (iv). To verify (iii) we will exploit the approximation properties of $\widehat{\Pi}_p^{\mathrm{MS}}$ given by Proposition 3.4.5 together with the fact that $\widehat{\Pi}_p u$ is the solution to a minimization problem. To prove (iv) we use the affine projection representation (3.17) of $\widehat{\Pi}_p$ together with the approximation properties of polynomials satisfying the restriction property given in Proposition 3.4.8.

Proof. Assertion (i) is trivially satisfied due to the construction in Definition 3.4.1 and Remark 3.4.3.

Assertion (ii) is also trivially satisfied, since for a given polynomial $q \in \mathcal{P}_p(\widehat{K})$ the norms in Definition 3.4.1 are minimized at q.

To prove Assertion (iii) recall that Step (iv) in Definition 3.4.1 is exactly the minimization of the norm in question, constrained to all polynomials satisfying the restriction property for u. Since $\widehat{\Pi}_p^{\mathrm{MS}} u$ given by Proposition 3.4.5 also satisfies the restriction property we can immediately conclude for $p \geq s-1$ that

$$\begin{aligned} p\|u - \widehat{\Pi}_p u\|_{0,\widehat{K}} + \|u - \widehat{\Pi}_p u\|_{1,\widehat{K}} &\leq p\|u - \widehat{\Pi}_p^{\mathrm{MS}} u\|_{0,\widehat{K}} + \|u - \widehat{\Pi}_p^{\mathrm{MS}} u\|_{1,\widehat{K}} \\ &\leq C_s p^{-(s-1)} |u|_{s,\widehat{K}}. \end{aligned}$$

We turn to Assertion (iv). Since polynomials up to degree p are preserved under $\widehat{\Pi}_p$, we immediately have

$$\|u - \widehat{\Pi}_p u\|_{W^{1,\infty}(\widehat{K})} \leq \|u - q\|_{W^{1,\infty}(\widehat{K})} + \|\widehat{\Pi}_p q - \widehat{\Pi}_p u\|_{W^{1,\infty}(\widehat{K})}, \tag{3.19}$$

for any $q \in \mathcal{P}_p(\widehat{K})$. We estimate the second term in (3.19). We have seen in (3.17) that the operator $\widehat{\Pi}_p$ can be written as $\widehat{\Pi}_p u = \pi^u + \widehat{\Pi}_{\mathcal{P}_p^0}(u - \pi^u)$ for any $\pi^u \in \mathcal{V}_p^u$ (the affine space of polynomials with restriction property for u), where $\widehat{\Pi}_{\mathcal{P}_p^0}$ is the orthogonal projection onto $\mathcal{P}_p^0(\widehat{K}) \leq \mathcal{P}_p(\widehat{K})$, the space of polynomials vanishing on $\partial\widehat{K}$, with respect to the norm $|||\cdot|||$. Therefore, we have

$$\widehat{\Pi}_p q - \widehat{\Pi}_p u = \pi^q - \pi^u + \widehat{\Pi}_{\mathcal{P}_p^0}(q - u + \pi^u - \pi^q)$$

for any $\pi^u \in \mathcal{V}_p^u$ and $\pi^q \in \mathcal{V}_p^q$. Selecting $q \in \mathcal{V}_p^u$ allows us to choose $\pi^u = \pi^q = q$, which immediately gives

$$\widehat{\Pi}_p q - \widehat{\Pi}_p u = \widehat{\Pi}_{\mathcal{P}_p^0}(q-u)$$

for all $q \in \mathcal{V}_p^u$. Using the polynomial inverse estimates $\|\pi\|_{L^\infty(\Omega)} \leq Cp^d\|\pi\|_{0,\Omega}$ for all $\pi \in \mathcal{P}_p(\widehat{K})$, (see, e.g., [Sch98, Thm. 4.76] for the case $d=2$), we find

$$\|\widehat{\Pi}_p q - \widehat{\Pi}_p u\|_{W^{1,\infty}(\widehat{K})} = \|\widehat{\Pi}_{\mathcal{P}_p^0}(q-u)\|_{W^{1,\infty}(\widehat{K})} \lesssim p^d\|\widehat{\Pi}_{\mathcal{P}_p^0}(q-u)\|_{1,\widehat{K}}.$$

Since $\widehat{\Pi}_{\mathcal{P}_p^0}$ is the orthogonal projection with respect to the norm $|||\cdot|||$ we obtain

$$p^d\|\widehat{\Pi}_{\mathcal{P}_p^0}(q-u)\|_{1,\widehat{K}} \leq p^d|||q-u||| \lesssim p^{d+1}\|q-u\|_{W^{1,\infty}(\widehat{K})}.$$

We therefore conclude that

$$\|u - \widehat{\Pi}_p u\|_{W^{1,\infty}(\widehat{K})} \lesssim p^{d+1}\|u-q\|_{W^{1,\infty}(\widehat{K})}$$

for all $q \in \mathcal{V}_p^u$. Proposition 3.4.8 provides a polynomial $q \in \mathcal{V}_p^u$ with the desired approximation properties. Absorbing the algebraic factor p^{d+1} into the exponential factor then yields the result.

□

3.4.3. $\boldsymbol{H}(\Omega, \operatorname{div})$-conforming approximation operators

In the following we will construct an approximation operator $\boldsymbol{\Pi}_p^{\operatorname{div},s}\colon \boldsymbol{H}^s(\Omega) \to \mathbf{BDM}_p(\mathcal{T}_h) \subset \mathbf{RT}_p(\mathcal{T}_h)$ that features the optimal convergence rates in p simultaneously in $L^2(\Omega)$ and $\boldsymbol{H}(\Omega, \operatorname{div})$ for $s > d/2$. The operator will act elementwise. First we consider any operator $\widehat{\boldsymbol{\Pi}}_p^{\operatorname{div},s}\colon \boldsymbol{H}^s(\widehat{K}) \to \mathbf{BDM}_p(\widehat{K}) \subset \mathbf{RT}_p(\widehat{K})$ and define $\boldsymbol{\Pi}_p^{\operatorname{div},s}$ on $\boldsymbol{H}^s(\Omega)$ elementwise using the Piola transformation by

$$\left(\boldsymbol{\Pi}_p^{\operatorname{div},s}\boldsymbol{\varphi}\right)\Big|_K := \left[(\det F_K')^{-1}F_K'\widehat{\boldsymbol{\Pi}}_p^{\operatorname{div},s}\left[(\det F_K')(F_K')^{-1}\boldsymbol{\varphi}\circ F_K\right]\right]\circ F_K^{-1}. \tag{3.20}$$

In order for $\boldsymbol{\Pi}_p^{\operatorname{div},s}$ to map into the conforming finite element space one has to select the operator $\widehat{\boldsymbol{\Pi}}_p^{\operatorname{div},s}$ correctly. We choose $\widehat{\boldsymbol{\Pi}}_p^{\operatorname{div},s}\colon \boldsymbol{H}^s(\widehat{K}) \to \mathcal{P}_p(\widehat{K})^d = \mathbf{BDM}_p(\widehat{K}) \subset \mathbf{RT}_p(\widehat{K})$ to be the componentwise application of $\widehat{\Pi}_p$ from Definition 3.4.1 and analyzed in Theorem 3.4.9:

$$\left(\widehat{\boldsymbol{\Pi}}_p^{\operatorname{div},s}\boldsymbol{\varphi}\right)_i := \widehat{\Pi}_p\boldsymbol{\varphi}_i, \qquad \text{for } i = 1,\dots,d. \tag{3.21}$$

This choice will ensure the desired approximation properties, and will also map into the conforming finite element space due to the restriction property. We will summarize and prove certain properties of the above constructed operators $\widehat{\boldsymbol{\Pi}}_p^{\operatorname{div},s}$ and $\boldsymbol{\Pi}_p^{\operatorname{div},s}$. See [MS21] for a similar construction concerning the space $\boldsymbol{H}(\Omega, \operatorname{curl})$.

Lemma 3.4.10. *Let $s > d/2$ and let the operators $\widehat{\boldsymbol{\Pi}}_p^{\operatorname{div},s}$ and $\boldsymbol{\Pi}_p^{\operatorname{div},s}$ be defined as above. Then there holds:*

(i) The operator $\widehat{\mathbf{\Pi}}_p^{\mathrm{div},s} : \boldsymbol{H}^s(\widehat{K}) \to \mathbf{BDM}_p(\widehat{K}) \subset \mathbf{RT}_p(\widehat{K})$ *satisfies for* $p \geq s-1$

$$p\|\boldsymbol{\varphi} - \widehat{\mathbf{\Pi}}_p^{\mathrm{div},s}\boldsymbol{\varphi}\|_{0,\widehat{K}} + \|\boldsymbol{\varphi} - \widehat{\mathbf{\Pi}}_p^{\mathrm{div},s}\boldsymbol{\varphi}\|_{1,\widehat{K}} \lesssim p^{-(s-1)}|\boldsymbol{\varphi}|_{s,\widehat{K}}. \tag{3.22}$$

(ii) Under the assumptions Theorem 3.4.9, (iv) there holds for some constants C_A, $\sigma > 0$ *independent of* p, h, R

$$\|\boldsymbol{\varphi} - \widehat{\mathbf{\Pi}}_p^{\mathrm{div},s}\boldsymbol{\varphi}\|_{W^{1,\infty}(\widehat{K})} \leq C_A C_{\boldsymbol{\varphi}} \left[\left(\frac{h/R}{\sigma + h/R} \right)^{p+1} + \left(\frac{h\kappa}{\sigma p} \right)^{p+1} \right].$$

(iii) The operator $\mathbf{\Pi}_p^{\mathrm{div},s}$ *defined on* $\boldsymbol{H}^s(\Omega)$ *maps to the conforming space* $\mathbf{BDM}_p(\mathcal{T}_h) \subset \mathbf{RT}_p(\mathcal{T}_h)$.

Proof. The first two assertions hold by construction as well as Theorem 3.4.9, see properties (iii), (iv). To prove the third assertion, note that $\widehat{\mathbf{\Pi}}_p^{\mathrm{div},s}$ maps to $\mathbf{BDM}_p(\widehat{K})$ so that

$$(\det F_K')(F_K')^{-1} \left(\mathbf{\Pi}_p^{\mathrm{div},s}\boldsymbol{\varphi} \right)\Big|_K \circ F_K \in \mathbf{BDM}_p(\widehat{K}) \qquad \text{for all } K \in \mathcal{T}_h, \tag{3.23}$$

by construction. We are therefore left with verifying that $\mathbf{\Pi}_p^{\mathrm{div},s}\boldsymbol{\varphi} \in \boldsymbol{H}(\Omega, \mathrm{div})$. Since $\mathbf{\Pi}_p^{\mathrm{div},s}\boldsymbol{\varphi}$ is piecewise smooth it suffices to show inter element continuity of the normal trace. We will first show that the normal trace of $\widehat{\mathbf{\Pi}}_p^{\mathrm{div},s}\boldsymbol{\varphi}$ in fact only depends on the normal trace of $\boldsymbol{\varphi}$. Consider a face $\hat{f}$ of $\widehat{K}$. Let $\gamma_{\hat{\boldsymbol{n}}_{\hat{f}}}$ denote the normal trace for the face $\hat{f}$. We calculate

$$\gamma_{\hat{\boldsymbol{n}}_{\hat{f}}} \left(\widehat{\mathbf{\Pi}}_p^{\mathrm{div},s}\boldsymbol{\varphi} \right) = \left(\widehat{\mathbf{\Pi}}_p^{\mathrm{div},s}\boldsymbol{\varphi} \right)\Big|_{\hat{f}} \cdot \hat{\boldsymbol{n}}_{\hat{f}} = \left. \begin{pmatrix} \widehat{\Pi}_p \boldsymbol{\varphi}_1 \\ \vdots \\ \widehat{\Pi}_p \boldsymbol{\varphi}_d \end{pmatrix} \right|_{\hat{f}} \cdot \hat{\boldsymbol{n}}_{\hat{f}}$$

$$= \begin{pmatrix} \widehat{\Pi}_p(\boldsymbol{\varphi}_1|_{\hat{f}}) \\ \vdots \\ \widehat{\Pi}_p(\boldsymbol{\varphi}_d|_{\hat{f}}) \end{pmatrix} \cdot \hat{\boldsymbol{n}}_{\hat{f}} = \widehat{\Pi}_p(\boldsymbol{\varphi} \cdot \hat{\boldsymbol{n}}_{\hat{f}}) = \widehat{\Pi}_p(\gamma_{\hat{\boldsymbol{n}}_{\hat{f}}}\boldsymbol{\varphi}).$$

Here we used that the operator $\widehat{\Pi}_p$ satisfies the restriction property and the fact that $\hat{\boldsymbol{n}}_{\hat{f}}$ is constant on $\hat{f}$. Furthermore, note that we abused notation in that the symbol $\widehat{\Pi}_p$ is used both for the d dimensional as well as the $d-1$ dimensional version. We conclude the proof using the fact that if $\hat{\boldsymbol{n}}$ is the unit outward normal to $\widehat{K}$ the vector $\boldsymbol{n}$ on K given by

$$\boldsymbol{n} \circ F_K = \frac{1}{\|(F_K')^{-T}\hat{\boldsymbol{n}}\|}(F_K')^{-T}\hat{\boldsymbol{n}}$$

is a unit normal to K, see, e.g., [Mon03, Sec. 3.9 and 5.4].

□

We have p-optimal approximation properties on the reference element $\widehat{K}$ by the operator $\widehat{\mathbf{\Pi}}_p^{\mathrm{div},s}$.

Corollary 3.4.11 (Approximation of $H^s(\Omega)$ functions)**.** *For $d = 2, 3$ and $s > d/2$ the operator $\mathbf{\Pi}_p^{\mathrm{div},s} \colon \boldsymbol{H}^s(\Omega) \to \mathbf{BDM}_p(\mathcal{T}_h) \subset \mathbf{RT}_p(\mathcal{T}_h)$ satisfies*

$$\frac{p}{h}\|\boldsymbol{\varphi} - \mathbf{\Pi}_p^{\mathrm{div},s}\boldsymbol{\varphi}\|_{0,\Omega} + \|\boldsymbol{\varphi} - \mathbf{\Pi}_p^{\mathrm{div},s}\boldsymbol{\varphi}\|_{1,\mathcal{T}_h} \lesssim \left(\frac{h}{p}\right)^{s-1} \|\boldsymbol{\varphi}\|_{s,\Omega} \quad \forall p \geq s-1,$$

where $\|\cdot\|_{1,\mathcal{T}_h}$ denotes the broken H^1-norm.

Proof. The proof follows from Lemma 3.4.10 together with a scaling argument. □

Corollary 3.4.12 (Approximation of analytic functions)**.** *Let $\boldsymbol{\varphi}$ satisfy, for some $C_{\boldsymbol{\varphi}}$, $\gamma > 0$,*

$$\|\nabla^n \boldsymbol{\varphi}\|_{0,\Omega} \leq C_{\boldsymbol{\varphi}} \gamma^n \max\{n,k\}^n \qquad \forall n \in \mathbb{N}_0.$$

Then there exist C, $\sigma > 0$ independent of h, p, and k such that

$$\begin{aligned}\|\boldsymbol{\varphi} - \mathbf{\Pi}_p^{\mathrm{div},s}\boldsymbol{\varphi}\|_{1,\mathcal{T}_h} &+ k\|\boldsymbol{\varphi} - \mathbf{\Pi}_p^{\mathrm{div},s}\boldsymbol{\varphi}\|_{0,\Omega} \\ &\leq CC_{\boldsymbol{\varphi}}\left[\left(\frac{h}{h+\sigma}\right)^p \left(1 + \frac{hk}{h+\sigma}\right) + k\left(\frac{kh}{\sigma p}\right)^p \left(\frac{1}{p} + \frac{kh}{\sigma p}\right)\right].\end{aligned}$$

Proof. We mimic the procedure of [MS10, Thm. 5.5] and [CQ17, Lemma 4.7]. First consider for each element $K \in \mathcal{T}_h$ the constant C_K given by

$$C_K^2 := \sum_{n \geq 0} \frac{\|\nabla^n \boldsymbol{\varphi}\|_{0,K}^2}{(2\gamma \max\{n,k\})^{2n}},$$

which is finite by assumption. Note that we immediately have

$$\begin{aligned}\|\nabla^n \boldsymbol{\varphi}\|_{0,K} &\leq 2^n \gamma^n \max\{n,k\}^n C_K, \\ \sum_{K \in \mathcal{T}_h} C_K^2 &\leq \frac{4}{3} C_{\boldsymbol{\varphi}}^2.\end{aligned}$$

We write $\widehat{\boldsymbol{\varphi}}$ as

$$\begin{aligned}\widehat{\boldsymbol{\varphi}} &= \det(F_K')(F_K')^{-1}\boldsymbol{\varphi} \circ F_K = \det(R_K' \circ A_K A_K')(R_K' \circ A_K A_K')^{-1}\boldsymbol{\varphi} \circ F_K \\ &= \det(A_K')(A_K')^{-1}\tilde{\boldsymbol{\varphi}} \circ A_K,\end{aligned}$$

with

$$\tilde{\boldsymbol{\varphi}} = \det(R_K')(R_K')^{-1}\boldsymbol{\varphi} \circ R_K.$$

As in [MS10, Lemma C.1] for simple changes of variables, we apply [Mel02, Lemma 4.3.1] to the function $\tilde{\boldsymbol{\varphi}}$ and obtain the existence of constants $\overline{\gamma}$, $C > 0$ depending additionally on the constants describing the analyticity of the map R_K such that

$$\|\nabla^n \tilde{\boldsymbol{\varphi}}\|_{0,\tilde{K}} \leq C\overline{\gamma}^n \max\{n,k\}^n C_K \qquad \forall n \in \mathbb{N}_0.$$

Since A_K is affine we immediately deduce that

$$\|\nabla^n \widehat{\varphi}\|_{0,\widehat{K}} \lesssim h^{d/2-1} h^n \|\nabla^n \tilde{\varphi}\|_{0,\tilde{K}} \leq h^{d/2-1} (\overline{\gamma} h)^n \max\{n,k\}^n C_K \quad \forall n \in \mathbb{N}_{\geq 1}.$$

Hence, by Lemma 3.4.10 with $R = 1$ we have

$$\|\widehat{\varphi} - \widehat{\mathbf{\Pi}}_p^{\mathrm{div},s} \widehat{\varphi}\|_{W^{1,\infty}(\widehat{K})} \lesssim C_K h^{d/2-1} \left[\left(\frac{h}{\sigma + h} \right)^{p+1} + \left(\frac{hk}{\sigma p} \right)^{p+1} \right]$$

for some $\sigma > 0$. By a change of variables there holds for $q = 0, 1$

$$\begin{aligned} \|\boldsymbol{\varphi} - \mathbf{\Pi}_p^{\mathrm{div},s} \boldsymbol{\varphi}\|_{q,K} &\lesssim h^{-d/2+1-q} \|\widehat{\varphi} - \widehat{\mathbf{\Pi}}_p^{\mathrm{div},s} \widehat{\varphi}\|_{q,\widehat{K}} \\ &\lesssim h^{-q} C_K \left[\left(\frac{h}{\sigma + h} \right)^{p+1} + \left(\frac{hk}{\sigma p} \right)^{p+1} \right]. \end{aligned}$$

Summation over all elements gives

$$\begin{aligned} &\|\boldsymbol{\varphi} - \mathbf{\Pi}_p^{\mathrm{div},s} \boldsymbol{\varphi}\|_{1,\mathcal{T}_h} + k \|\boldsymbol{\varphi} - \mathbf{\Pi}_p^{\mathrm{div},s} \boldsymbol{\varphi}\|_{0,\Omega} \\ &\lesssim \left[\left(\frac{h}{\sigma + h} \right)^{p} + k \left(\frac{h}{\sigma + h} \right)^{p+1} + \frac{k}{p} \left(\frac{hk}{\sigma p} \right)^{p} + k \left(\frac{hk}{\sigma p} \right)^{p+1} \right] \sqrt{\sum_{K \in \mathcal{T}_h} C_K^2} \\ &\lesssim \left[\left(\frac{h}{h + \sigma} \right)^{p} \left(1 + \frac{hk}{h + \sigma} \right) + k \left(\frac{kh}{\sigma p} \right)^{p} \left(\frac{1}{p} + \frac{kh}{\sigma p} \right) \right] C_{\boldsymbol{\varphi}}, \end{aligned}$$

which completes the proof.

□

3.5. A priori estimate

We turn to an *a priori* estimate of the FOSLS method. Again the proof follows the ideas of [CQ17, Lemma 5.1], resting, however, on the refined duality argument given in Lemma 3.3.1 and the approximation properties derived in Section 3.4 to obtain the factor h/p. For the readers' convenience we recapitulate the important steps. As in [MS10] we show that this can be achieved under the conditions kh/p sufficiently small and p of order $\log k$.

Theorem 3.5.1 (A priori estimate). *Let Assumptions 3.1.1, 2.0.1 be valid. Then there exist constants c_1, $c_2 > 0$ that are independent of h, p, and k such that the conditions*

$$\frac{kh}{p} \leq c_1 \qquad \text{and} \qquad p \geq c_2 (\log k + 1) \tag{3.24}$$

imply that the approximation $(\boldsymbol{\varphi}_h, u_h)$ of the FOSLS method satisfies the following: For any $(\boldsymbol{\psi}_h, v_h) \in \boldsymbol{V}_h \times W_h$ there holds

$$\begin{aligned} \|u - u_h\|_{0,\Omega} \lesssim \frac{h}{p} \Big(&\|\nabla(u - v_h)\|_{0,\Omega} + k \|u - v_h\|_{0,\Omega} + \\ &\|\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{\psi}_h)\|_{0,\Omega} + k \|\boldsymbol{\varphi} - \boldsymbol{\psi}_h\|_{0,\Omega} + k^{1/2} \|(\boldsymbol{\varphi} - \boldsymbol{\psi}_h) \cdot \boldsymbol{n}\|_{0,\Gamma} \Big). \end{aligned}$$

Proof. Let $e^u = u - u_h$ and $\boldsymbol{e}^{\boldsymbol{\varphi}} = \boldsymbol{\varphi} - \boldsymbol{\varphi}_h$ denote the errors of the two components. We apply the duality argument from Lemma 3.3.1 with $w = e^u$ and also apply the corresponding splitting:

$$\|e^u\|_{0,\Omega}^2 = b((\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u), (\boldsymbol{\psi}, v)) = b((\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u), (\boldsymbol{\psi}_A, v_A)) + b((\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u), (\boldsymbol{\psi}_{H^2}, v_{H^2})).$$

Exploiting the Galerkin orthogonality we have

$$\|e^u\|_{0,\Omega}^2 = b((\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u), (\boldsymbol{\psi}_A - \tilde{\boldsymbol{\psi}}_A, v_A - \tilde{v}_A)) + b((\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u), (\boldsymbol{\psi}_{H^2} - \tilde{\boldsymbol{\psi}}_{H^2}, v_{H^2} - \tilde{v}_{H^2})),$$

for any $(\tilde{\boldsymbol{\psi}}_A, \tilde{v}_A), (\tilde{\boldsymbol{\psi}}_{H^2}, \tilde{v}_{H^2}) \in \boldsymbol{V}_h \times W_h$. Using Cauchy-Schwarz we arrive at

$$\begin{aligned}
&\|e^u\|_{0,\Omega}^2 \lesssim \left[\|ik\boldsymbol{e}^{\boldsymbol{\varphi}} + \nabla e^u\|_{0,\Omega} + \|ike^u + \nabla \cdot \boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega} + k^{1/2}\|\boldsymbol{e}^{\boldsymbol{\varphi}} \cdot \boldsymbol{n} + e^u\|_{0,\Gamma}\right] \cdot \\
&\Big(\|\nabla \cdot (\boldsymbol{\psi}_A - \tilde{\boldsymbol{\psi}}_A)\|_{0,\Omega} + k\|\boldsymbol{\psi}_A - \tilde{\boldsymbol{\psi}}_A\|_{0,\Omega} + k^{1/2}\|(\boldsymbol{\psi}_A - \tilde{\boldsymbol{\psi}}_A) \cdot \boldsymbol{n}\|_{0,\Omega} + \\
&\quad \|\nabla \cdot (\boldsymbol{\psi}_{H^2} - \tilde{\boldsymbol{\psi}}_{H^2})\|_{0,\Omega} + k\|\boldsymbol{\psi}_{H^2} - \tilde{\boldsymbol{\psi}}_{H^2}\|_{0,\Omega} + k^{1/2}\|(\boldsymbol{\psi}_{H^2} - \tilde{\boldsymbol{\psi}}_{H^2}) \cdot \boldsymbol{n}\|_{0,\Gamma} + \\
&\quad \|\nabla(v_A - \tilde{v}_A)\|_{0,\Omega} + k\|v_A - \tilde{v}_A\|_{0,\Omega} + k^{1/2}\|v_A - \tilde{v}_A\|_{0,\Gamma} + \\
&\quad \|\nabla(v_{H^2} - \tilde{v}_{H^2})\|_{0,\Omega} + k\|v_{H^2} - \tilde{v}_{H^2}\|_{0,\Omega} + k^{1/2}\|v_{H^2} - \tilde{v}_{H^2}\|_{0,\Gamma}\Big).
\end{aligned} \tag{3.25}$$

We are going to exploit the approximation properties in the corresponding norms and spaces.

Approximation of v_A and v_{H^2}: We may apply [CQ17, Lemma 4.10], which is essentially the procedure of [MS10, Thm. 5.5] together with a multiplicative trace inequality. Using the estimates (3.7b), (3.7c), and (3.7e) in Lemma 3.3.1 as well as [MS10, Thm. B.4] to find appropriate approximations $\tilde{v}_{H^2}$ and $\tilde{v}_A$, we have

$$\begin{aligned}
&\|\nabla(v_A - \tilde{v}_A)\|_{0,\Omega} + k\|v_A - \tilde{v}_A\|_{0,\Omega} + k^{1/2}\|v_A - \tilde{v}_A\|_{0,\Gamma} \\
&\quad \lesssim \left[\left(\frac{h}{h+\sigma}\right)^p \left(1 + \frac{hk}{h+\sigma}\right) + k\left(\frac{kh}{\sigma p}\right)^p \left(\frac{1}{p} + \frac{kh}{\sigma p}\right)\right] \|e^u\|_{0,\Omega} \\
&\quad \lesssim \frac{h}{p}\|e^u\|_{0,\Omega}
\end{aligned}$$

as well as

$$\begin{aligned}
&\|\nabla(v_{H^2} - \tilde{v}_{H^2})\|_{0,\Omega} + k\|v_{H^2} - \tilde{v}_{H^2}\|_{0,\Omega} + k^{1/2}\|v_{H^2} - \tilde{v}_{H^2}\|_{0,\Gamma} \\
&\quad \lesssim \frac{1}{k}\left(\frac{kh}{p} + \left(\frac{kh}{p}\right)^2\right) \|e^u\|_{0,\Omega} \lesssim \frac{h}{p}\|e^u\|_{0,\Omega},
\end{aligned}$$

where the latter estimates are due to the boundedness of Ω, $\sigma > 0$, and choosing c_1 small and c_2 sufficiently large as well as elementary but tedious calculations.

Approximation of $\boldsymbol{\psi}_A$: To approximate $\boldsymbol{\psi}_A$ we choose $\tilde{\boldsymbol{\psi}}_A = \boldsymbol{\Pi}_p^{\mathrm{div},2}\boldsymbol{\psi}_A$ with $\boldsymbol{\Pi}_p^{\mathrm{div},2}$ as in Corollary 3.4.12 and apply the results therein. Furthermore, we apply the estimates (3.7a) and (3.7c) of Lemma 3.3.1. Proceeding as above together with a multiplicative trace inequality, again after tedious calculations, gives

$$\begin{aligned}&\|\nabla\cdot(\boldsymbol{\psi}_A-\tilde{\boldsymbol{\psi}}_A)\|_{0,\Omega}+k\|\boldsymbol{\psi}_A-\tilde{\boldsymbol{\psi}}_A\|_{0,\Omega}+k^{1/2}\|(\boldsymbol{\psi}_A-\tilde{\boldsymbol{\psi}}_A)\cdot\boldsymbol{n}\|_{0,\Gamma}\\&\quad\lesssim\frac{h}{p}\|e^u\|_{0,\Omega}.\end{aligned}$$

Approximation of $\boldsymbol{\psi}_{H^2}$: To approximate $\boldsymbol{\psi}_{H^2}$ we choose $\tilde{\boldsymbol{\psi}}_{H^2}=\boldsymbol{\Pi}_p^{\mathrm{div},2}\boldsymbol{\psi}_{H^2}$ with $\boldsymbol{\Pi}_p^{\mathrm{div},2}$ as in Corollary 3.4.11 and apply the results therein. We apply the estimate (3.7d) of Lemma 3.3.1. Due to the multiplicative trace inequality we also have

$$\|(\boldsymbol{\psi}_{H^2}-\tilde{\boldsymbol{\psi}}_{H^2})\cdot\boldsymbol{n}\|_{0,\Gamma}\le\left(\frac{h}{p}\right)^{3/2}\|\boldsymbol{\psi}_{H^2}\|_{2,\Omega}. \tag{3.26}$$

Therefore, we arrive at

$$\begin{aligned}&\|\nabla\cdot(\boldsymbol{\psi}_{H^2}-\tilde{\boldsymbol{\psi}}_{H^2})\|_{0,\Omega}+k\|\boldsymbol{\psi}_{H^2}-\tilde{\boldsymbol{\psi}}_{H^2}\|_{0,\Omega}+k^{1/2}\|(\boldsymbol{\psi}_{H^2}-\tilde{\boldsymbol{\psi}}_{H^2})\cdot\boldsymbol{n}\|_{0,\Gamma}\\&\quad\lesssim\frac{h}{p}\|\boldsymbol{\psi}_{H^2}\|_{2,\Omega}\lesssim\frac{h}{p}\|e^u\|_{0,\Omega},\end{aligned}$$

where we used the estimate (3.7d) of Lemma 3.3.1. Putting it all together we have

$$\begin{aligned}\|e^u\|_{0,\Omega}&\lesssim\frac{h}{p}(\|ik\boldsymbol{e}^{\boldsymbol{\varphi}}+\nabla e^u\|_{0,\Omega}+\|ike^u+\nabla\cdot\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}+k^{1/2}\|\boldsymbol{e}^{\boldsymbol{\varphi}}\cdot\boldsymbol{n}+e^u\|_{0,\Gamma})\\&\lesssim\frac{h}{p}\sqrt{b((\boldsymbol{e}^{\boldsymbol{\varphi}},e^u),(\boldsymbol{e}^{\boldsymbol{\varphi}},e^u))}.\end{aligned}$$

Applying again the Galerkin orthogonality and using the multiplicative trace inequality to absorb the term $k^{1/2}\|u-v_h\|_{0,\Gamma}$ into the L^2 norms of the volume yields the result. □

We conclude this section with a simple consequence of standard regularity theory and approximation properties of the employed finite element spaces in higher order Sobolev norms.

Corollary 3.5.2. *For $s\ge 0$, $f\in H^s(\Omega)$ and $g\in H^{s+1/2}(\partial\Omega)$ we have $u\in H^{s+2}(\Omega)$, $u\in H^{s+3/2}(\Gamma)$, $\partial_n u\in H^{s+1/2}(\Gamma)$, $\boldsymbol{\varphi}\in\boldsymbol{H}^{s+1}(\Omega)$, $\nabla\cdot\boldsymbol{\varphi}\in H^s(\Omega)$ and $\boldsymbol{\varphi}\cdot\boldsymbol{n}\in\boldsymbol{H}^{s+1/2}(\Gamma)$. Furthermore, there exist constants c_1, $c_2>0$ that are independent of h, p, and k such that the conditions*

$$\frac{kh}{p}\le c_1\qquad\text{and}\qquad p\ge c_2(\log k+1)+s \tag{3.27}$$

imply that the solution $(\boldsymbol{\varphi}_h,u_h)$ satisfies

$$\|u-u_h\|_{0,\Omega}\lesssim\left(\frac{h}{p}\right)^{s+1}(\|f\|_{s,\Omega}+\|g\|_{s+1/2,\Gamma}),$$

for $p\ge s$ with a wavenumber-independent constant.

Proof. The first assertion follows immediately from standard regularity theory. Consider the case $s > 0$. Theorem 3.5.1 together with a multiplicative trace inequality, which is applicable due to the already derived regularity of $\boldsymbol{\varphi}$, gives

$$\|u - u_h\|_{0,\Omega} \lesssim \frac{h}{p}\Big(\|u - v_h\|_{1,\Omega} + k\|u - v_h\|_{0,\Omega} + \|\boldsymbol{\varphi} - \boldsymbol{\psi}_h\|_{1,\mathcal{T}_h} + k\|\boldsymbol{\varphi} - \boldsymbol{\psi}_h\|_{0,\Omega}\Big).$$

Applying the higher order splitting of Theorem 3.2.2 and using the fact that $\boldsymbol{\varphi} = ik^{-1}\nabla u$, one can easily estimate, as in the proof of Theorem 3.5.1 together with the Corollaries 3.4.11 and 3.4.12,

$$\|\boldsymbol{\varphi} - \boldsymbol{\psi}_h\|_{1,\mathcal{T}_h} + k\|\boldsymbol{\varphi} - \boldsymbol{\psi}_h\|_{0,\Omega} \lesssim \left(\frac{h}{p}\right)^s (\|f\|_{s,\Omega} + \|g\|_{s+1/2,\Gamma}).$$

Note the exponent s, since $\boldsymbol{\varphi}$ is only in $\boldsymbol{H}^{s+1}(\Omega)$. Furthermore, again as in the proof of Theorem 3.5.1, see also [MPS13, Thm. 4.8], we have

$$\|u - v_h\|_{1,\Omega} + k\|u - v_h\|_{0,\Omega} \lesssim \left(\frac{h}{p}\right)^{s+1} (\|f\|_{s,\Omega} + \|g\|_{s+1/2,\Gamma}),$$

now with the exponent $s + 1$ since $u \in H^{s+2}(\Omega)$, which yields the result for $s > 0$. In the case $s = 0$ one simply sets $v_h = 0$ as well as $\boldsymbol{\psi}_h = 0$ and uses the wavenumber-explicit estimates of Theorem 3.2.2.

□

Remark 3.5.3. Note that although we assume $f \in H^s(\Omega)$ and $g \in H^{s+1/2}(\Gamma)$ in Corollary 3.5.2, we only obtained a convergence rate $s+1$. This seems suboptimal when compared with classical FEM where, given sufficient regularity of the data and the geometry, one can expect a rate of $s + 2$ for the convergence in the $L^2(\Omega)$-norm. Especially for $f \in L^2(\Omega)$ and $g \in H^{1/2}(\Gamma)$ one can only expect h/p for the FOSLS method compared to h^2/p^2 for the FEM. The proof of Corollary 3.5.2 is in that sense sharp since the leading error term in the *a priori* estimate is

$$\|\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{\psi}_h)\|_{0,\Omega} = \|ik^{-1}f + iku - \nabla \cdot \boldsymbol{\psi}_h\|_{0,\Omega},$$

where we used the fact $\boldsymbol{\varphi} = ik^{-1}\nabla u$. The essential part is therefore to approximate an f that is just in $L^2(\Omega)$ and therefore no further powers of h can be gained. Assuming more regularity on f would resolve this problem, however, the boundary data would restrict a further lifting of $\boldsymbol{\varphi}$ in classical Sobolev spaces, but not in $\boldsymbol{H}(\Omega, \mathrm{div})$ spaces. This in turn would make it necessary to directly estimate $\|\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{\psi}_h)\|_{0,\Omega}$ instead of generously bounding it by $\|\boldsymbol{\varphi} - \boldsymbol{\psi}_h\|_{1,\mathcal{T}_h}$. Last but not least there is the boundary term

$$\|(\boldsymbol{\varphi} - \boldsymbol{\psi}_h) \cdot \boldsymbol{n}\|_{0,\Gamma} = \|ik^{-1}g - u - \boldsymbol{\psi}_h \cdot \boldsymbol{n}\|_{0,\Gamma}.$$

Again if g is only $H^{1/2}(\Gamma)$ one can only expect $\sqrt{h/p}$, but favorable in terms of k. ■

3.6. Numerical examples

All our calculations are performed with the hp-FEM code NETGEN / NGSOLVE by J. Schöberl, [Sch, Sch97]. We plot the error against N_λ, the number of degrees of freedom per wavelength,

$$N_\lambda = \frac{2\pi \sqrt[d]{\mathrm{DOF}}}{k \sqrt[d]{|\Omega|}},$$

where the wavelength λ and the wavenumber k are related via $k = 2\pi/\lambda$ and DOF denotes the size of the linear system to be solved. We compare the results of the classical FEM with the FOSLS method, measured in the relative $L^2(\Omega)$ error. For the classical FEM we use the standard space $S_p(\mathcal{T}_h)$. For the FOSLS method we employ the pairing $\boldsymbol{V}_h \times W_h = \mathbf{BDM}_p(\mathcal{T}_h) \times S_p(\mathcal{T}_h)$.

Example 3.6.1. Let Ω be the unit circle in $\mathbb{R}^2$ and consider the problem

$$\begin{aligned} -\Delta u - k^2 u &= 0 \qquad \text{in } \Omega, \\ \partial_n u - iku &= g \qquad \text{on } \Gamma, \end{aligned}$$

where the data g is such that the exact solution is given by $u(x, y) = \mathrm{e}^{i(k_1 x + k_2 y)}$ with $k_1 = -k_2 = \frac{1}{\sqrt{2}} k$. For the numerical studies, this problem will be solved using h-FEM and h-FOSLS with polynomial degrees $p = 1, 2, 3, 4$. The results are visualized in Figure 3.1. For both methods we observe the expected convergence $O(h^{p+1})$ in the relative $L^2(\Omega)$ error. Note that for both methods higher order versions are less prone to the pollution effect. At the same number of degrees of freedom per wavelength we also observe that the classical FEM is superior to FOSLS, when measured in achieved accuracy in $L^2(\Omega)$. This is not surprising since, for the same mesh and polynomial degree p, the number of degrees of freedom of the FOSLS is roughly three times as large as for the classical FEM. Note, however, that we do not consider any solver aspects of the employed methods, where FOSLS might have advantages over the classical FEM since its system matrix is positive definite.

Example 3.6.2. For $\pi < \omega < 2\pi$ let $\Omega = \{(r\cos\varphi, r\sin\varphi) \colon r \in (0,1), \varphi \in (0,\omega)\} \subset \mathbb{R}^2$ and consider

$$\begin{aligned} -\Delta u - k^2 u &= 0 \qquad \text{in } \Omega, \\ \partial_n u - iku &= g \qquad \text{on } \Gamma. \end{aligned}$$

The data g is such that the exact solution is given by $u(x, y) = J_\alpha(kr)\cos(\alpha\varphi)$, with $\alpha = 3\pi/2$. Standard regularity theory gives $u \in H^{1+\alpha-\varepsilon}(\Omega)$ for every $\varepsilon > 0$. In the numerical experiments we keep $kh = 5$ and perform a p-FEM and a p-FOSLS method up to $p = 46$ and $p = 29$, respectively. The results are visualized in Figure 3.2. We observe that the FEM has significantly smaller errors than the FOSLS. For a discussion of the expected $L^2(\Omega)$-convergence rates of the p-FEM, we refer the reader to [JS92, Remark after Thm. 3 and Sec. 3].

The next example focuses on the Helmholtz equation with right-hand side f with finite Sobolev regularity.

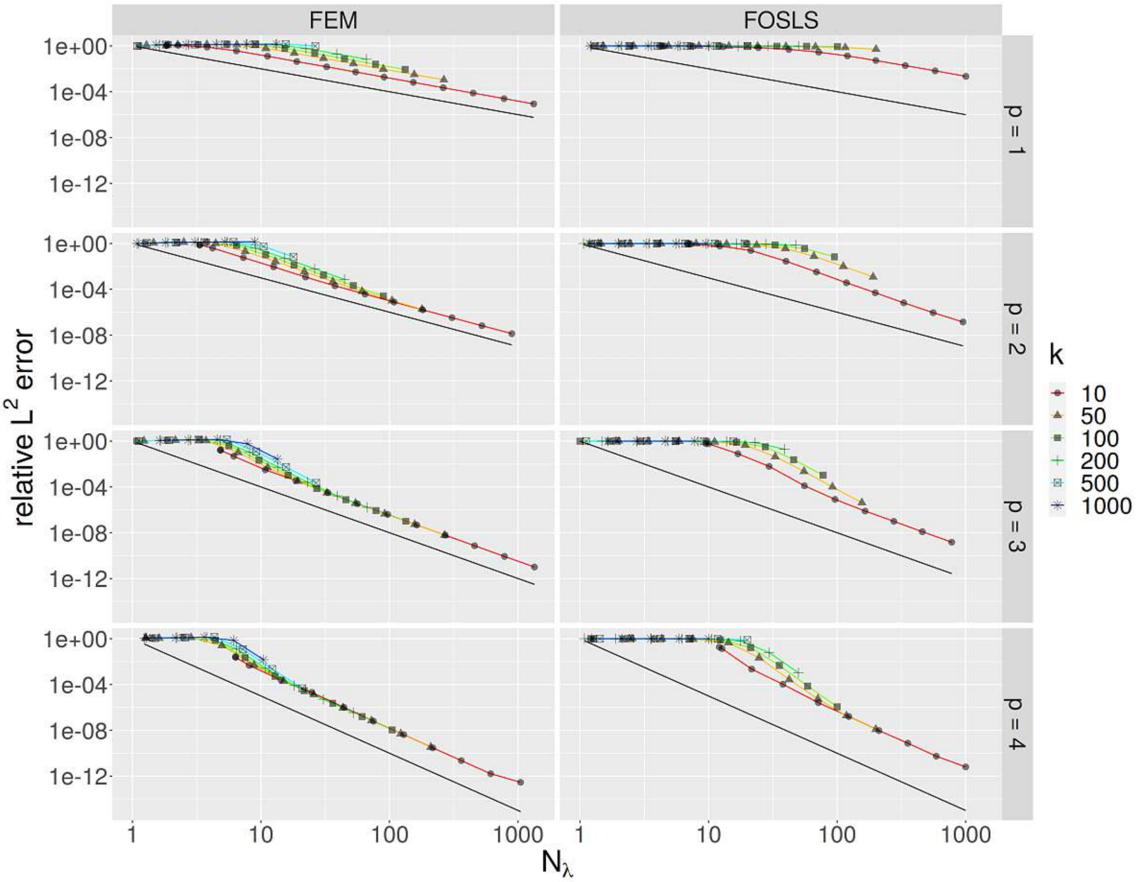

Figure 3.1.: Comparison between the h-FEM (left) and h-FOSLS (right) for $p = 1, 2, 3, 4$ as described in Example 3.6.1. The reference line in black corresponds to h^{p+1}.

Example 3.6.3. Let $\Omega = (-1, 1) \subset \mathbb{R}$ and $f = -\chi_{(-1,0]} + \chi_{(0,1)}$, where χ_A denotes the indicator function on $A \subset \mathbb{R}$. The function f is in $H^{1/2-\varepsilon}(\Omega)$ for every $\varepsilon > 0$. We consider uniform meshes $\mathcal{T}_h$ on Ω such that the break point zero is *not* a node, as otherwise the piecewise smooth solution could be approximated very well. We study

$$\begin{aligned} -u'' - k^2 u &= f && \text{in } \Omega, \\ \partial_n u - iku &= g && \text{on } \Gamma, \end{aligned}$$

where the data g is such that the exact solution is given by

$$u(x) = \begin{cases} \cos(kx) + \frac{1}{k^2} & x \leq 0, \\ (1 + \frac{2}{k^2}) \cos(kx) - \frac{1}{k^2} & x > 0. \end{cases}$$

Standard regularity theory gives $u \in H^{2.5-\varepsilon}(\Omega)$ for every $\varepsilon > 0$. For the h-FEM we expect $O(h^{\min\{2+0.5,p+1\}})$. In fact for $p > 1$ one can show (cf. [EM14, Cor. 4.6]) that $k\|u - u_h^{\text{FEM}}\|_{0,\Omega} \lesssim h^{2.5}$ and, by inspection, $\|u\|_{0,\Omega} = O(1)$ (uniformly in k). It is therefore

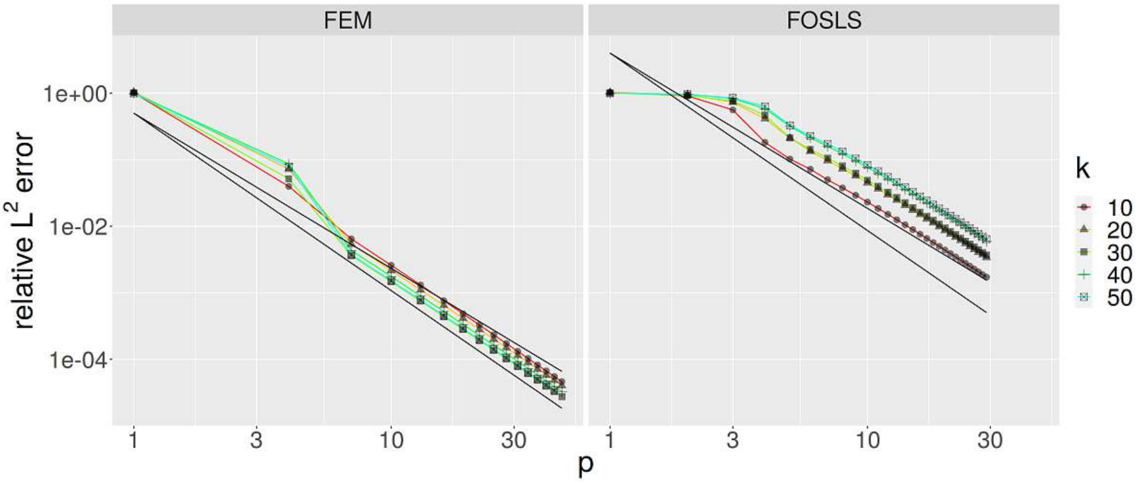

Figure 3.2.: Comparison between the p-FEM (left) and p-FOSLS (right) for $kh = 5$ as described in Example 3.6.2. We include the reference lines $p^{-4\cdot 2/3} = p^{-8/3}$ and $p^{-(2\cdot 2/3+1)} = p^{-7/3}$.

expedient to plot $k^{3.5}\|u - u_h^{\text{FEM}}\|_{0,\Omega}/\|u\|_{0,\Omega}$ versus $N_\lambda \sim (kh)$. For the h-FOSLS Corollary 3.5.2 predicts only $O(h^{\min\{1+0.5,p+1\}})$. The numerical results show, however, for both methods convergence $O(h^{\min\{2.5,p+1\}})$. The results are visualized in Figure 3.3.

Remark 3.6.4. The numerical results of Example 3.6.3 visualized in Figure 3.3 indicate that Corollary 3.5.2 is in fact suboptimal as it predicts only a convergence $O(h^{1.5})$ while we observe $O(h^{\min\{2.5,p+1\}})$. A starting point for understanding this better convergence behavior could be two observations: first, the duality argument in Theorem 3.5.1 is based on the regularity $(\boldsymbol{\psi}, v) \in \boldsymbol{H}^2(\Omega) \times H^2(\Omega)$ of the dual solution $(\boldsymbol{\psi}, v)$ whereas in fact (see the proof of Lemma 3.3.1) $(\boldsymbol{\psi}, v) \in \boldsymbol{H}^2(\text{div}, \Omega) \times H^2(\Omega)$. Second, a more careful application of the Cauchy-Schwarz inequality (3.25) at the beginning of the proof of Theorem 3.5.1 is advisable. In this connection, we point to the fact that the terms in the square brackets in (3.25) are not of the same order. To illustrate this, we plot the components

$$e_1 := ik\boldsymbol{e}^{\boldsymbol{\varphi}} + \nabla e^u \quad \text{and} \quad e_2 := ike^u + \nabla \cdot \boldsymbol{e}^{\boldsymbol{\varphi}} \tag{3.28}$$

in Figure 3.4 for the problem studied in Example 3.6.3. We investigate these improved convergence rates in the Chapters 4 and 5. ■

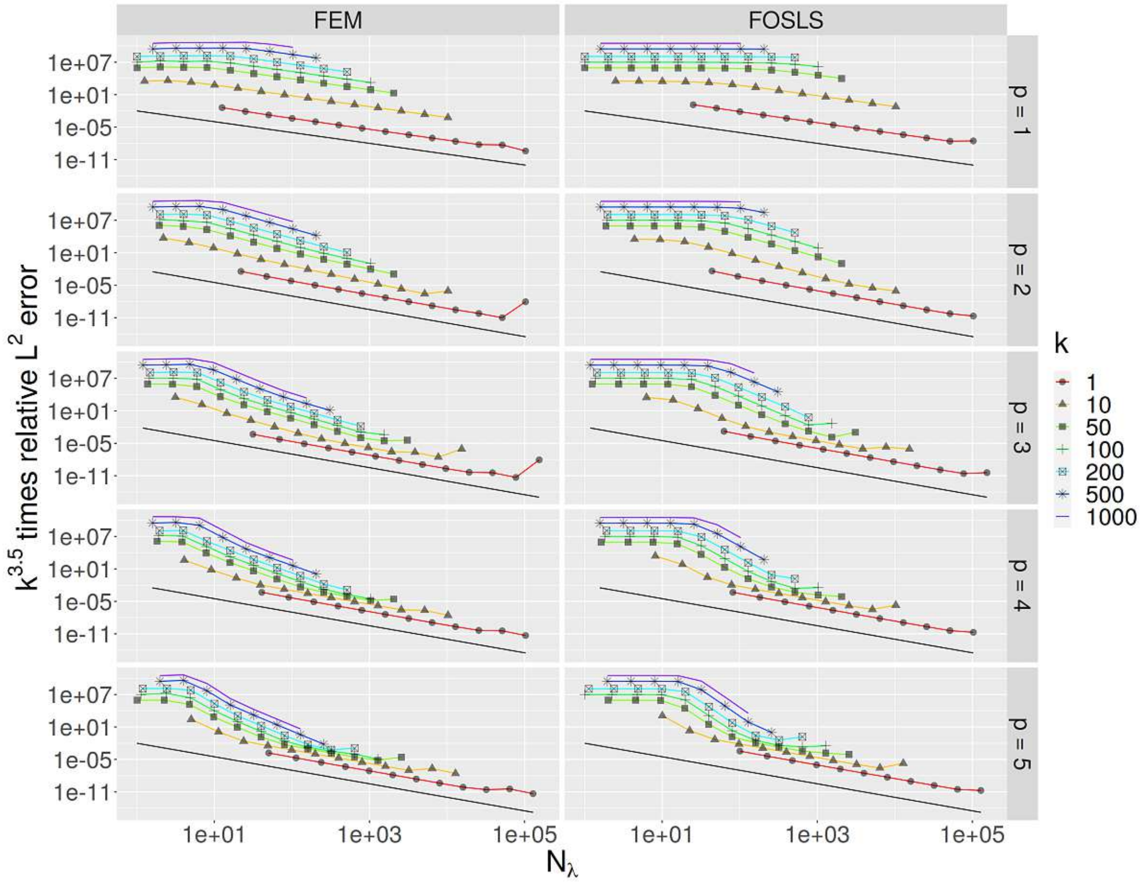

Figure 3.3.: Comparison between the h-FEM (left) and h-FOSLS (right) for $p = 1, \ldots, 5$ as described in Example 3.6.3. The reference line in black corresponds to $h^{\min\{2.5,p+1\}}$.

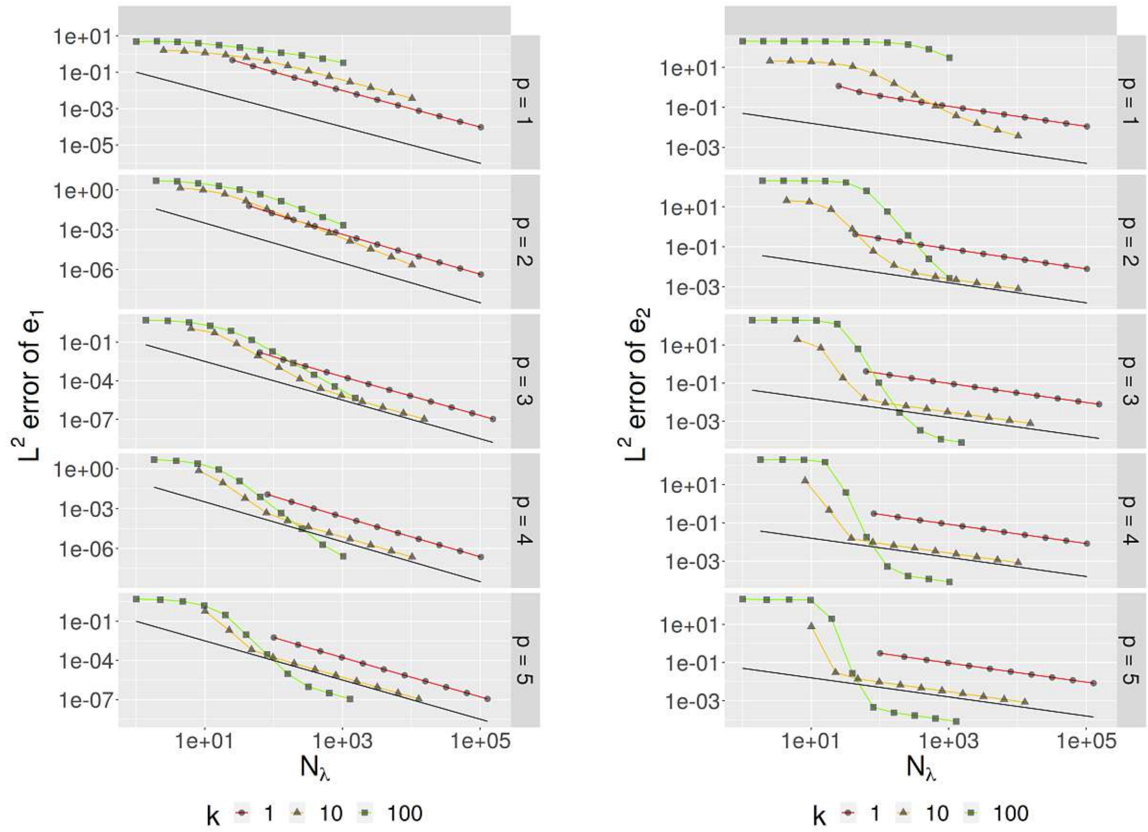

Figure 3.4.: Comparison between the error terms $e_1 := ik\boldsymbol{e}^{\boldsymbol{\varphi}} + \nabla e^u$ (left) and $e_2 := ike^u + \nabla \cdot \boldsymbol{e}^{\boldsymbol{\varphi}}$ (right) for $p = 1, \ldots, 5$ as described in Remark 3.6.4 and Example 3.6.3. The reference line on the left corresponds to h^1 for $p = 1$ and $h^{1.5}$ for $p > 1$. The reference line on the right corresponds to $h^{1/2}$.

4. FOSLS I - homogeneous boundary conditions

In the present chapter we analyze the *hp* version of a first order system least squares method applied to a Poisson-like model problem with homogeneous boundary conditions. Similar to the methodology of Chapter 3 we reformulate the second order model problem into a system of first order equations. The observed convergence rates discussed in Remark 3.6.4 in Chapter 3 motivate the analysis carried out in this chapter. The results of the current chapter are part of [BM20].

The outline is as follows. In Section 4.1 we introduce the model problem, the first order system least squares (FOSLS) method itself and prove norm equivalence results, which in turn guarantee unique solvability of the continuous as well as the discrete least squares formulation. Section 4.2 is devoted to the proof of duality results for the scalar variable, the gradient of the scalar variable as well as the vector variable. In the beginning of Section 4.3 we first exploit the duality result of Section 4.2 in order to prove $L^2(\Omega)$ error estimates for the scalar variable of the primal as well as the dual problem. We then argue first heuristically that these results are actually suboptimal and can be further improved. To that end, we introduce an approximation operator that also satisfies certain orthogonality relations and prove best approximation results for this operator, which are then used to prove our main result (Theorem 4.3.12). Furthermore, we derive $L^2(\Omega)$ error estimates for the gradient of the scalar variable as well as the vector variable. In Section 4.4 we present numerical examples showcasing the proved convergence rates, focusing especially on the case of finite Sobolev regularity.

4.1. Model problem

Throughout the present chapter the notation of Chapter 2 applies. Furthermore, let Ω be simply connected. Let $\Gamma = \partial\Omega$ consist of two disjoint parts Γ_D and Γ_N and let $f \in L^2(\Omega)$. (Later, we will focus on the special cases $\Gamma = \Gamma_D$ and $\Gamma = \Gamma_N$.) For $\gamma > 0$ fixed we consider the following model problem

$$\begin{aligned} -\Delta u + \gamma u &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \Gamma_D, \\ \partial_n u &= 0 \quad \text{on } \Gamma_N. \end{aligned} \tag{4.1}$$

We formulate (4.1) as a first order system. For a general discussion and presentation of current results concerning the numerical discretization of (4.1) using a least squares approach we refer to Section 1.2. Introducing the new variable $\boldsymbol{\varphi} = -\nabla u$ we formally arrive at the system

$$\nabla \cdot \boldsymbol{\varphi} + \gamma u = f \quad \text{in } \Omega, \tag{4.2a}$$
$$\nabla u + \boldsymbol{\varphi} = \mathbf{0} \quad \text{in } \Omega, \tag{4.2b}$$
$$u = 0 \quad \text{on } \Gamma_D, \tag{4.2c}$$
$$\boldsymbol{\varphi} \cdot \boldsymbol{n} = 0 \quad \text{on } \Gamma_N. \tag{4.2d}$$

Introducing the differential operator $\mathcal{L}\colon \boldsymbol{H}_N(\Omega, \mathrm{div}) \times H^1_D(\Omega) \to L^2(\Omega) \times \boldsymbol{L}^2(\Omega)$, given by

$$\mathcal{L}\begin{pmatrix} \boldsymbol{\varphi} \\ u \end{pmatrix} = \begin{pmatrix} \nabla\cdot & \gamma \\ 1 & \nabla \end{pmatrix}\begin{pmatrix} \boldsymbol{\varphi} \\ u \end{pmatrix} = \begin{pmatrix} \nabla \cdot \boldsymbol{\varphi} + \gamma u \\ \nabla u + \boldsymbol{\varphi} \end{pmatrix},$$

we want to solve the equation

$$\mathcal{L}\begin{pmatrix} \boldsymbol{\varphi} \\ u \end{pmatrix} = \begin{pmatrix} f \\ \mathbf{0} \end{pmatrix}.$$

The least squares approach to this problem is to find $(\boldsymbol{\varphi}, u) \in \boldsymbol{H}_N(\Omega, \mathrm{div}) \times H^1_D(\Omega)$ such that

$$\left(\mathcal{L}\begin{pmatrix} \boldsymbol{\varphi} \\ u \end{pmatrix}, \mathcal{L}\begin{pmatrix} \boldsymbol{\psi} \\ v \end{pmatrix}\right)_\Omega = \left(\begin{pmatrix} f \\ \mathbf{0} \end{pmatrix}, \mathcal{L}\begin{pmatrix} \boldsymbol{\psi} \\ v \end{pmatrix}\right)_\Omega \qquad \forall\, (\boldsymbol{\psi}, v) \in \boldsymbol{H}_N(\Omega, \mathrm{div}) \times H^1_D(\Omega),$$

where $(\cdot,\cdot)_\Omega$ denotes the usual $L^2(\Omega)$ scalar product. Introducing now the bilinear form b and the linear functional F by

$$b((\boldsymbol{\varphi}, u), (\boldsymbol{\psi}, v)) := (\nabla \cdot \boldsymbol{\varphi} + \gamma u, \nabla \cdot \boldsymbol{\psi} + \gamma v)_\Omega + (\nabla u + \boldsymbol{\varphi}, \nabla v + \boldsymbol{\psi})_\Omega, \tag{4.3}$$
$$F((\boldsymbol{\psi}, v)) := (f, \nabla \cdot \boldsymbol{\psi} + \gamma v)_\Omega, \tag{4.4}$$

we can state the mixed weak least squares formulation: Find $(\boldsymbol{\varphi}, u) \in \boldsymbol{H}_N(\Omega, \mathrm{div}) \times H^1_D(\Omega)$ such that

$$b((\boldsymbol{\varphi}, u), (\boldsymbol{\psi}, v)) = F((\boldsymbol{\psi}, v)) \quad \forall\, (\boldsymbol{\psi}, v) \in \boldsymbol{H}_N(\Omega, \mathrm{div}) \times H^1_D(\Omega). \tag{4.5}$$

To see solvability of (4.5), let $u \in H^1_D(\Omega)$ be the unique solution of (4.1). In view of $f \in L^2(\Omega)$ the pair $(-\nabla u, u)$ is a solution of (4.5). Uniqueness follows if one can show that $b((\boldsymbol{\varphi}, u), (\boldsymbol{\psi}, v)) = 0$ for all $(\boldsymbol{\psi}, v) \in \boldsymbol{H}_N(\Omega, \mathrm{div}) \times H^1_D(\Omega)$ implies $(\boldsymbol{\varphi}, u) = (\mathbf{0}, 0)$. To that end, we introduce the (yet to be verified) norm $\|\cdot\|_b$ induced by b:

$$\|(\boldsymbol{\varphi}, u)\|_b := \sqrt{b((\boldsymbol{\varphi}, u), (\boldsymbol{\varphi}, u))}. \tag{4.6}$$

A general approach would be to show norm equivalence. In our case:

$$\|u\|_{1,\Omega} + \|\boldsymbol{\varphi}\|_{\boldsymbol{H}(\Omega,\mathrm{div})} \lesssim \|(\boldsymbol{\varphi}, u)\|_b \lesssim \|u\|_{1,\Omega} + \|\boldsymbol{\varphi}\|_{\boldsymbol{H}(\Omega,\mathrm{div})}.$$

We will employ methods similar to a duality argument in the following Theorem 4.1.1 to prove such a norm equivalence.

Theorem 4.1.1 (Norm equivalence)**.** *For all* $(\boldsymbol{\varphi}, u) \in \boldsymbol{H}_N(\Omega, \mathrm{div}) \times H^1_D(\Omega)$ *there holds the norm equivalence*

$$\|u\|^2_{1,\Omega} + \|\boldsymbol{\varphi}\|^2_{\boldsymbol{H}(\Omega,\mathrm{div})} \lesssim b((\boldsymbol{\varphi}, u), (\boldsymbol{\varphi}, u)) \lesssim \|u\|^2_{1,\Omega} + \|\boldsymbol{\varphi}\|^2_{\boldsymbol{H}(\Omega,\mathrm{div})}. \tag{4.7}$$

Proof. First note that by definition

$$b((\boldsymbol{\varphi}, u), (\boldsymbol{\varphi}, u)) = \| \underbrace{\nabla \cdot \boldsymbol{\varphi} + \gamma u}_{=:w} \|^2_{0,\Omega} + \| \underbrace{\nabla u + \boldsymbol{\varphi}}_{=:\boldsymbol{\eta}} \|^2_{0,\Omega},$$

from which the second inequality in (4.7) is obvious. For the first inequality, we will now split $\boldsymbol{\varphi}$ and u as follows:

$$\begin{aligned} \nabla \cdot \boldsymbol{\varphi}_1 + \gamma u_1 &= w && \text{in } \Omega, & \nabla \cdot \boldsymbol{\varphi}_2 + \gamma u_2 &= 0 && \text{in } \Omega,\\ \nabla u_1 + \boldsymbol{\varphi}_1 &= \boldsymbol{0} && \text{in } \Omega, & \nabla u_2 + \boldsymbol{\varphi}_2 &= \boldsymbol{\eta} && \text{in } \Omega,\\ u_1 &= 0 && \text{on } \Gamma_D, & u_2 &= 0 && \text{on } \Gamma_D,\\ \boldsymbol{\varphi}_1 \cdot \boldsymbol{n} &= 0 && \text{on } \Gamma_N, & \boldsymbol{\varphi}_2 \cdot \boldsymbol{n} &= 0 && \text{on } \Gamma_N, \end{aligned}$$

with yet to be determined functions $\boldsymbol{\varphi}_1$, $\boldsymbol{\varphi}_2$, u_1 and u_2. We observe that $\boldsymbol{\varphi} = \boldsymbol{\varphi}_1 + \boldsymbol{\varphi}_2$ and $u = u_1 + u_2$ since the difference solves (4.2) with zero right-hand side, which is only solved by the trivial solution. Simply eliminating $\boldsymbol{\varphi}_1$ and $\boldsymbol{\varphi}_2$ in the above equations, we expect u_1 and u_2 to be solutions to

$$\begin{aligned} -\Delta u_1 + \gamma u_1 &= w && \text{in } \Omega, & -\Delta u_2 + \gamma u_2 &= -\nabla \cdot \boldsymbol{\eta} && \text{in } \Omega,\\ u_1 &= 0 && \text{on } \Gamma_D, & u_2 &= 0 && \text{on } \Gamma_D,\\ \partial_n u_1 &= 0 && \text{on } \Gamma_N, & \partial_n u_2 &= 0 && \text{on } \Gamma_N, \end{aligned}$$

where $-\nabla \cdot \boldsymbol{\eta}$ is to be understood as an element of $(H^1_D(\Omega))'$ given by $F : v \mapsto (\boldsymbol{\eta}, \nabla v)_\Omega$. Both equations are therefore uniquely solvable. This then determines the desired functions u_1, u_2 and consequently the functions $\boldsymbol{\varphi}_1$, $\boldsymbol{\varphi}_2$, using the second equation in the first order systems.

Let us show that $(\boldsymbol{\varphi}_1, u_1)$ solves the above system. By construction it satisfies the differential equations and furthermore, since $\boldsymbol{\varphi}_1 = -\nabla u_1$, we have by standard regularity theory $\boldsymbol{\varphi}_1 \cdot \boldsymbol{n} = -\nabla u_1 \cdot \boldsymbol{n} = -\partial_n u_1 = 0$.

Let us show that $(\boldsymbol{\varphi}_2, u_2)$ satisfies the above system. Let $v \in C^\infty_0(\Omega)$ be arbitrary. Integration by parts and exploiting the weak formulation gives

$$(\nabla \cdot \boldsymbol{\varphi}_2, v)_\Omega = -(\boldsymbol{\varphi}_2, \nabla v)_\Omega = -(\boldsymbol{\eta}, \nabla v)_\Omega + (\nabla u_2, \nabla v)_\Omega = -(\gamma u_2, v)_\Omega.$$

Therefore the div-equation is satisfied. To verify the boundary conditions we calculate for any $v \in H^1_D(\Omega)$

$$\begin{aligned} \langle \boldsymbol{\varphi}_2 \cdot \boldsymbol{n}, v \rangle_{H^{-1/2}(\Gamma) \times H^{1/2}(\Gamma)} &= (\boldsymbol{\varphi}_2, \nabla v)_\Omega + (\nabla \cdot \boldsymbol{\varphi}_2, v)_\Omega \\ &= (-\nabla u_2 + \boldsymbol{\eta}, \nabla v)_\Omega + (\nabla \cdot \boldsymbol{\varphi}_2, v)_\Omega = 0, \end{aligned}$$

where we first used Green's theorem, then the equations of the first order system and at last the weak formulation for u_2. The *a priori* estimate of the Lax-Milgram theorem gives

$$\|u_1\|_{1,\Omega} \lesssim \|w\|_{(H^1_D(\Omega))'} \leq \|w\|_{0,\Omega}\,,$$
$$\|u_2\|_{1,\Omega} \lesssim \|F\|_{(H^1_D(\Omega))'} \leq \|\boldsymbol{\eta}\|_{0,\Omega}\,.$$

Due to the splitting $u = u_1 + u_2$ it is now obvious that

$$\|u\|^2_{1,\Omega} \lesssim \|w\|^2_{0,\Omega} + \|\boldsymbol{\eta}\|^2_{0,\Omega}\,.$$

We now estimate the $\boldsymbol{H}(\Omega, \mathrm{div})$ norms of $\boldsymbol{\varphi}_1$ and $\boldsymbol{\varphi}_2$ as follows

$$\|\boldsymbol{\varphi}_1\|^2_{\boldsymbol{H}(\Omega,\mathrm{div})} = \|\boldsymbol{\varphi}_1\|^2_{0,\Omega} + \|\nabla\cdot\boldsymbol{\varphi}_1\|^2_{0,\Omega} = \|-\nabla u_1\|^2_{0,\Omega} + \|w - \gamma u_1\|^2_{0,\Omega} \lesssim \|w\|^2_{0,\Omega}\,,$$
$$\|\boldsymbol{\varphi}_2\|^2_{\boldsymbol{H}(\Omega,\mathrm{div})} = \|\boldsymbol{\varphi}_2\|^2_{0,\Omega} + \|\nabla\cdot\boldsymbol{\varphi}_2\|^2_{0,\Omega} = \|\boldsymbol{\eta} - \nabla u_2\|^2_{0,\Omega} + \|-\gamma u_2\|^2_{0,\Omega} \lesssim \|\boldsymbol{\eta}\|^2_{0,\Omega}\,,$$

which completes the proof. □

Remark 4.1.2. Theorem 4.1.1 (norm equivalence) does not hold on all of $\boldsymbol{H}(\Omega, \mathrm{div}) \times H^1(\Omega)$ since one can construct nontrivial solutions to the system

$$\begin{aligned} \nabla\cdot\boldsymbol{\varphi} + \gamma u &= 0 \qquad \text{in } \Omega,\\ \nabla u + \boldsymbol{\varphi} &= \mathbf{0} \qquad \text{in } \Omega, \end{aligned}$$

due to the missing boundary conditions, even though $\|(\boldsymbol{\varphi}, u)\|_b = 0$ by construction. ■

Remark 4.1.3. Theorem 4.1.1 (norm equivalence) is in fact much stronger than what we need to establish unique solvability of the system (4.5): The weaker coercivity estimate $\|u\|^2_{0,\Omega} + \|\boldsymbol{\varphi}\|^2_{0,\Omega} \lesssim b((\boldsymbol{\varphi}, u), (\boldsymbol{\varphi}, u))$ suffices to establish uniqueness. ■

Remark 4.1.4. In the literature there are two main ideas for showing unique solvability when working in a least squares setting concerning a first order system derived from a second order equation:

- The first one deduces solvability from the second order equation and uses some weaker coercivity estimates to establish uniqueness, as sketched in Remark 4.1.3. See also [CQ17, BM19] for these kinds of arguments for the Helmholtz equation.

- The second approach is to establish a stronger coercivity estimate as in Theorem 4.1.1 and directly apply the Lax-Milgram theorem to (4.5), where the right-hand side is a suitable continuous linear functional. See also [CLMM94, CMM97a] concerning the model problem in question and also [CMM97b] for the Stokes equation.

■

4.2. Duality argument

The current section is devoted to duality arguments that are later used for the analysis of the $L^2(\Omega)$ norms of $u - u_h$, $\nabla(u - u_h)$, and $\boldsymbol{\varphi} - \boldsymbol{\varphi}_h$. Since these duality arguments rely heavily on the elliptic shift theorem, we restrict ourself to either the pure Neumann or Dirichlet boundary conditions, i.e., $\Gamma = \Gamma_N$ or $\Gamma = \Gamma_D$, respectively. In contrast, when considering mixed boundary conditions one has to expect a singularity at the interface between the Dirichlet and Neumann condition, which has to be properly accounted for in the numerical analysis by graded meshes for both the primal and dual problem. This is beyond the scope of the present work. Our overall agenda is to derive regularity results for the dual solutions, always denoted by $(\boldsymbol{\psi}, v)$. For $w \in H^1(\Omega)$ and $\boldsymbol{\eta} \in \boldsymbol{H}_0(\Omega, \mathrm{div})$ we prove the existence of dual solutions such that:

- $\|w\|_{0,\Omega}^2 = b((\boldsymbol{\varphi}, w), (\boldsymbol{\psi}, v))$, see Theorem 4.2.1,
- $\|\nabla w\|_{0,\Omega}^2 = b((\boldsymbol{\varphi}, w), (\boldsymbol{\psi}, v))$, see Theorem 4.2.2,
- $\|\boldsymbol{\eta}\|_{0,\Omega}^2 = b((\boldsymbol{\eta}, u), (\boldsymbol{\psi}, v))$, see Theorem 4.2.3.

These results are exploited in Section 4.3 with the special choices of $w = u - u_h$ and $\boldsymbol{\eta} = \boldsymbol{\varphi} - \boldsymbol{\varphi}_h$, respectively.

Theorem 4.2.1 (Duality argument for the scalar variable)**.** *Let Γ be smooth. Then there holds:*

(i) For $\Gamma = \Gamma_N$ and any $(\boldsymbol{\varphi}, w) \in \boldsymbol{H}_0(\Omega, \mathrm{div}) \times H^1(\Omega)$ there exists $(\boldsymbol{\psi}, v) \in \boldsymbol{H}_0(\Omega, \mathrm{div}) \times H^1(\Omega)$ such that $\|w\|_{0,\Omega}^2 = b((\boldsymbol{\varphi}, w), (\boldsymbol{\psi}, v))$. Furthermore, $\boldsymbol{\psi} \in \boldsymbol{H}^3(\Omega)$, $\nabla \cdot \boldsymbol{\psi} \in H^2(\Omega)$, and $v \in H^2(\Omega)$. Additionally the following estimates hold:

$$\begin{aligned} \|v\|_{2,\Omega} &\lesssim \|w\|_{0,\Omega}\,, \\ \|\boldsymbol{\psi}\|_{3,\Omega} &\lesssim \|w\|_{0,\Omega}\,, \\ \|\nabla \cdot \boldsymbol{\psi}\|_{2,\Omega} &\lesssim \|w\|_{0,\Omega}\,. \end{aligned}$$

(ii) For $\Gamma = \Gamma_D$ and any $(\boldsymbol{\varphi}, w) \in \boldsymbol{H}(\Omega, \mathrm{div}) \times H_0^1(\Omega)$ there exists $(\boldsymbol{\psi}, v) \in \boldsymbol{H}(\Omega, \mathrm{div}) \times H_0^1(\Omega)$ such that $\|w\|_{0,\Omega}^2 = b((\boldsymbol{\varphi}, w), (\boldsymbol{\psi}, v))$. The same regularity results and estimates as in (i) hold.

Proof. We prove (i). Theorem 4.1.1 gives the existence of a unique $(\boldsymbol{\psi}, v) \in \boldsymbol{H}_0(\Omega, \mathrm{div}) \times H^1(\Omega)$ satisfying

$$(u, w)_\Omega = b((\boldsymbol{\varphi}, u), (\boldsymbol{\psi}, v)) \quad \forall\, (\boldsymbol{\varphi}, u) \in \boldsymbol{H}_0(\Omega, \mathrm{div}) \times H^1(\Omega). \tag{4.8}$$

For the regularity assertions, we introduce the auxiliary functions z and $\boldsymbol{\mu}$ by

$$\begin{aligned} \nabla \cdot \boldsymbol{\psi} + \gamma v &= z && \text{in } \Omega, \\ \nabla v + \boldsymbol{\psi} &= \boldsymbol{\mu} && \text{in } \Omega, \\ \boldsymbol{\psi} \cdot \boldsymbol{n} &= 0 && \text{on } \Gamma. \end{aligned} \tag{4.9}$$

Regularity properties of z and $\boldsymbol{\mu}$: Regularity properties of z are inferred from a scalar elliptic equation satisfied by z. To that end, we note that (4.8) is equivalent to

$$(u,w)_\Omega = (\nabla u + \boldsymbol{\varphi}, \boldsymbol{\mu})_\Omega + (\nabla\cdot\boldsymbol{\varphi} + \gamma u, z)_\Omega \quad \forall\,(\boldsymbol{\varphi},u) \in \boldsymbol{H}_0(\Omega,\mathrm{div}) \times H^1(\Omega). \tag{4.10}$$

For $u = 0$ and integrating by parts we find

$$0 = (\boldsymbol{\varphi},\boldsymbol{\mu})_\Omega + (\nabla\cdot\boldsymbol{\varphi}, z)_\Omega = (\boldsymbol{\varphi}, \boldsymbol{\mu} - \nabla z)_\Omega \qquad \forall \boldsymbol{\varphi} \in \boldsymbol{H}_0(\Omega,\mathrm{div}),$$

which gives $z \in H^1(\Omega)$ as well as $\boldsymbol{\mu} = \nabla z$. Inserting $\boldsymbol{\mu} = \nabla z$ and setting $\boldsymbol{\varphi} = 0$ in (4.10) we find

$$(u,w)_\Omega = (\nabla u, \nabla z)_\Omega + (\gamma u, z)_\Omega \quad \forall\, u \in H^1(\Omega).$$

Therefore z satisfies, in strong form,

$$\begin{aligned} -\Delta z + \gamma z &= w \quad \text{in } \Omega,\\ \partial_n z &= 0 \quad \text{on } \Gamma, \end{aligned} \tag{4.11}$$

and the shift theorem immediately gives $z \in H^2(\Omega)$ with the estimate $\|z\|_{2,\Omega} \lesssim \|w\|_{0,\Omega}$.

Regularity properties of v: Eliminating $\boldsymbol{\psi}$ in (4.9), we discover that v satisfies

$$\begin{aligned} -\Delta v + \gamma v &= w + (1-\gamma)z \quad && \text{in } \Omega,\\ \partial_n v &= 0 && \text{on } \Gamma. \end{aligned} \tag{4.12}$$

By elliptic regularity $v \in H^2(\Omega)$ with the *a priori* estimate

$$\|v\|_{2,\Omega} \lesssim \|w + (1-\gamma)z\|_{0,\Omega} \lesssim \|w\|_{0,\Omega}\,.$$

Regularity properties of $\boldsymbol{\psi}$: Setting $\boldsymbol{\psi} = \nabla(z - v)$, we have found the desired pair $(\boldsymbol{\psi}, v) \in \boldsymbol{H}_0(\Omega,\mathrm{div}) \times H^1(\Omega)$. Since $\boldsymbol{\psi} = \nabla(z-v)$, we first look at the regularity of $z - v$. Subtracting the equations (4.11), (4.12) satisfied by z and v, respectively we obtain

$$\begin{aligned} -\Delta(z-v) + \gamma(z-v) &= (\gamma - 1)z \quad && \text{in } \Omega,\\ \partial_n(z-v) &= 0 && \text{on } \Gamma, \end{aligned}$$

which gives $z - v \in H^4(\Omega)$ with the estimate

$$\|z - v\|_{4,\Omega} \lesssim \|(\gamma-1)z\|_{2,\Omega} \lesssim \|w\|_{0,\Omega}\,.$$

We can therefore deduce

$$\|\boldsymbol{\psi}\|_{3,\Omega} = \|\nabla(z-v)\|_{3,\Omega} \le \|z-v\|_{4,\Omega} \lesssim \|w\|_{0,\Omega}\,,$$

and since $\nabla\cdot\boldsymbol{\psi} = z - \gamma v$, we have

$$\|\nabla\cdot\boldsymbol{\psi}\|_{2,\Omega} = \|z - \gamma v\|_{2,\Omega} \lesssim \|w\|_{0,\Omega}\,,$$

which concludes the proof of (i). For the Dirichlet case (ii) the proof is completely analogous by replacing every Neumann boundary condition with a Dirichlet one. □

Theorem 4.2.2 (Duality argument for the gradient of the scalar variable)**.** *Let Γ be smooth. Then there holds:*

(i) For $\Gamma = \Gamma_N$ and any $(\boldsymbol{\varphi}, w) \in \boldsymbol{H}_0(\Omega, \mathrm{div}) \times H^1(\Omega)$ there exists $(\boldsymbol{\psi}, v) \in \boldsymbol{H}_0(\Omega, \mathrm{div}) \times H^1(\Omega)$ such that $\|\nabla w\|_{0,\Omega}^2 = b((\boldsymbol{\varphi}, w), (\boldsymbol{\psi}, v))$. Furthermore, $\boldsymbol{\psi} \in \boldsymbol{H}^2(\Omega)$, $\nabla \cdot \boldsymbol{\psi} \in H^1(\Omega)$, and $v \in H^1(\Omega)$. Additionally the following estimates hold:

$$\|v\|_{1,\Omega} \lesssim \|\nabla w\|_{0,\Omega}\,,$$
$$\|\boldsymbol{\psi}\|_{2,\Omega} \lesssim \|\nabla w\|_{0,\Omega}\,,$$
$$\|\nabla \cdot \boldsymbol{\psi}\|_{1,\Omega} \lesssim \|\nabla w\|_{0,\Omega}\,.$$

(ii) For $\Gamma = \Gamma_D$ and any $(\boldsymbol{\varphi}, w) \in \boldsymbol{H}(\Omega, \mathrm{div}) \times H_0^1(\Omega)$ there exists $(\boldsymbol{\psi}, v) \in \boldsymbol{H}(\Omega, \mathrm{div}) \times H_0^1(\Omega)$ such that $\|\nabla w\|_{0,\Omega}^2 = b((\boldsymbol{\varphi}, w), (\boldsymbol{\psi}, v))$. The same regularity results and estimates as in (i) hold.

Proof. We prove (i). Theorem 4.1.1 gives the existence of a unique $(\boldsymbol{\psi}, v) \in \boldsymbol{H}_0(\Omega, \mathrm{div}) \times H^1(\Omega)$ satisfying

$$(\nabla u, \nabla w)_\Omega = b((\boldsymbol{\varphi}, u), (\boldsymbol{\psi}, v)) \quad \forall\, (\boldsymbol{\varphi}, u) \in \boldsymbol{H}_0(\Omega, \mathrm{div}) \times H^1(\Omega). \tag{4.13}$$

For the regularity assertion, we introduce the auxiliary functions z and $\boldsymbol{\mu}$ by

$$\begin{aligned} \nabla \cdot \boldsymbol{\psi} + \gamma v &= z && \text{in } \Omega,\\ \nabla v + \boldsymbol{\psi} &= \boldsymbol{\mu} && \text{in } \Omega,\\ \boldsymbol{\psi} \cdot \boldsymbol{n} &= 0 && \text{on } \Gamma. \end{aligned} \tag{4.14}$$

Regularity properties of z and $\boldsymbol{\mu}$: We note that (4.13) is equivalent to

$$(\nabla u, \nabla w)_\Omega = (\nabla u + \boldsymbol{\varphi}, \boldsymbol{\mu})_\Omega + (\nabla \cdot \boldsymbol{\varphi} + \gamma u, z)_\Omega \quad \forall\, (\boldsymbol{\varphi}, u) \in \boldsymbol{H}_0(\Omega, \mathrm{div}) \times H^1(\Omega). \tag{4.15}$$

For $u = 0$ and integrating by parts we find

$$0 = (\boldsymbol{\varphi}, \boldsymbol{\mu})_\Omega + (\nabla \cdot \boldsymbol{\varphi}, z)_\Omega = (\boldsymbol{\varphi}, \boldsymbol{\mu} - \nabla z)_\Omega,$$

which gives $\boldsymbol{\mu} = \nabla z$. Inserting $\boldsymbol{\mu} = \nabla z$ and setting $\boldsymbol{\varphi} = 0$ in (4.15) we find

$$(\nabla u, \nabla w)_\Omega = (\nabla u, \nabla z)_\Omega + (\gamma u, z)_\Omega \quad \forall\, u \in H^1(\Omega),$$

which can be solved for $z \in H^1(\Omega)$ with the *a priori* estimate $\|z\|_{1,\Omega} \lesssim \|\nabla w\|_{0,\Omega}$. Formally, z satisfies

$$\begin{aligned} -\Delta z + \gamma z &= -\nabla \cdot \nabla w && \text{in } \Omega,\\ \partial_n z &= 0 && \text{on } \Gamma, \end{aligned} \tag{4.16}$$

where $-\nabla \cdot \nabla w \in (H^1(\Omega))'$ is to be understood as the mapping $u \mapsto (\nabla u, \nabla w)_\Omega$.

Regularity of v: Eliminating $\boldsymbol{\psi}$ from (4.14) and using $\boldsymbol{\mu} = \nabla z$, we discover that v satisfies

$$\begin{aligned} -\Delta v + \gamma v &= (1 - \gamma) z - \nabla \cdot \nabla w && \text{in } \Omega,\\ \partial_n v &= 0 && \text{on } \Gamma. \end{aligned}$$

By the Lax-Milgram theorem we find that $v \in H^1(\Omega)$ as well as

$$\|v\|_{1,\Omega} \lesssim \|(1-\gamma)z - \nabla\cdot\nabla w\|_{(H^1(\Omega))'} \lesssim \|\nabla w\|_{0,\Omega}\,.$$

Regularity of $\boldsymbol{\psi}$: Upon setting $\boldsymbol{\psi} = \nabla(z-v)$, we have found the solution $(\boldsymbol{\psi}, v) \in \boldsymbol{H}_0(\Omega,\mathrm{div}) \times H^1(\Omega)$ of (4.13). To prove the estimates and regularity results for $\boldsymbol{\psi}$ first note that

$$\begin{aligned} -\Delta(z-v) + \gamma(z-v) &= (1-\gamma)z \quad &&\text{in } \Omega,\\ \partial_n(z-v) &= 0 &&\text{on } \Gamma, \end{aligned}$$

and therefore by elliptic regularity $z - v \in H^3(\Omega)$ with the estimate

$$\|z-v\|_{3,\Omega} \lesssim \|(1-\gamma)z\|_{1,\Omega} \lesssim \|\nabla w\|_{0,\Omega}\,.$$

Finally since $\boldsymbol{\psi} = \nabla(z-v)$ the regularity assertion for $\boldsymbol{\psi} \in \boldsymbol{H}^2(\Omega)$ follows. For the Dirichlet case (ii) the proof is completely analogous by replacing every Neumann boundary condition with a Dirichlet one. □

Theorem 4.2.3 (Duality argument for the vector valued variable)**.** *Let Γ be smooth. Then there holds:*

(i) For $\Gamma = \Gamma_N$ and any $(\boldsymbol{\eta}, u) \in \boldsymbol{H}_0(\Omega,\mathrm{div}) \times H^1(\Omega)$ there exists $(\boldsymbol{\psi}, v) \in \boldsymbol{H}_0(\Omega,\mathrm{div}) \times H^1(\Omega)$ such that $\|\boldsymbol{\eta}\|_{0,\Omega}^2 = b((\boldsymbol{\eta},u),(\boldsymbol{\psi},v))$. Furthermore, $\boldsymbol{\psi} \in \boldsymbol{L}^2(\Omega)$, $\nabla\cdot\boldsymbol{\psi} \in H^1(\Omega)$ and $v \in H^3(\Omega)$. Additionally the following estimates hold:

$$\begin{aligned} \|v\|_{3,\Omega} &\lesssim \|\boldsymbol{\eta}\|_{0,\Omega}\,,\\ \|\boldsymbol{\psi}\|_{0,\Omega} &\lesssim \|\boldsymbol{\eta}\|_{0,\Omega}\,,\\ \|\nabla\cdot\boldsymbol{\psi}\|_{1,\Omega} &\lesssim \|\boldsymbol{\eta}\|_{0,\Omega}\,. \end{aligned}$$

(ii) For $\Gamma = \Gamma_D$ and any $(\boldsymbol{\eta}, u) \in \boldsymbol{H}(\Omega,\mathrm{div}) \times H_0^1(\Omega)$ there exists $(\boldsymbol{\psi}, v) \in \boldsymbol{H}(\Omega,\mathrm{div}) \times H_0^1(\Omega)$ such that $\|\boldsymbol{\eta}\|_{0,\Omega}^2 = b((\boldsymbol{\eta},u),(\boldsymbol{\psi},v))$. The same regularity results and estimates as in (i) hold.

Proof. We prove (i). Theorem 4.1.1 gives the existence of a unique $(\boldsymbol{\psi}, v) \in \boldsymbol{H}_0(\Omega,\mathrm{div}) \times H^1(\Omega)$ such that

$$(\boldsymbol{\varphi},\boldsymbol{\eta})_\Omega = b((\boldsymbol{\varphi},u),(\boldsymbol{\psi},v)) \quad \forall\,(\boldsymbol{\varphi},u) \in \boldsymbol{H}_0(\Omega,\mathrm{div}) \times H^1(\Omega). \tag{4.17}$$

For the regularity assertions, we introduce the auxiliary functions z and $\boldsymbol{\mu}$ by

$$\begin{aligned} \nabla\cdot\boldsymbol{\psi} + \gamma v &= z \qquad &&\text{in } \Omega,\\ \nabla v + \boldsymbol{\psi} &= \boldsymbol{\mu} &&\text{in } \Omega,\\ \boldsymbol{\psi}\cdot\boldsymbol{n} &= 0 &&\text{on } \Gamma. \end{aligned} \tag{4.18}$$

Regularity of z and $\boldsymbol{\mu}$: (4.17) is equivalent to

$$(\boldsymbol{\varphi},\boldsymbol{\eta})_\Omega = (\nabla u + \boldsymbol{\varphi}, \boldsymbol{\mu})_\Omega + (\nabla\cdot\boldsymbol{\varphi} + \gamma u, z)_\Omega \quad \forall\,(\boldsymbol{\varphi},u) \in \boldsymbol{H}_0(\Omega,\mathrm{div}) \times H^1(\Omega). \tag{4.19}$$

For $u = 0$ and integrating by parts we find

$$(\boldsymbol{\varphi}, \boldsymbol{\eta})_\Omega = (\boldsymbol{\varphi}, \boldsymbol{\mu})_\Omega + (\nabla \cdot \boldsymbol{\varphi}, z)_\Omega = (\boldsymbol{\varphi}, \boldsymbol{\mu} - \nabla z)_\Omega,$$

which gives $\boldsymbol{\mu} - \nabla z = \boldsymbol{\eta}$. Inserting $\boldsymbol{\mu} = \boldsymbol{\eta} + \nabla z$ and setting $\boldsymbol{\varphi} = 0$ in (4.17) we find

$$0 = (\nabla u, \boldsymbol{\eta} + \nabla z)_\Omega + (\gamma u, z)_\Omega \quad \forall\, u \in H^1(\Omega).$$

Hence, with the understanding that $\nabla \cdot \boldsymbol{\eta}$ means $u \mapsto (\nabla u, \boldsymbol{\eta})$, the function z solves

$$\begin{aligned} -\Delta z + \gamma z &= \nabla \cdot \boldsymbol{\eta} && \text{in } \Omega, \\ \partial_n z &= 0 && \text{on } \Gamma. \end{aligned} \tag{4.20}$$

Thus, $z \in H^1(\Omega)$ and setting $\boldsymbol{\mu} = \boldsymbol{\eta} + \nabla z$ we find (4.19) to be satisfied. Furthermore, note that

$$\|z\|_{1,\Omega} \lesssim \|\nabla \cdot \eta\|_{(H^1(\Omega))'} \leq \|\eta\|_{0,\Omega}\,,$$

where the last inequality following from integration by parts and exploiting the boundary condition $\eta \in \boldsymbol{H}_0(\Omega, \mathrm{div})$.

Regularity of v: By eliminating $\boldsymbol{\psi}$ we find that v solves

$$\begin{aligned} -\Delta v + \gamma v &= (1-\gamma) z && \text{in } \Omega, \\ \partial_n v &= 0 && \text{on } \Gamma. \end{aligned}$$

Again by elliptic regularity we find that $v \in H^3(\Omega)$ as well as

$$\|v\|_{3,\Omega} \lesssim \|(1-\gamma)z\|_{1,\Omega} \lesssim \|\eta\|_{0,\Omega}\,.$$

Regularity of $\boldsymbol{\psi}$: We have $\boldsymbol{\psi} = \boldsymbol{\eta} + \nabla(z - v)$, and the regularity of $\boldsymbol{\psi}$ follows from that of z of v. For the Dirichlet case (ii) the proof is completely analogous by replacing every Neumann boundary condition with a Dirichlet one. □

4.3. Error analysis

The goal of the present section is to establish optimal convergence rates for an *hp* version of the FOSLS method for the scalar variable, the gradient of the scalar variable as well as the vector variable, all measured in the $L^2(\Omega)$ norm, as long as the polynomial degree of the other variable is chosen appropriately.

4.3.1. Notation, assumptions, and road map of the current section

Throughout we denote by $(\boldsymbol{\varphi}_h, u_h)$ the least squares approximation of $(\boldsymbol{\varphi}, u)$. Furthermore, let $e^u = u - u_h$ and $\boldsymbol{e}^{\boldsymbol{\varphi}} = \boldsymbol{\varphi} - \boldsymbol{\varphi}_h$ denote the corresponding error terms. For simplicity we also assume $\Gamma = \Gamma_N$, i.e., $\Gamma_D = \emptyset$. Furthermore, p will denote the minimum of the two polynomial degrees p_s and p_v, i.e., $p = \min\{p_s, p_v\}$. The overall agenda of the present section is as follows:

1. We start off by proving [BG05, Lemma 3.4] in an hp setting using our duality argument, i.e., the (in our sense) suboptimal $L^2(\Omega)$ estimate

$$\|e^u\|_{0,\Omega} \lesssim h/p \, \|(\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u)\|_b \, .$$

   This is done in Lemma 4.3.1. In Remark 4.3.2 we present heuristic arguments that suggest the possibility of optimal $L^2(\Omega)$ convergence rates. These arguments suggest to construct an $\boldsymbol{H}_0(\Omega, \operatorname{div})$ conforming approximation operator $\boldsymbol{I}_h^0$ with additional orthogonality properties.

2. In Lemma 4.3.3 we prove that the operator $\boldsymbol{I}_h^0$ is in fact well-defined. As a tool of independent interest we derive certain continuous and discrete Helmholtz decompositions in Lemmas 4.3.4 and 4.3.5. These decompositions are then used in Lemma 4.3.6 to analyze the $L^2(\Omega)$ error of the operator $\boldsymbol{I}_h^0$.

3. Next we prove an hp version of [BG05, Lemma 3.6] (an h analysis of $\boldsymbol{e}^{\boldsymbol{\varphi}}$ in the $L^2(\Omega)$ norm).

4. In Theorem 4.3.10 we exploit the results of Lemma 4.3.9, which analyzes the convergence rate of the FOSLS approximation of the dual solution for the gradient of the scalar variable, in order to prove new optimal $L^2(\Omega)$ error estimates for ∇e^u.

5. We analyze the convergence rate of the FOSLS approximation of the dual solution in various norms in Lemma 4.3.11. Finally we prove our main result, Theorem 4.3.12, which analyzes the convergence of e^u in the $L^2(\Omega)$ norm.

6. Closing this section we derive Corollary 4.3.14, which summarizes the results for general right-hand side $f \in H^s(\Omega)$, by exploiting the estimates given by the Theorems 4.3.8, 4.3.10 and 4.3.12 together with the approximation properties of the employed finite element spaces.

Since we are dealing with smooth boundaries we employ curved elements. We assume the triangulation $\mathcal{T}_h$ to satisfy Assumption 2.0.1. We employ the scalar and vector valued finite element spaces as discussed in Chapter 2. For the approximation properties of the $\boldsymbol{H}(\Omega, \operatorname{div})$ conforming finite element spaces see [BBF13, Prop. 2.5.4] as a standard reference for noncurved elements and without the p-aspect. For an analysis of the hp-version under Assumption 2.0.1 we refer to Section 3.4.

4.3.2. The standard duality argument

Before formulating various duality arguments, we recall that the conforming least squares approximation $(\boldsymbol{\varphi}_h, u_h)$ is the best approximation in the $\|\cdot\|_b$ norm:

$$\|(\boldsymbol{\varphi} - \boldsymbol{\varphi}_h, u - u_h)\|_b = \min_{\substack{\tilde{u}_h \in S_{p_s}(\mathcal{T}_h), \\ \tilde{\boldsymbol{\varphi}}_h \in \mathbf{V}^0_{p_v}(\mathcal{T}_h)}} \|(\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h, u - \tilde{u}_h)\|_b. \tag{4.21}$$

Lemma 4.3.1. *Let Γ be smooth and $(\boldsymbol{\varphi}_h, u_h)$ be the least squares approximation of $(\boldsymbol{\varphi}, u)$. Furthermore, let $e^u = u - u_h$ and $\boldsymbol{e^{\varphi}} = \boldsymbol{\varphi} - \boldsymbol{\varphi}_h$. Then, for any $\tilde{u}_h \in S_{p_s}(\mathcal{T}_h)$, $\tilde{\boldsymbol{\varphi}}_h \in \mathbf{V}^0_{p_v}(\mathcal{T}_h)$,*

$$\begin{aligned}\|e^u\|_{0,\Omega} &\lesssim \frac{h}{p}\|(\boldsymbol{e^{\varphi}}, e^u)\|_b \\ &\lesssim \frac{h}{p}\|u - \tilde{u}_h\|_{1,\Omega} + \frac{h}{p}\|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h\|_{0,\Omega} + \frac{h}{p}\|\nabla \cdot (\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega}.\end{aligned}$$

Proof. Apply Theorem 4.2.1 (duality argument for the scalar variable) with $w = e^u$. For any $\tilde{v}_h \in S_{p_s}(\mathcal{T}_h)$, $\tilde{\boldsymbol{\psi}}_h \in \mathbf{V}^0_{p_v}(\mathcal{T}_h)$, we find due to the Galerkin orthogonality and the Cauchy-Schwarz inequality:

$$\begin{aligned}\|e^u\|^2_{0,\Omega} &= b((\boldsymbol{e^{\varphi}}, e^u), (\boldsymbol{\psi}, v)) \\ &= b((\boldsymbol{e^{\varphi}}, e^u), (\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h, v - \tilde{v}_h)) \\ &\leq \|(\boldsymbol{e^{\varphi}}, e^u)\|_b \|(\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h, v - \tilde{v}_h)\|_b.\end{aligned} \tag{4.22}$$

Using Theorem 4.1.1 (norm equivalence), and exploiting the regularity results and estimates of Theorem 4.2.1 as well as the $H^1(\Omega)$ and $\boldsymbol{H}(\Omega, \mathrm{div})$ conforming operators in [MR20], we can find $\tilde{v}_h \in S_{p_s}(\mathcal{T}_h)$, $\tilde{\boldsymbol{\psi}}_h \in \mathbf{V}^0_{p_v}(\mathcal{T}_h)$, such that

$$\begin{aligned}\|(\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h, v - \tilde{v}_h)\|_b &\lesssim \|v - \tilde{v}_h\|_{1,\Omega} + \|\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h\|_{\boldsymbol{H}(\Omega,\mathrm{div})} \\ &\lesssim h/p \left(\|v\|_{2,\Omega} + \|\boldsymbol{\psi}\|_{\boldsymbol{H}^1(\Omega,\mathrm{div})}\right) \\ &\lesssim h/p \, \|e^u\|_{0,\Omega},\end{aligned}$$

where we exploited the regularity for $(\boldsymbol{\psi}, v)$ and the *a priori* estimates of Theorem 4.2.1, which proves the first estimate. The second one follows by the fact that the least squares solution is the projection with respect to the scalar product b. Therefore,

$$\|(\boldsymbol{e^{\varphi}}, e^u)\|_b \leq \|(\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h, u - \tilde{u}_h)\|_b$$

holds. The result follows by applying the norm equivalence given in Theorem 4.1.1. □

Remark 4.3.2 (Heuristic arguments for improved $L^2(\Omega)$ convergence)**.** We present an argument why improved convergence of the scalar variable u can be expected. We again start by applying our duality argument and exploit the Galerkin orthogonality as in (4.22) in the proof of Lemma 4.3.1. Instead of immediately applying the Cauchy-Schwarz inequality we investigate the terms in the b scalar product and analyze the best rate we can expect from the regularity of the dual problem:

$$\begin{aligned}\|e^u\|^2_{0,\Omega} &= b((\boldsymbol{e^{\varphi}}, e^u), (\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h, v - \tilde{v}_h)) \\ &= (\underbrace{\nabla \cdot \boldsymbol{e^{\varphi}} + \gamma e^u}_{\frown}, \underbrace{\nabla \cdot (\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h)}_{\sim h^2} + \gamma \underbrace{(v - \tilde{v}_h)}_{\sim h^2})_\Omega + (\underbrace{\nabla e^u + \boldsymbol{e^{\varphi}}}_{\frown}, \underbrace{\nabla(v - \tilde{v}_h)}_{\sim h} + \underbrace{\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h}_{\sim h^3})_\Omega.\end{aligned}$$

Note that the terms are not equilibrated and we cannot expect any rate from the terms marked by ☹. However, choosing $(\tilde{\boldsymbol{\psi}}_h, \tilde{v}_h)$ to be the least squares approximation $(\boldsymbol{\psi}_h, v_h)$

of $(\boldsymbol{\psi}, v)$ and again exploiting the Galerkin orthogonality we have for any $(\tilde{\boldsymbol{\varphi}}_h, \tilde{u}_h)$:

$$\begin{aligned}
\|e^u\|_{0,\Omega}^2 &= b((\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u), (\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)) \\
&= b((\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h, u - \tilde{u}_h), (\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)) \\
&= (\underbrace{\nabla \cdot (\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h)}_{☹} + \gamma \underbrace{(u - \tilde{u}_h)}_{\sim h^2}, \underbrace{\nabla \cdot \boldsymbol{e}^{\boldsymbol{\psi}}}_{\sim h} + \gamma \underbrace{e^v}_{\sim h^2})_\Omega + (\underbrace{\nabla (u - \tilde{u}_h)}_{\sim h} + \underbrace{\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h}_{\sim h}, \underbrace{\nabla e^v + \boldsymbol{e}^{\boldsymbol{\psi}}}_{\sim h})_\Omega.
\end{aligned}$$

The improved convergence of the dual solution will be shown in Lemma 4.3.11. From a best approximation viewpoint the $\nabla\cdot$ term involving $\boldsymbol{\varphi}$ still has no rate. To be more precise, the second term has the right powers of h resulting in an overall h^2. Since the term $\gamma(u - \tilde{u}_h)$ already has order h^2 we have no problem with that one. The term with the worst rate is

$$(\nabla \cdot (\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h), \nabla \cdot \boldsymbol{e}^{\boldsymbol{\psi}})_\Omega \sim h.$$

Out of the box we cannot find an extra h to get optimal convergence, even though $\boldsymbol{\psi}$ has far more regularity, which we did not exploit yet. We now want to construct an operator $\boldsymbol{I}_h^0$ mapping into the conforming finite element space of the vector variable. To exploit the regularity of $\boldsymbol{\psi}$ we insert any $\tilde{\boldsymbol{\psi}}_h \in \mathbf{V}_{p_v}^0(\mathcal{T}_h)$. We have

$$(\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{I}_h^0\boldsymbol{\varphi}), \nabla \cdot \boldsymbol{e}^{\boldsymbol{\psi}})_\Omega = (\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{I}_h^0\boldsymbol{\varphi}), \nabla \cdot (\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h))_\Omega + (\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{I}_h^0\boldsymbol{\varphi}), \nabla \cdot (\tilde{\boldsymbol{\psi}}_h - \boldsymbol{\psi}_h))_\Omega.$$

Note that $\tilde{\boldsymbol{\psi}}_h - \boldsymbol{\psi}_h$ is a discrete object. If we assume $\boldsymbol{I}_h^0$ to satisfy the orthogonality condition

$$(\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{I}_h^0\boldsymbol{\varphi}), \nabla \cdot \boldsymbol{\chi}_h)_\Omega = 0 \qquad \forall \boldsymbol{\chi}_h \in \mathbf{V}_{p_v}^0(\mathcal{T}_h),$$

we arrive at

$$(\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{I}_h^0\boldsymbol{\varphi}), \nabla \cdot \boldsymbol{e}^{\boldsymbol{\psi}})_\Omega = (\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{I}_h^0\boldsymbol{\varphi}), \underbrace{\nabla \cdot (\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h)}_{h^2})_\Omega \sim h^2.$$

Therefore the operator $\boldsymbol{I}_h^0$ should satisfy the aforementioned orthogonality condition and have good approximation properties in $L^2(\Omega)$, as needed above. In the following we will construct operators $\boldsymbol{I}_h^0$ and $\boldsymbol{I}_h$ acting on $\boldsymbol{H}_0(\Omega, \mathrm{div})$ and $\boldsymbol{H}(\Omega, \mathrm{div})$, respectively. ∎

4.3.3. The operators $\boldsymbol{I}_h^0$ and $\boldsymbol{I}_h$

In the spirit of Remark 4.3.2 a natural choice for the operator $\boldsymbol{I}_h^0$ is the following constrained minimization problem

$$\boldsymbol{I}_h^0\boldsymbol{\varphi} = \underset{\boldsymbol{\varphi}_h \in \mathbf{V}_{p_v}^0(\mathcal{T}_h)}{\operatorname{argmin}} \frac{1}{2} \|\boldsymbol{\varphi} - \boldsymbol{\varphi}_h\|_{0,\Omega}^2 \quad \text{s.t.} \quad (\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{I}_h^0\boldsymbol{\varphi}), \nabla \cdot \boldsymbol{\chi}_h)_\Omega = 0 \quad \forall \boldsymbol{\chi}_h \in \mathbf{V}_{p_v}^0(\mathcal{T}_h).$$

The corresponding Lagrange function is

$$L(\boldsymbol{\varphi}_h, \boldsymbol{\lambda}_h) = \frac{1}{2} \|\boldsymbol{\varphi}_h - \boldsymbol{\varphi}\|_{0,\Omega}^2 + (\nabla \cdot (\boldsymbol{\varphi}_h - \boldsymbol{\varphi}), \nabla \cdot \boldsymbol{\lambda}_h)_\Omega$$

and the associated saddle point problem is to find $(\boldsymbol{\varphi}_h, \boldsymbol{\lambda}_h) \in \mathbf{V}^0_{p_v}(\mathcal{T}_h) \times \mathbf{V}^0_{p_v}(\mathcal{T}_h)$ such that

$$
\begin{aligned}
(\boldsymbol{\varphi}_h - \boldsymbol{\varphi}, \boldsymbol{\mu}_h)_\Omega + (\nabla \cdot \boldsymbol{\mu}_h, \nabla \cdot \boldsymbol{\lambda}_h)_\Omega &= 0 \qquad \forall \boldsymbol{\mu}_h \in \mathbf{V}^0_{p_v}(\mathcal{T}_h), &\text{(4.23a)}\\
(\nabla \cdot (\boldsymbol{\varphi}_h - \boldsymbol{\varphi}), \nabla \cdot \boldsymbol{\eta}_h)_\Omega &= 0 \qquad \forall \boldsymbol{\eta}_h \in \mathbf{V}^0_{p_v}(\mathcal{T}_h). &\text{(4.23b)}
\end{aligned}
$$

Uniqueness is not given since only the divergence of the Lagrange parameter appears. However, by focussing on the divergence of the Lagrange parameter, we can formulate it in the following way: Find $(\boldsymbol{\varphi}_h, \lambda_h) \in \mathbf{V}^0_{p_v}(\mathcal{T}_h) \times \nabla \cdot \mathbf{V}^0_{p_v}(\mathcal{T}_h)$ such that

$$
\begin{aligned}
(\boldsymbol{\varphi}_h, \boldsymbol{\mu}_h)_\Omega + (\nabla \cdot \boldsymbol{\mu}_h, \lambda_h)_\Omega &= (\boldsymbol{\varphi}, \boldsymbol{\mu}_h)_\Omega \qquad \forall \boldsymbol{\mu}_h \in \mathbf{V}^0_{p_v}(\mathcal{T}_h), &\text{(4.24a)}\\
(\nabla \cdot \boldsymbol{\varphi}_h, \eta_h)_\Omega &= (\nabla \cdot \boldsymbol{\varphi}, \eta_h)_\Omega \qquad \forall \eta_h \in \nabla \cdot \mathbf{V}^0_{p_v}(\mathcal{T}_h). &\text{(4.24b)}
\end{aligned}
$$

The construction of $\boldsymbol{I}_h$ is completely analogous, one just drops the zero boundary conditions everywhere.

To see that the operator $\boldsymbol{I}^0_h$ is well-defined, we have to check the Babuška–Brezzi conditions, see [BBF13]. First, let us verify solvability on the continuous level.

Coercivity on the kernel: Let $\boldsymbol{\mu} \in \boldsymbol{H}_0(\Omega, \mathrm{div})$ with $(\nabla \cdot \boldsymbol{\mu}, \eta)_\Omega = 0$ for all $\eta \in \nabla \cdot \boldsymbol{H}_0(\Omega, \mathrm{div})$ be given. The coercivity is trivial since by construction $(\nabla \cdot \boldsymbol{\mu}, \nabla \cdot \boldsymbol{\mu})_\Omega = 0$ and therefore

$$
(\boldsymbol{\mu}, \boldsymbol{\mu})_\Omega = \|\boldsymbol{\mu}\|^2_{0,\Omega} = \|\boldsymbol{\mu}\|^2_{0,\Omega} + \|\nabla \cdot \boldsymbol{\mu}\|^2_{0,\Omega} = \|\boldsymbol{\mu}\|^2_{\boldsymbol{H}(\Omega,\mathrm{div})}.
$$

inf-sup condition: Let $\eta \in \nabla \cdot \boldsymbol{H}_0(\Omega, \mathrm{div})$ be given. First let $u \in H^1(\Omega)$ with zero average solve

$$
\begin{aligned}
-\Delta u &= \eta \quad \text{in } \Omega,\\
\partial_n u &= 0 \quad \text{on } \Gamma.
\end{aligned}
$$

By elliptic regularity we have $\|u\|_{2,\Omega} \lesssim \|\eta\|_{0,\Omega}$ and upon defining $\boldsymbol{\mu} = -\nabla u$ we also have $\|\boldsymbol{\mu}\|_{\boldsymbol{H}(\Omega,\mathrm{div})} \lesssim \|\eta\|_{0,\Omega}$. Note that by construction $\boldsymbol{\mu} \in \boldsymbol{H}_0(\Omega, \mathrm{div})$ as well as

$$
(\nabla \cdot \boldsymbol{\mu}, \eta)_\Omega = (\eta, \eta)_\Omega = \|\eta\|_{0,\Omega} \|\eta\|_{0,\Omega} \gtrsim \|\eta\|_{0,\Omega} \|\boldsymbol{\mu}\|_{\boldsymbol{H}(\Omega,\mathrm{div})},
$$

which proves the inf-sup condition.

Coercivity on the kernel - discrete: The coercivity is again trivial by the same argument as above.

inf-sup condition - discrete: Let $\lambda_h \in \nabla \cdot \mathbf{V}^0_{p_v}(\mathcal{T}_h)$ be given. As above in the continuous case we solve the Poisson problem

$$
\begin{aligned}
-\Delta u &= \lambda_h \quad \text{in } \Omega,\\
\partial_n u &= 0 \quad \text{on } \Gamma.
\end{aligned}
$$

Let $\boldsymbol{\Lambda} = -\nabla u$ and again we have $\|\boldsymbol{\Lambda}\|_{\boldsymbol{H}(\Omega,\mathrm{div})} \leq \|\boldsymbol{\Lambda}\|_{1,\Omega} \leq \|u\|_{2,\Omega} \lesssim \|\lambda_h\|_{0,\Omega}$. We now employ the commuting projection based interpolation operators defined in [MR20], especially the global operator $\boldsymbol{\Pi}^{\mathrm{div}}_p$ given in [MR20, Remark 2.9], see also [Roj19, Sec. 4.8] in the case $\mathbf{V}^0_{p_v}(\mathcal{T}_h) = \mathbf{BDM}^0_{p_v}(\mathcal{T}_h)$. Let therefore $\boldsymbol{\Pi}^{\mathrm{div},\star}_{p_v}$ denote either the operator $\boldsymbol{\Pi}^{\mathrm{div}}_{p_v-1}$ if $\mathbf{V}^0_{p_v}(\mathcal{T}_h) =$

$\mathbf{RT}^0_{p_v-1}(\mathcal{T}_h)$ or the analogous operator $\boldsymbol{\Pi}^{\mathrm{div}}_{p_v}$ in the case $\mathbf{V}^0_{p_v}(\mathcal{T}_h) = \mathbf{BDM}^0_{p_v}(\mathcal{T}_h)$. We use this operator to project $\boldsymbol{\Lambda}$ onto the conforming subspace. With $\boldsymbol{\Lambda}_h := \boldsymbol{\Pi}^{\mathrm{div},\star}_{p_v}\boldsymbol{\Lambda}$ we find

$$\nabla \cdot \boldsymbol{\Lambda}_h = \nabla \cdot \boldsymbol{\Pi}^{\mathrm{div},\star}_{p_v}\boldsymbol{\Lambda} = \Pi^{L^2}_{p_v}\nabla \cdot \boldsymbol{\Lambda} = \Pi^{L^2}_{p_v}\lambda_h = \lambda_h,$$

where $\Pi^{L^2}_{p_v}$ denotes the L^2 orthogonal projection on $\nabla \cdot \mathbf{V}^0_{p_v}(\mathcal{T}_h)$. Using [MR20, Thm 2.10, (vi)] we can estimate

$$\|\boldsymbol{\Lambda} - \boldsymbol{\Pi}^{\mathrm{div},\star}_{p_v}\boldsymbol{\Lambda}\|_{\boldsymbol{H}(\Omega,\mathrm{div})} \lesssim \|\boldsymbol{\Lambda}\|_{1,\Omega} \lesssim \|\lambda_h\|_{0,\Omega}\,,$$

which finally leads to

$$\|\boldsymbol{\Lambda}_h\|_{\boldsymbol{H}(\Omega,\mathrm{div})} = \|\boldsymbol{\Pi}^{\mathrm{div},\star}_{p_v}\boldsymbol{\Lambda}\|_{\boldsymbol{H}(\Omega,\mathrm{div})} \lesssim \|\boldsymbol{\Lambda} - \boldsymbol{\Pi}^{\mathrm{div},\star}_{p_v}\boldsymbol{\Lambda}\|_{\boldsymbol{H}(\Omega,\mathrm{div})} + \|\boldsymbol{\Lambda}\|_{\boldsymbol{H}(\Omega,\mathrm{div})} \lesssim \|\lambda_h\|_{0,\Omega}\,.$$

For any $\lambda_h \in \nabla \cdot \mathbf{V}^0_{p_v}(\mathcal{T}_h)$ we estimate

$$\sup_{\boldsymbol{\varphi}_h \in \mathbf{V}^0_{p_v}(\mathcal{T}_h)} \frac{(\nabla \cdot \boldsymbol{\varphi}_h, \lambda_h)_\Omega}{\|\boldsymbol{\varphi}_h\|_{\boldsymbol{H}(\Omega,\mathrm{div})}\,\|\lambda_h\|_{0,\Omega}} \geq \frac{(\nabla \cdot \boldsymbol{\Lambda}_h, \lambda_h)_\Omega}{\|\boldsymbol{\Lambda}_h\|_{\boldsymbol{H}(\Omega,\mathrm{div})}\,\|\lambda_h\|_{0,\Omega}} = \frac{\|\lambda_h\|_{0,\Omega}}{\|\boldsymbol{\Lambda}_h\|_{\boldsymbol{H}(\Omega,\mathrm{div})}} \gtrsim 1,$$

which proves the discrete inf-sup condition. The above arguments can be modified in a straightforward manner when replacing the discrete space $\mathbf{V}^0_{p_v}(\mathcal{T}_h)$ with $\mathbf{V}_{p_v}(\mathcal{T}_h)$ and $\boldsymbol{H}_0(\Omega,\mathrm{div})$ with $\boldsymbol{H}(\Omega,\mathrm{div})$. The only caveat is the fact that one has to replace the homogeneous Neumann boundary condition in the auxiliary problem, used in the verification of the inf-sup condition, by a homogeneous Dirichlet boundary condition. We have therefore proven

Lemma 4.3.3. *For any mesh $\mathcal{T}_h$ satisfying Assumption 2.0.1, the operators $\boldsymbol{I}^0_h : \boldsymbol{H}_0(\Omega,\mathrm{div}) \to \mathbf{V}^0_{p_v}(\mathcal{T}_h)$ and $\boldsymbol{I}_h : \boldsymbol{H}(\Omega,\mathrm{div}) \to \mathbf{V}_{p_v}(\mathcal{T}_h)$ are well-defined with bounds independent of the mesh size h and the polynomial degree p_v. They are projections.*

We are now going to analyze the approximation properties of the operator $\boldsymbol{I}^0_h$ and $\boldsymbol{I}_h$ in the $L^2(\Omega)$ norm. To that end, we need certain decompositions on a continuous as well as a discrete level.

Lemma 4.3.4 (Continuous and discrete Helmholtz-like decomposition - no boundary conditions)**.** *The operators $\boldsymbol{\Pi}^{\mathrm{curl}} : \boldsymbol{H}(\Omega,\mathrm{div}) \to \nabla \times \boldsymbol{H}(\Omega,\mathrm{curl})$ and $\boldsymbol{\Pi}^{\mathrm{curl}}_h : \mathbf{V}_{p_v}(\mathcal{T}_h) \to \nabla \times \mathbf{N}_{p_v}(\mathcal{T}_h)$ given by*

$$(\boldsymbol{\Pi}^{\mathrm{curl}}\boldsymbol{\varphi}, \nabla \times \boldsymbol{\mu})_\Omega = (\boldsymbol{\varphi}, \nabla \times \boldsymbol{\mu})_\Omega \quad \forall \boldsymbol{\mu} \in \boldsymbol{H}(\Omega,\mathrm{curl}), \tag{4.25}$$

$$(\boldsymbol{\Pi}^{\mathrm{curl}}_h\boldsymbol{\varphi}_h, \nabla \times \boldsymbol{\mu})_\Omega = (\boldsymbol{\varphi}_h, \nabla \times \boldsymbol{\mu})_\Omega \quad \forall \boldsymbol{\mu} \in \mathbf{N}_{p_v}(\mathcal{T}_h) \tag{4.26}$$

are well-defined. Furthermore, the remainder $\boldsymbol{r}$ of the continuous decomposition $\boldsymbol{\varphi} = \boldsymbol{\Pi}^{\mathrm{curl}}\boldsymbol{\varphi} + \boldsymbol{r}$ satisfies

$$\begin{aligned} \nabla \cdot \boldsymbol{r} &= \nabla \cdot \boldsymbol{\varphi} && \textit{in } \Omega, \\ \nabla \times \boldsymbol{r} &= 0 && \textit{in } \Omega, \\ \boldsymbol{n} \times \boldsymbol{r} &= 0 && \textit{on } \Gamma, \end{aligned}$$

as well as $\boldsymbol{r} \in \boldsymbol{H}^1(\Omega)$*. Additionally there exists* $R \in H^2(\Omega) \cap H_0^1(\Omega)$ *such that* $\boldsymbol{r} = \nabla R$*, where* R *satisfies*

$$\begin{aligned} \Delta R &= \nabla \cdot \boldsymbol{\varphi} && \text{in } \Omega, \\ R &= 0 && \text{on } \Gamma. \end{aligned}$$

Finally, the estimate $\|R\|_{2,\Omega} \lesssim \|\boldsymbol{r}\|_{1,\Omega} \lesssim \|\nabla \cdot \boldsymbol{\varphi}\|_{0,\Omega}$ *holds.*

Proof. For unique solvability of the variational definition of the operators, just note that they are the $L^2(\Omega)$ orthogonal projection on $\nabla \times \boldsymbol{H}(\Omega, \mathrm{curl})$ and $\nabla \times \mathbf{N}_{p_v}(\mathcal{T}_h)$, respectively. By construction we have

$$(\boldsymbol{r}, \nabla \times \boldsymbol{\mu})_\Omega = 0 \quad \forall \boldsymbol{\mu} \in \boldsymbol{H}(\Omega, \mathrm{curl}),$$

which by definition gives $\nabla \times \boldsymbol{r} = 0$. Furthermore, by the characterization of $\boldsymbol{H}_0(\Omega, \mathrm{curl})$ given in [Mon03, Thm. 3.33] we have $\boldsymbol{n} \times \boldsymbol{r} = 0$. Since $\boldsymbol{\Pi}^{\mathrm{curl}}\boldsymbol{\varphi} \in \nabla \times \boldsymbol{H}(\Omega, \mathrm{curl})$ we immediately have $\nabla \cdot \boldsymbol{r} = \nabla \cdot \boldsymbol{\varphi}$. Exploiting the exact sequence property of the following de Rham complex

$$\{0\} \xrightarrow{\mathrm{id}} H_0^1(\Omega) \xrightarrow{\nabla} \boldsymbol{H}_0(\Omega, \mathrm{curl}) \xrightarrow{\nabla\times} \boldsymbol{H}_0(\Omega, \mathrm{div}) \xrightarrow{\nabla\cdot} L_0^2(\Omega) \xrightarrow{0} \{0\}$$

in the case that both Ω and Γ are simply connected, we can find $R \in H_0^1(\Omega)$ such that $\boldsymbol{r} = \nabla R$. Therefore R solves the asserted equation. The Friedrichs inequality and elliptic regularity theory then give the desired results. ☐

By nearly the same arguments we also have a version for zero boundary conditions:

Lemma 4.3.5 (Continuous and discrete Helmholtz-like decomposition - zero boundary conditions)**.** *The operators* $\boldsymbol{\Pi}^{\mathrm{curl},0} \colon \boldsymbol{H}_0(\Omega, \mathrm{div}) \to \nabla \times \boldsymbol{H}_0(\Omega, \mathrm{curl})$ *and* $\boldsymbol{\Pi}_h^{\mathrm{curl},0} \colon \mathbf{V}_{p_v}^0(\mathcal{T}_h) \to \nabla \times \mathbf{N}_{p_v}^0(\mathcal{T}_h)$ *given by*

$$(\boldsymbol{\Pi}^{\mathrm{curl},0}\boldsymbol{\varphi}, \nabla \times \boldsymbol{\mu})_\Omega = (\boldsymbol{\varphi}, \nabla \times \boldsymbol{\mu})_\Omega \quad \forall \boldsymbol{\mu} \in \boldsymbol{H}_0(\Omega, \mathrm{curl}), \tag{4.27}$$

$$(\boldsymbol{\Pi}_h^{\mathrm{curl},0}\boldsymbol{\varphi}_h, \nabla \times \boldsymbol{\mu})_\Omega = (\boldsymbol{\varphi}_h, \nabla \times \boldsymbol{\mu})_\Omega \quad \forall \boldsymbol{\mu} \in \mathbf{N}_{p_v}^0(\mathcal{T}_h) \tag{4.28}$$

are well-defined. Furthermore, the remainder $\boldsymbol{r}$ *of the continuous decomposition* $\boldsymbol{\varphi} = \boldsymbol{\Pi}^{\mathrm{curl},0}\boldsymbol{\varphi} + \boldsymbol{r}$ *satisfies*

$$\begin{aligned} \nabla \cdot \boldsymbol{r} &= \nabla \cdot \boldsymbol{\varphi} && \text{in } \Omega, \\ \nabla \times \boldsymbol{r} &= 0 && \text{in } \Omega, \\ \boldsymbol{r} \cdot \boldsymbol{n} &= 0 && \text{on } \Gamma, \end{aligned}$$

as well as $\boldsymbol{r} \in \boldsymbol{H}^1(\Omega)$*. Additionally there exists an* $R \in H^2(\Omega) \cap H^1(\Omega)/\mathbb{R}$ *such that* $\boldsymbol{r} = \nabla R$*, where* R *satisfies*

$$\begin{aligned} \Delta R &= \nabla \cdot \boldsymbol{\varphi} && \text{in } \Omega, \\ \partial_n R &= 0 && \text{on } \Gamma. \end{aligned}$$

Finally, the estimate $\|R\|_{2,\Omega} \lesssim \|\boldsymbol{r}\|_{1,\Omega} \lesssim \|\nabla \cdot \boldsymbol{\varphi}\|_{0,\Omega}$ *holds.*

Proof. Unique solvability, $\nabla \times \boldsymbol{r} = 0$ and $\nabla \cdot \boldsymbol{r} = \nabla \cdot \boldsymbol{\varphi}$ follows by the same arguments as in the proof of Lemma 4.3.4. Since $\boldsymbol{\varphi} \in \boldsymbol{H}_0(\Omega, \mathrm{div})$ and $\boldsymbol{\Pi}^{\mathrm{curl},0}\boldsymbol{\varphi} \in \nabla \times \boldsymbol{H}_0(\Omega, \mathrm{curl}) \subset \boldsymbol{H}_0(\Omega, \mathrm{div})$ we find

$$\boldsymbol{r} \cdot \boldsymbol{n} = \boldsymbol{\varphi} \cdot \boldsymbol{n} - \boldsymbol{\Pi}^{\mathrm{curl},0}\boldsymbol{\varphi} \cdot \boldsymbol{n} = 0.$$

Again by the exact sequence

$$\mathbb{R} \xrightarrow{\mathrm{id}} H^1(\Omega) \xrightarrow{\nabla} \boldsymbol{H}(\Omega, \mathrm{curl}) \xrightarrow{\nabla\times} \boldsymbol{H}(\Omega, \mathrm{div}) \xrightarrow{\nabla\cdot} L^2(\Omega) \xrightarrow{0} \{0\}$$

we can find $R \in H^1(\Omega)$ such that $\boldsymbol{r} = \nabla R$. Finally since $\partial_n R = \nabla R \cdot \boldsymbol{n} = \boldsymbol{r} \cdot \boldsymbol{n} = 0$, we find that R solves the asserted equation. The Poincaré inequality and elliptic regularity theory then give the desired results. □

Lemma 4.3.6. *The operator $\boldsymbol{I}_h^0$ satisfies for arbitrary $\tilde{\boldsymbol{\varphi}}_h \in \mathbf{V}_{p_v}^0(\mathcal{T}_h)$ the estimates*

$$\left\|\boldsymbol{\varphi} - \boldsymbol{I}_h^0\boldsymbol{\varphi}\right\|_{0,\Omega} \lesssim \|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h\|_{0,\Omega} + \frac{h}{p_v}\|\nabla \cdot (\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega}, \tag{4.29}$$

$$\left\|\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{I}_h^0\boldsymbol{\varphi})\right\|_{0,\Omega} \leq \|\nabla \cdot (\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega}. \tag{4.30}$$

The same estimates hold true for the operator $\boldsymbol{I}_h$ for arbitrary $\tilde{\boldsymbol{\varphi}}_h \in \mathbf{V}_{p_v}(\mathcal{T}_h)$.

Proof. Let $\tilde{\boldsymbol{\varphi}}_h \in \mathbf{V}_{p_v}^0(\mathcal{T}_h)$ be arbitrary. Due to the orthogonality relation satisfied by the operator $\boldsymbol{I}_h^0$ the estimate (4.30) is obvious. We have with $\boldsymbol{e} = \boldsymbol{\varphi} - \boldsymbol{I}_h^0\boldsymbol{\varphi}$

$$\|\boldsymbol{e}\|_{0,\Omega}^2 = (\boldsymbol{e}, \boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h)_\Omega + (\boldsymbol{e}, \tilde{\boldsymbol{\varphi}}_h - \boldsymbol{I}_h^0\boldsymbol{\varphi})_\Omega.$$

In order to treat the second term we apply Lemma 4.3.5 and split the discrete object $\tilde{\boldsymbol{\varphi}}_h - \boldsymbol{I}_h^0\boldsymbol{\varphi} \in \mathbf{V}_{p_v}^0(\mathcal{T}_h)$ on a discrete and a continuous level. That is,

$$\tilde{\boldsymbol{\varphi}}_h - \boldsymbol{I}_h^0\boldsymbol{\varphi} = \nabla \times \boldsymbol{\mu} + \boldsymbol{r},$$
$$\tilde{\boldsymbol{\varphi}}_h - \boldsymbol{I}_h^0\boldsymbol{\varphi} = \nabla \times \boldsymbol{\mu}_h + \boldsymbol{r}_h$$

for certain $\boldsymbol{\mu} \in \boldsymbol{H}_0(\Omega, \mathrm{curl})$, $\boldsymbol{r} \in \boldsymbol{H}_0(\Omega, \mathrm{div})$, $\boldsymbol{\mu}_h \in \mathbf{N}_{p_v}^0(\mathcal{T}_h)$, and $\boldsymbol{r}_h \in \mathbf{V}_{p_v}^0(\mathcal{T}_h)$. Since $\nabla \cdot \nabla\times = 0$ we have

$$(\boldsymbol{\varphi} - \boldsymbol{I}_h^0\boldsymbol{\varphi}, \nabla \times \boldsymbol{\mu}_h)_\Omega = 0$$

by definition of the operator $\boldsymbol{I}_h^0$ and consequently

$$(\boldsymbol{e}, \tilde{\boldsymbol{\varphi}}_h - \boldsymbol{I}_h^0\boldsymbol{\varphi})_\Omega = (\boldsymbol{e}, \nabla \times \boldsymbol{\mu}_h + \boldsymbol{r}_h)_\Omega = (\boldsymbol{e}, \boldsymbol{r}_h)_\Omega = (\boldsymbol{e}, \boldsymbol{r}_h - \boldsymbol{r})_\Omega + (\boldsymbol{e}, \boldsymbol{r})_\Omega =: T_1 + T_2.$$

Treatment of T_1: To estimate T_1 we first need one of the commuting projection based interpolation operators defined in [MR20]. Specifically we employ the global operator $\boldsymbol{\Pi}_p^{\mathrm{div}}$ given in [MR20, Remark 2.9], see also [Roj19]. Let therefore $\boldsymbol{\Pi}_{p_v}^{\mathrm{div},\star}$ denote either the operator $\boldsymbol{\Pi}_{p_v-1}^{\mathrm{div}}$ if $\mathbf{V}_{p_v}^0(\mathcal{T}_h) = \mathbf{RT}_{p_v-1}^0(\mathcal{T}_h)$ or the analogous operator $\boldsymbol{\Pi}_{p_v}^{\mathrm{div}}$ in the case $\mathbf{V}_{p_v}^0(\mathcal{T}_h) = \mathbf{BDM}_{p_v}^0(\mathcal{T}_h)$. First note that $\nabla \cdot \boldsymbol{r} = \nabla \cdot \boldsymbol{r}_h \in \nabla \cdot \mathbf{V}_{p_v}^0(\mathcal{T}_h)$. By the commuting diagram property of the operator $\boldsymbol{\Pi}_{p_v}^{\mathrm{div},\star}$ as well as the projection property we therefore have

$$\nabla \cdot (\boldsymbol{\Pi}_{p_v}^{\mathrm{div},\star}\boldsymbol{r} - \boldsymbol{r}_h) = \Pi_{p_v}^{L^2}(\nabla \cdot \boldsymbol{r}) - \nabla \cdot \boldsymbol{r}_h = \nabla \cdot \boldsymbol{r} - \nabla \cdot \boldsymbol{r}_h = 0.$$

By the exact sequence property we therefore have $\mathbf{\Pi}_{p_v}^{\mathrm{div},\star}\boldsymbol{r}-\boldsymbol{r}_h \in \nabla\times\mathbf{N}_{p_v}^0(\mathcal{T}_h)$. Furthermore, the definition of $\boldsymbol{r}$ and $\boldsymbol{r}_h$ in Lemma 4.3.5 gives the orthogonality relation $\boldsymbol{r}-\boldsymbol{r}_h \perp \nabla \times \mathbf{N}_{p_v}^0(\mathcal{T}_h)$. Putting it all together we have

$$\|\boldsymbol{r}-\boldsymbol{r}_h\|_{0,\Omega}^2 = (\boldsymbol{r}-\boldsymbol{r}_h, \boldsymbol{r}-\mathbf{\Pi}_{p_v}^{\mathrm{div},\star}\boldsymbol{r})_\Omega + (\boldsymbol{r}-\boldsymbol{r}_h, \mathbf{\Pi}_{p_v}^{\mathrm{div},\star}\boldsymbol{r}-\boldsymbol{r}_h)_\Omega = (\boldsymbol{r}-\boldsymbol{r}_h, \boldsymbol{r}-\mathbf{\Pi}_{p_v}^{\mathrm{div},\star}\boldsymbol{r})_\Omega,$$

which by the Cauchy-Schwarz inequality gives

$$\|\boldsymbol{r}-\boldsymbol{r}_h\|_{0,\Omega} \leq \|\boldsymbol{r}-\mathbf{\Pi}_{p_v}^{\mathrm{div},\star}\boldsymbol{r}\|_{0,\Omega}.$$

Since $\nabla\cdot\boldsymbol{r} = \nabla\cdot\boldsymbol{r}_h$ is discrete we may apply [MR20, Thm. 2.10, (vi)] as well as perform a simple scaling argument to arrive at

$$\|\boldsymbol{r}-\mathbf{\Pi}_{p_v}^{\mathrm{div},\star}\boldsymbol{r}\|_{0,\Omega} \lesssim \frac{h}{p_v}\|\boldsymbol{r}\|_{1,\Omega} \lesssim \frac{h}{p_v}\left\|\nabla\cdot(\tilde{\boldsymbol{\varphi}}_h - \boldsymbol{I}_h^0\boldsymbol{\varphi})\right\|_{0,\Omega},$$

where the last estimate is due to the *a priori* estimate of Lemma 4.3.5. Summarizing we have

$$T_1 \lesssim \frac{h}{p_v}\|\boldsymbol{e}\|_{0,\Omega}\left\|\nabla\cdot(\tilde{\boldsymbol{\varphi}}_h - \boldsymbol{I}_h^0\boldsymbol{\varphi})\right\|_{0,\Omega} \lesssim \frac{h}{p_v}\|\boldsymbol{e}\|_{0,\Omega}\|\nabla\cdot(\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega},$$

where the last estimate follows by adding and subtracting $\boldsymbol{\varphi}$, the triangle inequality as well as the second inequality of the present lemma.

Treatment of T_2: The term T_2 is treated with a duality argument. We select $\boldsymbol{\psi} \in \boldsymbol{H}(\Omega,\mathrm{div})$ such that

$$(\nabla\cdot\boldsymbol{v}, \nabla\cdot\boldsymbol{\psi})_\Omega = (\boldsymbol{v},\boldsymbol{r})_\Omega \qquad \forall \boldsymbol{v}\in\boldsymbol{H}_0(\Omega,\mathrm{div}).$$

To that end, we note that by Lemma 4.3.5 we have $\boldsymbol{r} = \nabla R$ for some $R \in H^2(\Omega)$. Therefore for $\boldsymbol{v}\in\boldsymbol{H}_0(\Omega,\mathrm{div})$ we have

$$(\nabla\cdot\boldsymbol{v}, \nabla\cdot\boldsymbol{\psi})_\Omega = (\boldsymbol{v},\boldsymbol{r})_\Omega = (\boldsymbol{v},\nabla R)_\Omega = -(\nabla\cdot\boldsymbol{v}, R)_\Omega,$$

so that the desired $\boldsymbol{\psi}$ is found as $\boldsymbol{\psi} = \nabla w$ with w solving

$$\begin{aligned} -\Delta w &= R \quad \text{in } \Omega, \\ w &= 0 \quad \text{on } \Gamma. \end{aligned}$$

Furthermore, since $R \in H^2(\Omega)$, elliptic regularity gives $w \in H^4(\Omega)$ and therefore $\boldsymbol{\psi} \in \boldsymbol{H}^3(\Omega)$. Finally the following estimates hold

$$\|\nabla\cdot\boldsymbol{\psi}\|_{2,\Omega} \leq \|\boldsymbol{\psi}\|_{3,\Omega} \leq \|w\|_{4,\Omega} \lesssim \|R\|_{2,\Omega} \lesssim \|\boldsymbol{r}\|_{1,\Omega} \lesssim \left\|\nabla\cdot(\tilde{\boldsymbol{\varphi}}_h - \boldsymbol{I}_h^0\boldsymbol{\varphi})\right\|_{0,\Omega}, \tag{4.31}$$

due to elliptic regularity and the results of Lemma 4.3.5. We therefore have for any $\boldsymbol{\psi}_h \in \mathbf{V}_{p_v}^0(\mathcal{T}_h)$

$$T_2 = (\boldsymbol{e},\boldsymbol{r})_\Omega = (\nabla\cdot\boldsymbol{e}, \nabla\cdot\boldsymbol{\psi})_\Omega = (\nabla\cdot\boldsymbol{e}, \nabla\cdot(\boldsymbol{\psi}-\boldsymbol{\psi}_h))_\Omega \leq \|\nabla\cdot\boldsymbol{e}\|_{0,\Omega}\|\nabla\cdot(\boldsymbol{\psi}-\boldsymbol{\psi}_h)\|_{0,\Omega},$$

where we used the definition of T_2, the duality argument elaborated above, the orthogonality relation of $\boldsymbol{I}_h^0$ to insert any $\boldsymbol{\psi}_h \in \mathbf{V}_{p_v}^0(\mathcal{T}_h)$, and the Cauchy-Schwarz inequality. Finally exploiting the *a priori* estimate of $\boldsymbol{\psi}$ in (4.31) we find for $p_v > 1$ that

$$\begin{aligned} T_2 &\leq \|\nabla\cdot\boldsymbol{e}\|_{0,\Omega} \cdot \inf_{\boldsymbol{\psi}_h\in\mathbf{V}_{p_v}^0(\mathcal{T}_h)} \|\nabla\cdot(\boldsymbol{\psi}-\boldsymbol{\psi}_h)\|_{0,\Omega} \lesssim \|\nabla\cdot\boldsymbol{e}\|_{0,\Omega}\,(h/p_v)^2\,\|\nabla\cdot\boldsymbol{\psi}\|_{2,\Omega} \\ &\lesssim \|\nabla\cdot\boldsymbol{e}\|_{0,\Omega}\,(h/p_v)^2\left\|\nabla\cdot(\tilde{\boldsymbol{\varphi}}_h-\boldsymbol{I}_h^0\boldsymbol{\varphi})\right\|_{0,\Omega}. \end{aligned}$$

In the lowest order case $p_v = 1$ we cannot fully exploit the regularity. However, we find

$$\|\nabla\cdot\boldsymbol{\psi}\|_{1,\Omega} \leq \|\boldsymbol{\psi}\|_{2,\Omega} \leq \|w\|_{3,\Omega} \lesssim \|R\|_{1,\Omega} \lesssim \left\|\nabla\cdot(\tilde{\boldsymbol{\varphi}}_h-\boldsymbol{I}_h^0\boldsymbol{\varphi})\right\|_{(H^1(\Omega))'}. \tag{4.32}$$

Proceeding as above and using estimate (4.32) we find

$$\begin{aligned} T_2 &\leq \|\nabla\cdot\boldsymbol{e}\|_{0,\Omega} \cdot \inf_{\boldsymbol{\psi}_h\in\mathbf{V}_{p_v}^0(\mathcal{T}_h)} \|\nabla\cdot(\boldsymbol{\psi}-\boldsymbol{\psi}_h)\|_{0,\Omega} \lesssim \|\nabla\cdot\boldsymbol{e}\|_{0,\Omega}\,h/p_v\,\|\nabla\cdot\boldsymbol{\psi}\|_{1,\Omega} \\ &\lesssim \|\nabla\cdot\boldsymbol{e}\|_{0,\Omega}\,h/p_v\left\|\nabla\cdot(\tilde{\boldsymbol{\varphi}}_h-\boldsymbol{I}_h^0\boldsymbol{\varphi})\right\|_{(H^1(\Omega))'} \lesssim \|\nabla\cdot\boldsymbol{e}\|_{0,\Omega}\,h/p_v\left\|\tilde{\boldsymbol{\varphi}}_h-\boldsymbol{I}_h^0\boldsymbol{\varphi}\right\|_{0,\Omega}. \end{aligned}$$

The last last estimate is due to integration by parts and the boundary condition of $\tilde{\boldsymbol{\varphi}}_h - \boldsymbol{I}_h^0\boldsymbol{\varphi}$; in fact

$$\begin{aligned} \left\|\nabla\cdot(\tilde{\boldsymbol{\varphi}}_h-\boldsymbol{I}_h^0\boldsymbol{\varphi})\right\|_{(H^1(\Omega))'} &= \sup_{v\in H^1(\Omega)} \frac{|(\nabla\cdot(\tilde{\boldsymbol{\varphi}}_h-\boldsymbol{I}_h^0\boldsymbol{\varphi}),v)_\Omega|}{\|v\|_{1,\Omega}} = \sup_{v\in H^1(\Omega)} \frac{|(\tilde{\boldsymbol{\varphi}}_h-\boldsymbol{I}_h^0\boldsymbol{\varphi},\nabla v)_\Omega|}{\|v\|_{1,\Omega}} \\ &\leq \left\|\tilde{\boldsymbol{\varphi}}_h-\boldsymbol{I}_h^0\boldsymbol{\varphi}\right\|_{0,\Omega} \end{aligned}$$

holds. Putting everything together we have for $p_v > 1$

$$\begin{aligned} \|\boldsymbol{e}\|_{0,\Omega}^2 &= (\boldsymbol{e},\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h)_\Omega + (\boldsymbol{e},\tilde{\boldsymbol{\varphi}}_h-\boldsymbol{I}_h^0\boldsymbol{\varphi})_\Omega \\ &= (\boldsymbol{e},\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h)_\Omega + T_1 + T_2 \\ &\lesssim \|\boldsymbol{e}\|_{0,\Omega}\,\|\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h\|_{0,\Omega} + \frac{h}{p_v}\,\|\boldsymbol{e}\|_{0,\Omega}\,\|\nabla\cdot(\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega} \\ &\quad + \frac{h^2}{p_v^2}\,\|\nabla\cdot\boldsymbol{e}\|_{0,\Omega}\left\|\nabla\cdot(\tilde{\boldsymbol{\varphi}}_h-\boldsymbol{I}_h^0\boldsymbol{\varphi})\right\|_{0,\Omega} \\ &\lesssim \|\boldsymbol{e}\|_{0,\Omega}\,\|\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h\|_{0,\Omega} + \frac{h}{p_v}\,\|\boldsymbol{e}\|_{0,\Omega}\,\|\nabla\cdot(\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega} + \frac{h^2}{p_v^2}\,\|\nabla\cdot(\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega}^2, \end{aligned}$$

where the last estimate again follows from inserting $\boldsymbol{\varphi}$ and using the second estimate of the present lemma. Young's inequality then yields the result for the operator $\boldsymbol{I}_h^0$. The lowest order case is treated analogous. For the operator $\boldsymbol{I}_h$ the only difference is that one applies Lemma 4.3.4 instead of Lemma 4.3.5 and perform the duality argument on all of $\boldsymbol{H}(\Omega, \operatorname{div})$ instead of $\boldsymbol{H}_0(\Omega, \operatorname{div})$. Here it is important to note that the potential R given by Lemma 4.3.4 satisfies homogeneous boundary conditions, so that the boundary term vanishes in the partial integration. □

Remark 4.3.7. $\boldsymbol{H}(\Omega, \operatorname{div})$-conforming approximation operators similar to $\boldsymbol{I}_h$ and $\boldsymbol{I}_h^0$ are presented in [EGSV21], where the focus is on a patchwise construction rather than the (global) orthogonalities (4.23b), (4.24b). ■

Theorem 4.3.8. *Let Γ be smooth and $(\boldsymbol{\varphi}_h, u_h)$ be the least squares approximation of $(\boldsymbol{\varphi}, u)$. Furthermore, let $e^u = u - u_h$ and $\boldsymbol{e}^{\boldsymbol{\varphi}} = \boldsymbol{\varphi} - \boldsymbol{\varphi}_h$. Then, for any $\tilde{u}_h \in S_{p_s}(\mathcal{T}_h)$, $\tilde{\boldsymbol{\varphi}}_h \in \mathbf{V}^0_{p_v}(\mathcal{T}_h)$,*

$$\begin{aligned}\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega} &\lesssim \frac{h}{p}\|(\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u)\|_b + \|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h\|_{0,\Omega} + \frac{h}{p}\|\nabla \cdot (\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega} \\ &\lesssim \frac{h}{p}\|u - \tilde{u}_h\|_{1,\Omega} + \|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h\|_{0,\Omega} + \frac{h}{p}\|\nabla \cdot (\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega}.\end{aligned}$$

Proof. Let $(\boldsymbol{\psi}, v) \in \boldsymbol{H}_0(\Omega, \mathrm{div}) \times H^1(\Omega)$ denote the dual solution given by Theorem 4.2.3 applied to $\boldsymbol{\eta} = \boldsymbol{e}^{\boldsymbol{\varphi}}$. Theorem 4.2.3 gives $\boldsymbol{\psi} \in \boldsymbol{L}^2(\Omega)$, $\nabla \cdot \boldsymbol{\psi} \in H^1(\Omega)$, and $v \in H^3(\Omega)$. Due to the Galerkin orthogonality we have for any $(\tilde{\boldsymbol{\psi}}_h, \tilde{v}_h)$

$$\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|^2_{0,\Omega} = b((\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u), (\boldsymbol{\psi}, v)) = b((\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u), (\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h, v - \tilde{v}_h)).$$

We now estimate all terms in the above:

$$\begin{aligned}(\nabla e^u + \boldsymbol{e}^{\boldsymbol{\varphi}}, \nabla(v - \tilde{v}_h))_\Omega &\le \|(\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u)\|_b \|\nabla(v - \tilde{v}_h)\|_{0,\Omega}, \\ (\nabla \cdot \boldsymbol{e}^{\boldsymbol{\varphi}} + \gamma e^u, \nabla \cdot (\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h) + \gamma(v - \tilde{v}_h))_\Omega &\lesssim \|(\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u)\|_b \left[\|\nabla \cdot (\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h)\|_{0,\Omega} + \|v - \tilde{v}_h\|_{0,\Omega}\right], \\ (\nabla e^u, \boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h)_\Omega = -(e^u, \nabla \cdot (\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h))_\Omega &\le \|e^u\|_{0,\Omega} \|\nabla \cdot (\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h)\|_{0,\Omega}.\end{aligned}$$

Therefore, we conclude that

$$\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|^2_{0,\Omega} \lesssim \|(\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u)\|_b \left[\|\nabla \cdot (\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h)\|_{0,\Omega} + \|v - \tilde{v}_h\|_{1,\Omega}\right] + (\boldsymbol{e}^{\boldsymbol{\varphi}}, \boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h)_\Omega, \tag{4.33}$$

the limiting term being for now the last one. To overcome the lack of regularity of $\boldsymbol{\psi}$ we perform a Helmholtz decomposition. In fact since $\boldsymbol{\psi} \in \boldsymbol{H}_0(\Omega, \mathrm{div})$ as well as $\nabla \cdot \boldsymbol{\psi} \in H^1(\Omega)$ there exist $\boldsymbol{\rho} \in \boldsymbol{H}_0(\Omega, \mathrm{curl})$ and $z \in H^3(\Omega)$ such that $\boldsymbol{\psi} = \nabla \times \boldsymbol{\rho} + \nabla z$. The construction is as follows: Let $z \in H^1(\Omega)$ solve

$$\begin{aligned}-\Delta z &= -\nabla \cdot \boldsymbol{\psi} && \text{in } \Omega, \\ \partial_n z &= 0 && \text{on } \Gamma.\end{aligned}$$

Since $\nabla \cdot (\boldsymbol{\psi} - \nabla z) = 0$ as well as $(\boldsymbol{\psi} - \nabla z) \cdot \boldsymbol{n} = 0$ by construction, the exact sequence property of the employed spaces allows for the existence of $\boldsymbol{\rho} \in \boldsymbol{H}_0(\Omega, \mathrm{curl})$ such that $\boldsymbol{\psi} - \nabla z = \nabla \times \boldsymbol{\rho}$. Finally the following estimates hold due to the *a priori* estimate of the Lax-Milgram theorem and partial integration for the first estimate, elliptic regularity theory for the second, and the triangle inequality together with the first estimate for the third one:

$$\begin{aligned}\|z\|_{1,\Omega} &\lesssim \|\nabla \cdot \boldsymbol{\psi}\|_{(H^1(\Omega))'} \le \|\boldsymbol{\psi}\|_{0,\Omega}\,, \\ \|z\|_{3,\Omega} &\lesssim \|\nabla \cdot \boldsymbol{\psi}\|_{1,\Omega}\,, \\ \|\nabla \times \boldsymbol{\rho}\|_{0,\Omega} &\le \|\boldsymbol{\psi}\|_{0,\Omega} + \|\nabla z\|_{0,\Omega} \lesssim \|\boldsymbol{\psi}\|_{0,\Omega}\,.\end{aligned}$$

We now continue estimating (4.33) by applying the Helmholtz decomposition. For any $\tilde{\boldsymbol{\psi}}^c_h$, $\tilde{\boldsymbol{\psi}}^g_h \in \mathbf{V}^0_{p_v}(\mathcal{T}_h)$ we have with $\tilde{\boldsymbol{\psi}}_h = \tilde{\boldsymbol{\psi}}^c_h + \tilde{\boldsymbol{\psi}}^g_h$

$$(\boldsymbol{e}^{\boldsymbol{\varphi}}, \boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h)_\Omega = (\boldsymbol{e}^{\boldsymbol{\varphi}}, \nabla \times \boldsymbol{\rho} - \tilde{\boldsymbol{\psi}}^c_h)_\Omega + (\boldsymbol{e}^{\boldsymbol{\varphi}}, \nabla z - \tilde{\boldsymbol{\psi}}^g_h)_\Omega =: T^c + T^g.$$

Treatment of T^g: By the Cauchy-Schwarz inequality we have

$$T^g = (\boldsymbol{e}^{\boldsymbol{\varphi}}, \nabla z - \tilde{\boldsymbol{\psi}}_h^g)_\Omega \leq \|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega} \, \|\nabla z - \tilde{\boldsymbol{\psi}}_h^g\|_{0,\Omega}.$$

Treatment of T^c: For any $\tilde{\boldsymbol{\varphi}}_h \in \mathbf{V}^0_{p_v}(\mathcal{T}_h)$ we have

$$\begin{aligned} T^c &= (\boldsymbol{e}^{\boldsymbol{\varphi}}, \nabla \times \boldsymbol{\rho} - \tilde{\boldsymbol{\psi}}_h^c)_\Omega \\ &= (\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h, \nabla \times \boldsymbol{\rho} - \tilde{\boldsymbol{\psi}}_h^c)_\Omega + (\tilde{\boldsymbol{\varphi}}_h - \boldsymbol{\varphi}_h, \nabla \times \boldsymbol{\rho} - \tilde{\boldsymbol{\psi}}_h^c)_\Omega =: T_1^c + T_2^c. \end{aligned}$$

Treatment of T_1^c: By the Cauchy-Schwarz inequality we have

$$T_1^c = (\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h, \nabla \times \boldsymbol{\rho} - \tilde{\boldsymbol{\psi}}_h^c)_\Omega \leq \|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h\|_{0,\Omega} \|\nabla \times \boldsymbol{\rho} - \tilde{\boldsymbol{\psi}}_h^c\|_{0,\Omega}.$$

Treatment of T_2^c: In order to treat T_2^c we proceed as in the proof of Lemma 4.3.6 and apply Lemma 4.3.5 to split the discrete object $\tilde{\boldsymbol{\varphi}}_h - \boldsymbol{\varphi}_h \in \mathbf{V}^0_{p_v}(\mathcal{T}_h)$ on a discrete and a continuous level:

$$\begin{aligned} \tilde{\boldsymbol{\varphi}}_h - \boldsymbol{\varphi}_h &= \nabla \times \boldsymbol{\mu} + \boldsymbol{r}, \\ \tilde{\boldsymbol{\varphi}}_h - \boldsymbol{\varphi}_h &= \nabla \times \boldsymbol{\mu}_h + \boldsymbol{r}_h, \end{aligned}$$

for certain $\boldsymbol{\mu} \in \boldsymbol{H}_0(\Omega, \mathrm{curl})$, $\boldsymbol{r} \in \boldsymbol{H}_0(\Omega, \mathrm{div})$, $\boldsymbol{\mu}_h \in \mathbf{N}^0_{p_v}(\mathcal{T}_h)$, and $\boldsymbol{r}_h \in \mathbf{V}^0_{p_v}(\mathcal{T}_h)$. We now choose $\tilde{\boldsymbol{\psi}}_h^c = \mathbf{\Pi}_h^{\mathrm{curl},0} \nabla \times \boldsymbol{\rho}$ given by Lemma 4.3.5. Exploiting the definition of the operator $\mathbf{\Pi}_h^{\mathrm{curl},0}$ we find

$$\begin{aligned} T_2^c &= (\tilde{\boldsymbol{\varphi}}_h - \boldsymbol{\varphi}_h, \nabla \times \boldsymbol{\rho} - \tilde{\boldsymbol{\psi}}_h^c)_\Omega \\ &= \underbrace{(\nabla \times \boldsymbol{\mu}_h, \nabla \times \boldsymbol{\rho} - \mathbf{\Pi}_h^{\mathrm{curl},0} \nabla \times \boldsymbol{\rho})_\Omega}_{=0} + (\boldsymbol{r}_h, \nabla \times \boldsymbol{\rho} - \mathbf{\Pi}_h^{\mathrm{curl},0} \nabla \times \boldsymbol{\rho})_\Omega \\ &= (\boldsymbol{r}_h - \boldsymbol{r}, \nabla \times \boldsymbol{\rho} - \mathbf{\Pi}_h^{\mathrm{curl},0} \nabla \times \boldsymbol{\rho})_\Omega + (\boldsymbol{r}, \nabla \times \boldsymbol{\rho} - \mathbf{\Pi}_h^{\mathrm{curl},0} \nabla \times \boldsymbol{\rho})_\Omega \\ &=: T_1 + T_2. \end{aligned}$$

Treatment of T_1: With the same notation as in the proof of Lemma 4.3.6 and with exactly the same arguments we have

$$\|\boldsymbol{r} - \boldsymbol{r}_h\|_{0,\Omega} \lesssim \frac{h}{p_v} \|\boldsymbol{r}\|_{1,\Omega} \lesssim \frac{h}{p_v} \|\nabla \cdot (\tilde{\boldsymbol{\varphi}}_h - \boldsymbol{\varphi}_h)\|_{0,\Omega}.$$

By the Cauchy-Schwarz inequality we have

$$T_1 \lesssim \frac{h}{p_v} \|\nabla \cdot (\tilde{\boldsymbol{\varphi}}_h - \boldsymbol{\varphi}_h)\|_{0,\Omega} \, \|\nabla \times \boldsymbol{\rho} - \mathbf{\Pi}_h^{\mathrm{curl},0} \nabla \times \boldsymbol{\rho}\|_{0,\Omega} \lesssim \frac{h}{p_v} \|\nabla \cdot (\tilde{\boldsymbol{\varphi}}_h - \boldsymbol{\varphi}_h)\|_{0,\Omega} \, \|\nabla \times \boldsymbol{\rho}\|_{0,\Omega},$$

where the last estimate follows from the fact that

$$\|\nabla \times \boldsymbol{\rho} - \mathbf{\Pi}_h^{\mathrm{curl},0} \nabla \times \boldsymbol{\rho}\|_{0,\Omega} \leq \|\nabla \times \boldsymbol{\rho} - \nabla \times \tilde{\boldsymbol{\rho}}_h\|_{0,\Omega}$$

for any $\tilde{\boldsymbol{\rho}}_h \in \mathbf{N}^0_{p_v}(\mathcal{T}_h)$ since it is a projection. Finally inserting $\boldsymbol{\varphi}$ and applying the triangle inequality as well as estimating $\|\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{\varphi}_h)\|_{0,\Omega}$ by $\|(e^u, \boldsymbol{e}^{\boldsymbol{\varphi}})\|_b$ we find

$$T_1 \lesssim \frac{h}{p_v} \|\nabla \cdot (\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega} \|\nabla \times \boldsymbol{\rho}\|_{0,\Omega} + \frac{h}{p_v} \|(\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u)\|_b \|\nabla \times \boldsymbol{\rho}\|_{0,\Omega}.$$

Treatment of T_2: Note again that $\boldsymbol{\rho} \in \boldsymbol{H}_0(\Omega, \mathrm{curl})$ and the fact that $\boldsymbol{\Pi}_h^{\mathrm{curl},0}$ maps into $\nabla \times \mathbf{N}^0_{p_v}(\mathcal{T}_h)$. Therefore, we can write $\nabla \times \boldsymbol{\rho} - \boldsymbol{\Pi}_h^{\mathrm{curl},0} \nabla \times \boldsymbol{\rho} = \nabla \times \widehat{\boldsymbol{\rho}}$ for some $\widehat{\boldsymbol{\rho}} \in \boldsymbol{H}_0(\Omega, \mathrm{curl})$ and the boundary terms consequently vanish in the following integration by parts

$$T_2 = (\boldsymbol{r}, \nabla \times \widehat{\boldsymbol{\rho}})_\Omega = (\nabla \times \boldsymbol{r}, \widehat{\boldsymbol{\rho}})_\Omega.$$

Finally, $T_2 = 0$, since $\nabla \times \boldsymbol{r} = 0$ by Lemma 4.3.5.
Collecting all the terms: Collecting the terms together with the estimate $\|\nabla \times \boldsymbol{\rho}\|_{0,\Omega} \lesssim \|\boldsymbol{\psi}\|_{0,\Omega} \lesssim \|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}$ from the Helmholtz decomposition and the regularity estimates given in Lemma 4.2.3 we find

$$\begin{aligned}(\boldsymbol{e}^{\boldsymbol{\varphi}}, \boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h)_\Omega \lesssim \Bigg[&\|\nabla z - \tilde{\boldsymbol{\psi}}_h^g\|_{0,\Omega} + \|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h\|_{0,\Omega} \\ &+ \frac{h}{p_v} \|\nabla \cdot (\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega} + \frac{h}{p_v} \|(\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u)\|_b \Bigg] \|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega} .\end{aligned} \tag{4.34}$$

Since $\tilde{\boldsymbol{\psi}}_h^c = \boldsymbol{\Pi}_h^{\mathrm{curl},0} \nabla \times \boldsymbol{\rho} \in \nabla \times \mathbf{N}^0_{p_v}(\mathcal{T}_h)$ we have

$$\|\nabla \cdot (\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h))\|_{0,\Omega} = \|\nabla \cdot (\nabla z - \tilde{\boldsymbol{\psi}}_h^g))\|_{0,\Omega}.$$

Due to the regularity of $z \in H^3(\Omega)$ we can find $\tilde{\boldsymbol{\psi}}_h^g \in \mathbf{V}^0_{p_v}(\mathcal{T}_h)$ such that

$$\|\nabla z - \tilde{\boldsymbol{\psi}}_h^g\|_{\boldsymbol{H}(\Omega,\mathrm{div})} \lesssim \frac{h}{p_v} \|\nabla z\|_{\boldsymbol{H}^1(\Omega,\mathrm{div})} \lesssim \frac{h}{p_v} \|\nabla \cdot \boldsymbol{\psi}\|_{1,\Omega} \lesssim \frac{h}{p_v} \|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega} \lesssim \frac{h}{p_v} \|(\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u)\|_b.$$

Therefore, estimate (4.34) can be summarized as follows:

$$(\boldsymbol{e}^{\boldsymbol{\varphi}}, \boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h)_\Omega \lesssim \left[\frac{h}{p_v} \|(e^u, \boldsymbol{e}^{\boldsymbol{\varphi}})\|_b + \|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h\|_{0,\Omega} + \frac{h}{p_v} \|\nabla \cdot (\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega} \right] \|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega} . \tag{4.35}$$

Again due to the regularity of $v \in H^3(\Omega)$ we can find $\tilde{v}_h \in S_{p_s}(\mathcal{T}_h)$ such that

$$\|v - \tilde{v}_h\|_{1,\Omega} \lesssim \frac{h}{p_s} \|v\|_{2,\Omega} \lesssim \frac{h}{p_s} \|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega} .$$

Finally, summarizing the estimates (4.33) and (4.35) and again using

$$\|\nabla \cdot (\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h))\|_{0,\Omega} = \|\nabla \cdot (\nabla z - \tilde{\boldsymbol{\psi}}_h^g))\|_{0,\Omega} \lesssim \frac{h}{p_v} \|(\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u)\|_b$$

we find

$$\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|^2_{0,\Omega} \lesssim \left[\frac{h}{p} \|(\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u)\|_b + \|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h\|_{0,\Omega} + \frac{h}{p} \|\nabla \cdot (\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega} \right] \|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega} .$$

Canceling one power of $\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}$ then yields the first estimate. The second one follows again by the fact that the least squares approximation is the projection with respect to b and the norm equivalence given in Theorem 4.1.1. □

Lemma 4.3.9. *Let Γ be smooth and $(\boldsymbol{\varphi}_h, u_h)$ be the least squares approximation of $(\boldsymbol{\varphi}, u)$. Furthermore, let $e^u = u - u_h$ and $\boldsymbol{e}^{\boldsymbol{\varphi}} = \boldsymbol{\varphi} - \boldsymbol{\varphi}_h$. Let $(\boldsymbol{\psi}, v) \in \boldsymbol{H}_0(\Omega, \operatorname{div}) \times H^1(\Omega)$ be the solution of the dual problem given by Theorem 4.2.2 with $w = e^u$. Additionally, let $(\boldsymbol{\psi}_h, v_h)$ be the least squares approximation of $(\boldsymbol{\psi}, v)$ and denote $e^v = v - v_h$ and $\boldsymbol{e}^{\boldsymbol{\psi}} = \boldsymbol{\psi} - \boldsymbol{\psi}_h$. Then,*

$$\|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b \lesssim \|\nabla e^u\|_{0,\Omega} \quad \text{and} \quad \|e^v\|_{0,\Omega} \lesssim \frac{h}{p} \|\nabla e^u\|_{0,\Omega} \quad \text{and} \quad \|\boldsymbol{e}^{\boldsymbol{\psi}}\|_{0,\Omega} \lesssim \frac{h}{p} \|\nabla e^u\|_{0,\Omega}.$$

Proof. Theorem 4.2.2 provides $\|\boldsymbol{\psi}\|_{2,\Omega} + \|\nabla \cdot \boldsymbol{\psi}\|_{1,\Omega} + \|v\|_{1,\Omega} \lesssim \|\nabla e^u\|_{0,\Omega}$. Stability of the least squares method (cf. (4.21)) yields

$$\|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b \lesssim \|\nabla e^u\|_{0,\Omega}.$$

By Lemma 4.3.1 we have

$$\|e^v\|_{0,\Omega} \lesssim h/p \|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b,$$

which together with the above gives the second estimate. By Theorem 4.3.8 we have

$$\|\boldsymbol{e}^{\boldsymbol{\psi}}\|_{0,\Omega} \lesssim \frac{h}{p} \|v - \tilde{v}_h\|_{1,\Omega} + \|\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h\|_{0,\Omega} + \frac{h}{p} \|\nabla \cdot (\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h)\|_{0,\Omega}$$

for any $\tilde{v}_h \in S_{p_s}(\mathcal{T}_h)$, $\tilde{\boldsymbol{\psi}}_h \in \mathbf{V}^0_{p_v}(\mathcal{T}_h)$. The result follows immediately by again exploiting the regularity of the dual solution and the approximation properties of the employed spaces. □

Theorem 4.3.10. *Let Γ be smooth and $(\boldsymbol{\varphi}_h, u_h)$ be the least squares approximation of $(\boldsymbol{\varphi}, u)$. Furthermore, let $e^u = u - u_h$. Then, for any $\tilde{\boldsymbol{\varphi}}_h \in \mathbf{V}^0_{p_v}(\mathcal{T}_h)$, $\tilde{u}_h \in S_{p_s}(\mathcal{T}_h)$,*

$$\|\nabla e^u\|_{0,\Omega} \lesssim \|u - \tilde{u}_h\|_{1,\Omega} + \frac{h}{p} \|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h\|_{0,\Omega} + \frac{h}{p} \|\nabla \cdot (\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega}.$$

Proof. As in Remark 4.3.2 with $(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)$ denoting the error of the FOSLS approximation of the dual solution given by Theorem 4.2.2 (duality argument for the gradient of the scalar variable) applied to $w = e^u$ we have for any $\tilde{\boldsymbol{\varphi}}_h \in \mathbf{V}^0_{p_v}(\mathcal{T}_h)$, $\tilde{u}_h \in S_{p_s}(\mathcal{T}_h)$

$$\begin{aligned}
\|e^u\|^2_{0,\Omega} &= b((\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h, u - \tilde{u}_h), (\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)) \\
&= (\nabla \cdot (\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h) + \gamma(u - \tilde{u}_h), \nabla \cdot \boldsymbol{e}^{\boldsymbol{\psi}} + \gamma e^v)_\Omega + (\nabla(u - \tilde{u}_h) + \boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h, \nabla e^v + \boldsymbol{e}^{\boldsymbol{\psi}})_\Omega.
\end{aligned}$$

We specifically choose $\tilde{\boldsymbol{\varphi}}_h = \boldsymbol{I}^0_h \boldsymbol{\varphi}$. In the following we heavily use the properties of the operator $\boldsymbol{I}^0_h$ given in Lemma 4.3.6. First we exploit the regularity of the dual solution using

Lemma 4.3.9 as well as the estimates of Theorem 4.2.2:

$$\begin{aligned}
(\gamma(u-\tilde u_h), \nabla\cdot \boldsymbol{e}^{\boldsymbol{\psi}} + \gamma e^v)_\Omega &\lesssim \|u-\tilde u_h\|_{0,\Omega}\, \|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b \\
&\lesssim \|u-\tilde u_h\|_{1,\Omega}\, \|\nabla e^u\|_{0,\Omega}\,, \\
(\nabla(u-\tilde u_h), \nabla e^v + \boldsymbol{e}^{\boldsymbol{\psi}})_\Omega &\lesssim \|\nabla(u-\tilde u_h)\|_{0,\Omega}\, \|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b \\
&\lesssim \|u-\tilde u_h\|_{1,\Omega}\, \|\nabla e^u\|_{0,\Omega}\,, \\
(\boldsymbol{\varphi} - \boldsymbol{I}_h^0\boldsymbol{\varphi}, \nabla e^v)_\Omega &= -(\nabla\cdot(\boldsymbol{\varphi} - \boldsymbol{I}_h^0\boldsymbol{\varphi}), e^v)_\Omega \\
&\leq \left\|\nabla\cdot(\boldsymbol{\varphi} - \boldsymbol{I}_h^0\boldsymbol{\varphi})\right\|_{0,\Omega} \|e^v\|_{0,\Omega} \\
&\lesssim h/p\left\|\nabla\cdot(\boldsymbol{\varphi} - \boldsymbol{I}_h^0\boldsymbol{\varphi})\right\|_{0,\Omega} \|\nabla e^u\|_{0,\Omega}\,, \\
(\nabla\cdot(\boldsymbol{\varphi} - \boldsymbol{I}_h^0\boldsymbol{\varphi}), \gamma e^v)_\Omega &\leq \left\|\nabla\cdot(\boldsymbol{\varphi} - \boldsymbol{I}_h^0\boldsymbol{\varphi})\right\|_{0,\Omega} \|e^v\|_{0,\Omega} \\
&\lesssim h/p\left\|\nabla\cdot(\boldsymbol{\varphi} - \boldsymbol{I}_h^0\boldsymbol{\varphi})\right\|_{0,\Omega} \|\nabla e^u\|_{0,\Omega}\,, \\
(\boldsymbol{\varphi} - \boldsymbol{I}_h^0\boldsymbol{\varphi}, \boldsymbol{e}^{\boldsymbol{\psi}})_\Omega &\lesssim \left\|\boldsymbol{\varphi} - \boldsymbol{I}_h^0\boldsymbol{\varphi}\right\|_{0,\Omega} \|\boldsymbol{e}^{\boldsymbol{\psi}}\|_{0,\Omega} \\
&\lesssim h/p\left\|\boldsymbol{\varphi} - \boldsymbol{I}_h^0\boldsymbol{\varphi}\right\|_{0,\Omega} \|\nabla e^u\|_{0,\Omega}\,, \\
(\nabla\cdot(\boldsymbol{\varphi} - \boldsymbol{I}_h^0\boldsymbol{\varphi}), \nabla\cdot\boldsymbol{e}^{\boldsymbol{\psi}})_\Omega &= (\nabla\cdot(\boldsymbol{\varphi} - \boldsymbol{I}_h^0\boldsymbol{\varphi}), \nabla\cdot(\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h))_\Omega \\
&\leq \left\|\nabla\cdot(\boldsymbol{\varphi} - \boldsymbol{I}_h^0\boldsymbol{\varphi})\right\|_{0,\Omega} \|\nabla\cdot(\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h)\|_{0,\Omega} \\
&\lesssim h/p\left\|\nabla\cdot(\boldsymbol{\varphi} - \boldsymbol{I}_h^0\boldsymbol{\varphi})\right\|_{0,\Omega} \|\nabla e^u\|_{0,\Omega}\,.
\end{aligned}$$

Canceling one power of $\|\nabla e^u\|_{0,\Omega}$, collecting the terms, and using the estimate for $\boldsymbol{I}_h^0$ we arrive at the asserted estimate. □

As a tool in the proof of our main theorem (Theorem 4.3.12) we need to analyze the error of the FOSLS approximation of the dual solution. This is summarized in

Lemma 4.3.11. *Let Γ be smooth and $(\boldsymbol{\varphi}_h, u_h)$ be the least squares approximation of $(\boldsymbol{\varphi}, u)$. Furthermore, let $e^u = u - u_h$ and $\boldsymbol{e}^{\boldsymbol{\varphi}} = \boldsymbol{\varphi} - \boldsymbol{\varphi}_h$. Let $(\boldsymbol{\psi}, v) \in \boldsymbol{H}_0(\Omega, \mathrm{div}) \times H^1(\Omega)$ be the solution of the dual problem given by Theorem 4.2.1 with $w = e^u$. Additionally, let $(\boldsymbol{\psi}_h, v_h)$ be the least squares approximation of $(\boldsymbol{\psi}, v)$ and denote $e^v = v - v_h$ and $\boldsymbol{e}^{\boldsymbol{\psi}} = \boldsymbol{\psi} - \boldsymbol{\psi}_h$. Then,*

$$\|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b \lesssim \frac{h}{p}\|e^u\|_{0,\Omega} \quad \text{and} \quad \|e^v\|_{0,\Omega} \lesssim \left(\frac{h}{p}\right)^2 \|e^u\|_{0,\Omega}\,.$$

Furthermore,

$$\|\boldsymbol{e}^{\boldsymbol{\psi}}\|_{0,\Omega} \lesssim \begin{cases} h\,\|e^u\|_{0,\Omega} & \text{if } \mathbf{V}^0_{p_v}(\mathcal{T}_h) = \mathbf{RT}^0_0(\mathcal{T}_h), \\ \left(\frac{h}{p}\right)^2 \|e^u\|_{0,\Omega} & \text{else.} \end{cases}$$

Proof. Theorem 4.2.1 gives $\boldsymbol{\psi} \in \boldsymbol{H}^3(\Omega)$, $\nabla \cdot \boldsymbol{\psi} \in H^2(\Omega)$ and $v \in H^2(\Omega)$ with norms bounded by $\|e^v\|_{0,\Omega}$. Therefore we have in view of optimality of the FOSLS method in the b-norm

$$\|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b \overset{(4.21)}{\leq} \|(\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h, v - \tilde{v}_h)\|_b \lesssim h/p\, \|e^v\|_{0,\Omega}\,,$$

where the first estimate holds for any $\tilde{v}_h \in S_p$, $\tilde{\boldsymbol{\psi}}_h \in \mathbf{V}^0_{p_v}(\mathcal{T}_h)$ and the second one follows with the same arguments as in the proof of Lemma 4.3.1. By Lemma 4.3.1 we have

$$\|e^v\|_{0,\Omega} \lesssim h/p \|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b,$$

which together with the above gives the second estimate. By Theorem 4.3.8 we have

$$\|\boldsymbol{e}^{\boldsymbol{\psi}}\|_{0,\Omega} \lesssim \frac{h}{p}\, \|v - \tilde{v}_h\|_{1,\Omega} + \|\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h\|_{0,\Omega} + \frac{h}{p} \|\nabla \cdot (\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h)\|_{0,\Omega}$$

for any $\tilde{v}_h \in S_{p_s}(\mathcal{T}_h)$, $\tilde{\boldsymbol{\psi}}_h \in \mathbf{V}^0_{p_v}(\mathcal{T}_h)$. The result follows immediately by again exploiting the regularity of the dual solution and the approximation properties of the employed spaces. □

Theorem 4.3.12. *Let Γ be smooth and $(\boldsymbol{\varphi}_h, u_h)$ be the least squares approximation of $(\boldsymbol{\varphi}, u)$. Furthermore, let $e^u = u - u_h$. Then, for any $\tilde{\boldsymbol{\varphi}}_h \in \mathbf{V}^0_{p_v}(\mathcal{T}_h)$, $\tilde{u}_h \in S_{p_s}(\mathcal{T}_h)$,*

$$\|e^u\|_{0,\Omega} \lesssim \begin{cases} h\, \|u - \tilde{u}_h\|_{1,\Omega} + h\|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h\|_{0,\Omega} + h\|\nabla \cdot (\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega} & \text{for } \mathbf{RT}^0_0(\mathcal{T}_h), \\ h\, \|u - \tilde{u}_h\|_{1,\Omega} + h^2\|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h\|_{0,\Omega} + h\|\nabla \cdot (\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega} & \text{for } \mathbf{BDM}^0_1(\mathcal{T}_h), \\ \frac{h}{p}\, \|u - \tilde{u}_h\|_{1,\Omega} + \left(\frac{h}{p}\right)^2 \|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h\|_{0,\Omega} + \left(\frac{h}{p}\right)^2 \|\nabla \cdot (\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega} & \text{else.} \end{cases}$$

Proof. As in Remark 4.3.2 with $(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)$ denoting the FOSLS approximation of the dual solution given by Theorem 4.2.1 applied to $w = e^u$ we have for any $\tilde{\boldsymbol{\varphi}}_h \in \mathbf{V}^0_{p_v}(\mathcal{T}_h)$, $\tilde{u}_h \in S_{p_s}(\mathcal{T}_h)$

$$\begin{aligned} \|e^u\|^2_{0,\Omega} &= b((\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h, u - \tilde{u}_h), (\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)) \\ &= (\nabla \cdot (\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h) + \gamma(u - \tilde{u}_h), \nabla \cdot \boldsymbol{e}^{\boldsymbol{\psi}} + \gamma e^v)_\Omega + (\nabla(u - \tilde{u}_h) + \boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h, \nabla e^v + \boldsymbol{e}^{\boldsymbol{\psi}})_\Omega. \end{aligned}$$

We specifically choose $\tilde{\boldsymbol{\varphi}}_h = \boldsymbol{I}^0_h \boldsymbol{\varphi}$. In the following we heavily use the properties of the operator $\boldsymbol{I}^0_h$ given in Lemma 4.3.6. First we exploit the regularity of the dual solution using

Lemma 4.3.11 as well as the estimates of Theorem 4.2.1:

$$
\begin{aligned}
(\gamma(u-\tilde{u}_h), \nabla\cdot \boldsymbol{e}^{\boldsymbol{\psi}} + \gamma e^v)_\Omega &\lesssim \|u-\tilde{u}_h\|_{0,\Omega}\,\|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b \\
&\lesssim h/p\,\|u-\tilde{u}_h\|_{1,\Omega}\,\|e^u\|_{0,\Omega}\,, \\
(\nabla(u-\tilde{u}_h), \nabla e^v + \boldsymbol{e}^{\boldsymbol{\psi}})_\Omega &\lesssim \|\nabla(u-\tilde{u}_h)\|_{0,\Omega}\,\|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b \\
&\lesssim h/p\,\|u-\tilde{u}_h\|_{1,\Omega}\,\|e^u\|_{0,\Omega}\,, \\
(\boldsymbol{\varphi}-\boldsymbol{I}_h^0\boldsymbol{\varphi}, \nabla e^v)_\Omega &= -(\nabla\cdot(\boldsymbol{\varphi}-\boldsymbol{I}_h^0\boldsymbol{\varphi}), e^v)_\Omega \\
&\le \left\|\nabla\cdot(\boldsymbol{\varphi}-\boldsymbol{I}_h^0\boldsymbol{\varphi})\right\|_{0,\Omega}\|e^v\|_{0,\Omega} \\
&\lesssim (h/p)^2\left\|\nabla\cdot(\boldsymbol{\varphi}-\boldsymbol{I}_h^0\boldsymbol{\varphi})\right\|_{0,\Omega}\|e^u\|_{0,\Omega}\,, \\
(\nabla\cdot(\boldsymbol{\varphi}-\boldsymbol{I}_h^0\boldsymbol{\varphi}), \gamma e^v)_\Omega &\le \left\|\nabla\cdot(\boldsymbol{\varphi}-\boldsymbol{I}_h^0\boldsymbol{\varphi})\right\|_{0,\Omega}\|e^v\|_{0,\Omega} \\
&\lesssim (h/p)^2\left\|\nabla\cdot(\boldsymbol{\varphi}-\boldsymbol{I}_h^0\boldsymbol{\varphi})\right\|_{0,\Omega}\|e^u\|_{0,\Omega}\,, \\
(\boldsymbol{\varphi}-\boldsymbol{I}_h^0\boldsymbol{\varphi}, \boldsymbol{e}^{\boldsymbol{\psi}})_\Omega &\lesssim \left\|\boldsymbol{\varphi}-\boldsymbol{I}_h^0\boldsymbol{\varphi}\right\|_{0,\Omega}\|\boldsymbol{e}^{\boldsymbol{\psi}}\|_{0,\Omega} \\
&\lesssim \begin{cases} h\left\|\boldsymbol{\varphi}-\boldsymbol{I}_h^0\boldsymbol{\varphi}\right\|_{0,\Omega}\|e^u\|_{0,\Omega} & \text{if } \mathbf{V}_{p_v}^0(\mathcal{T}_h) = \mathbf{RT}_0^0(\mathcal{T}_h), \\ \left(\frac{h}{p}\right)^2\left\|\boldsymbol{\varphi}-\boldsymbol{I}_h^0\boldsymbol{\varphi}\right\|_{0,\Omega}\|e^u\|_{0,\Omega} & \textit{else}, \end{cases} \\
(\nabla\cdot(\boldsymbol{\varphi}-\boldsymbol{I}_h^0\boldsymbol{\varphi}), \nabla\cdot\boldsymbol{e}^{\boldsymbol{\psi}})_\Omega &= (\nabla\cdot(\boldsymbol{\varphi}-\boldsymbol{I}_h^0\boldsymbol{\varphi}), \nabla\cdot(\boldsymbol{\psi}-\tilde{\boldsymbol{\psi}}_h))_\Omega \\
&\le \left\|\nabla\cdot(\boldsymbol{\varphi}-\boldsymbol{I}_h^0\boldsymbol{\varphi})\right\|_{0,\Omega}\|\nabla\cdot(\boldsymbol{\psi}-\tilde{\boldsymbol{\psi}}_h)\|_{0,\Omega} \\
&\lesssim \begin{cases} h\left\|\nabla\cdot(\boldsymbol{\varphi}-\boldsymbol{I}_h^0\boldsymbol{\varphi})\right\|_{0,\Omega}\|e^u\|_{0,\Omega} & \text{if } p_v = 1, \\ \left(\frac{h}{p}\right)^2\left\|\nabla\cdot(\boldsymbol{\varphi}-\boldsymbol{I}_h^0\boldsymbol{\varphi})\right\|_{0,\Omega}\|e^u\|_{0,\Omega} & \textit{else}. \end{cases}
\end{aligned}
$$

Canceling one power of $\|e^u\|_{0,\Omega}$, collecting the terms, and using the estimate for $\boldsymbol{I}_h^0$ we arrive at the asserted estimate. □

Remark 4.3.13. Before stating the general corollary with prescribed right-hand side $f \in H^s(\Omega)$ we highlight the improved convergence result. Consider $f \in L^2(\Omega)$. For the classical conforming finite element method one observes convergence $O(h^2)$ due to the Aubin-Nitsche trick. More precisely, let u_h^{FEM} be the solution to the model problem obtained by classical FEM, then there holds

$$\|u - u_h^{\mathrm{FEM}}\|_{0,\Omega} \lesssim h^2\,\|u\|_{2,\Omega} \lesssim h^2\,\|f\|_{0,\Omega}\,.$$

As elaborated in Section 1.2 this rate could not be obtained for the FOSLS method by previous results, since further regularity of the vector variable $\boldsymbol{\varphi}$ would be necessary. Results like [BG05, Lemma 3.4] and [Jes77, Thm. 4.1] are essentially a duality argument like Theorem 4.2.1 and the strategy of Lemma 4.3.1. Without further analysis the estimate of Lemma 4.3.1, does not give any further powers of h, since the b-norm is equivalent to the $\boldsymbol{H}(\Omega,\mathrm{div}) \times H^1(\Omega)$ norm. Theorem 4.3.12 ensures, at least if the space $\mathbf{V}_{p_v}^0(\mathcal{T}_h)$ is not of lowest order, i.e. $p_v > 1$, that the FOSLS method converges also as $O(h^2)$. More precisely, the estimate in Theorem 4.3.12 together with the approximation properties of the employed finite element spaces and $p_v > 1$ and $p_s \ge 1$ gives

$$\|e^u\|_{0,\Omega} \lesssim h^2\,\|u\|_{2,\Omega} + h^2\|\boldsymbol{\varphi}\|_{1,\Omega} + h^2\|\nabla\cdot\boldsymbol{\varphi}\|_{0,\Omega} \lesssim h^2\,\|f\|_{0,\Omega}\,.$$

So in fact the optimal rate in the sense of the beginning of Section 4.3 is achieved. If the lowest order case $p_v = 1$ also achieves optimal order is yet to be answered. Numerical experiments in Section 4.4, however, indicate it to be true. ■

We summarize the results for general right-hand side $f \in H^s(\Omega)$. This summary is essentially the estimates given by the Theorems 4.3.8, 4.3.10, and 4.3.12 together with the approximation properties of the employed finite element spaces.

Corollary 4.3.14. *Let Γ be smooth and $f \in H^s(\Omega)$ for some $s \geq 0$. Then the solution to (4.2) satisfies $u \in H^{s+2}(\Omega)$, $\boldsymbol{\varphi} \in \boldsymbol{H}^{s+1}(\Omega)$ and $\nabla \cdot \boldsymbol{\varphi} \in H^s(\Omega)$. Let $(\boldsymbol{\varphi}_h, u_h)$ be the least squares approximation of $(\boldsymbol{\varphi}, u)$. Furthermore, let $e^u = u - u_h$ and $\boldsymbol{e}^{\boldsymbol{\varphi}} = \boldsymbol{\varphi} - \boldsymbol{\varphi}_h$. Then, for the lowest order case $p_v = 1$,*

$$\|e^u\|_{0,\Omega} \lesssim h^{\min\{s+1,2\}} \|f\|_{s,\Omega} .$$

For $p_v > 1$ there holds

$$\|e^u\|_{0,\Omega} \lesssim \left(\frac{h}{p}\right)^{\min\{s+1,p_s,p_v+1\}+1} \|f\|_{s,\Omega} .$$

Furthermore, the estimate

$$\|\nabla e^u\|_{0,\Omega} \lesssim \left(\frac{h}{p}\right)^{\min\{s+1,p_s,p_v+1\}} \|f\|_{s,\Omega} .$$

holds. Finally, we have

$\mathbf{V}^0_{p_v}(\mathcal{T}_h) = \mathbf{RT}^0_{p_v-1}(\mathcal{T}_h)$	$\mathbf{V}^0_{p_v}(\mathcal{T}_h) = \mathbf{BDM}^0_{p_v}(\mathcal{T}_h)$
$\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega} \lesssim \left(\frac{h}{p}\right)^{\min\{s+1,p_s+1,p_v\}} \|f\|_{s,\Omega}$	$\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega} \lesssim \left(\frac{h}{p}\right)^{\min\{s+1,p_s+1,p_v+1\}} \|f\|_{s,\Omega}.$

Proof. The regularity result follows immediately by standard arguments together with the fact that $\boldsymbol{\varphi} = -\nabla u$. We now analyze the quantities in the estimates of the Theorems 4.3.8, 4.3.10 and 4.3.12:

$$\|u - \tilde{u}_h\|_{1,\Omega} \lesssim (h/p)^{\min\{s+1,p_s\}} \|u\|_{H^{s+2}(\Omega)} \lesssim (h/p)^{\min\{s+1,p_s\}} \|f\|_{s,\Omega} ,$$

$$\|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h\|_{0,\Omega} \lesssim \begin{cases} (h/p)^{\min\{s+1,p_v\}} \|\boldsymbol{\varphi}\|_{s+1,\Omega} \lesssim (h/p)^{\min\{s+1,p_v\}} \|f\|_{s,\Omega} & \text{for } \mathbf{RT}^0_{p_v-1}(\mathcal{T}_h), \\ (h/p)^{\min\{s+1,p_v+1\}} \|\boldsymbol{\varphi}\|_{s+1,\Omega} \lesssim (h/p)^{\min\{s+1,p_v+1\}} \|f\|_{s,\Omega} & \text{for } \mathbf{BDM}^0_{p_v}(\mathcal{T}_h), \end{cases}$$

$$\|\nabla \cdot (\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega} \lesssim (h/p)^{\min\{s,p_v\}} \|\nabla \cdot \boldsymbol{\varphi}\|_{s,\Omega} \lesssim (h/p)^{\min\{s,p_v\}} \|f\|_{s,\Omega} .$$

The estimates of the Theorems 4.3.8, 4.3.10, and 4.3.12 together with the above estimates give, after straightforward calculations, the asserted rates. □

We close this section with some remarks concerning sharpness of the estimates of Corollary 4.3.14:

Remark 4.3.15. Let the assumptions of Corollary 4.3.14 be satisfied. From a best approximation point of view, since $u \in H^{s+2}(\Omega)$, we have

$$\inf_{\tilde{u}_h \in S_{p_s}(\mathcal{T}_h)} \|u - \tilde{u}_h\|_{0,\Omega} = O(h^{\min\{s+1,p_s\}+1}),$$
$$\inf_{\tilde{u}_h \in S_{p_s}(\mathcal{T}_h)} \|\nabla(u - \tilde{u}_h)\|_{0,\Omega} = O(h^{\min\{s+1,p_s\}}),$$
$$\inf_{\tilde{\boldsymbol{\varphi}}_h \in \mathbf{V}^0_{p_v}(\mathcal{T}_h)} \|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h\|_{0,\Omega} = \begin{cases} O(h^{\min\{s+1,p_v\}}) & \text{if } \mathbf{V}^0_{p_v}(\mathcal{T}_h) = \mathbf{RT}^0_{p_v-1}(\mathcal{T}_h), \\ O(h^{\min\{s+1,p_v+1\}}) & \text{if } \mathbf{V}^0_{p_v}(\mathcal{T}_h) = \mathbf{BDM}^0_{p_v}(\mathcal{T}_h). \end{cases}$$

Excluding the lowest order case $p_v = 1$ we have, choosing $p_v \geq p_s - 1$, sharpness of the estimates for e^u and ∇e^u. This can be easily seen, since the rates guaranteed by Corollary 4.3.14 for $\|e^u\|_{0,\Omega}$ and $\|\nabla e^u\|_{0,\Omega}$ are the same as the ones from a best approximation argument. The estimates are therefore sharp. The lowest order case $p_v = 1$ seems to be suboptimal, as the numerical examples in Section 4.4 suggest. In all other cases, i.e., $p_v > 1$ and $p_v < p_s - 1$, our numerical examples suggest sharpness of the estimates, in both the setting of a smooth solution as well as one with finite Sobolev regularity, but not achieving the best approximation rate. Since in the least squares functional the term $\|\nabla u_h + \boldsymbol{\varphi}_h\|_{0,\Omega}$ enforces ∇u_h and $\boldsymbol{\varphi}_h$ to be *close*, it is to be expected that an insufficient choice of p_v limits the convergence rate. A theoretical justification concerning the observed rates in the cases $p_v = 1$ as well as $p_v > 1$ and $p_v < p_s - 1$ is yet to be studied. In conclusion, when the application in question is concerned with approximation of u or ∇u in the $L^2(\Omega)$ norm, the best possible rate with the smallest number of degrees of freedom is achieved with the choice $p_v = p_s - 1$ regardless of the choice of $\mathbf{V}^0_{p_v}(\mathcal{T}_h)$. Therefore, it is computationally favorable to choose Raviart-Thomas elements over Brezzi-Douglas-Marini elements. Turning now to $\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}$ similar arguments guarantee sharpness of the estimates. In this case when $p_s + 1 \geq p_v$ and $p_s + 1 \geq p_v + 1$, for the choice of Raviart-Thomas elements and Brezzi-Douglas-Marini elements, respectively. Again the other cases are open for theoretical justification. However, both theoretical as well as the numerical examples in Section 4.4 suggest the choice of Brezzi-Douglas-Marini elements over Raviart-Thomas elements, when application is concerned with approximation of $\boldsymbol{\varphi}$ in the $L^2(\Omega)$ norm. ■

4.4. Numerical examples

All our calculations are performed with the *hp*-FEM code NETGEN / NGSOLVE by J. Schöberl, [Sch, Sch97]. The curved boundaries are implemented using second order rational splines.

In the following we will perform two different numerical experiments.

1. For the first one we choose $f \in C^\infty(\overline{\Omega})$. Since the data is sufficiently smooth the suboptimal estimate $\|e^u\|_{0,\Omega} \lesssim h/p\|(\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u)\|_b$ of Lemma 4.3.1 suffices to deduce optimal rates. Therefore, we only present three graphs in this section in order to highlight two aspects of the least squares approach: On the one hand the optimal choice of the employed polynomial degrees p_s and p_v. On the other hand the superiority of Brezzi-Douglas-Marini elements over Raviart-Thomas elements when approximating

the vector valued variable. For completeness we present other convergence plots in Appendix A.

2. To showcase our new convergence result we then choose $f \in \cap_{\varepsilon>0} H^{1/2-\varepsilon}(\Omega)$, but $f \notin H^{1/2}(\Omega)$ with $u \in \cap_{\varepsilon>0} H^{5/2-\varepsilon}(\Omega)$ and $\boldsymbol{\varphi} \in \cap_{\varepsilon>0} \boldsymbol{H}^{3/2-\varepsilon}(\Omega)$. We again only present a selection of graphs focusing on the new convergence results, other convergence plots can be found in Appendix A.

In all graphs, the actual numerical results are given by red dots. The rate that is guaranteed by Corollary 4.3.14 is plotted in black together with the number written out near the bottom right. Furthermore, in blue the reference line for the best rate possible with the employed space $S_{p_s}(\mathcal{T}_h)$ or $\mathbf{V}^0_{p_v}(\mathcal{T}_h)$ is plotted, depending on the quantity of interest, i.e., for $\|e^u\|_{0,\Omega}$ the blue reference line corresponds to $h^{\min\{s+1,p_s\}+1}$, for $\|\nabla e^u\|_{0,\Omega}$ the blue reference line corresponds to $h^{\min\{s+1,p_s\}}$ and for $\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}$ the blue reference line corresponds to $h^{\min\{s+1,p_v\}}$ for $\mathbf{V}^0_{p_v}(\mathcal{T}_h) = \mathbf{RT}^0_{p_v-1}(\mathcal{T}_h)$ and $h^{\min\{s+1,p_v+1\}}$ for $\mathbf{V}^0_{p_v}(\mathcal{T}_h) = \mathbf{BDM}^0_{p_v}(\mathcal{T}_h)$.

Example 4.4.1. We consider as the domain Ω the unit sphere in $\mathbb{R}^2$. The exact solution is the smooth function $u(x,y) = \cos(2\pi(x^2+y^2))$. The numerical results are plotted in Figures 4.1 and A.1 for $\|e^u\|_{0,\Omega}$, in Figures A.2 and A.3 for $\|\nabla e^u\|_{0,\Omega}$, and in Figures 4.2 and 4.3 for $\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}$. There are some remarks to be made:

- Consider Figure 4.1 depicting $\|e^u\|_{0,\Omega}$ using Raviart-Thomas elements. The rates guaranteed by Corollary 4.3.14 are achieved in the numerical experiment. The important subfigures are the ones in the subdiagonal of the discussed figure, i.e., corresponding to the choice $p_v = p_s - 1$. Here, apart from the lowest order case, the best possible rate with the smallest number of degrees of freedom is achieved. Above this subdiagonal, i.e., $p_v \geq p_s$, additional degrees of freedom will not increase the rate of convergence, since by pure best approximation arguments the rate of convergence is limited by the polynomial degree p_s of the scalar variable. Below this subdiagonal, i.e., $p_v < p_s - 1$, we notice that the rate of convergence is also limited by the polynomial degree p_v of the vector variable. Note that the results for $\|e^u\|_{0,\Omega}$ in Corollary 4.3.14 are independent of the choice of the vector valued finite element space, which is also confirmed by our experiments. Additional convergence plots can be found in Appendix A.

- Consider Figures 4.2 and 4.3 depicting $\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}$. Apart from similar observations as for the scalar variable, it is notable that a difference in the approximation properties of the different spaces for the vector variable is observed, as predicted by Corollary 4.3.14. Consider wanting to achieve a rate of h^5. The combination of spaces with the smallest number of degrees of freedom corresponds to $\mathbf{BDM}^0_4(\mathcal{T}_h) \times S_4(\mathcal{T}_h)$ and $\mathbf{RT}^0_4(\mathcal{T}_h) \times S_4(\mathcal{T}_h)$, respectively, highlighting the superiority of the Brezzi-Douglas-Marini elements when approximating $\boldsymbol{\varphi}$. For further discussion see again Remark 4.3.15.

Example 4.4.2. For our second example we consider again the case of Ω being the unit sphere in $\mathbb{R}^2$. The exact solution $u(x,y)$ is calculated corresponding to the right-hand side

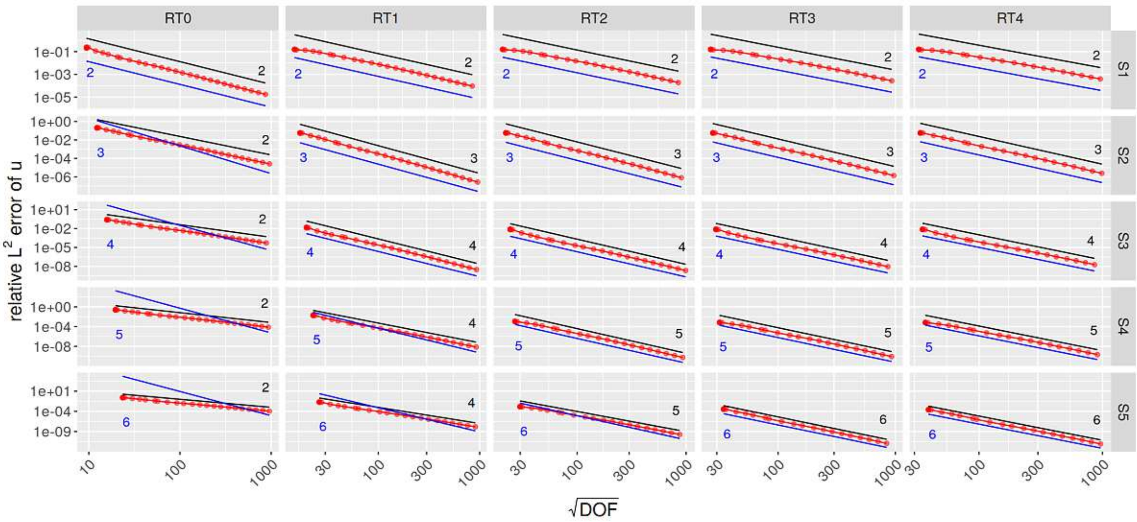

Figure 4.1.: (cf. Example 4.4.1) Convergence of $\|e^u\|_{0,\Omega}$ vs. $\sqrt{\mathrm{DOF}} \sim 1/h$ employing $\mathbf{V}^0_{p_v}(\mathcal{T}_h) = \mathbf{RT}^0_{p_v-1}(\mathcal{T}_h)$.

$f(x,y) = \mathbb{1}_{[0,1/2]}(\sqrt{x^2+y^2})$. Therefore $u \in \cap_{\varepsilon>0} H^{5/2-\varepsilon}(\Omega)$. The numerical results for the choice of Raviart-Thomas elements are plotted in Figure 4.4 for $\|e^u\|_{0,\Omega}$, in Figure 4.5 for $\|\nabla e^u\|_{0,\Omega}$ and in Figure 4.6 for $\|\boldsymbol{e}^{\varphi}\|_{0,\Omega}$. Apart from the remarks already made in Example 4.4.1 we note that we observe the improved convergence result when dealing with limited Sobolev regularity of the data. Furthermore, in the lowest order case $p_v = 1$ the guaranteed rate seems to be suboptimal. The plots for the choice of Brezzi-Douglas-Marini elements are presented in Appendix A.

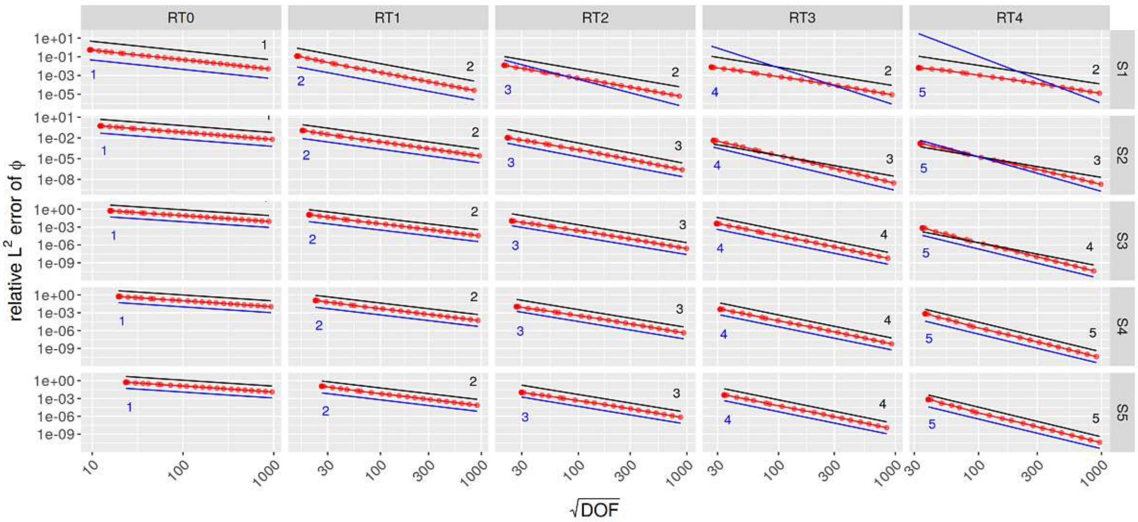

Figure 4.2.: (cf. Example 4.4.1) Convergence of $\|\boldsymbol{e}^{\varphi}\|_{0,\Omega}$ vs. $\sqrt{\mathrm{DOF}} \sim 1/h$ employing $\mathbf{V}^0_{p_v}(\mathcal{T}_h) = \mathbf{RT}^0_{p_v-1}(\mathcal{T}_h)$.

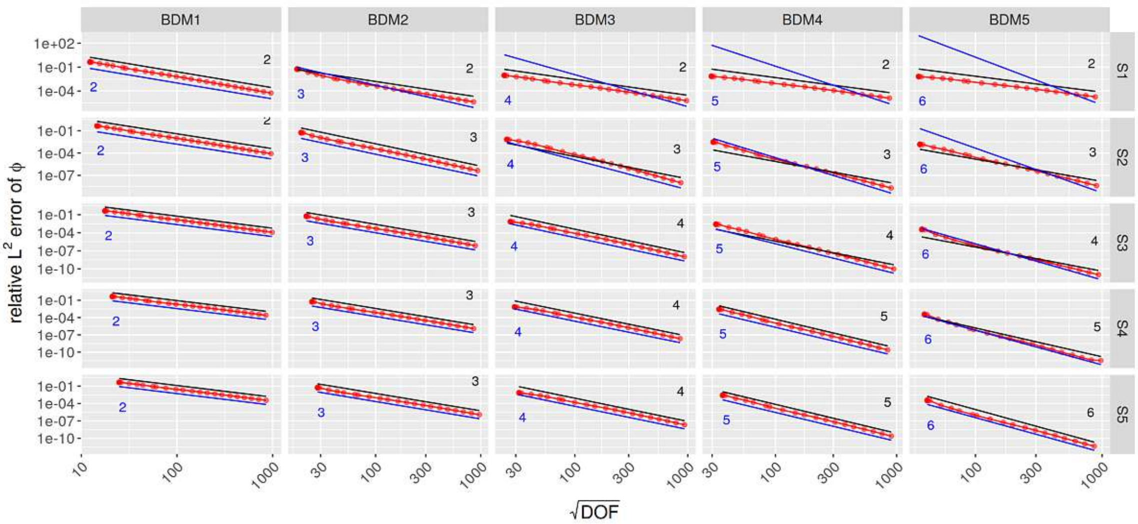

Figure 4.3.: (cf. Example 4.4.1) Convergence of $\|\boldsymbol{e}^{\varphi}\|_{0,\Omega}$ vs. $\sqrt{\mathrm{DOF}} \sim 1/h$ employing $\mathbf{V}^0_{p_v}(\mathcal{T}_h) = \mathbf{BDM}^0_{p_v}(\mathcal{T}_h)$.

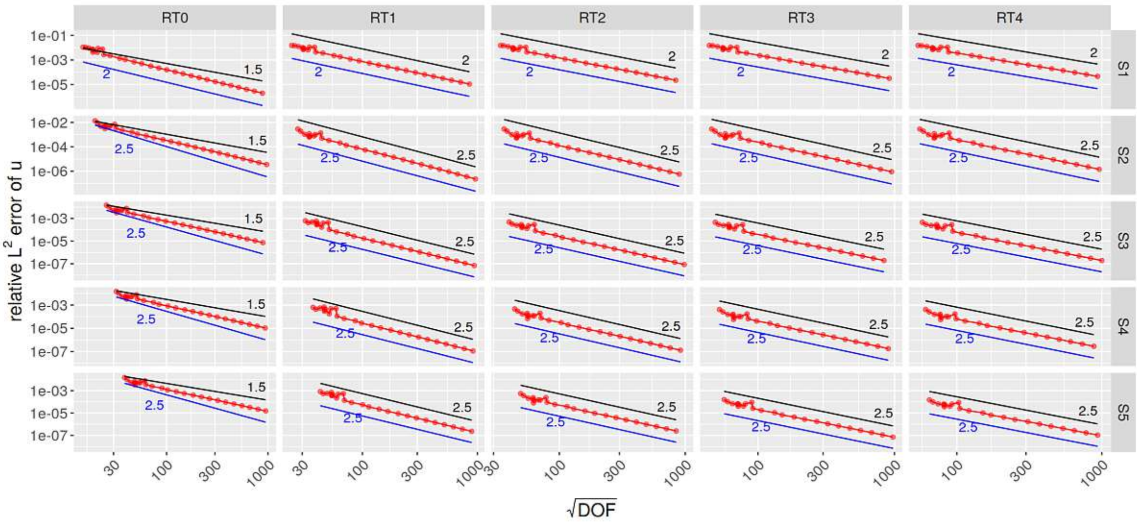

Figure 4.4.: (cf. Example 4.4.2) Convergence of $\|e^u\|_{0,\Omega}$ vs. $\sqrt{\mathrm{DOF}} \sim 1/h$ employing $\mathbf{V}^0_{p_v}(\mathcal{T}_h) = \mathbf{RT}^0_{p_v-1}(\mathcal{T}_h)$.

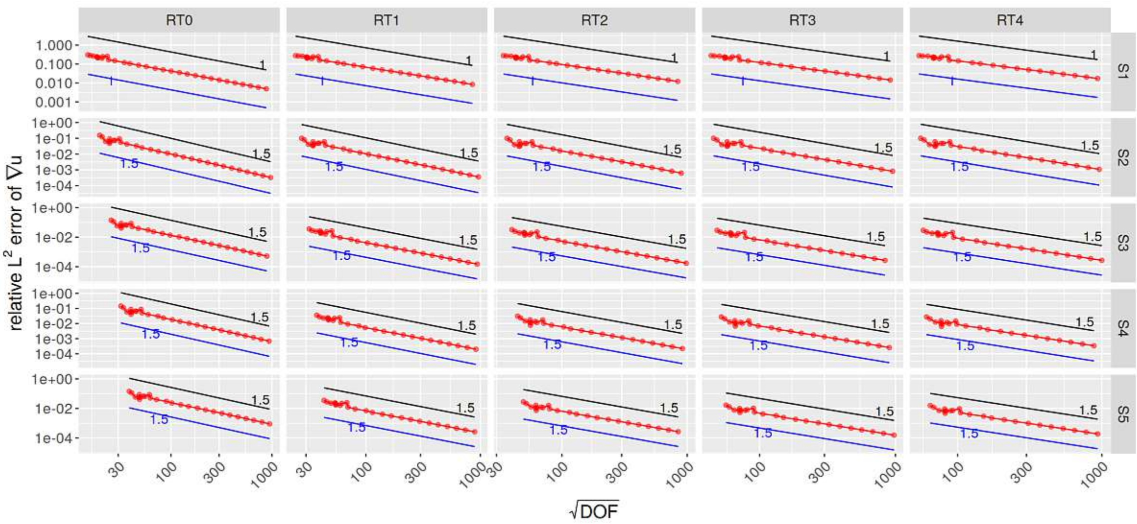

Figure 4.5.: (cf. Example 4.4.2) Convergence of $\|\nabla e^u\|_{0,\Omega}$ vs. $\sqrt{\mathrm{DOF}} \sim 1/h$ employing $\mathbf{V}^0_{p_v}(\mathcal{T}_h) = \mathbf{RT}^0_{p_v-1}(\mathcal{T}_h)$.

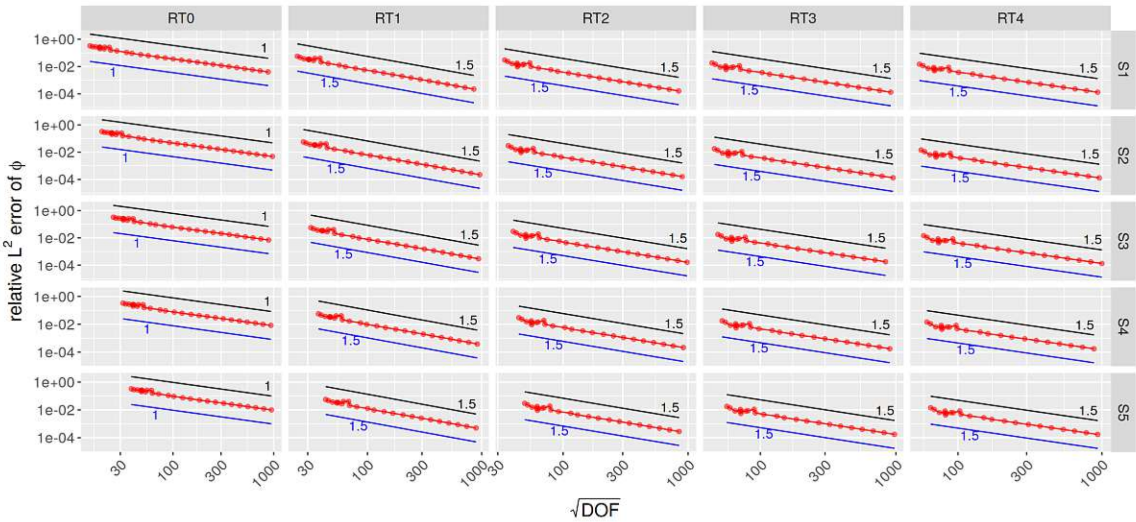

Figure 4.6.: (cf. Example 4.4.2) Convergence of $\|\boldsymbol{e}^{\varphi}\|_{0,\Omega}$ vs. $\sqrt{\text{DOF}} \sim 1/h$ employing $\mathbf{V}^0_{p_v}(\mathcal{T}_h) = \mathbf{RT}^0_{p_v-1}(\mathcal{T}_h)$.

5. FOSLS II - inhomogeneous boundary conditions

In the present chapter we extend our analysis performed in Chapter 4 to the setting of inhomogeneous Robin boundary conditions. These boundary conditions contribute to additional boundary terms in the bilinear form. This fact further limits the regularity of the dual solutions. Furthermore, the operator $\boldsymbol{I}_h$ needs to be adjusted in order to account for the additional boundary term. The results presented in the current chapter are part of the work [BM21].
The outline of this chapter is as follows. In Section 5.1 we first introduce the model problem, the FOSLS method itself and prove a norm equivalence result, which in turn guarantees unique solvability of the continuous as well as the discrete least squares formulation. Section 5.2 proves duality results for the scalar variable, the gradient of the scalar variable as well as the vector variable and corresponding traces. It is important to note that the additional boundary conditions result in limited regularity of the dual solutions, see e.g., Theorem 5.2.1 vs. Theorem 4.2.1, and also lead to the necessity of additional duality arguments, see Theorem 5.2.4. In Section 5.3 we present several error estimates for different quantities of interest, which in a bootstrapping fashion then yield optimal convergence rates for the scalar variable. Closing with Section 5.4 we present numerical examples showcasing the proved convergence rates, focusing especially on the case of finite Sobolev regularity.

5.1. Extensions to Robin boundary value problems

Throughout the present chapter again the notation of Chapter 2 applies. For $\gamma, \alpha > 0$ fixed as well as $f \in L^2(\Omega)$ and $g \in L^2(\Gamma)$ we consider the following model problem

$$\begin{aligned} -\Delta u + \gamma u &= f \quad \text{in } \Omega, \\ \partial_n u + \alpha u &= g \quad \text{on } \Gamma. \end{aligned} \tag{5.1}$$

As in Chapter 4 with the variable $\boldsymbol{\varphi} = -\nabla u$ we arrive at the system

$$\begin{aligned} \nabla \cdot \boldsymbol{\varphi} + \gamma u &= f \qquad \text{in } \Omega, \\ \nabla u + \boldsymbol{\varphi} &= 0 \qquad \text{in } \Omega, \\ \boldsymbol{\varphi} \cdot \boldsymbol{n} - \alpha u &= -g \quad \text{on } \Gamma. \end{aligned} \tag{5.2}$$

Furthermore, we introduce the Hilbert spaces

$$\boldsymbol{V} := \{\boldsymbol{\varphi} \in \boldsymbol{H}(\Omega, \operatorname{div}) \colon \boldsymbol{\varphi} \cdot \boldsymbol{n} \in L^2(\Gamma)\} \text{ and } W := H^1(\Omega),$$

where $\boldsymbol{V}$ is equipped with the graph norm $(\|\boldsymbol{\varphi}\|^2_{\boldsymbol{H}(\Omega,\operatorname{div})} + \|\boldsymbol{\varphi} \cdot \boldsymbol{n}\|^2_{0,\Gamma})^{1/2}$, in order to control the $L^2(\Gamma)$ normal trace. This is necessary since for general $\boldsymbol{\varphi} \in \boldsymbol{H}(\Omega, \operatorname{div})$ one only has

$\boldsymbol{\varphi}\cdot\boldsymbol{n}\in H^{-1/2}(\Gamma)$. In order to verify that $\boldsymbol{V}$ is in fact a Hilbert space, consider any Cauchy sequence $\boldsymbol{\varphi}_n$ in $\boldsymbol{V}$. Therefore $\boldsymbol{\varphi}_n$ is also a Cauchy sequence in $\boldsymbol{H}(\Omega,\mathrm{div})$ as well as $\boldsymbol{\varphi}_n\cdot\boldsymbol{n}$ in $L^2(\Gamma)$. Consequently $\boldsymbol{\varphi}_n\to\boldsymbol{\varphi}$ in $\boldsymbol{H}(\Omega,\mathrm{div})$ and $\boldsymbol{\varphi}_n\cdot\boldsymbol{n}\to g$ in $L^2(\Gamma)$, for some $\boldsymbol{\varphi}\in\boldsymbol{H}(\Omega,\mathrm{div})$ and some $g\in L^2(\Gamma)$. We therefore need to verify $\boldsymbol{\varphi}\cdot\boldsymbol{n}=g$. To that end, we calculate via Green's theorem for any $v\in H^1(\Omega)$

$$\langle\boldsymbol{\varphi}_n\cdot\boldsymbol{n},v\rangle_\Gamma=\langle\boldsymbol{\varphi}_n\cdot\boldsymbol{n},v\rangle_{H^{-1/2}(\Gamma)\times H^{1/2}(\Gamma)}=(\nabla\cdot\boldsymbol{\varphi}_n,v)_\Omega+(\boldsymbol{\varphi}_n,\nabla v)_\Omega.$$

Taking the limit in the above we arrive at

$$\langle g,v\rangle_\Gamma=(\nabla\cdot\boldsymbol{\varphi},v)_\Omega+(\boldsymbol{\varphi},\nabla v)_\Omega$$

for any $v\in H^1(\Omega)$, which proves $\boldsymbol{\varphi}\cdot\boldsymbol{n}=g$. The bilinear form b and linear functional F are given as in the homogeneous boundary case, just with additional boundary terms, by

$$\begin{aligned}b((\boldsymbol{\varphi},u),(\boldsymbol{\psi},v))&:=(\nabla\cdot\boldsymbol{\varphi}+\gamma u,\nabla\cdot\boldsymbol{\psi}+\gamma v)_\Omega+(\nabla u+\boldsymbol{\varphi},\nabla v+\boldsymbol{\psi})_\Omega\\&\quad+\langle\boldsymbol{\varphi}\cdot\boldsymbol{n}-\alpha u,\boldsymbol{\psi}\cdot\boldsymbol{n}-\alpha v\rangle_\Gamma,\\F((\boldsymbol{\varphi},v))&:=(f,\nabla\cdot\boldsymbol{\psi}+\gamma v)_\Omega+\langle -g,\boldsymbol{\psi}\cdot\boldsymbol{n}-\alpha v\rangle_\Gamma.\end{aligned}$$

We start our analysis with a norm equivalence theorem.

Theorem 5.1.1 (Norm equivalence - Robin version of Theorem 4.1.1)**.** *For all* $(\boldsymbol{\varphi},u)\in\boldsymbol{V}\times W$ *there holds the norm equivalence*

$$\|u\|_{1,\Omega}^2+\|\boldsymbol{\varphi}\|_{\boldsymbol{H}(\Omega,\mathrm{div})}^2+\|\boldsymbol{\varphi}\cdot\boldsymbol{n}\|_{0,\Gamma}^2\lesssim b((\boldsymbol{\varphi},u),(\boldsymbol{\varphi},u))\lesssim\|u\|_{1,\Omega}^2+\|\boldsymbol{\varphi}\|_{\boldsymbol{H}(\Omega,\mathrm{div})}^2+\|\boldsymbol{\varphi}\cdot\boldsymbol{n}\|_{0,\Gamma}^2.$$

Proof. Apart from constructing the correct splitting the proof is completely analogous to the proof of Theorem 4.1.1. We will therefore only write down the splitting and omit the rest. By definition we have

$$b((\boldsymbol{\varphi},u),(\boldsymbol{\varphi},u))=\|\underbrace{\nabla\cdot\boldsymbol{\varphi}+\gamma u}_{=:w}\|_{0,\Omega}^2+\|\underbrace{\nabla u+\boldsymbol{\varphi}}_{=:\boldsymbol{\eta}}\|_{0,\Omega}^2+\|\underbrace{\boldsymbol{\varphi}\cdot\boldsymbol{n}-\alpha u}_{=:\mu}\|_{0,\Gamma}^2,$$

from which the second inequality follows immediately by the triangle inequality and a trace estimate. To prove the first estimate, the correct system of equations to look at is given by

$$\begin{aligned}\nabla\cdot\boldsymbol{\varphi}_1+\gamma u_1&=w &&\text{in }\Omega, &\qquad \nabla\cdot\boldsymbol{\varphi}_2+\gamma u_2&=0 &&\text{in }\Omega,\\ \nabla u_1+\boldsymbol{\varphi}_1&=0 &&\text{in }\Omega, & \nabla u_2+\boldsymbol{\varphi}_2&=\boldsymbol{\eta} &&\text{in }\Omega,\\ \boldsymbol{\varphi}_1\cdot\boldsymbol{n}-\alpha u_1&=0 &&\text{on }\Gamma, & \boldsymbol{\varphi}_2\cdot\boldsymbol{n}-\alpha u_2&=\mu &&\text{on }\Gamma.\end{aligned}$$

In terms of second order equations we have

$$\begin{aligned}-\Delta u_1+\gamma u_1&=w &&\text{in }\Omega, &\qquad -\Delta u_2+\gamma u_2&=-\nabla\cdot\boldsymbol{\eta} &&\text{in }\Omega,\\ \partial_n u_1+\alpha u_1&=0 &&\text{on }\Gamma, & \partial_n u_2+\alpha u_2&=-\mu &&\text{on }\Gamma.\end{aligned}$$

From this point onward the proof is completely analogous to the proof of Theorem 4.1.1. □

5.2. Duality argument

In the following we perform duality arguments for several quantities of interest. The procedure is very similar to Section 4.2, however we note that the additional boundary term in the bilinear form b further limits the regularity of the dual solutions.

Theorem 5.2.1 (Duality argument for the scalar variable - Robin version of Thm. 4.2.1)**.** *Let Γ be smooth. For any $(\boldsymbol{\varphi}, w) \in \boldsymbol{V} \times W$ there exists $(\boldsymbol{\psi}, v) \in \boldsymbol{V} \times W$ such that $\|w\|_{0,\Omega}^2 = b((\boldsymbol{\varphi}, w), (\boldsymbol{\psi}, v))$. Furthermore, $\boldsymbol{\psi} \in \boldsymbol{H}^2(\Omega)$, $\nabla \cdot \boldsymbol{\psi} \in H^2(\Omega)$ and $v \in H^2(\Omega)$. Additionally the following estimates hold:*

$$\begin{aligned} \|v\|_{2,\Omega} &\lesssim \|w\|_{0,\Omega}\,, \\ \|\boldsymbol{\psi}\|_{2,\Omega} &\lesssim \|w\|_{0,\Omega}\,, \\ \|\nabla \cdot \boldsymbol{\psi}\|_{2,\Omega} &\lesssim \|w\|_{0,\Omega}\,. \end{aligned}$$

Proof. Theorem 5.1.1 gives the existence of a unique $(\boldsymbol{\psi}, v) \in \boldsymbol{V} \times W$ satisfying

$$(u, w)_\Omega = b((\boldsymbol{\varphi}, u), (\boldsymbol{\psi}, v)) \quad \forall\, (\boldsymbol{\varphi}, u) \in \boldsymbol{V} \times W. \tag{5.3}$$

We introduce the additional unknowns z, $\boldsymbol{\mu}$ and σ by

$$\begin{aligned} \nabla \cdot \boldsymbol{\psi} + \gamma v &= z && \text{in } \Omega, \\ \nabla v + \boldsymbol{\psi} &= \boldsymbol{\mu} && \text{in } \Omega, \\ \boldsymbol{\psi} \cdot \boldsymbol{n} - \alpha v &= \sigma && \text{on } \Gamma. \end{aligned}$$

Hence, (5.3) is equivalent to

$$(u, w)_\Omega = (\nabla u + \boldsymbol{\varphi}, \boldsymbol{\mu})_\Omega + (\nabla \cdot \boldsymbol{\varphi} + \gamma u, z)_\Omega + \langle \boldsymbol{\varphi} \cdot \boldsymbol{n} - \alpha u, \sigma \rangle_\Gamma \quad \forall\, (\boldsymbol{\varphi}, u) \in \boldsymbol{V} \times W. \tag{5.4}$$

Choosing $u = 0$ in (5.4) and integrating by parts we find

$$0 = (\boldsymbol{\varphi}, \boldsymbol{\mu})_\Omega + (\nabla \cdot \boldsymbol{\varphi}, z)_\Omega + \langle \boldsymbol{\varphi} \cdot \boldsymbol{n}, \sigma \rangle_\Gamma = (\boldsymbol{\varphi}, \boldsymbol{\mu} - \nabla z)_\Omega + \langle \boldsymbol{\varphi} \cdot \boldsymbol{n}, \sigma + z \rangle_\Gamma,$$

which gives $\boldsymbol{\mu} = \nabla z$ as well as $\sigma = -z$. Therefore we find with $\boldsymbol{\varphi} = 0$ in (5.4)

$$(u, w)_\Omega = (\nabla u, \nabla z)_\Omega + (\gamma u, z)_\Omega + \langle \alpha u, z \rangle_\Gamma \quad \forall\, u \in H^1(\Omega).$$

Hence, z satisfies

$$\begin{aligned} -\Delta z + \gamma z &= w && \text{in } \Omega, \\ \partial_n z + \alpha z &= 0 && \text{on } \Gamma, \end{aligned} \tag{5.5}$$

and the shift theorem immediately gives $z \in H^2(\Omega)$ with the estimate $\|z\|_{2,\Omega} \lesssim \|w\|_{0,\Omega}$. We now proceed as in the proof of Theorem 4.2.1. To highlight the fact that $\boldsymbol{\psi}$ is only in $\boldsymbol{H}^2(\Omega)$ compared to Theorem 4.2.1 we write down the equations for v and $z - v$:

$$\begin{aligned} -\Delta v + \gamma v &= w + (1-\gamma) z && \text{in } \Omega, & \qquad -\Delta(z - v) + \gamma(z - v) &= (\gamma - 1) z && \text{in } \Omega, \\ \partial_n v + \alpha v &= (1-\alpha) z && \text{on } \Gamma, & \qquad \partial_n(z - v) + \alpha(z - v) &= (\alpha - 1) z && \text{on } \Gamma. \end{aligned}$$

Again standard regularity theory gives $v \in H^2(\Omega)$. However, the regularity of $z - v$ is limited by the exploitable regularity of the boundary data $(\alpha - 1)z \in H^{3/2}(\Gamma)$. Therefore we have $z - v \in H^3(\Omega)$ with the estimate

$$\|z - v\|_{3,\Omega} \lesssim \|(\gamma - 1)z\|_{1,\Omega} + \|(\alpha - 1)z\|_{3/2,\Gamma} \lesssim \|w\|_{0,\Omega}\,,$$

and consequently $\boldsymbol{\psi} = \nabla(z - v) \in \boldsymbol{H}^2(\Omega)$. The regularity of $\nabla \cdot \boldsymbol{\psi}$ as well as the remaining estimates are obvious. □

Theorem 5.2.2 (Duality argument for the gradient of the scalar variable - Robin version of Thm. 4.2.2)**.** *Let Γ be smooth. For any $(\boldsymbol{\varphi}, w) \in \boldsymbol{V} \times W$ there exists $(\boldsymbol{\psi}, v) \in \boldsymbol{V} \times W$ such that $\|\nabla w\|_{0,\Omega}^2 = b((\boldsymbol{\varphi}, w), (\boldsymbol{\psi}, v))$. Furthermore, $\boldsymbol{\psi} \in \boldsymbol{H}^1(\Omega)$, $\nabla \cdot \boldsymbol{\psi} \in H^1(\Omega)$ and $v \in H^1(\Omega)$. Additionally the following estimates hold:*

$$\|v\|_{1,\Omega} \lesssim \|\nabla w\|_{0,\Omega}\,,$$
$$\|\boldsymbol{\psi}\|_{1,\Omega} \lesssim \|\nabla w\|_{0,\Omega}\,,$$
$$\|\nabla \cdot \boldsymbol{\psi}\|_{1,\Omega} \lesssim \|\nabla w\|_{0,\Omega}\,.$$

Proof. Theorem 5.1.1 gives the existence of a unique $(\boldsymbol{\psi}, v) \in \boldsymbol{V} \times W$ satisfying

$$(\nabla u, \nabla w)_\Omega = b((\boldsymbol{\varphi}, u), (\boldsymbol{\psi}, v)) \quad \forall\, (\boldsymbol{\varphi}, u) \in \boldsymbol{V} \times W. \tag{5.6}$$

We introduce the additional unknowns z, $\boldsymbol{\mu}$ and σ by

$$\begin{aligned} \nabla \cdot \boldsymbol{\psi} + \gamma v &= z && \text{in } \Omega,\\ \nabla v + \boldsymbol{\psi} &= \boldsymbol{\mu} && \text{in } \Omega,\\ \boldsymbol{\psi} \cdot \boldsymbol{n} - \alpha v &= \sigma && \text{on } \Gamma. \end{aligned}$$

Hence, (5.6) is equivalent to

$$(\nabla u, \nabla w)_\Omega = (\nabla u + \boldsymbol{\varphi}, \boldsymbol{\mu})_\Omega + (\nabla \cdot \boldsymbol{\varphi} + \gamma u, z)_\Omega + \langle \boldsymbol{\varphi} \cdot \boldsymbol{n} - \alpha u, \sigma \rangle_\Gamma \quad \forall\, (\boldsymbol{\varphi}, u) \in \boldsymbol{V} \times W. \tag{5.7}$$

For $u = 0$ in (5.7) and integrating by parts we find

$$0 = (\boldsymbol{\varphi}, \boldsymbol{\mu})_\Omega + (\nabla \cdot \boldsymbol{\varphi}, z)_\Omega + \langle \boldsymbol{\varphi} \cdot \boldsymbol{n}, \sigma \rangle_\Gamma = (\boldsymbol{\varphi}, \boldsymbol{\mu} - \nabla z)_\Omega + \langle \boldsymbol{\varphi} \cdot \boldsymbol{n}, \sigma + z \rangle_\Gamma$$

which gives $\boldsymbol{\mu} = \nabla z$ as well as $\sigma = -z$. Therefore we find with $\boldsymbol{\varphi} = 0$ in (5.7)

$$(\nabla u, \nabla w)_\Omega = (\nabla u, \nabla z)_\Omega + (\gamma u, z)_\Omega + \langle \alpha u, z \rangle_\Gamma \quad \forall\, u \in H^1(\Omega), \tag{5.8}$$

which is uniquely solvable by the Lax-Milgram theorem. Furthermore, z satisfies the estimate $\|z\|_{1,\Omega} \lesssim \|\nabla w\|_{0,\Omega}$. Formally z satisfies

$$\begin{aligned} -\Delta z + \gamma z &= -\nabla \cdot \nabla w && \text{in } \Omega,\\ \partial_n z + \alpha z &= \nabla w \cdot \boldsymbol{n} && \text{on } \Gamma. \end{aligned} \tag{5.9}$$

The right-hand side in (5.9) is understood in accordance with (5.8) as the mapping $u \mapsto (\nabla w, \nabla u)_\Omega$, see the proof of Theorem 4.2.2. We now proceed as in the proof of Theorem 4.2.2. The equations satisfied by v and $z - v$ are easily derived:

$$\begin{aligned} -\Delta v + \gamma v &= (1-\gamma)z - \nabla \cdot \nabla w \quad \text{in } \Omega, & -\Delta(z-v) + \gamma(z-v) &= (\gamma - 1)z \quad \text{in } \Omega, \\ \partial_n v + \alpha v &= (1-\alpha)z + \nabla w \cdot \boldsymbol{n} \quad \text{on } \Gamma, & \partial_n(z-v) + \alpha(z-v) &= (\alpha-1)z \quad \text{on } \Gamma. \end{aligned}$$

Again by the Lax-Milgram theorem we have $v \in H^1(\Omega)$ with $\|v\|_{1,\Omega} \lesssim \|\nabla w\|_{0,\Omega}$. The regularity of $z - v$ is limited by the exploitable regularity of the boundary data $(\alpha - 1)z \in H^{1/2}(\Gamma)$. Therefore we have $z - v \in H^2(\Omega)$ with the estimate

$$\|z - v\|_{2,\Omega} \lesssim \|(\gamma - 1)z\|_{0,\Omega} + \|(\alpha - 1)z\|_{1/2,\Gamma} \lesssim \|\nabla w\|_{0,\Omega}\,,$$

and consequently $\boldsymbol{\psi} = \nabla(z - v) \in \boldsymbol{H}^1(\Omega)$. The regularity of $\nabla \cdot \boldsymbol{\psi}$ as well as the remaining estimates are obvious. □

Theorem 5.2.3 (Duality argument for the vector valued variable - Robin version of Theorem 4.2.3)**.** *Let Γ be smooth. For any $(\boldsymbol{\eta}, u) \in \boldsymbol{V} \times W$ there exists $(\boldsymbol{\psi}, v) \in \boldsymbol{V} \times W$ such that $\|\boldsymbol{\eta}\|_{0,\Omega}^2 = b((\boldsymbol{\eta}, u), (\boldsymbol{\psi}, v))$. Furthermore, $\boldsymbol{\psi} \in \boldsymbol{L}^2(\Omega)$, $\nabla \cdot \boldsymbol{\psi} \in H^1(\Omega)$, $\boldsymbol{\psi} \cdot \boldsymbol{n} \in H^{1/2}(\Gamma)$ and $v \in H^2(\Omega)$. Additionally the following estimates hold:*

$$\begin{aligned} \|v\|_{2,\Omega} &\lesssim \|\boldsymbol{\eta}\|_{0,\Omega}\,, \\ \|\boldsymbol{\psi}\|_{0,\Omega} &\lesssim \|\boldsymbol{\eta}\|_{0,\Omega}\,, \\ \|\nabla \cdot \boldsymbol{\psi}\|_{1,\Omega} &\lesssim \|\boldsymbol{\eta}\|_{0,\Omega}\,, \\ \|\boldsymbol{\psi} \cdot \boldsymbol{n}\|_{1/2,\Gamma} &\lesssim \|\boldsymbol{\eta}\|_{0,\Omega}\,. \end{aligned}$$

Proof. Theorem 5.1.1 gives the existence of a unique $(\boldsymbol{\psi}, v) \in \boldsymbol{V} \times W$ satisfying

$$(\boldsymbol{\varphi}, \boldsymbol{\eta})_\Omega = b((\boldsymbol{\varphi}, u), (\boldsymbol{\psi}, v)) \quad \forall\, (\boldsymbol{\varphi}, u) \in \boldsymbol{V} \times W. \tag{5.10}$$

We introduce the additional unknowns z, $\boldsymbol{\mu}$ and σ by

$$\begin{aligned} \nabla \cdot \boldsymbol{\psi} + \gamma v &= z \qquad \text{in } \Omega, \\ \nabla v + \boldsymbol{\psi} &= \boldsymbol{\mu} \qquad \text{in } \Omega, \\ \boldsymbol{\psi} \cdot \boldsymbol{n} - \alpha v &= \sigma \qquad \text{on } \Gamma. \end{aligned}$$

Hence, (5.10) is equivalent to

$$(\boldsymbol{\varphi}, \boldsymbol{\eta})_\Omega = (\nabla u + \boldsymbol{\varphi}, \boldsymbol{\mu})_\Omega + (\nabla \cdot \boldsymbol{\varphi} + \gamma u, z)_\Omega + \langle \boldsymbol{\varphi} \cdot \boldsymbol{n} - \alpha u, \sigma \rangle_\Gamma \quad \forall\, (\boldsymbol{\varphi}, u) \in \boldsymbol{V} \times W. \tag{5.11}$$

For $u = 0$ and integrating by parts we find

$$(\boldsymbol{\varphi}, \boldsymbol{\eta})_\Omega = (\boldsymbol{\varphi}, \boldsymbol{\mu})_\Omega + (\nabla \cdot \boldsymbol{\varphi}, z)_\Omega + \langle \boldsymbol{\varphi} \cdot \boldsymbol{n}, \sigma \rangle_\Gamma = (\boldsymbol{\varphi}, \boldsymbol{\mu} - \nabla z)_\Omega + \langle \boldsymbol{\varphi} \cdot \boldsymbol{n}, \sigma + z \rangle_\Gamma$$

which gives $\boldsymbol{\mu} - \nabla z = \boldsymbol{\eta}$ as well as $\sigma = -z$. Therefore we find with $\boldsymbol{\varphi} = 0$

$$0 = (\nabla u, \boldsymbol{\eta} + \nabla z)_\Omega + (\gamma u, z)_\Omega + \langle \alpha u, z \rangle_\Gamma \quad \forall\, u \in H^1(\Omega),$$

which is uniquely solvable by the Lax-Milgram theorem with the estimate $\|z\|_{1,\Omega} \lesssim \|\boldsymbol{\eta}\|_{0,\Omega}$. In fact z satisfies

$$\begin{aligned} -\Delta z + \gamma z &= \nabla \cdot \boldsymbol{\eta} \quad \text{in } \Omega, \\ \partial_n z + \alpha z &= -\boldsymbol{\eta} \cdot \boldsymbol{n} \ \text{ on } \Gamma. \end{aligned} \tag{5.12}$$

The equations satisfied by v are easily derived:

$$\begin{aligned} -\Delta v + \gamma v &= (1-\gamma) z \quad \text{in } \Omega, \\ \partial_n v + \alpha v &= (1-\alpha) z \quad \text{on } \Gamma. \end{aligned}$$

By elliptic regularity we find $v \in H^2(\Omega)$ with

$$\|v\|_{2,\Omega} \lesssim \|(1-\gamma)z\|_{0,\Omega} + \|(1-\alpha)z\|_{1/2,\Gamma} \lesssim \|z\|_{1,\Omega} \lesssim \|\boldsymbol{\eta}\|_{0,\Omega}$$

Finally $\boldsymbol{\psi} = \boldsymbol{\eta} + \nabla(z-v) \in \boldsymbol{L}^2(\Omega)$. The regularity of $\nabla \cdot \boldsymbol{\psi}$ and $\boldsymbol{\psi} \cdot \boldsymbol{n}$ as well as the remaining estimates are trivial. □

Theorem 5.2.4 (Duality argument for the normal trace of the vector valued variable)**.** *Let Γ be smooth. For any $(\boldsymbol{\eta}, u) \in \boldsymbol{V} \times W$ there exists $(\boldsymbol{\psi}, v) \in \boldsymbol{V} \times W$ such that $\|\boldsymbol{\eta} \cdot \boldsymbol{n}\|_{0,\Gamma}^2 = b((\boldsymbol{\eta}, u), (\boldsymbol{\psi}, v))$. Furthermore, $\boldsymbol{\psi} \in \boldsymbol{H}^{1/2}(\Omega)$, $\nabla \cdot \boldsymbol{\psi} \in H^{3/2}(\Omega)$, $\boldsymbol{\psi} \cdot \boldsymbol{n} \in L^2(\Gamma)$ and $v \in H^{3/2}(\Omega)$. Additionally the following estimates hold:*

$$\begin{aligned} \|v\|_{3/2,\Omega} &\lesssim \|\boldsymbol{\eta} \cdot \boldsymbol{n}\|_{0,\Gamma}\,, \\ \|\boldsymbol{\psi}\|_{1/2,\Omega} &\lesssim \|\boldsymbol{\eta} \cdot \boldsymbol{n}\|_{0,\Gamma}\,, \\ \|\nabla \cdot \boldsymbol{\psi}\|_{3/2,\Omega} &\lesssim \|\boldsymbol{\eta} \cdot \boldsymbol{n}\|_{0,\Gamma}\,, \\ \|\boldsymbol{\psi} \cdot \boldsymbol{n}\|_{0,\Gamma} &\lesssim \|\boldsymbol{\eta} \cdot \boldsymbol{n}\|_{0,\Gamma}\,. \end{aligned}$$

Proof. Theorem 5.1.1 gives the existence of a unique $(\boldsymbol{\psi}, v) \in \boldsymbol{V} \times W$ satisfying

$$\langle \boldsymbol{\varphi} \cdot \boldsymbol{n}, \boldsymbol{\eta} \cdot \boldsymbol{n} \rangle_\Gamma = b((\boldsymbol{\varphi}, u), (\boldsymbol{\psi}, v)) \quad \forall\, (\boldsymbol{\varphi}, u) \in \boldsymbol{V} \times W. \tag{5.13}$$

We introduce the additional unknowns z, $\boldsymbol{\mu}$ and σ by

$$\begin{aligned} \nabla \cdot \boldsymbol{\psi} + \gamma v &= z \qquad \text{in } \Omega, \\ \nabla v + \boldsymbol{\psi} &= \boldsymbol{\mu} \qquad \text{in } \Omega, \\ \boldsymbol{\psi} \cdot \boldsymbol{n} - \alpha v &= \sigma \qquad \text{on } \Gamma. \end{aligned}$$

Hence, (5.13) is equivalent to

$$\langle \boldsymbol{\varphi} \cdot \boldsymbol{n}, \boldsymbol{\eta} \cdot \boldsymbol{n} \rangle_\Gamma = (\nabla u + \boldsymbol{\varphi}, \boldsymbol{\mu})_\Omega + (\nabla \cdot \boldsymbol{\varphi} + \gamma u, z)_\Omega + \langle \boldsymbol{\varphi} \cdot \boldsymbol{n} - \alpha u, \sigma \rangle_\Gamma \quad \forall\, (\boldsymbol{\varphi}, u) \in \boldsymbol{V} \times W. \tag{5.14}$$

For $u = 0$ in (5.14) and integrating by parts we find

$$\langle \boldsymbol{\varphi} \cdot \boldsymbol{n}, \boldsymbol{\eta} \cdot \boldsymbol{n} \rangle_\Gamma = (\boldsymbol{\varphi}, \boldsymbol{\mu})_\Omega + (\nabla \cdot \boldsymbol{\varphi}, z)_\Omega + \langle \boldsymbol{\varphi} \cdot \boldsymbol{n}, \sigma \rangle_\Gamma = (\boldsymbol{\varphi}, \boldsymbol{\mu} - \nabla z)_\Omega + \langle \boldsymbol{\varphi} \cdot \boldsymbol{n}, \sigma + z \rangle_\Gamma$$

which gives $\boldsymbol{\mu} = \nabla z$ as well as $\sigma = \boldsymbol{\eta} \cdot \boldsymbol{n} - z$. Therefore we find with $\boldsymbol{\varphi} = 0$

$$0 = (\nabla u, \nabla z)_\Omega + (\gamma u, z)_\Omega + \langle \alpha u, z \rangle_\Gamma - \alpha \langle u, \boldsymbol{\eta} \cdot \boldsymbol{n} \rangle_\Gamma \quad \forall\, u \in H^1(\Omega).$$

Hence, $z \in H^1(\Omega)$ satisfies

$$\begin{aligned} -\Delta z + \gamma z &= 0 && \text{in } \Omega, \\ \partial_n z + \alpha z &= \alpha \boldsymbol{\eta} \cdot \boldsymbol{n} && \text{on } \Gamma. \end{aligned}$$

Standard regularity theory gives $z \in H^{3/2}(\Omega)$ with $\|z\|_{3/2,\Omega} \lesssim \|\boldsymbol{\eta} \cdot \boldsymbol{n}\|_{0,\Gamma}$. The equation for v reads

$$\begin{aligned} -\Delta v + \gamma v &= (1-\gamma) z && \text{in } \Omega, \\ \partial_n v + \alpha v &= (\alpha - 1)(\boldsymbol{\eta} \cdot \boldsymbol{n} - z) && \text{on } \Gamma, \end{aligned}$$

which immediately gives $v \in H^{3/2}(\Omega)$ with $\|v\|_{3/2,\Omega} \lesssim \|\boldsymbol{\eta} \cdot \boldsymbol{n}\|_{0,\Gamma}$. The remaining regularity results and estimates follow immediately from

$$\begin{aligned} \nabla \cdot \boldsymbol{\psi} + \gamma v &= z && \text{in } \Omega, \\ \nabla v + \boldsymbol{\psi} &= \nabla z && \text{in } \Omega, \\ \boldsymbol{\psi} \cdot \boldsymbol{n} - \alpha v &= \boldsymbol{\eta} \cdot \boldsymbol{n} - z && \text{on } \Gamma, \end{aligned}$$

which concludes the proof. □

Remark 5.2.5. Note that usually a duality argument results in a dual solution with higher order Sobolev regularity. However, this is not the case in Theorem 5.2.4, wherein the regularity is *not* improved, since $\boldsymbol{\psi} \cdot \boldsymbol{n}$ is still only in $L^2(\Gamma)$. The sole purpose of this duality argument is to again exploit Galerkin orthogonality, this time to overcome the limiting regularity of the boundary data. ■

5.3. Error analysis

In the error analysis it is crucial to understand the approximation properties of the vector valued finite element space in the classical $\boldsymbol{H}(\Omega, \mathrm{div})$ norm as well as the $L^2(\Gamma)$ norm of the normal trace simultaneously. We are therefore interested in quantifying

$$\inf_{\tilde{\boldsymbol{\psi}}_h \in \mathbf{V}_{p_v}(\mathcal{T}_h)} \|\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h\|_{\boldsymbol{H}(\Omega,\mathrm{div})} + \|(\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h) \cdot \boldsymbol{n}\|_{0,\Gamma}$$

for $\boldsymbol{\psi} \in \boldsymbol{V}$. For the readers' convenience we quickly summarize some results of [MR20] concerning the $\boldsymbol{H}(\Omega, \mathrm{div})$ conforming approximation operator constructed therein. A simple scaling argument gives the desired h estimates of the global operator.

Proposition 5.3.1 (Definition 2.3, Theorem 2.10 & Remark 2.9 in [MR20])**.** *The global operator $\boldsymbol{\Pi}_{p_v}^{\mathrm{div}}$ satisfies for every $\boldsymbol{\varphi} \in \boldsymbol{H}^{1/2}(\Omega, \mathrm{div})$ and $\tilde{\boldsymbol{\varphi}}_h \in \mathbf{V}_{p_v}(\mathcal{T}_h)$,*

(i) $(\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{\Pi}_{p_v}^{\mathrm{div}} \boldsymbol{\varphi}), \nabla \cdot \tilde{\boldsymbol{\varphi}}_h)_\Omega = 0$ *and consequently* $\|\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{\Pi}_{p_v}^{\mathrm{div}} \boldsymbol{\varphi})\|_{0,\Omega} \leq \|\nabla \cdot (\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega}$,

(ii) $\langle (\boldsymbol{\varphi} - \boldsymbol{\Pi}_{p_v}^{\mathrm{div}} \boldsymbol{\varphi}) \cdot \boldsymbol{n}, \tilde{\boldsymbol{\varphi}}_h \cdot \boldsymbol{n} \rangle_\Gamma = 0$ *and consequently* $\|(\boldsymbol{\psi} - \boldsymbol{\Pi}_{p_v}^{\mathrm{div}} \boldsymbol{\varphi}) \cdot \boldsymbol{n}\|_{0,\Gamma} \leq \|(\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h) \cdot \boldsymbol{n}\|_{0,\Gamma}$,

(iii) $\|\boldsymbol{\varphi} - \boldsymbol{\Pi}_{p_v}^{\mathrm{div}} \boldsymbol{\varphi}\|_{\boldsymbol{H}(\Omega,\mathrm{div})} \lesssim \left(\frac{h}{p_v}\right)^{1/2} \|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h\|_{\boldsymbol{H}^{1/2}(\mathrm{div},\Omega)}$.

Proof. For the orthogonality relation in Item (ii) see [MR20, Def. 2.3, Eq. (2.18a), (2.18b)]. The approximation properties for the normal trace follow immediately. For Item (iii) see [MR20, Thm. 2.10, (v)]. A scaling argument yields the result. Finally, concerning Item (i) note that the definition of the operator $\mathbf{\Pi}^{\mathrm{div}}_{p_v}$ is such that

$$(\nabla \cdot (\boldsymbol{\varphi} - \mathbf{\Pi}^{\mathrm{div}}_{p_v}\boldsymbol{\varphi}), \nabla \cdot \tilde{\boldsymbol{\varphi}}_h)_\Omega = 0$$

for any $\tilde{\boldsymbol{\varphi}}_h \in \mathbf{V}_{p_v}(\mathcal{T}_h)$ with $\tilde{\boldsymbol{\varphi}}_h \cdot \boldsymbol{n} = 0$ on Γ, see [MR20, Def. 2.3, Eq. (2.18d)]. However, due to the commuting diagram property we can calculate for any $\tilde{\boldsymbol{\varphi}}_h \in \mathbf{V}_{p_v}(\mathcal{T}_h)$

$$(\nabla \cdot (\boldsymbol{\varphi} - \mathbf{\Pi}^{\mathrm{div}}_{p_v}\boldsymbol{\varphi}), \nabla \cdot \tilde{\boldsymbol{\varphi}}_h)_\Omega = (\nabla \cdot \boldsymbol{\varphi} - \Pi^{L^2}_{p_v}\nabla \cdot \boldsymbol{\varphi}), \nabla \cdot \tilde{\boldsymbol{\varphi}}_h)_\Omega = 0,$$

where $\Pi^{L^2}_{p_v}$ denotes the L^2 orthogonal projection, which then gives the orthogonality relation in Item (i). The approximation properties for the divergence follow immediately. □

Lemma 5.3.2 (Suboptimal estimate for $\|e^u\|_{0,\Omega}$ - Robin version of Lemma 4.3.1)**.** *Let Γ be smooth and $(\boldsymbol{\varphi}_h, u_h)$ be the least squares approximation of $(\boldsymbol{\varphi}, u)$. Furthermore, let $e^u = u - u_h$ and $\boldsymbol{e}^{\boldsymbol{\varphi}} = \boldsymbol{\varphi} - \boldsymbol{\varphi}_h$. Then, for any $\tilde{u}_h \in S_{p_s}(\mathcal{T}_h)$, $\tilde{\boldsymbol{\varphi}}_h \in \mathbf{V}_{p_v}(\mathcal{T}_h)$,*

$$\begin{aligned}\|e^u\|_{0,\Omega} &\lesssim \frac{h}{p}\|(\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u)\|_b \\ &\lesssim \frac{h}{p}\|u - \tilde{u}_h\|_{1,\Omega} + \frac{h}{p}\|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h\|_{0,\Omega} + \frac{h}{p}\|(\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h)\cdot\boldsymbol{n}\|_{0,\Gamma} + \frac{h}{p}\|\nabla\cdot(\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega}.\end{aligned}$$

Proof. We apply the duality argument of Theorem 5.2.1 with $w = e^u$. As in Lemma 4.3.1 we find

$$\|e^u\|^2_{0,\Omega} \leq \|(\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u)\|_b \|(\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h, v - \tilde{v}_h)\|_b$$

for any $\tilde{\boldsymbol{\psi}}_h \in \mathbf{V}_{p_v}(\mathcal{T}_h)$ and $\tilde{v}_h \in S_{p_s}(\mathcal{T}_h)$, due to the Galerkin orthogonality and the Cauchy-Schwarz inequality. The norm equivalence in Theorem 5.1.1 gives

$$\|(\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h, v - \tilde{v}_h)\|_b \lesssim \|v - \tilde{v}_h\|_{1,\Omega} + \|\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h\|_{\boldsymbol{H}(\Omega,\mathrm{div})} + \|(\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h)\cdot\boldsymbol{n}\|_{0,\Gamma}.$$

Using Proposition 5.3.1 and exploiting the regularity estimates given by Theorem 5.2.1 yields the result. □

We are going to need an approximation operator satisfying certain orthogonality relations, i.e., a similar operator to $\boldsymbol{I}^0_h$ and $\boldsymbol{I}_h$ as constructed in Section 4.3. Even though the operator $\boldsymbol{I}_h$ is applicable to derive improved convergence results, they are only optimal in a pure h version of the FOSLS method. The p version is however suboptimal. This is due to the fact that the analysis requires the approximation properties of $\boldsymbol{I}_h$ in the $L^2(\Gamma)$ norm of the normal trace, which hinges on an inverse estimate. It is therefore natural to introduce the normal trace into the definition of the operator:

Construction of $\boldsymbol{I}_h^\Gamma$: In the following we construct an operator which *sees* the $L^2(\Gamma)$ normal trace. We define $\boldsymbol{I}_h^\Gamma$ again as a constrained minimization problem:

$$\boldsymbol{I}_h^\Gamma \boldsymbol{\varphi} = \underset{\boldsymbol{\varphi}_h \in \mathbf{V}_{p_v}(\mathcal{T}_h)}{\operatorname{argmin}} \frac{1}{2} \|\boldsymbol{\varphi} - \boldsymbol{\varphi}_h\|_{0,\Omega}^2 + \frac{1}{2} \|(\boldsymbol{\varphi} - \boldsymbol{\varphi}_h) \cdot \boldsymbol{n}\|_{0,\Gamma}^2$$

$$\text{s.t.} \quad (\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}), \nabla \cdot \boldsymbol{\chi}_h)_\Omega = 0 \quad \forall \boldsymbol{\chi}_h \in \mathbf{V}_{p_v}(\mathcal{T}_h).$$

To simplify notation we introduce the scalar product $\langle\langle \cdot, \cdot \rangle\rangle$ and the induced norm $|||\cdot|||$ on $\boldsymbol{V} = \{\boldsymbol{\varphi} \in \boldsymbol{H}(\Omega, \operatorname{div}) \colon \boldsymbol{\varphi} \cdot \boldsymbol{n} \in L^2(\Gamma)\}$:

$$\langle\langle \boldsymbol{\varphi}, \boldsymbol{\psi} \rangle\rangle \coloneqq (\boldsymbol{\varphi}, \boldsymbol{\psi})_\Omega + \langle \boldsymbol{\varphi} \cdot \boldsymbol{n}, \boldsymbol{\psi} \cdot \boldsymbol{n} \rangle_\Gamma.$$

Therefore we can write the operator $\boldsymbol{I}_h^\Gamma$ as

$$\boldsymbol{I}_h^\Gamma \boldsymbol{\varphi} = \underset{\boldsymbol{\varphi}_h \in \mathbf{V}_{p_v}(\mathcal{T}_h)}{\operatorname{argmin}} \frac{1}{2} |||\boldsymbol{\varphi} - \boldsymbol{\varphi}_h|||^2 \qquad \text{s.t.} \qquad (\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}), \nabla \cdot \boldsymbol{\chi}_h)_\Omega = 0 \quad \forall \boldsymbol{\chi}_h \in \mathbf{V}_{p_v}(\mathcal{T}_h).$$

The variational formulation is now given by: Find $(\boldsymbol{\varphi}_h, \lambda_h) \in \mathbf{V}_{p_v}(\mathcal{T}_h) \times \nabla \cdot \mathbf{V}_{p_v}(\mathcal{T}_h)$ such that

$$\begin{aligned} \langle\langle \boldsymbol{\varphi}_h, \boldsymbol{\mu}_h \rangle\rangle + (\nabla \cdot \boldsymbol{\mu}_h, \lambda_h)_\Omega &= \langle\langle \boldsymbol{\varphi}, \boldsymbol{\mu}_h \rangle\rangle && \forall \boldsymbol{\mu}_h \in \mathbf{V}_{p_v}(\mathcal{T}_h), \\ (\nabla \cdot \boldsymbol{\varphi}_h, \eta_h)_\Omega &= (\nabla \cdot \boldsymbol{\varphi}, \eta_h)_\Omega && \forall \eta_h \in \nabla \cdot \mathbf{V}_{p_v}(\mathcal{T}_h). \end{aligned}$$

Coercivity on kernel: Let $\boldsymbol{\mu} \in \{\boldsymbol{\psi} \in \boldsymbol{V} \colon (\nabla \cdot \boldsymbol{\psi}, \eta)_\Omega = 0 \quad \forall \eta \in \nabla \cdot \boldsymbol{V}\}$ be given. The coercivity is trivial since by construction $(\nabla \cdot \boldsymbol{\mu}, \nabla \cdot \boldsymbol{\mu})_\Omega = 0$ and therefore

$$\langle\langle \boldsymbol{\mu}, \boldsymbol{\mu} \rangle\rangle = |||\boldsymbol{\mu}|||^2 = |||\boldsymbol{\mu}|||^2 + \|\nabla \cdot \boldsymbol{\mu}\|_{0,\Omega}^2 = \|\boldsymbol{\mu}\|_{\boldsymbol{V}}^2 .$$

inf-sup condition: Let $\eta \in \nabla \cdot \boldsymbol{V}$ be given. First let $u \in H_0^1(\Omega)$ solve

$$\begin{aligned} -\Delta u &= \eta \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \Gamma. \end{aligned}$$

By elliptic regularity we have $\|u\|_{2,\Omega} \lesssim \|\eta\|_{0,\Omega}$. Let $\boldsymbol{\mu} \coloneqq -\nabla u$, which gives due to regularity $\boldsymbol{\mu} \in \boldsymbol{V}$. We therefore have $\|\boldsymbol{\mu}\|_{\boldsymbol{H}(\Omega,\operatorname{div})} \lesssim \|\eta\|_{0,\Omega}$. Furthermore, due to the smoothness of Γ as well as due to a multiplicative trace inequality we find

$$\|\boldsymbol{\mu} \cdot \boldsymbol{n}\|_{0,\Gamma} = \|\nabla u \cdot \boldsymbol{n}\|_{0,\Gamma} \lesssim \|\nabla u\|_{0,\Gamma} \lesssim \|\nabla u\|_{0,\Omega}^{1/2} \|\nabla u\|_{1,\Omega}^{1/2} \leq \|u\|_{2,\Omega} \lesssim \|\eta\|_{0,\Omega} .$$

Consequently we find $\|\boldsymbol{\mu}\|_{\boldsymbol{V}} \lesssim \|\eta\|_{0,\Omega}$. Finally we have

$$(\nabla \cdot \boldsymbol{\mu}, \eta)_\Omega = (\eta, \eta)_\Omega = \|\eta\|_{0,\Omega} \|\eta\|_{0,\Omega} \gtrsim \|\eta\|_{0,\Omega} \|\boldsymbol{\mu}\|_{\boldsymbol{V}} ,$$

which proves the inf-sup condition.
Coercivity on kernel - discrete: The coercivity is again trivial with the same argument as above.

inf-sup condition - discrete: Let $\lambda_h \in \nabla \cdot \mathbf{V}_{p_v}(\mathcal{T}_h)$ be given. As above in the continuous case we solve the Poisson problem

$$\begin{aligned} -\Delta u &= \lambda_h \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \Gamma. \end{aligned}$$

Let $\boldsymbol{\Lambda} = -\nabla u$ and again we have $\|\boldsymbol{\Lambda}\|_{\boldsymbol{V}} \lesssim \|\boldsymbol{\Lambda}\|_{\boldsymbol{H}(\Omega,\mathrm{div})} + \|\boldsymbol{\Lambda}\cdot\boldsymbol{n}\|_{0,\Gamma} \leq \|u\|_{2,\Omega} \lesssim \|\lambda_h\|_{0,\Omega}$. We now employ the commuting diagram projection based interpolation operators defined in [MR20], see also Proposition 5.3.1. We use this operator to project $\boldsymbol{\Lambda}$ onto the conforming subspace. With $\boldsymbol{\Lambda}_h \coloneqq \boldsymbol{\Pi}^{\mathrm{div}}_{p_v}\boldsymbol{\Lambda}$ we find

$$\nabla\cdot\boldsymbol{\Lambda}_h = \nabla\cdot\boldsymbol{\Pi}^{\mathrm{div}}_{p_v}\boldsymbol{\Lambda} = \Pi^{L^2}_{p_v}\nabla\cdot\boldsymbol{\Lambda} = \Pi^{L^2}_{p_v}\lambda_h = \lambda_h,$$

where $\Pi^{L^2}_{p_v}$ denotes the L^2 orthogonal projection on $\nabla\ \cdot\mathbf{V}_{p_v}(\mathcal{T}_h)$. Using [MR20, Thm. 2.10, (vi)] we can estimate

$$\|\boldsymbol{\Lambda} - \boldsymbol{\Pi}^{\mathrm{div}}_{p_v}\boldsymbol{\Lambda}\|_{\boldsymbol{H}(\Omega,\mathrm{div})} \lesssim \|\boldsymbol{\Lambda}\|_{1,\Omega} \lesssim \|\lambda_h\|_{0,\Omega}\,.$$

Furthermore, since $\boldsymbol{\Pi}^{\mathrm{div}}_{p_v}$ realizes the $L^2(\Gamma)$ orthogonal projection of the normal trace, we find

$$\|(\boldsymbol{\Lambda} - \boldsymbol{\Pi}^{\mathrm{div}}_{p_v}\boldsymbol{\Lambda})\cdot\boldsymbol{n}\|_{0,\Gamma} \leq \|\boldsymbol{\Lambda}\cdot\boldsymbol{n}\|_{0,\Gamma} \lesssim \|\lambda_h\|_{0,\Omega}\,,$$

which finally leads to

$$\|\boldsymbol{\Lambda}_h\|_{\boldsymbol{V}} = \|\boldsymbol{\Pi}^{\mathrm{div}}_{p_v}\boldsymbol{\Lambda}\|_{\boldsymbol{V}} \lesssim \|\boldsymbol{\Lambda} - \boldsymbol{\Pi}^{\mathrm{div}}_{p_v}\boldsymbol{\Lambda}\|_{\boldsymbol{V}} + \|\boldsymbol{\Lambda}\|_{\boldsymbol{V}} \lesssim \|\lambda_h\|_{0,\Omega}\,.$$

For any $\lambda_h \in \nabla\cdot\mathbf{V}_{p_v}(\mathcal{T}_h)$ we estimate

$$\sup_{\boldsymbol{\varphi}_h\in\mathbf{V}_{p_v}(\mathcal{T}_h)} \frac{(\nabla\cdot\boldsymbol{\varphi}_h,\lambda_h)_\Omega}{\|\boldsymbol{\varphi}_h\|_{\boldsymbol{V}}\,\|\lambda_h\|_{0,\Omega}} \geq \frac{(\nabla\cdot\boldsymbol{\Lambda}_h,\lambda_h)_\Omega}{\|\boldsymbol{\Lambda}_h\|_{\boldsymbol{V}}\,\|\lambda_h\|_{0,\Omega}} = \frac{\|\lambda_h\|_{0,\Omega}}{\|\boldsymbol{\Lambda}_h\|_{\boldsymbol{V}}} \gtrsim 1,$$

which proves the discrete inf-sup condition. We have therefore proven

Lemma 5.3.3. *For any mesh $\mathcal{T}_h$ satisfying Assumption 2.0.1, the operator $\boldsymbol{I}^\Gamma_h : \boldsymbol{V} \to \mathbf{V}_{p_v}(\mathcal{T}_h)$ is well-defined with bounds independent of the mesh size h and the polynomial degree p_v.*

As a tool in the $L^2(\Omega)$ analysis of the operator $\boldsymbol{I}^\Gamma_h$ we need the following decomposition. Compared to Section 4.3 we need a Helmholtz-like decomposition accounting for the regularity of the normal trace:

Lemma 5.3.4 (Continuous and discrete Helmholtz-like decomposition - $L^2(\Gamma)$ normal trace)**.** *Let $\boldsymbol{Y} \subset \boldsymbol{H}(\Omega,\mathrm{curl})$ be given by*

$$\boldsymbol{Y} \coloneqq \left\{\boldsymbol{\mu}\in\boldsymbol{H}(\Omega,\mathrm{curl})\colon (\nabla\times\boldsymbol{\mu})\cdot\boldsymbol{n}\in L^2(\Gamma)\right\}.$$

The operators $\boldsymbol{\Pi}^{\mathrm{curl},\Gamma}\colon \boldsymbol{V}\to\nabla\times\boldsymbol{Y}$ and $\boldsymbol{\Pi}^{\mathrm{curl},\Gamma}_h\colon \mathbf{V}_{p_v}(\mathcal{T}_h)\to\nabla\times\mathbf{N}_{p_v}(\mathcal{T}_h)$ given by

$$\begin{aligned} \langle\langle\boldsymbol{\Pi}^{\mathrm{curl},\Gamma}\boldsymbol{\varphi},\nabla\times\boldsymbol{\mu}\rangle\rangle &= \langle\langle\boldsymbol{\varphi},\nabla\times\boldsymbol{\mu}\rangle\rangle \quad \forall\boldsymbol{\mu}\in\boldsymbol{Y}, \\ \langle\langle\boldsymbol{\Pi}^{\mathrm{curl},\Gamma}_h\boldsymbol{\varphi}_h,\nabla\times\boldsymbol{\mu}_h\rangle\rangle &= \langle\langle\boldsymbol{\varphi}_h,\nabla\times\boldsymbol{\mu}_h\rangle\rangle \quad \forall\boldsymbol{\mu}_h\in\mathbf{N}_{p_v}(\mathcal{T}_h) \end{aligned}$$

are well-defined. Furthermore, the remainder $\boldsymbol{r}$ *of the continuous decomposition* $\boldsymbol{\varphi} = \boldsymbol{\Pi}^{\mathrm{curl},\Gamma}\boldsymbol{\varphi} + \boldsymbol{r}$ *satisfies* $\boldsymbol{r} \in \boldsymbol{H}^1(\Omega)$ *with the a priori estimate* $\|\boldsymbol{r}\|_{1,\Omega} \lesssim \|\nabla \cdot \boldsymbol{\varphi}\|_{0,\Omega}$. *Additionally there exists an* $R \in H^2(\Omega)$ *such that* $\boldsymbol{r} = \nabla R$, *where* R *satisfies*

$$\begin{aligned} -\Delta R &= -\nabla \cdot \boldsymbol{\varphi} && \text{in } \Omega, \\ \partial_n R + R &= 0 && \text{on } \Gamma. \end{aligned}$$

Furthermore, $\boldsymbol{r}$ *satisfies*

$$\begin{aligned} \nabla \cdot \boldsymbol{r} &= \nabla \cdot \boldsymbol{\varphi} && \text{in } \Omega, \\ \nabla \times \boldsymbol{r} &= 0 && \text{in } \Omega, \\ \boldsymbol{r} \cdot \boldsymbol{n} &= -R && \text{on } \Gamma. \end{aligned}$$

Finally, the estimate $\|R\|_{2,\Omega} \lesssim \|\boldsymbol{r}\|_{1,\Omega} \lesssim \|\nabla \cdot \boldsymbol{\varphi}\|_{0,\Omega}$ *holds.*

Proof. The unique solvability on a discrete and continuous level follows immediately from the fact that the variational formulations are just the definition of the orthogonal projections onto $\nabla \times \boldsymbol{Y}$ and $\nabla \times \mathbf{N}_{p_v}(\mathcal{T}_h)$, respectively. For any $\boldsymbol{\mu} \in \boldsymbol{C}_0^\infty(\Omega)$ we find

$$\langle\langle \boldsymbol{r}, \nabla \times \boldsymbol{\mu} \rangle\rangle = (\boldsymbol{r}, \nabla \times \boldsymbol{\mu})_\Omega = 0,$$

which gives $\nabla \times \boldsymbol{r} = 0$. Since $\boldsymbol{\Pi}^{\mathrm{curl},\Gamma}\boldsymbol{\varphi} \in \nabla \times \boldsymbol{Y}$ we can conclude $\nabla \cdot \boldsymbol{r} = \nabla \cdot \boldsymbol{\varphi}$. The fact that $\nabla \times \boldsymbol{r} = 0$ gives via the exact sequence property of the following spaces

$$\mathbb{R} \xrightarrow{\mathrm{id}} H^1(\Omega) \xrightarrow{\nabla} \boldsymbol{H}(\Omega, \mathrm{curl}) \xrightarrow{\nabla\times} \boldsymbol{H}(\Omega, \mathrm{div}) \xrightarrow{\nabla\cdot} L^2(\Omega) \xrightarrow{0} \{0\}$$

the existence of a potential $R \in H^1(\Omega)$ such that $\boldsymbol{r} = \nabla R$. Therefore, we immediately have $-\Delta R = -\nabla \cdot \nabla R = -\nabla \cdot \boldsymbol{r} = -\nabla \cdot \boldsymbol{\varphi}$. To analyze the boundary conditions satisfied by R we insert $\boldsymbol{r} = \nabla R$ into the variational formulation and integrate by parts:

$$0 = \langle\langle \nabla R, \nabla \times \boldsymbol{\mu} \rangle\rangle = (\nabla R, \nabla \times \boldsymbol{\mu})_\Omega + \langle \partial_n R, (\nabla \times \boldsymbol{\mu}) \cdot \boldsymbol{n} \rangle_\Gamma = \langle R + \partial_n R, (\nabla \times \boldsymbol{\mu}) \cdot \boldsymbol{n} \rangle_\Gamma.$$

Since $(\nabla \times \boldsymbol{\mu}) \cdot \boldsymbol{n} = \nabla_\Gamma \cdot (\boldsymbol{\mu} \times \boldsymbol{n})$ we conclude $\partial_n R + R = c$ for some $c \in \mathbb{R}$. We can however choose $c = 0$. This is due to the fact that the family of solutions

$$\begin{aligned} -\Delta R_c &= -\nabla \cdot \boldsymbol{\varphi} && \text{in } \Omega, \\ \partial_n R_c + R_c &= c && \text{on } \Gamma, \end{aligned}$$

for $c \in \mathbb{R}$ is uniquely determined up to a constant, since the difference $D = R_a - R_b$ satisfies

$$\begin{aligned} -\Delta D &= 0 && \text{in } \Omega, \\ \partial_n D + D &= a - b && \text{on } \Gamma, \end{aligned}$$

to which the constant solution $D \equiv a - b$ is the unique solution in $H^1(\Omega)$. In order to prove the final estimates, first note that due to elliptic regularity we have $R \in H^2(\Omega)$ with $\|R\|_{2,\Omega} \lesssim \|\nabla \cdot \boldsymbol{\varphi}\|_{0,\Omega}$. Note that $\|\nabla \cdot\|_{0,\Omega} + \|\cdot\|_{0,\Gamma}$ defines an equivalent norm to the $H^1(\Omega)$

norm. Using norm equivalence, the boundary condition satisfied by R, as well as a trace inequality we find:

$$\begin{aligned}\|R\|_{2,\Omega} &\lesssim \|R\|_{1,\Omega} + \|\nabla R\|_{1,\Omega} \\ &\lesssim \|\nabla R\|_{0,\Omega} + \|R\|_{0,\Gamma} + \|\nabla R\|_{1,\Omega} \\ &= \|\nabla R\|_{0,\Omega} + \|\partial_n R\|_{0,\Gamma} + \|\nabla R\|_{1,\Omega} \\ &= \|\nabla R\|_{0,\Omega} + \|\nabla R \cdot \boldsymbol{n}\|_{0,\Gamma} + \|\nabla R\|_{1,\Omega} \\ &\lesssim \|\nabla R\|_{0,\Omega} + \|\nabla R\|_{0,\Gamma} + \|\nabla R\|_{1,\Omega} \\ &\lesssim \|\nabla R\|_{1,\Omega} = \|\boldsymbol{r}\|_{1,\Omega} \lesssim \|\nabla \cdot \boldsymbol{\varphi}\|_{0,\Omega}\,,\end{aligned}$$

which concludes the proof. □

Lemma 5.3.5. *The operator $\boldsymbol{I}_h^\Gamma$ satisfies for arbitrary $\tilde{\boldsymbol{\varphi}}_h \in \mathbf{V}_{p_v}(\mathcal{T}_h)$ the estimates*

$$\begin{aligned}|||\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}||| &\lesssim |||\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h||| + \frac{h}{p_v}\|\nabla \cdot (\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega}\,, \\ \|\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi})\|_{0,\Omega} &\leq \|\nabla \cdot (\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega}\,.\end{aligned}$$

Proof. The proof is very similar to the one of Lemma 4.3.6, where a similar operator is analyzed. In essence the arguments are the same by replacing $\|\cdot\|_{0,\Omega}$ with $|||\cdot|||$. Let $\tilde{\boldsymbol{\varphi}}_h \in \mathbf{V}_{p_v}(\mathcal{T}_h)$ be arbitrary. Due to the orthogonality relation satisfied by the operator $\boldsymbol{I}_h^\Gamma$ the second estimate is obvious. We have with $\boldsymbol{e} = \boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}$

$$|||\boldsymbol{e}|||^2 = \langle\langle \boldsymbol{e}, \boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h \rangle\rangle + \langle\langle \boldsymbol{e}, \tilde{\boldsymbol{\varphi}}_h - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi} \rangle\rangle.$$

Lemma 5.3.4 enables us to split the discrete object $\tilde{\boldsymbol{\varphi}}_h - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi} \in \mathbf{V}_{p_v}(\mathcal{T}_h)$ on a discrete and a continuous level:

$$\begin{aligned}\tilde{\boldsymbol{\varphi}}_h - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi} &= \nabla \times \boldsymbol{\mu} + \boldsymbol{r}, \\ \tilde{\boldsymbol{\varphi}}_h - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi} &= \nabla \times \boldsymbol{\mu}_h + \boldsymbol{r}_h\end{aligned}$$

for certain $\boldsymbol{\mu} \in \boldsymbol{Y}$, $\boldsymbol{r} \in \boldsymbol{V}$, $\boldsymbol{\mu}_h \in \mathbf{N}_{p_v}(\mathcal{T}_h)$ and $\boldsymbol{r}_h \in \mathbf{V}_{p_v}(\mathcal{T}_h)$. Since $\nabla \cdot \nabla \times = 0$, the definition of $\boldsymbol{I}_h^\Gamma$ immediately gives

$$\langle\langle \boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}, \nabla \times \boldsymbol{\mu}_h \rangle\rangle = 0. \tag{5.15}$$

With (5.15) we therefore have

$$\langle\langle \boldsymbol{e}, \tilde{\boldsymbol{\varphi}}_h - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi} \rangle\rangle = \langle\langle \boldsymbol{e}, \nabla \times \boldsymbol{\mu}_h + \boldsymbol{r}_h \rangle\rangle = \langle\langle \boldsymbol{e}, \boldsymbol{r}_h \rangle\rangle = \langle\langle \boldsymbol{e}, \boldsymbol{r}_h - \boldsymbol{r} \rangle\rangle + \langle\langle \boldsymbol{e}, \boldsymbol{r} \rangle\rangle \coloneqq T_1 + T_2.$$

Treatment of T_1: See the proof of Lemma 4.3.6 for completely analogous arguments and more details. Since $\nabla \cdot \boldsymbol{r} = \nabla \cdot \boldsymbol{r}_h \in \nabla \cdot \mathbf{V}_{p_v}(\mathcal{T}_h)$ we find using the commuting diagram as well as the projection property of the operator $\mathbf{\Pi}_{p_v}^{\mathrm{div}}$

$$\nabla \cdot (\mathbf{\Pi}_{p_v}^{\mathrm{div}} \boldsymbol{r} - \boldsymbol{r}_h) = \Pi_{p_v}^{L^2}(\nabla \cdot \boldsymbol{r}) - \nabla \cdot \boldsymbol{r}_h = \nabla \cdot \boldsymbol{r} - \nabla \cdot \boldsymbol{r}_h = 0.$$

By the exact sequence property we therefore have $\mathbf{\Pi}_{p_v}^{\mathrm{div}}\boldsymbol{r}-\boldsymbol{r}_h \in \nabla\times\mathbf{N}_{p_v}(\mathcal{T}_h)$. The definition of $\boldsymbol{r}$ and $\boldsymbol{r}_h$ in Lemma 5.3.4 gives the orthogonality

$$\langle\langle \boldsymbol{r}-\boldsymbol{r}_h, \nabla\times\tilde{\boldsymbol{\mu}}_h\rangle\rangle = 0 \qquad \forall\tilde{\boldsymbol{\mu}}_h\in\mathbf{N}_{p_v}(\mathcal{T}_h).$$

Putting it all together we have

$$|||\boldsymbol{r}-\boldsymbol{r}_h|||^2 = \langle\langle \boldsymbol{r}-\boldsymbol{r}_h, \boldsymbol{r}-\mathbf{\Pi}_{p_v}^{\mathrm{div}}\boldsymbol{r}\rangle\rangle + \langle\langle \boldsymbol{r}-\boldsymbol{r}_h, \mathbf{\Pi}_{p_v}^{\mathrm{div}}\boldsymbol{r}-\boldsymbol{r}_h\rangle\rangle = \langle\langle \boldsymbol{r}-\boldsymbol{r}_h, \boldsymbol{r}-\mathbf{\Pi}_{p_v}^{\mathrm{div}}\boldsymbol{r}\rangle\rangle.$$

Applying the Cauchy-Schwarz inequality and the definition of $|||\cdot|||$ we find

$$|||\boldsymbol{r}-\boldsymbol{r}_h||| \le |||\boldsymbol{r}-\mathbf{\Pi}_{p_v}^{\mathrm{div}}\boldsymbol{r}||| \lesssim \|\boldsymbol{r}-\mathbf{\Pi}_{p_v}^{\mathrm{div}}\boldsymbol{r}\|_{0,\Omega} + \|(\boldsymbol{r}-\mathbf{\Pi}_{p_v}^{\mathrm{div}}\boldsymbol{r})\cdot\boldsymbol{n}\|_{0,\Gamma}.$$

In order to treat the volume term we invoke [MR20, Thm. 2.10, (vi)], which is applicable since $\nabla\cdot\boldsymbol{r} = \nabla\cdot\boldsymbol{r}_h$ is discrete. The estimate of Lemma 5.3.4 therefore gives

$$\|\boldsymbol{r}-\mathbf{\Pi}_{p_v}^{\mathrm{div}}\boldsymbol{r}\|_{0,\Omega} \lesssim \frac{h}{p_v}\|\boldsymbol{r}\|_{1,\Omega} \lesssim \frac{h}{p_v}\|\nabla\cdot(\tilde{\boldsymbol{\varphi}}_h - \boldsymbol{I}_h^\Gamma\boldsymbol{\varphi})\|_{0,\Omega}.$$

To estimate the boundary term we apply Proposition 5.3.1 to conclude

$$\begin{aligned}\|(\boldsymbol{r}-\mathbf{\Pi}_{p_v}^{\mathrm{div}}\boldsymbol{r})\cdot\boldsymbol{n}\|_{0,\Gamma} &= \|\boldsymbol{r}\cdot\boldsymbol{n} - \Pi_{p_v}^{L^2(\Gamma)}\boldsymbol{r}\cdot\boldsymbol{n}\|_{0,\Gamma} \lesssim \frac{h}{p_v}\|\boldsymbol{r}\cdot\boldsymbol{n}\|_{1,\Gamma} \lesssim \frac{h}{p_v}\|\nabla R\|_{1,\Gamma}\\ &\lesssim \frac{h}{p_v}\|R\|_{2,\Omega} \lesssim \frac{h}{p_v}\|\nabla\cdot(\tilde{\boldsymbol{\varphi}}_h - \boldsymbol{I}_h^\Gamma\boldsymbol{\varphi})\|_{0,\Omega}.\end{aligned}$$

Summarizing the above we have

$$|||\boldsymbol{r}-\boldsymbol{r}_h||| \lesssim \frac{h}{p_v}\|\nabla\cdot(\tilde{\boldsymbol{\varphi}}_h - \boldsymbol{I}_h^\Gamma\boldsymbol{\varphi})\|_{0,\Omega}. \tag{5.16}$$

Adding and subtracting $\boldsymbol{\varphi}$, applying the triangle inequality as well as the second inequality of the present lemma we find

$$T_1 \le |||\boldsymbol{e}|||\cdot|||\boldsymbol{r}-\boldsymbol{r}_h||| \lesssim \frac{h}{p_v}|||\boldsymbol{e}|||\cdot\|\nabla\cdot(\tilde{\boldsymbol{\varphi}}_h - \boldsymbol{I}_h^\Gamma\boldsymbol{\varphi})\|_{0,\Omega} \lesssim \frac{h}{p_v}|||\boldsymbol{e}|||\cdot\|\nabla\cdot(\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega}.$$

Treatment of T_2: The term T_2 is treated with a duality argument. We seek to find $\boldsymbol{\psi}\in\boldsymbol{H}(\Omega,\mathrm{div})$ such that

$$\langle\langle \boldsymbol{v},\boldsymbol{r}\rangle\rangle = (\nabla\cdot\boldsymbol{v}, \nabla\cdot\boldsymbol{\psi})_\Omega \qquad \forall\boldsymbol{v}\in\boldsymbol{V}.$$

Since $\boldsymbol{r} = \nabla R$ for some $R\in H^2(\Omega)$, see Lemma 5.3.4, we have

$$\begin{aligned}(\nabla\cdot\boldsymbol{v}, \nabla\cdot\boldsymbol{\psi})_\Omega &= \langle\langle \boldsymbol{v},\boldsymbol{r}\rangle\rangle = \langle\langle \boldsymbol{v},\nabla R\rangle\rangle = (\boldsymbol{v},\nabla R)_\Omega + \langle \boldsymbol{v}\cdot\boldsymbol{n}, \partial_n R\rangle_\Omega\\ &= -(\nabla\cdot\boldsymbol{v}, R)_\Omega + \langle \boldsymbol{v}\cdot\boldsymbol{n}, \partial_n R + R\rangle_\Gamma = -(\nabla\cdot\boldsymbol{v}, R)_\Omega.\end{aligned}$$

Upon solving the problem

$$\begin{aligned}-\Delta w &= R \quad \text{in } \Omega,\\ w &= 0 \quad \text{on } \Gamma,\end{aligned}$$

and setting $\boldsymbol{\psi} = \nabla w$ we found the desired $\boldsymbol{\psi}$. Furthermore, elliptic regularity gives $w \in H^4(\Omega)$ and therefore $\boldsymbol{\psi} \in \boldsymbol{H}^3(\Omega)$. Finally the following estimates hold

$$\|\nabla \cdot \boldsymbol{\psi}\|_{1,\Omega} \leq \|\boldsymbol{\psi}\|_{2,\Omega} \leq \|w\|_{3,\Omega} \leq \|R\|_{1,\Omega} \lesssim \|\nabla \cdot (\tilde{\boldsymbol{\varphi}}_h - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi})\|_{(H^1(\Omega))'},$$

due to elliptic regularity and the results of Lemma 5.3.4. We therefore have for any $\boldsymbol{\psi}_h \in \mathbf{V}_{p_v}(\mathcal{T}_h)$

$$T_2 = \langle\langle \boldsymbol{e}, \boldsymbol{r} \rangle\rangle = (\nabla \cdot \boldsymbol{e}, \nabla \cdot \boldsymbol{\psi})_\Omega = (\nabla \cdot \boldsymbol{e}, \nabla \cdot (\boldsymbol{\psi} - \boldsymbol{\psi}_h))_\Omega \leq \|\nabla \cdot \mathbf{e}\|_{0,\Omega} \|\nabla \cdot (\boldsymbol{\psi} - \boldsymbol{\psi}_h)\|_{0,\Omega},$$

where we used the definition of T_2, the duality argument elaborated above, the orthogonality relation of $\boldsymbol{I}_h^\Gamma$ to insert any $\boldsymbol{\psi}_h \in \mathbf{V}_{p_v}(\mathcal{T}_h)$, and the Cauchy-Schwarz inequality. Finally, exploiting the *a priori* estimate of $\boldsymbol{\psi}$ we find

$$T_2 \leq \|\nabla \cdot \mathbf{e}\|_{0,\Omega} \cdot \inf_{\boldsymbol{\psi}_h \in \mathbf{V}_{p_v}(\mathcal{T}_h)} \|\nabla \cdot (\boldsymbol{\psi} - \boldsymbol{\psi}_h)\|_{0,\Omega} \lesssim \frac{h}{p_v} \|\nabla \cdot \mathbf{e}\|_{0,\Omega} \|\nabla \cdot (\tilde{\boldsymbol{\varphi}}_h - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi})\|_{(H^1(\Omega))'}.$$

We now estimate using partial integration

$$\begin{aligned}
\|\nabla \cdot (\tilde{\boldsymbol{\varphi}}_h - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi})\|_{(H^1(\Omega))'} &= \sup_{f \in H^1(\Omega)} \frac{|(\nabla \cdot (\tilde{\boldsymbol{\varphi}}_h - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}), f)_\Omega|}{\|f\|_{1,\Omega}} \\
&= \sup_{f \in H^1(\Omega)} \frac{|-(\tilde{\boldsymbol{\varphi}}_h - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}, \nabla f)_\Omega + \langle (\tilde{\boldsymbol{\varphi}}_h - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}) \cdot \boldsymbol{n}, f \rangle_\Gamma|}{\|f\|_{1,\Omega}} \\
&\lesssim |||\tilde{\boldsymbol{\varphi}}_h - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}|||.
\end{aligned}$$

Hence, we find

$$\begin{aligned}
|||\mathbf{e}|||^2 &= \langle\langle \boldsymbol{e}, \boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h \rangle\rangle + \langle\langle \boldsymbol{e}, \tilde{\boldsymbol{\varphi}}_h - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi} \rangle\rangle \\
&= \langle\langle \boldsymbol{e}, \boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h \rangle\rangle + T_1 + T_2 \\
&\lesssim |||\mathbf{e}||| \cdot |||\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h|||_{0,\Omega} + \frac{h}{p_v} |||\mathbf{e}||| \cdot \|\nabla \cdot (\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega} + \frac{h}{p_v} \|\nabla \cdot \mathbf{e}\|_{0,\Omega} \cdot |||\tilde{\boldsymbol{\varphi}}_h - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}|||.
\end{aligned}$$

Adding and subtracting $\boldsymbol{\varphi}$ in the last term, applying the triangle inequality, the second estimate of the present lemma as well as the Young inequality yields the result. □

Theorem 5.3.6 (Suboptimal estimate for $\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}$ - Robin version of Theorem 4.3.8)**.** *Let Γ be smooth and $(\boldsymbol{\varphi}_h, u_h)$ be the least squares approximation of $(\boldsymbol{\varphi}, u)$. Furthermore, let $e^u = u - u_h$ and $\boldsymbol{e}^{\boldsymbol{\varphi}} = \boldsymbol{\varphi} - \boldsymbol{\varphi}_h$. Then, for any $\tilde{u}_h \in S_{p_s}(\mathcal{T}_h)$, $\tilde{\boldsymbol{\varphi}}_h \in \mathbf{V}_{p_v}(\mathcal{T}_h)$,*

$$\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega} \lesssim \left(\frac{h}{p}\right)^{1/2} \|u - \tilde{u}_h\|_{1,\Omega} + \|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h\|_{0,\Omega} + \|(\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h) \cdot \boldsymbol{n}\|_{0,\Gamma} + \left(\frac{h}{p}\right)^{1/2} \|\nabla \cdot (\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega}.$$

Proof. Let $(\boldsymbol{\psi}, v) \in \boldsymbol{V} \times W$ denote the dual solution given by Theorem 5.2.3 applied to $\boldsymbol{\eta} = \boldsymbol{e}^{\boldsymbol{\varphi}}$. Theorem 5.2.3 gives $\boldsymbol{\psi} \in \boldsymbol{L}^2(\Omega)$, $\nabla \cdot \boldsymbol{\psi} \in H^1(\Omega)$, $\boldsymbol{\psi} \cdot \boldsymbol{n} \in H^{1/2}(\Gamma)$ and $v \in H^2(\Omega)$. Due to the Galerkin orthogonality we have for any $(\tilde{\boldsymbol{\psi}}_h, \tilde{v}_h)$

$$\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}^2 = b((\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u), (\boldsymbol{\psi}, v)) = b((\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u), (\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h, v - \tilde{v}_h)).$$

We now estimate all terms, except for $\langle\langle \boldsymbol{e}^{\boldsymbol{\varphi}}, \boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h \rangle\rangle = (\boldsymbol{e}^{\boldsymbol{\varphi}}, \boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h)_\Omega + \langle \boldsymbol{e}^{\boldsymbol{\varphi}} \cdot \boldsymbol{n}, (\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h) \cdot \boldsymbol{n} \rangle_\Gamma$, in the above:

$$
\begin{aligned}
(\nabla e^u + \boldsymbol{e}^{\boldsymbol{\varphi}}, \nabla(v - \tilde{v}_h))_\Omega &\leq \|(\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u)\|_b \|\nabla(v - \tilde{v}_h)\|_{0,\Omega}, \\
\langle -\alpha e^u, (\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h) \cdot \boldsymbol{n} \rangle_\Gamma &\leq \|e^u\|_{0,\Gamma} \, \|(\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h) \cdot \boldsymbol{n}\|_{0,\Gamma} \\
&\leq (h/p)^{1/2} \|(\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u)\|_b \|(\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h) \cdot \boldsymbol{n}\|_{0,\Gamma}, \\
(\nabla \cdot \boldsymbol{e}^{\boldsymbol{\varphi}} + \gamma e^u, \nabla \cdot (\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h) + \gamma(v - \tilde{v}_h))_\Omega &\lesssim \|(\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u)\|_b \Big[\|\nabla \cdot (\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h))\|_{0,\Omega} + \|v - \tilde{v}_h\|_{0,\Omega} \Big], \\
(\nabla e^u, \boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h)_\Omega &= -(e^u, \nabla \cdot (\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h))_\Omega + \langle e^u, (\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h) \cdot \boldsymbol{n} \rangle_\Gamma \\
&\lesssim \|(\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u)\|_b \Big[\|\nabla \cdot (\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h))\|_{0,\Omega} \\
&\qquad + (h/p)^{1/2} \|(\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h) \cdot \boldsymbol{n}\|_{0,\Gamma} \Big], \\
\langle \boldsymbol{e}^{\boldsymbol{\varphi}} \cdot \boldsymbol{n} - \alpha e^u, -\alpha(v - \tilde{v}_h) \rangle_\Gamma &\lesssim \|(\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u)\|_b \|v - \tilde{v}_h\|_{1,\Omega}.
\end{aligned}
\tag{5.17}
$$

Therefore, we conclude that

$$
\begin{aligned}
\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}^2 \lesssim \|(\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u)\|_b \Big[\|\nabla \cdot (\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h))\|_{0,\Omega} + (h/p)^{1/2} \|(\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h) \cdot \boldsymbol{n}\|_{0,\Gamma} + \|v - \tilde{v}_h\|_{1,\Omega} \Big] \\
+ \langle\langle \boldsymbol{e}^{\boldsymbol{\varphi}}, \boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h \rangle\rangle.
\end{aligned}
\tag{5.18}
$$

To analyze the term $\langle\langle \boldsymbol{e}^{\boldsymbol{\varphi}}, \boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h \rangle\rangle$ we follow a similar procedure as in the proof of Theorem 4.3.8. Therefore we first perform a Helmholtz decomposition of the vector field $\boldsymbol{\psi}$. Since $\boldsymbol{\psi} \in \boldsymbol{H}(\Omega, \mathrm{div})$ with $\nabla \cdot \boldsymbol{\psi} \in H^1(\Omega)$ and $\boldsymbol{\psi} \cdot \boldsymbol{n} \in H^{1/2}(\Gamma)$ there exist $\boldsymbol{\rho} \in \boldsymbol{H}_0(\Omega, \mathrm{curl})$ and $z \in H^2(\Omega)$ such that $\boldsymbol{\psi} = \nabla \times \boldsymbol{\rho} + \nabla z$. To that end, let $z \in H^1(\Omega)$ with zero average solve

$$
\begin{aligned}
-\Delta z &= -\nabla \cdot \boldsymbol{\psi} \quad && \text{in } \Omega, \\
\partial_n z &= \boldsymbol{\psi} \cdot \boldsymbol{n} && \text{on } \Gamma.
\end{aligned}
$$

Since $\nabla \cdot (\boldsymbol{\psi} - \nabla z) = 0$ as well as $(\boldsymbol{\psi} - \nabla z) \cdot \boldsymbol{n} = 0$ by construction, the exact sequence property of the employed spaces allows for the existence of $\boldsymbol{\rho} \in \boldsymbol{H}_0(\Omega, \mathrm{curl})$ such that $\boldsymbol{\psi} - \nabla z = \nabla \times \boldsymbol{\rho}$. Elliptic regularity furthermore gives $z \in H^2(\Omega)$ with the estimate

$$
\|z\|_{2,\Omega} \lesssim \|\nabla \cdot \boldsymbol{\psi}\|_{0,\Omega} + \|\boldsymbol{\psi} \cdot \boldsymbol{n}\|_{1/2,\Gamma} .
$$

To estimate $\|z\|_{1,\Omega}$ we use the *a priori* estimate of the Lax-Milgram theorem applied to the weak formulation of the above problem, which is given by

$$
(\nabla z, \nabla w)_\Omega = (-\nabla \cdot \boldsymbol{\psi}, w)_\Omega + \langle \boldsymbol{\psi} \cdot \boldsymbol{n}, w \rangle_\Gamma = (\boldsymbol{\psi}, \nabla w)_\Omega,
$$

due to partial integration. Therefore we have $\|z\|_{1,\Omega} \lesssim \|\boldsymbol{\psi}\|_{0,\Omega}$. Finally, we have the estimate $\|\nabla \times \boldsymbol{\rho}\|_{0,\Omega} \leq \|\boldsymbol{\psi}\|_{0,\Omega} + \|\nabla z\|_{0,\Omega} \lesssim \|\boldsymbol{\psi}\|_{0,\Omega}$. We now continue estimating (5.18) by applying the Helmholtz decomposition. In essence this is again the procedure of Theorem 4.3.8 by replacing $\|\cdot\|_{0,\Omega}$ with $|||\cdot|||$. For the readers' convenience we recall the important steps. For any $\tilde{\boldsymbol{\psi}}_h^c, \tilde{\boldsymbol{\psi}}_h^g \in \mathbf{V}_{p_v}(\mathcal{T}_h)$ we have with $\tilde{\boldsymbol{\psi}}_h = \tilde{\boldsymbol{\psi}}_h^c + \tilde{\boldsymbol{\psi}}_h^g$

$$
\langle\langle \boldsymbol{e}^{\boldsymbol{\varphi}}, \boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h \rangle\rangle = \langle\langle \boldsymbol{e}^{\boldsymbol{\varphi}}, \nabla \times \boldsymbol{\rho} - \tilde{\boldsymbol{\psi}}_h^c \rangle\rangle + \langle\langle \boldsymbol{e}^{\boldsymbol{\varphi}}, \nabla z - \tilde{\boldsymbol{\psi}}_h^g \rangle\rangle =: T^c + T^g.
$$

Treatment of T^g: By the Cauchy-Schwarz inequality we have

$$T^g = \langle\langle \boldsymbol{e}^{\boldsymbol{\varphi}}, \nabla z - \tilde{\boldsymbol{\psi}}_h^g \rangle\rangle \leq |||\boldsymbol{e}^{\boldsymbol{\varphi}}||| \cdot |||\nabla z - \tilde{\boldsymbol{\psi}}_h^g|||.$$

Treatment of T^c: For any $\tilde{\boldsymbol{\varphi}}_h \in \mathbf{V}_{p_v}(\mathcal{T}_h)$ we have

$$\begin{aligned} T^c &= \langle\langle \boldsymbol{e}^{\boldsymbol{\varphi}}, \nabla \times \boldsymbol{\rho} - \tilde{\boldsymbol{\psi}}_h^c \rangle\rangle \\ &= \langle\langle \boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h, \nabla \times \boldsymbol{\rho} - \tilde{\boldsymbol{\psi}}_h^c \rangle\rangle + \langle\langle \tilde{\boldsymbol{\varphi}}_h - \boldsymbol{\varphi}_h, \nabla \times \boldsymbol{\rho} - \tilde{\boldsymbol{\psi}}_h^c \rangle\rangle =: T_1^c + T_2^c. \end{aligned}$$

Treatment of T_1^c: By the Cauchy-Schwarz inequality we have

$$T_1^c = \langle\langle \boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h, \nabla \times \boldsymbol{\rho} - \tilde{\boldsymbol{\psi}}_h^c \rangle\rangle \leq |||\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h||| \cdot |||\nabla \times \boldsymbol{\rho} - \tilde{\boldsymbol{\psi}}_h^c|||.$$

Treatment of T_2^c: To treat T_2^c we apply Lemma 5.3.4 to split the discrete object $\tilde{\boldsymbol{\varphi}}_h - \boldsymbol{\varphi}_h \in \mathbf{V}_{p_v}(\mathcal{T}_h)$ on a discrete and a continuous level:

$$\begin{aligned} \tilde{\boldsymbol{\varphi}}_h - \boldsymbol{\varphi}_h &= \nabla \times \boldsymbol{\mu} + \boldsymbol{r}, \\ \tilde{\boldsymbol{\varphi}}_h - \boldsymbol{\varphi}_h &= \nabla \times \boldsymbol{\mu}_h + \boldsymbol{r}_h \end{aligned}$$

for certain $\boldsymbol{\mu} \in \boldsymbol{Y}$, $\boldsymbol{r} \in \boldsymbol{V}$, $\boldsymbol{\mu}_h \in \mathbf{N}_{p_v}(\mathcal{T}_h)$ and $\boldsymbol{r}_h \in \mathbf{V}_{p_v}(\mathcal{T}_h)$. We now choose $\tilde{\boldsymbol{\psi}}_h^c = \boldsymbol{\Pi}_h^{\mathrm{curl},\Gamma} \nabla \times \boldsymbol{\rho}$ given by Lemma 5.3.4. Exploiting the definition of the operator $\boldsymbol{\Pi}_h^{\mathrm{curl},\Gamma}$ we find

$$\begin{aligned} T_2^c &= \langle\langle \tilde{\boldsymbol{\varphi}}_h - \boldsymbol{\varphi}_h, \nabla \times \boldsymbol{\rho} - \tilde{\boldsymbol{\psi}}_h^c \rangle\rangle \\ &= \underbrace{\langle\langle \nabla \times \boldsymbol{\mu}_h, \nabla \times \boldsymbol{\rho} - \boldsymbol{\Pi}_h^{\mathrm{curl},\Gamma} \nabla \times \boldsymbol{\rho} \rangle\rangle}_{=0} + \langle\langle \boldsymbol{r}_h, \nabla \times \boldsymbol{\rho} - \boldsymbol{\Pi}_h^{\mathrm{curl},\Gamma} \nabla \times \boldsymbol{\rho} \rangle\rangle \\ &= \langle\langle \boldsymbol{r}_h - \boldsymbol{r}, \nabla \times \boldsymbol{\rho} - \boldsymbol{\Pi}_h^{\mathrm{curl},\Gamma} \nabla \times \boldsymbol{\rho} \rangle\rangle + \langle\langle \boldsymbol{r}, \nabla \times \boldsymbol{\rho} - \boldsymbol{\Pi}_h^{\mathrm{curl},\Gamma} \nabla \times \boldsymbol{\rho} \rangle\rangle \\ &=: T_1 + T_2. \end{aligned}$$

Treatment of T_1: As in the estimate (5.16) we have

$$|||\boldsymbol{r} - \boldsymbol{r}_h||| \lesssim \frac{h}{p_v} \|\nabla \cdot (\tilde{\boldsymbol{\varphi}}_h - \boldsymbol{\varphi}_h)\|_{0,\Omega},$$

which gives after applying the Cauchy-Schwarz inequality

$$T_1 \lesssim \frac{h}{p_v} \|\nabla \cdot (\tilde{\boldsymbol{\varphi}}_h - \boldsymbol{\varphi}_h)\|_{0,\Omega} \, |||\nabla \times \boldsymbol{\rho} - \boldsymbol{\Pi}_h^{\mathrm{curl},\Gamma} \nabla \times \boldsymbol{\rho}||| \lesssim \frac{h}{p_v} \|\nabla \cdot (\tilde{\boldsymbol{\varphi}}_h - \boldsymbol{\varphi}_h)\|_{0,\Omega} \, |||\nabla \times \boldsymbol{\rho}|||,$$

where the last estimate follows immediately from the fact that $\boldsymbol{\Pi}_h^{\mathrm{curl},\Gamma}$ is by definition a projection. Finally, adding and subtracting $\boldsymbol{\varphi}$ and applying the triangle inequality as well as estimating $\|\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{\varphi}_h)\|_{0,\Omega}$ by $\|(\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u)\|_b$ we find

$$T_1 \lesssim \frac{h}{p_v} \|\nabla \cdot (\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega} \, \|\nabla \times \boldsymbol{\rho}\|_{0,\Omega} + \frac{h}{p_v} \|(\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u)\|_b \|\nabla \times \boldsymbol{\rho}\|_{0,\Omega}.$$

Treatment of T_2: Note again that $\boldsymbol{\rho} \in \boldsymbol{H}_0(\Omega, \mathrm{curl})$ as well as the fact that $\boldsymbol{\Pi}_h^{\mathrm{curl},\Gamma}$ maps into

$\nabla\times\mathbf{N}_{p_v}(\mathcal{T}_h)$. Therefore we can write $\nabla\times\boldsymbol{\rho}-\boldsymbol{\Pi}_h^{\mathrm{curl},\Gamma}\nabla\times\boldsymbol{\rho}=\nabla\times\widehat{\boldsymbol{\rho}}$ for some $\widehat{\boldsymbol{\rho}}\in\boldsymbol{H}(\Omega,\mathrm{curl})$. In fact $\widehat{\boldsymbol{\rho}}\in\boldsymbol{Y}$ since $(\nabla\times\widehat{\boldsymbol{\rho}})\cdot\boldsymbol{n}=(\nabla\times\boldsymbol{\rho}-\boldsymbol{\Pi}_h^{\mathrm{curl},\Gamma}\nabla\times\boldsymbol{\rho})\cdot\boldsymbol{n}=(\boldsymbol{\Pi}_h^{\mathrm{curl},\Gamma}\nabla\times\boldsymbol{\rho})\cdot\boldsymbol{n}\in L^2(\Gamma)$. Consequently the definition of the remainder $\boldsymbol{r}$ gives $T_2=\langle\langle\boldsymbol{r},\nabla\times\widehat{\boldsymbol{\rho}}\rangle\rangle=0$, see Lemma 5.3.4.
Collecting all the terms: Since $\boldsymbol{\rho}\in\boldsymbol{H}_0(\Omega,\mathrm{curl})$ and consequently $\nabla\times\boldsymbol{\rho}\in\boldsymbol{H}_0(\Omega,\mathrm{div})$ we can estimate $|||\nabla\times\boldsymbol{\rho}|||=\|\nabla\times\boldsymbol{\rho}\|_{0,\Omega}\lesssim\|\boldsymbol{\psi}\|_{0,\Omega}\lesssim\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}$, where we used the estimates of the Helmholtz decomposition as well as the regularity estimates of Lemma 5.2.3. We can now summarize

$$\begin{aligned}\langle\langle\boldsymbol{e}^{\boldsymbol{\varphi}},\boldsymbol{\psi}-\tilde{\boldsymbol{\psi}}_h\rangle\rangle\lesssim&|||\nabla z-\tilde{\boldsymbol{\psi}}_h^g|||\cdot|||\boldsymbol{e}^{\boldsymbol{\varphi}}|||\\&+\left[|||\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h|||+\frac{h}{p_v}\|\nabla\cdot(\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega}+\frac{h}{p_v}\|(\boldsymbol{e}^{\boldsymbol{\varphi}},e^u)\|_b\right]\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}\,.\end{aligned}\tag{5.19}$$

To conclude the proof we estimate the quantities arising in the estimates (5.18) and (5.19). To that end, note that $\nabla z\in\boldsymbol{H}^1(\Omega,\mathrm{div})$. Using the estimates of the Helmholtz decomposition, the equation satisfied by z as well as the regularity estimates given by Theorem 5.2.3 we find

$$\begin{aligned}\|\nabla z\|_{\boldsymbol{H}^1(\Omega,\mathrm{div})}&\lesssim\|z\|_{2,\Omega}+\|\underbrace{\Delta z}_{=\nabla\cdot\boldsymbol{\psi}}\|_{1,\Omega}\lesssim\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}\,,\\\|\nabla z\cdot\boldsymbol{n}\|_{1/2,\Gamma}&=\|\boldsymbol{\psi}\cdot\boldsymbol{n}\|_{1/2,\Gamma}\lesssim\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}\,.\end{aligned}$$

Exploiting these regularity estimates and employing the operator in Proposition 5.3.1 we can find $\tilde{\boldsymbol{\psi}}_h^g\in\mathbf{V}_{p_v}(\mathcal{T}_h)$ such that

$$\begin{aligned}\|\nabla z-\tilde{\boldsymbol{\psi}}_h^g\|_{\boldsymbol{H}(\Omega,\mathrm{div})}&\lesssim h/p_v\,\|\nabla z\|_{\boldsymbol{H}^1(\Omega,\mathrm{div})}\lesssim h/p_v\,\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}\,,\\\|(\nabla z-\tilde{\boldsymbol{\psi}}_h^g)\cdot\boldsymbol{n}\|_{0,\Gamma}&\lesssim(h/p_v)^{1/2}\,\|\nabla z\cdot\boldsymbol{n}\|_{1/2,\Gamma}\lesssim(h/p_v)^{1/2}\,\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}\,,\\|||\nabla z-\tilde{\boldsymbol{\psi}}_h^g|||&\lesssim(h/p_v)^{1/2}\,\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}\,,\end{aligned}$$

where the last one is just a combination of the previous ones. These estimates in turn give

$$\begin{aligned}\|\nabla\cdot(\boldsymbol{\psi}-\tilde{\boldsymbol{\psi}}_h))\|_{0,\Omega}&=\|\nabla\cdot(\nabla z-\tilde{\boldsymbol{\psi}}_h^g)\|_{0,\Omega}\lesssim h/p_v\,\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}\,,\\\|(\boldsymbol{\psi}-\tilde{\boldsymbol{\psi}}_h)\cdot\boldsymbol{n}\|_{0,\Gamma}&\leq\|(\nabla\times\boldsymbol{\rho}-\boldsymbol{\Pi}_h^{\mathrm{curl},\Gamma}\nabla\times\boldsymbol{\rho})\cdot\boldsymbol{n}\|_{0,\Gamma}+\|(\nabla z-\tilde{\boldsymbol{\psi}}_h^g)\cdot\boldsymbol{n}\|_{0,\Gamma}\lesssim\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}\,.\end{aligned}$$

Furthermore, there exists $\tilde{v}_h\in S_{p_s}(\mathcal{T}_h)$ such that $\|v-\tilde{v}_h\|_{1,\Omega}\lesssim h/p_s\,\|v\|_{2,\Omega}\lesssim h/p_s\,\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}$. Finally, we combine the estimates (5.18) and (5.19) to find

$$\begin{aligned}\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}^2\lesssim&(h/p)^{1/2}\|(\boldsymbol{e}^{\boldsymbol{\varphi}},e^u)\|_b\,\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}+(h/p)^{1/2}\cdot|||\boldsymbol{e}^{\boldsymbol{\varphi}}|||\,\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}\\&+\left[|||\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h|||+h/p\,\|\nabla\cdot(\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega}+h/p\|(\boldsymbol{e}^{\boldsymbol{\varphi}},e^u)\|_b\right]\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}\,.\end{aligned}$$

Canceling one power of $\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}$, estimating $|||\boldsymbol{e}^{\boldsymbol{\varphi}}|||$ by $\|(\boldsymbol{e}^{\boldsymbol{\varphi}},e^u)\|_b$ and summarizing the terms we find

$$\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}\lesssim(h/p)^{1/2}\|(\boldsymbol{e}^{\boldsymbol{\varphi}},e^u)\|_b+|||\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h|||+h/p\,\|\nabla\cdot(\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega}\,.$$

The result follows by using the fact that the FOSLS approximation is the projection with respect to the b scalar product, using the norm equivalence given in Theorem 5.1.1 and collecting the terms. □

Remark 5.3.7. Theorem 5.3.6 seems suboptimal in the following sense: Given $f \in L^2(\Omega)$ and $g \in H^{1/2}(\Gamma)$ the shift theorem gives $u \in H^2(\Omega)$ and consequently $\boldsymbol{\varphi} \in \boldsymbol{H}^1(\Omega)$. Theorem 5.3.6 gives

$$\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega} \lesssim h^{3/2}\,\|u\|_{2,\Omega} + h\,\|\boldsymbol{\varphi}\|_{1,\Omega} + h^{1/2}\|\boldsymbol{\varphi}\cdot\boldsymbol{n}\|_{1/2,\Gamma} + h^{1/2}\,\|\nabla\cdot\boldsymbol{\varphi}\|_{0,\Omega} \lesssim h^{1/2}(\|f\|_{0,\Omega} + \|g\|_{1/2,\Gamma}),$$

whereas from a best approximation viewpoint we could hope for $\mathcal{O}(h)$. ■

Lemma 5.3.8 (Convergence of dual solution for ∇e^u - Robin version of Lemma 4.3.9)**.** *Let Γ be smooth and $(\boldsymbol{\varphi}_h, u_h)$ be the least squares approximation of $(\boldsymbol{\varphi}, u)$. Let $e^u = u - u_h$ and $\boldsymbol{e}^{\boldsymbol{\varphi}} = \boldsymbol{\varphi} - \boldsymbol{\varphi}_h$. Let $(\boldsymbol{\psi}, v) \in \boldsymbol{V} \times W$ be the solution of the dual problem given by Theorem 5.2.2 with $w = e^u$. Furthermore, let $(\boldsymbol{\psi}_h, v_h)$ be the least squares approximation of $(\boldsymbol{\psi}, v)$ and denote $e^v = v - v_h$ and $\boldsymbol{e}^{\boldsymbol{\psi}} = \boldsymbol{\psi} - \boldsymbol{\psi}_h$. Then,*

$$\begin{aligned}
\|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b &\lesssim \|\nabla e^u\|_{0,\Omega}\,,\\
\|e^v\|_{0,\Omega} &\lesssim \frac{h}{p}\,\|\nabla e^u\|_{0,\Omega}\,,\\
\|e^v\|_{0,\Gamma} &\lesssim \left(\frac{h}{p}\right)^{1/2}\|\nabla e^u\|_{0,\Omega}\,,\\
\|\boldsymbol{e}^{\boldsymbol{\psi}}\|_{0,\Omega} &\lesssim \left(\frac{h}{p}\right)^{1/2}\|\nabla e^u\|_{0,\Omega}\,.
\end{aligned}$$

Proof. Theorem 5.2.2 gives $\boldsymbol{\psi} \in \boldsymbol{H}^1(\Omega)$, $\nabla\cdot\boldsymbol{\psi} \in H^1(\Omega)$ and $v \in H^1(\Omega)$ and exploiting the regularity estimates therein we find

$$\|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b \lesssim \|\nabla e^u\|_{0,\Omega}\,.$$

By Lemma 5.3.2 we have

$$\|e^v\|_{0,\Omega} \lesssim h/p\|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b,$$

which together with the above gives the second estimate. The third one follows by a multiplicative trace inequality together with the second estimate and the norm equivalence theorem in conjunction with the first estimate of the present lemma:

$$\|e^v\|_{0,\Gamma} \lesssim \|e^v\|_{0,\Omega}^{1/2}\,\|e^v\|_{1,\Omega}^{1/2} \lesssim (h/p)^{1/2}\|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b \lesssim (h/p)^{1/2}\,\|\nabla e^u\|_{0,\Omega}\,.$$

By Theorem 5.3.6 we have

$$\|\boldsymbol{e}^{\boldsymbol{\psi}}\|_{0,\Omega} \lesssim \left(\frac{h}{p}\right)^{1/2}\|v - \tilde{v}_h\|_{1,\Omega} + \|\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h\|_{0,\Omega} + \|(\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h)\cdot\boldsymbol{n}\|_{0,\Gamma} + \left(\frac{h}{p}\right)^{1/2}\|\nabla\cdot(\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h)\|_{0,\Omega}$$

for any $\tilde{v}_h \in S_{p_s}(\mathcal{T}_h)$, $\tilde{\boldsymbol{\psi}}_h \in \mathbf{V}_{p_v}(\mathcal{T}_h)$. The result follows immediately by again exploiting the regularity of the dual solution and the approximation properties of the employed spaces. □

Theorem 5.3.9 (Suboptimal estimate for $\|\nabla e^u\|_{0,\Omega}$ - Robin version of Theorem 4.3.10)**.** *Let Γ be smooth and $(\boldsymbol{\varphi}_h, u_h)$ be the least squares approximation of $(\boldsymbol{\varphi}, u)$. Furthermore, let $e^u = u - u_h$. Then, for any $\tilde{\boldsymbol{\varphi}}_h \in \mathbf{V}_{p_v}(\mathcal{T}_h)$, $\tilde{u}_h \in S_{p_s}(\mathcal{T}_h)$,*

$$\|\nabla e^u\|_{0,\Omega} \lesssim \|u - \tilde{u}_h\|_{1,\Omega} + \|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h\|_{0,\Omega} + \|(\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h)\cdot\boldsymbol{n}\|_{0,\Gamma} + \frac{h}{p}\|\nabla\cdot(\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega}.$$

Proof. We proceed as in Theorem 4.3.10 with $(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)$ denoting the FOSLS approximation of the dual solution given by Theorem 5.2.2 (duality argument for the gradient of the scalar variable) applied to $w = e^u$ we have for any $\tilde{\boldsymbol{\varphi}}_h \in \mathbf{V}_{p_v}(\mathcal{T}_h)$, $\tilde{u}_h \in S_{p_s}(\mathcal{T}_h)$

$$\|\nabla e^u\|_{0,\Omega}^2 = b((\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h, u - \tilde{u}_h), (\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)).$$

We specifically choose $\tilde{\boldsymbol{\varphi}}_h = \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}$. In the following we heavily use the properties of the operator $\boldsymbol{I}_h^\Gamma$ given in Lemma 5.3.5. We exploit the regularity of the dual solution using Lemma 5.3.8 as well as the estimates of Theorem 5.2.2:

$$\begin{aligned}
(\gamma(u - \tilde{u}_h), \nabla \cdot \boldsymbol{e}^{\boldsymbol{\psi}} + \gamma e^v)_\Omega &\lesssim \|u - \tilde{u}_h\|_{0,\Omega} \, \|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b \\
&\lesssim \|u - \tilde{u}_h\|_{1,\Omega} \, \|\nabla e^u\|_{0,\Omega} \,, \\
(\nabla(u - \tilde{u}_h), \nabla e^v + \boldsymbol{e}^{\boldsymbol{\psi}})_\Omega &\lesssim \|\nabla(u - \tilde{u}_h)\|_{0,\Omega} \, \|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b \\
&\lesssim \|u - \tilde{u}_h\|_{1,\Omega} \, \|\nabla e^u\|_{0,\Omega} \,, \\
\langle -\alpha(u - \tilde{u}_h), \boldsymbol{e}^{\boldsymbol{\psi}} \cdot \boldsymbol{n} - \alpha e^v \rangle_\Gamma &\lesssim \|u - \tilde{u}_h\|_{0,\Gamma} \, \|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b \\
&\lesssim \|u - \tilde{u}_h\|_{1,\Omega} \, \|\nabla e^u\|_{0,\Omega} \,, \\
(\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}, \nabla e^v)_\Omega &= -(\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}), e^v)_\Omega + \langle (\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}) \cdot \boldsymbol{n}, e^v \rangle_\Gamma \\
&\leq \|\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi})\|_{0,\Omega} \, \|e^v\|_{0,\Omega} + \|(\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}) \cdot \boldsymbol{n}\|_{0,\Gamma} \, \|e^v\|_{0,\Gamma} \\
&\lesssim \left[h/p \|\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi})\|_{0,\Omega} + (h/p)^{1/2} |||\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}||| \right] \|\nabla e^u\|_{0,\Omega} \,, \\
(\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}), \gamma e^v)_\Omega &\leq \|\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi})\|_{0,\Omega} \, \|e^v\|_{0,\Omega} \\
&\lesssim h/p \|\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi})\|_{0,\Omega} \, \|\nabla e^u\|_{0,\Omega} \,, \\
\langle (\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}) \cdot \boldsymbol{n}, -\alpha e^v \rangle_\Gamma &\leq \|(\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}) \cdot \boldsymbol{n}\|_{0,\Gamma} \, \|e^v\|_{0,\Gamma} \\
&\lesssim (h/p)^{1/2} |||\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}||| \, \|\nabla e^u\|_{0,\Omega} \,, \\
(\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}, \boldsymbol{e}^{\boldsymbol{\psi}})_\Omega &\lesssim \|\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}\|_{0,\Omega} \|\boldsymbol{e}^{\boldsymbol{\psi}}\|_{0,\Omega} \\
&\lesssim (h/p)^{1/2} |||\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}||| \, \|\nabla e^u\|_{0,\Omega} \,, \\
(\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}), \nabla \cdot \boldsymbol{e}^{\boldsymbol{\psi}})_\Omega &= (\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}), \nabla \cdot (\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h))_\Omega \\
&\leq \|\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi})\|_{0,\Omega} \|\nabla \cdot (\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h)\|_{0,\Omega} \\
&\lesssim h/p \|\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi})\|_{0,\Omega} \, \|\nabla e^u\|_{0,\Omega} \,, \\
\langle (\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}) \cdot \boldsymbol{n}, \boldsymbol{e}^{\boldsymbol{\psi}} \cdot \boldsymbol{n} \rangle_\Gamma &\leq \|(\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}) \cdot \boldsymbol{n}\|_{0,\Gamma} \|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b \\
&\lesssim |||\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}||| \, \|\nabla e^u\|_{0,\Omega} \,.
\end{aligned}$$

Canceling one power of $\|\nabla e^u\|_{0,\Omega}$ and collecting the terms yields

$$\|\nabla e^u\|_{0,\Omega} \lesssim \|u - \tilde{u}_h\|_{1,\Omega} + |||\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}||| + \frac{h}{p} \|\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi})\|_{0,\Omega}.$$

Finally, exploiting the estimates of the operator $\boldsymbol{I}_h^\Gamma$ given in Lemma 5.3.5 we arrive at the asserted estimate. □

Remark 5.3.10. Theorem 5.3.9 seems again suboptimal: Given $f \in L^2(\Omega)$ and $g \in H^{1/2}(\Gamma)$ the shift theorem gives $u \in H^2(\Omega)$ and consequently $\boldsymbol{\varphi} \in \boldsymbol{H}^1(\Omega)$. Theorem 5.3.9 gives

$$\|\nabla e^u\|_{0,\Omega} \lesssim h\,\|u\|_{2,\Omega} + h\,\|\boldsymbol{\varphi}\|_{1,\Omega} + h^{1/2}\|\boldsymbol{\varphi}\cdot\boldsymbol{n}\|_{H^{1/2}(\Gamma)} + h\,\|\nabla\cdot\boldsymbol{\varphi}\|_{0,\Omega} \lesssim h^{1/2}(\|f\|_{0,\Omega} + \|g\|_{1/2,\Gamma}),$$

whereas from a best approximation viewpoint we could hope for $\mathcal{O}(h)$. ■

Theorem 5.3.11 (Optimal estimate for $\|\boldsymbol{e}^{\boldsymbol{\varphi}}\cdot\boldsymbol{n}\|_{0,\Gamma}$)**.** *Let Γ be smooth and $(\boldsymbol{\varphi}_h, u_h)$ be the least squares approximation of $(\boldsymbol{\varphi}, u)$. Furthermore, let $e^u = u - u_h$ and $\boldsymbol{e}^{\boldsymbol{\varphi}} = \boldsymbol{\varphi} - \boldsymbol{\varphi}_h$. Then, for any $\tilde{u}_h \in S_{p_s}(\mathcal{T}_h)$, $\tilde{\boldsymbol{\varphi}}_h \in \mathbf{V}_{p_v}(\mathcal{T}_h)$,*

$$\begin{aligned}\|\boldsymbol{e}^{\boldsymbol{\varphi}}\cdot\boldsymbol{n}\|_{0,\Gamma} \lesssim \left(\frac{h}{p}\right)^{1/2}\|u-\tilde{u}_h\|_{1,\Omega} + \left(\frac{h}{p}\right)^{1/2}\|\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h\|_{0,\Omega} \\ + \|(\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h)\cdot\boldsymbol{n}\|_{0,\Gamma} + \frac{h}{p}\|\nabla\cdot(\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega}\,.\end{aligned}$$

Proof. Let $(\boldsymbol{\psi}, v) \in \boldsymbol{V} \times W$ denote the dual solution given by Theorem 5.2.4 applied to $\boldsymbol{\eta} = \boldsymbol{e}^{\boldsymbol{\varphi}}$. Theorem 5.2.4 gives $\boldsymbol{\psi} \in \boldsymbol{H}^{1/2}(\Omega)$, $\nabla\cdot\boldsymbol{\psi} \in H^{3/2}(\Omega)$, $\boldsymbol{\psi}\cdot\boldsymbol{n} \in L^2(\Gamma)$ and $v \in H^{3/2}(\Omega)$. For the analysis we employ the operator $\boldsymbol{\Pi}^{\mathrm{div}}_{p_v}$ from [MR20] and summarized in Proposition 5.3.1. The main features exploited in the proof, are that $\boldsymbol{\Pi}^{\mathrm{div}}_{p_v}$ realizes the L^2 orthogonal projections of the divergence as well as the normal trace. Due to the Galerkin orthogonality we have for any $(\tilde{\boldsymbol{\psi}}_h, \tilde{v}_h)$

$$\|\boldsymbol{e}^{\boldsymbol{\varphi}}\cdot\boldsymbol{n}\|^2_{0,\Gamma} = b((\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u), (\boldsymbol{\psi}, v)) = b((\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u), (\boldsymbol{\psi}-\tilde{\boldsymbol{\psi}}_h, v-\tilde{v}_h)).$$

Choosing $\tilde{\boldsymbol{\psi}}_h = \boldsymbol{\Pi}^{\mathrm{div}}_{p_v}\boldsymbol{\psi}$, exploiting norm equivalence, the orthogonality properties of $\boldsymbol{\Pi}^{\mathrm{div}}_{p_v}$ and the Cauchy-Schwarz inequality we find

$$\begin{aligned}(\nabla\cdot\boldsymbol{e}^{\boldsymbol{\varphi}} + \gamma e^u, \nabla\cdot(\boldsymbol{\psi}-\boldsymbol{\Pi}^{\mathrm{div}}_{p_v}\boldsymbol{\psi}) + \gamma(v-\tilde{v}_h))_\Omega &\lesssim \|(\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u)\|_b\Big[\|\nabla\cdot(\boldsymbol{\psi}-\boldsymbol{\Pi}^{\mathrm{div}}_{p_v}\boldsymbol{\psi})\|_{0,\Omega} \\ &\quad + \|v-\tilde{v}_h\|_{0,\Omega}\Big],\\ (\nabla e^u + \boldsymbol{e}^{\boldsymbol{\varphi}}, \nabla(v-\tilde{v}_h) + \boldsymbol{\psi}-\boldsymbol{\Pi}^{\mathrm{div}}_{p_v}\boldsymbol{\psi})_\Omega &\lesssim [\|\nabla e^u\|_{0,\Omega} + \|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}]\Big[\|v-\tilde{v}_h\|_{1,\Omega} \\ &\quad + \|\boldsymbol{\psi}-\boldsymbol{\Pi}^{\mathrm{div}}_{p_v}\boldsymbol{\psi}\|_{0,\Omega}\Big],\\ \langle -\alpha e^u, -\alpha(v-\tilde{v}_h)\rangle_\Gamma &\lesssim \|(\boldsymbol{e}^{\boldsymbol{\varphi}}, e^u)\|_b\|v-\tilde{v}_h\|_{0,\Gamma},\\ \langle \boldsymbol{e}^{\boldsymbol{\varphi}}\cdot\boldsymbol{n}, (\boldsymbol{\psi}-\tilde{\boldsymbol{\psi}}_h)\cdot\boldsymbol{n}\rangle_\Gamma &= \langle(\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h)\cdot\boldsymbol{n}, (\boldsymbol{\psi}-\boldsymbol{\Pi}^{\mathrm{div}}_{p_v}\boldsymbol{\psi})\cdot\boldsymbol{n}\rangle_\Gamma \\ &\lesssim \|(\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h)\cdot\boldsymbol{n}\|_{0,\Gamma}\|(\boldsymbol{\psi}-\boldsymbol{\Pi}^{\mathrm{div}}_{p_v}\boldsymbol{\psi})\cdot\boldsymbol{n}\|_{0,\Gamma}.\end{aligned}$$

The two missing boundary terms, i.e., $\langle \boldsymbol{e}^{\boldsymbol{\varphi}}\cdot\boldsymbol{n}, -\alpha(v-\tilde{v}_h)\rangle_\Gamma$ and $\langle -\alpha e^u, (\boldsymbol{\psi}-\tilde{\boldsymbol{\psi}}_h)\cdot\boldsymbol{n}\rangle_\Gamma$, can be absorbed into the first two estimates by means of partial integration. We now exploit the regularity estimates given in Theorem 5.2.4, the properties of $\boldsymbol{\Pi}^{\mathrm{div}}_{p_v}$ given in Proposition 5.3.1 as well as the approximation properties of the employed spaces to find $\tilde{v}_h$

such that

$$\|\nabla\cdot(\boldsymbol{\psi}-\boldsymbol{\Pi}^{\mathrm{div}}_{p_v}\boldsymbol{\psi})\|_{0,\Omega}\lesssim h/p_v\|\nabla\cdot\boldsymbol{\psi}\|_{1,\Omega}\lesssim h/p_v\,\|\boldsymbol{e}^{\boldsymbol{\varphi}}\cdot\boldsymbol{n}\|_{0,\Gamma}\,,$$
$$\|\boldsymbol{\psi}-\boldsymbol{\Pi}^{\mathrm{div}}_{p_v}\boldsymbol{\psi}\|_{0,\Omega}\lesssim (h/p_v)^{1/2}\|\boldsymbol{\psi}\|_{\boldsymbol{H}^{1/2}(\Omega,\mathrm{div})}\lesssim (h/p_v)^{1/2}\,\|\boldsymbol{e}^{\boldsymbol{\varphi}}\cdot\boldsymbol{n}\|_{0,\Gamma}\,,$$
$$\|(\boldsymbol{\psi}-\boldsymbol{\Pi}^{\mathrm{div}}_{p_v}\boldsymbol{\psi})\cdot\boldsymbol{n}\|_{0,\Gamma}\lesssim\|\boldsymbol{\psi}\cdot\boldsymbol{n}\|_{0,\Gamma}\lesssim\|\boldsymbol{e}^{\boldsymbol{\varphi}}\cdot\boldsymbol{n}\|_{0,\Gamma}\,,$$
$$\|v-\tilde v_h\|_{0,\Omega}\lesssim (h/p_s)^{3/2}\,\|v\|_{3/2,\Omega}\lesssim (h/p_s)^{3/2}\,\|\boldsymbol{e}^{\boldsymbol{\varphi}}\cdot\boldsymbol{n}\|_{0,\Gamma}\,,$$
$$\|v-\tilde v_h\|_{1,\Omega}\lesssim (h/p_s)^{1/2}\,\|v\|_{3/2,\Omega}\lesssim (h/p_s)^{1/2}\,\|\boldsymbol{e}^{\boldsymbol{\varphi}}\cdot\boldsymbol{n}\|_{0,\Gamma}\,,$$
$$\|v-\tilde v_h\|_{0,\Gamma}\lesssim h/p_s\,\|v\|_{3/2,\Omega}\lesssim h/p_s\,\|\boldsymbol{e}^{\boldsymbol{\varphi}}\cdot\boldsymbol{n}\|_{0,\Gamma}\,,$$

which in turn gives after summarizing and canceling one power of $\|\boldsymbol{e}^{\boldsymbol{\varphi}}\cdot\boldsymbol{n}\|_{0,\Gamma}$ the estimate

$$\|\boldsymbol{e}^{\boldsymbol{\varphi}}\cdot\boldsymbol{n}\|_{0,\Gamma}\lesssim h/p\|(\boldsymbol{e}^{\boldsymbol{\varphi}},e^u)\|_b+(h/p)^{1/2}\,[\|\nabla e^u\|_{0,\Omega}+\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}]+\|(\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h)\cdot\boldsymbol{n}\|_{0,\Gamma}.$$

Applying Theorems 5.3.6 and 5.3.9 to estimate $\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}$ and $\|\nabla e^u\|_{0,\Omega}$ yields the result. □

Remark 5.3.12. Theorem 5.3.11 seems optimal in the following sense: Given $f\in L^2(\Omega)$ and $g\in H^{1/2}(\Gamma)$ the shift theorem gives $u\in H^2(\Omega)$ and consequently $\boldsymbol{\varphi}\in\boldsymbol{H}^1(\Omega)$. Theorem 5.3.11 gives

$$\begin{aligned}\|\boldsymbol{e}^{\boldsymbol{\varphi}}\cdot\boldsymbol{n}\|_{0,\Gamma}&\lesssim h^{3/2}\,\|u\|_{2,\Omega}+h^{3/2}\,\|\boldsymbol{\varphi}\|_{1,\Omega}+h^{1/2}\|\boldsymbol{\varphi}\cdot\boldsymbol{n}\|_{1/2,\Gamma}+h\,\|\nabla\cdot\boldsymbol{\varphi}\|_{0,\Omega}\\&\lesssim h^{1/2}(\|f\|_{0,\Omega}+\|g\|_{1/2,\Gamma}),\end{aligned}$$

which is the rate expected from a best approximation argument. ■

We are in the position to derive an optimal estimate for $\|\nabla e^u\|_{0,\Omega}$ using the estimate given in Theorem 5.3.11.

Theorem 5.3.13 (Optimal estimate for $\|\nabla e^u\|_{0,\Omega}$ - Robin version of Theorem 4.3.10)**.** *Let Γ be smooth and $(\boldsymbol{\varphi}_h,u_h)$ be the least squares approximation of $(\boldsymbol{\varphi},u)$. Furthermore, let $e^u=u-u_h$. Then, for any $\tilde{\boldsymbol{\varphi}}_h\in\mathbf{V}_{p_v}(\mathcal{T}_h)$, $\tilde u_h\in S_{p_s}(\mathcal{T}_h)$,*

$$\|\nabla e^u\|_{0,\Omega}\lesssim\|u-\tilde u_h\|_{1,\Omega}+\left(\frac{h}{p}\right)^{1/2}\|\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h\|_{0,\Omega}+\left(\frac{h}{p}\right)^{1/2}\|(\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h)\cdot\boldsymbol{n}\|_{0,\Gamma}+\frac{h}{p}\|\nabla\cdot(\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega}.$$

Proof. Reentering the proof of Theorem 5.3.9 we therein estimated

$$\begin{aligned}\langle(\boldsymbol{\varphi}-\boldsymbol{I}^{\Gamma}_h\boldsymbol{\varphi})\cdot\boldsymbol{n},\boldsymbol{e}^{\boldsymbol{\psi}}\cdot\boldsymbol{n}\rangle_\Gamma&\le\|(\boldsymbol{\varphi}-\boldsymbol{I}^{\Gamma}_h\boldsymbol{\varphi})\cdot\boldsymbol{n}\|_{0,\Gamma}\|(\boldsymbol{e}^{\boldsymbol{\psi}},e^v)\|_b\\&\lesssim|||\boldsymbol{\varphi}-\boldsymbol{I}^{\Gamma}_h\boldsymbol{\varphi}|||\,\|\nabla e^u\|_{0,\Omega}\,.\end{aligned}$$

Theorem 5.3.11 however now gives together with the regularity of the dual solution the estimate

$$\begin{aligned}\|\boldsymbol{e}^{\boldsymbol{\psi}}\cdot\boldsymbol{n}\|_{0,\Gamma}&\lesssim\left(\frac{h}{p}\right)^{1/2}\|v-\tilde v_h\|_{1,\Omega}+\left(\frac{h}{p}\right)^{1/2}\|\boldsymbol{\psi}-\tilde{\boldsymbol{\psi}}_h\|_{0,\Omega}\\&\qquad+\|(\boldsymbol{\psi}-\tilde{\boldsymbol{\psi}}_h)\cdot\boldsymbol{n}\|_{0,\Gamma}+\frac{h}{p}\|\nabla\cdot(\boldsymbol{\psi}-\tilde{\boldsymbol{\psi}}_h)\|_{0,\Omega}\\&\lesssim\left(\frac{h}{p}\right)^{1/2}\|\nabla e^u\|_{0,\Omega}\,,\end{aligned}$$

which in turn enables us to improve the estimate as follows:

$$\langle(\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}) \cdot \boldsymbol{n}, \boldsymbol{e}^{\boldsymbol{\psi}} \cdot \boldsymbol{n}\rangle_\Gamma \lesssim (h/p)^{1/2} |||\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}||| \, \|\nabla e^u\|_{0,\Omega} .$$

All other estimates in the proof of 5.3.9 stay the same. Canceling one power of $\|\nabla e^u\|_{0,\Omega}$ and collecting the terms yields

$$\|\nabla e^u\|_{0,\Omega} \lesssim \|u - \tilde{u}_h\|_{1,\Omega} + \left(\frac{h}{p}\right)^{1/2} |||\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi}||| + \frac{h}{p} \|\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{I}_h^\Gamma \boldsymbol{\varphi})\|_{0,\Omega}.$$

Finally, exploiting the estimates of the operator $\boldsymbol{I}_h^\Gamma$ given in Lemma 5.3.5 we arrive at the asserted estimate. □

Before turning to the estimate for $\|e^u\|_{0,\Omega}$ we first derive a slightly better version of Theorem 5.3.6. To that end, we first analyze the convergence of the corresponding dual solution:

Lemma 5.3.14 (Convergence of dual solution for $\boldsymbol{e}^{\boldsymbol{\varphi}}$)**.** *Let Γ be smooth and $(\boldsymbol{\varphi}_h, u_h)$ be the least squares approximation of $(\boldsymbol{\varphi}, u)$. Let $e^u = u - u_h$ and $\boldsymbol{e}^{\boldsymbol{\varphi}} = \boldsymbol{\varphi} - \boldsymbol{\varphi}_h$. Let $(\boldsymbol{\psi}, v) \in \boldsymbol{V} \times W$ be the solution of the dual problem given by Theorem 5.2.3 with $\boldsymbol{\eta} = \boldsymbol{e}^{\boldsymbol{\varphi}}$. Furthermore, let $(\boldsymbol{\psi}_h, v_h)$ be the least squares approximation of $(\boldsymbol{\psi}, v)$ and denote $e^v = v - v_h$ and $\boldsymbol{e}^{\boldsymbol{\psi}} = \boldsymbol{\psi} - \boldsymbol{\psi}_h$. Then,*

$$\begin{aligned}
\|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b &\lesssim \|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega} ,\\
\|e^v\|_{0,\Omega} &\lesssim \frac{h}{p} \|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega} ,\\
\|e^v\|_{0,\Gamma} &\lesssim \left(\frac{h}{p}\right)^{1/2} \|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega} ,\\
\|\nabla e^v\|_{0,\Omega} &\lesssim \left(\frac{h}{p}\right)^{1/2} \|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega} ,\\
\|\boldsymbol{e}^{\boldsymbol{\psi}}\|_{0,\Omega} &\lesssim \|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega} ,\\
\|\boldsymbol{e}^{\boldsymbol{\psi}} \cdot \boldsymbol{n}\|_{0,\Gamma} &\lesssim \left(\frac{h}{p}\right)^{1/2} \|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega} .
\end{aligned}$$

Proof. Theorem 5.2.3 gives $\boldsymbol{\psi} \in \boldsymbol{L}^2(\Omega)$, $\nabla \cdot \boldsymbol{\psi} \in H^1(\Omega)$, $\boldsymbol{\psi} \cdot \boldsymbol{n} \in H^{1/2}(\Omega)$ and $v \in H^2(\Omega)$ and exploiting the regularity estimates therein we find

$$\|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b \lesssim \|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega} .$$

By Lemma 5.3.2 we have

$$\|e^v\|_{0,\Omega} \lesssim h/p \|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b,$$

which together with the above gives the second estimate. The third one follows by a multiplicative trace inequality together with the second estimate and the norm equivalence theorem in conjunction with the first estimate of the present lemma:

$$\|e^v\|_{0,\Gamma} \lesssim \|e^v\|_{0,\Omega}^{1/2} \|e^v\|_{1,\Omega}^{1/2} \lesssim (h/p)^{1/2} \|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b \lesssim (h/p)^{1/2} \|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega} .$$

The Theorems 5.3.13, 5.3.6 and 5.3.11 then yield the result by exploiting the regularity of the dual solution and the approximation properties of the employed spaces. □

Theorem 5.3.15 (Suboptimal but improved estimate for $\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}$ - Robin version of Theorem 4.3.8)**.** *Let Γ be smooth and $(\boldsymbol{\varphi}_h, u_h)$ be the least squares approximation of $(\boldsymbol{\varphi}, u)$. Furthermore, let $e^u = u - u_h$ and $\boldsymbol{e}^{\boldsymbol{\varphi}} = \boldsymbol{\varphi} - \boldsymbol{\varphi}_h$. Then, for any $\tilde{u}_h \in S_{p_s}(\mathcal{T}_h)$, $\tilde{\boldsymbol{\varphi}}_h \in \mathbf{V}_{p_v}(\mathcal{T}_h)$,*

$$\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega} \lesssim \|u - \tilde{u}_h\|_{0,\Omega} + \left(\frac{h}{p}\right)^{1/2} \|u - \tilde{u}_h\|_{1,\Omega} + \|\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h\|_{0,\Omega} + \|(\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h) \cdot \boldsymbol{n}\|_{0,\Gamma} + \frac{h}{p} \|\nabla \cdot (\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega} .$$

Proof. We proceed as in the proof of Theorem 5.3.9 with $(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)$ denoting the FOSLS approximation of the dual solution given by Theorem 5.2.3 (duality argument for the vector variable) applied to $\boldsymbol{\eta} = \boldsymbol{e}^{\boldsymbol{\varphi}}$. As before for any $\tilde{\boldsymbol{\varphi}}_h \in \mathbf{V}_{p_v}(\mathcal{T}_h)$, $\tilde{u}_h \in S_{p_s}(\mathcal{T}_h)$

$$\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|^2_{0,\Omega} = b((\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h, u - \tilde{u}_h), (\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)).$$

We again choose $\tilde{\boldsymbol{\varphi}}_h = \boldsymbol{I}_h^{\Gamma}\boldsymbol{\varphi}$, extensively use the properties of the operator $\boldsymbol{I}_h^{\Gamma}$ given in Lemma 5.3.5, exploit the regularity of the dual solution using Lemma 5.3.14 as well as the estimates of Theorem 5.2.3:

$$\begin{aligned}
(\gamma(u - \tilde{u}_h), \nabla \cdot \boldsymbol{e}^{\boldsymbol{\psi}})_\Omega &\lesssim \|u - \tilde{u}_h\|_{0,\Omega} \|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b \\
&\lesssim \|u - \tilde{u}_h\|_{0,\Omega} \|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega} , \\
\langle -\alpha(u - \tilde{u}_h), \boldsymbol{e}^{\boldsymbol{\psi}} \cdot \boldsymbol{n} - \alpha e^v \rangle_\Gamma &\lesssim \|u - \tilde{u}_h\|_{0,\Gamma} \left[\|\boldsymbol{e}^{\boldsymbol{\psi}} \cdot \boldsymbol{n}\|_{0,\Gamma} + \|e^v\|_{0,\Gamma} \right] \\
&\lesssim (h/p)^{1/2} \|u - \tilde{u}_h\|_{0,\Gamma} \|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega} , \\
(\nabla(u - \tilde{u}_h), \boldsymbol{e}^{\boldsymbol{\psi}})_\Omega &= -(u - \tilde{u}_h, \nabla \cdot \boldsymbol{e}^{\boldsymbol{\psi}})_\Omega + \langle u - \tilde{u}_h, \boldsymbol{e}^{\boldsymbol{\psi}} \cdot \boldsymbol{n} \rangle_\Gamma \\
&\lesssim \left[\|u - \tilde{u}_h\|_{0,\Omega} + (h/p)^{1/2} \|u - \tilde{u}_h\|_{0,\Gamma} \right] \|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega} , \\
(\gamma(u - \tilde{u}_h), \gamma e^v)_\Omega &\lesssim \|u - \tilde{u}_h\|_{0,\Omega} \|e^v\|_{0,\Omega} \\
&\lesssim h/p \|u - \tilde{u}_h\|_{1,\Omega} \|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega} , \\
(\nabla(u - \tilde{u}_h), \nabla e^v)_\Omega &\lesssim \|\nabla(u - \tilde{u}_h)\|_{0,\Omega} \|\nabla e^v\|_{0,\Omega} \\
&\lesssim (h/p)^{1/2} \|u - \tilde{u}_h\|_{1,\Omega} \|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega} , \\
(\boldsymbol{\varphi} - \boldsymbol{I}_h^{\Gamma}\boldsymbol{\varphi}, \nabla e^v + \boldsymbol{e}^{\boldsymbol{\psi}})_\Omega &\le \|\boldsymbol{\varphi} - \boldsymbol{I}_h^{\Gamma}\boldsymbol{\varphi}\|_{0,\Omega} \|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b \\
&\lesssim |||\boldsymbol{\varphi} - \boldsymbol{I}_h^{\Gamma}\boldsymbol{\varphi}||| \, \|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega} , \\
\langle (\boldsymbol{\varphi} - \boldsymbol{I}_h^{\Gamma}\boldsymbol{\varphi}) \cdot \boldsymbol{n}, \boldsymbol{e}^{\boldsymbol{\psi}} \cdot \boldsymbol{n} - \alpha e^v \rangle_\Gamma &\le \|(\boldsymbol{\varphi} - \boldsymbol{I}_h^{\Gamma}\boldsymbol{\varphi}) \cdot \boldsymbol{n}\|_{0,\Gamma} \|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b \\
&\lesssim |||\boldsymbol{\varphi} - \boldsymbol{I}_h^{\Gamma}\boldsymbol{\varphi}||| \, \|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega} , \\
(\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{I}_h^{\Gamma}\boldsymbol{\varphi}), \nabla \cdot \boldsymbol{e}^{\boldsymbol{\psi}})_\Omega &= (\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{I}_h^{\Gamma}\boldsymbol{\varphi}), \nabla \cdot (\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h))_\Omega \\
&\le \|\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{I}_h^{\Gamma}\boldsymbol{\varphi})\|_{0,\Omega} \|\nabla \cdot (\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h)\|_{0,\Omega} \\
&\lesssim h/p \|\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{I}_h^{\Gamma}\boldsymbol{\varphi})\|_{0,\Omega} \|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega} .
\end{aligned}$$

Canceling one power of $\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}$ and summarizing the estimates we find

$$\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega} \lesssim \|u - \tilde{u}_h\|_{0,\Omega} + (h/p)^{1/2} \|u - \tilde{u}_h\|_{0,\Gamma} + (h/p)^{1/2} \|u - \tilde{u}_h\|_{1,\Omega} + |||\boldsymbol{\varphi} - \boldsymbol{I}_h^{\Gamma}\boldsymbol{\varphi}||| + h/p \|\nabla \cdot (\boldsymbol{\varphi} - \boldsymbol{I}_h^{\Gamma}\boldsymbol{\varphi})\|_{0,\Omega}.$$

A trace estimate, using the estimates of the operator $\boldsymbol{I}_h^\Gamma$ given in Lemma 5.3.5 and collecting the terms yields the result. $\square$

Lemma 5.3.16 (Convergence of dual solution for e^u - Robin version of Lemma 4.3.11)**.** *Let Γ be smooth and $(\boldsymbol{\varphi}_h, u_h)$ be the least squares approximation of $(\boldsymbol{\varphi}, u)$. Let $e^u = u - u_h$ and $\boldsymbol{e}^{\boldsymbol{\varphi}} = \boldsymbol{\varphi} - \boldsymbol{\varphi}_h$. Let $(\boldsymbol{\psi}, v) \in \boldsymbol{V} \times W$ be the solution of the dual problem given by Theorem 5.2.1 with $w = e^u$. Furthermore, let $(\boldsymbol{\psi}_h, v_h)$ be the least squares approximation of $(\boldsymbol{\psi}, v)$ and denote $e^v = v - v_h$ and $\boldsymbol{e}^{\boldsymbol{\psi}} = \boldsymbol{\psi} - \boldsymbol{\psi}_h$. Then,*

$$\begin{aligned}
\|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b &\lesssim \frac{h}{p} \|e^u\|_{0,\Omega}\,, \\
\|e^v\|_{0,\Omega} &\lesssim \left(\frac{h}{p}\right)^2 \|e^u\|_{0,\Omega}\,, \\
\|e^v\|_{0,\Gamma} &\lesssim \left(\frac{h}{p}\right)^{3/2} \|e^u\|_{0,\Omega}\,, \\
\|\boldsymbol{e}^{\boldsymbol{\psi}}\|_{0,\Omega} &\lesssim \begin{cases} h\,\|e^u\|_{0,\Omega} & \mathbf{V}_{p_v}(\mathcal{T}_h) = \mathbf{RT}_0(\mathcal{T}_h), \\ \left(\frac{h}{p}\right)^{3/2} \|e^u\|_{0,\Omega} & \textit{else}, \end{cases} \\
\|\boldsymbol{e}^{\boldsymbol{\psi}} \cdot \boldsymbol{n}\|_{0,\Gamma} &\lesssim \begin{cases} h\,\|e^u\|_{0,\Omega} & \mathbf{V}_{p_v}(\mathcal{T}_h) = \mathbf{RT}_0(\mathcal{T}_h), \\ \left(\frac{h}{p}\right)^{3/2} \|e^u\|_{0,\Omega} & \textit{else.} \end{cases}
\end{aligned}$$

Proof. Theorem 5.2.1 gives $\boldsymbol{\psi} \in \boldsymbol{H}^2(\Omega)$, $\nabla \cdot \boldsymbol{\psi} \in H^2(\Omega)$ and $v \in H^2(\Omega)$ and exploiting the regularity estimates therein we find

$$\|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b \lesssim \frac{h}{p} \|e^u\|_{0,\Omega}\,.$$

By Lemma 5.3.2 we have

$$\|e^v\|_{0,\Omega} \lesssim h/p \|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b,$$

which together with the above gives the second estimate. The third one follows by a multiplicative trace inequality together with the second estimate and the norm equivalence theorem in conjunction with the first estimate of the present lemma:

$$\|e^v\|_{0,\Gamma} \lesssim \|e^v\|_{0,\Omega}^{1/2} \|e^v\|_{1,\Omega}^{1/2} \lesssim (h/p)^{3/2} \|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b \lesssim (h/p)^{3/2} \|e^u\|_{0,\Omega}\,.$$

The Theorems 5.3.6 and 5.3.13 then yield the result by exploiting the regularity of the dual solution and the approximation properties of the employed spaces. $\square$

Theorem 5.3.17 (Optimal estimate for $\|e^u\|_{0,\Omega}$ - Robin version of Theorem 4.3.12)**.** *Let Γ be smooth and $(\boldsymbol{\varphi}_h, u_h)$ be the least squares approximation of $(\boldsymbol{\varphi}, u)$. Furthermore, let*

$e^u = u - u_h$. *Then, for any* $\tilde{\boldsymbol{\varphi}}_h \in \mathbf{V}_{p_v}(\mathcal{T}_h)$, $\tilde{u}_h \in S_{p_s}(\mathcal{T}_h)$,

$$\|e^u\|_{0,\Omega} \lesssim \begin{cases} h\,\|u-\tilde{u}_h\|_{1,\Omega} + h\|\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h\|_{0,\Omega} & \\ +h\|(\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h)\cdot\boldsymbol{n}\|_{0,\Gamma} + h\|\nabla\cdot(\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega} & \mathbf{V}^0_{p_v}(\mathcal{T}_h) = \mathbf{RT}_0(\mathcal{T}_h), \\ h\,\|u-\tilde{u}_h\|_{1,\Omega} + h^{3/2}\|\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h\|_{0,\Omega} & \\ +h^{3/2}\|(\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h)\cdot\boldsymbol{n}\|_{0,\Gamma} + h\|\nabla\cdot(\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega} & \mathbf{V}^0_{p_v}(\mathcal{T}_h) = \mathbf{BDM}_1(\mathcal{T}_h), \\ \frac{h}{p}\,\|u-\tilde{u}_h\|_{1,\Omega} + \left(\frac{h}{p}\right)^{3/2}\|\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h\|_{0,\Omega} & \\ +\left(\frac{h}{p}\right)^{3/2}\|(\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h)\cdot\boldsymbol{n}\|_{0,\Gamma} + \left(\frac{h}{p}\right)^{2}\|\nabla\cdot(\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h)\|_{0,\Omega} & \textit{else.} \end{cases}$$

Proof. We proceed as in the proof of Theorem 5.3.9 with $(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)$ denoting the FOSLS approximation of the dual solution given by Theorem 5.2.1 (duality argument for the scalar variable) applied to $w = e^u$. As before for any $\tilde{\boldsymbol{\varphi}}_h \in \mathbf{V}_{p_v}(\mathcal{T}_h)$, $\tilde{u}_h \in S_{p_s}(\mathcal{T}_h)$

$$\|e^u\|^2_{0,\Omega} = b((\boldsymbol{\varphi}-\tilde{\boldsymbol{\varphi}}_h, u-\tilde{u}_h), (\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)).$$

We again choose $\tilde{\boldsymbol{\varphi}}_h = \boldsymbol{I}^\Gamma_h\boldsymbol{\varphi}$, extensively use the properties of the operator $\boldsymbol{I}^\Gamma_h$ given in Lemma 5.3.5, exploit the regularity of the dual solution using Lemma 5.3.16 as well as the estimates of Theorem 5.2.1:

$$\begin{aligned}
(\gamma(u-\tilde{u}_h), \nabla\cdot\boldsymbol{e}^{\boldsymbol{\psi}} + \gamma e^v)_\Omega &\lesssim \|u-\tilde{u}_h\|_{0,\Omega}\,\|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b \\
&\lesssim h/p\,\|u-\tilde{u}_h\|_{1,\Omega}\,\|e^u\|_{0,\Omega}\,, \\
(\nabla(u-\tilde{u}_h), \nabla e^v + \boldsymbol{e}^{\boldsymbol{\psi}})_\Omega &\lesssim \|\nabla(u-\tilde{u}_h)\|_{0,\Omega}\,\|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b \\
&\lesssim h/p\,\|u-\tilde{u}_h\|_{1,\Omega}\,\|e^u\|_{0,\Omega}\,, \\
\langle -\alpha(u-\tilde{u}_h), \boldsymbol{e}^{\boldsymbol{\psi}}\cdot\boldsymbol{n} - \alpha e^v\rangle_\Gamma &\lesssim \|u-\tilde{u}_h\|_{0,\Gamma}\,\|(\boldsymbol{e}^{\boldsymbol{\psi}}, e^v)\|_b \\
&\lesssim h/p\,\|u-\tilde{u}_h\|_{1,\Omega}\,\|e^u\|_{0,\Omega}\,, \\
(\boldsymbol{\varphi}-\boldsymbol{I}^\Gamma_h\boldsymbol{\varphi}, \nabla e^v)_\Omega &= -(\nabla\cdot(\boldsymbol{\varphi}-\boldsymbol{I}^\Gamma_h\boldsymbol{\varphi}), e^v)_\Omega + \langle(\boldsymbol{\varphi}-\boldsymbol{I}^\Gamma_h\boldsymbol{\varphi})\cdot\boldsymbol{n}, e^v\rangle_\Gamma \\
&\le \|\nabla\cdot(\boldsymbol{\varphi}-\boldsymbol{I}^\Gamma_h\boldsymbol{\varphi})\|_{0,\Omega}\,\|e^v\|_{0,\Omega} + \|(\boldsymbol{\varphi}-\boldsymbol{I}^\Gamma_h\boldsymbol{\varphi})\cdot\boldsymbol{n}\|_{0,\Gamma}\,\|e^v\|_{0,\Gamma} \\
&\lesssim \left[(h/p)^2\|\nabla\cdot(\boldsymbol{\varphi}-\boldsymbol{I}^\Gamma_h\boldsymbol{\varphi})\|_{0,\Omega} + (h/p)^{3/2}|||\boldsymbol{\varphi}-\boldsymbol{I}^\Gamma_h\boldsymbol{\varphi}|||\right]\|e^u\|_{0,\Omega}\,, \\
(\nabla\cdot(\boldsymbol{\varphi}-\boldsymbol{I}^\Gamma_h\boldsymbol{\varphi}), \gamma e^v)_\Omega &\le \|\nabla\cdot(\boldsymbol{\varphi}-\boldsymbol{I}^\Gamma_h\boldsymbol{\varphi})\|_{0,\Omega}\,\|e^v\|_{0,\Omega} \\
&\lesssim (h/p)^2\|\nabla\cdot(\boldsymbol{\varphi}-\boldsymbol{I}^\Gamma_h\boldsymbol{\varphi})\|_{0,\Omega}\,\|e^u\|_{0,\Omega}\,, \\
\langle(\boldsymbol{\varphi}-\boldsymbol{I}^\Gamma_h\boldsymbol{\varphi})\cdot\boldsymbol{n}, -\alpha e^v\rangle_\Gamma &\le \|(\boldsymbol{\varphi}-\boldsymbol{I}^\Gamma_h\boldsymbol{\varphi})\cdot\boldsymbol{n}\|_{0,\Gamma}\,\|e^v\|_{0,\Gamma} \\
&\lesssim (h/p)^{3/2}|||\boldsymbol{\varphi}-\boldsymbol{I}^\Gamma_h\boldsymbol{\varphi}|||\,\|e^u\|_{0,\Omega}\,, \\
(\boldsymbol{\varphi}-\boldsymbol{I}^\Gamma_h\boldsymbol{\varphi}, \boldsymbol{e}^{\boldsymbol{\psi}})_\Omega &\lesssim \|\boldsymbol{\varphi}-\boldsymbol{I}^\Gamma_h\boldsymbol{\varphi}\|_{0,\Omega}\|\boldsymbol{e}^{\boldsymbol{\psi}}\|_{0,\Omega} \\
&\lesssim \begin{cases} h|||\boldsymbol{\varphi}-\boldsymbol{I}^\Gamma_h\boldsymbol{\varphi}|||\,\|e^u\|_{0,\Omega} & \mathbf{V}_{p_v}(\mathcal{T}_h) = \mathbf{RT}_0(\mathcal{T}_h), \\ \left(\frac{h}{p}\right)^{3/2}|||\boldsymbol{\varphi}-\boldsymbol{I}^\Gamma_h\boldsymbol{\varphi}|||\,\|e^u\|_{0,\Omega} & \textit{else,} \end{cases}
\end{aligned}$$

$$
\begin{aligned}
(\nabla\cdot(\boldsymbol{\varphi}-\boldsymbol{I}_h^\Gamma\boldsymbol{\varphi}),\nabla\cdot\boldsymbol{e}^{\boldsymbol{\psi}})_\Omega &= (\nabla\cdot(\boldsymbol{\varphi}-\boldsymbol{I}_h^\Gamma\boldsymbol{\varphi}),\nabla\cdot(\boldsymbol{\psi}-\tilde{\boldsymbol{\psi}}_h))_\Omega \\
&\leq \|\nabla\cdot(\boldsymbol{\varphi}-\boldsymbol{I}_h^\Gamma\boldsymbol{\varphi})\|_{0,\Omega}\|\nabla\cdot(\boldsymbol{\psi}-\tilde{\boldsymbol{\psi}}_h)\|_{0,\Omega} \\
&\lesssim \begin{cases} h\|\nabla\cdot(\boldsymbol{\varphi}-\boldsymbol{I}_h^\Gamma\boldsymbol{\varphi})\|_{0,\Omega}\,\|e^u\|_{0,\Omega} & p_v=1, \\ \left(\frac{h}{p}\right)^2\|\nabla\cdot(\boldsymbol{\varphi}-\boldsymbol{I}_h^\Gamma\boldsymbol{\varphi})\|_{0,\Omega}\,\|e^u\|_{0,\Omega} & \textit{else}, \end{cases} \\
\langle(\boldsymbol{\varphi}-\boldsymbol{I}_h^\Gamma\boldsymbol{\varphi})\cdot\boldsymbol{n},\boldsymbol{e}^{\boldsymbol{\psi}}\cdot\boldsymbol{n}\rangle_\Gamma &\leq \|(\boldsymbol{\varphi}-\boldsymbol{I}_h^\Gamma\boldsymbol{\varphi})\cdot\boldsymbol{n}\|_{0,\Gamma}\|\boldsymbol{e}^{\boldsymbol{\psi}}\cdot\boldsymbol{n}\|_{0,\Gamma} \\
&\lesssim \begin{cases} h|||\boldsymbol{\varphi}-\boldsymbol{I}_h^\Gamma\boldsymbol{\varphi}|||\,\|e^u\|_{0,\Omega} & \mathbf{V}_{p_v}(\mathcal{T}_h)=\mathbf{RT}_0(\mathcal{T}_h), \\ \left(\frac{h}{p}\right)^{3/2}|||\boldsymbol{\varphi}-\boldsymbol{I}_h^\Gamma\boldsymbol{\varphi}|||\,\|e^u\|_{0,\Omega} & \textit{else}. \end{cases}
\end{aligned}
$$

Canceling one power of $\|e^u\|_{0,\Omega}$, using the estimates of the operator $\boldsymbol{I}_h^\Gamma$ given in Lemma 5.3.5 and collecting the terms yields the result. □

Corollary 5.3.18. *Let Γ be smooth, $f\in H^s(\Omega)$ and $g\in H^{s+1/2}(\Gamma)$ for some $s\geq 0$ and denote $C_{f,g}:=\|f\|_{H^s(\Omega)}+\|g\|_{H^{s+1/2}(\Gamma)}$. Then the solution to (5.2) satisfies $u\in H^{s+2}(\Omega)$, $\boldsymbol{\varphi}\in\boldsymbol{H}^{s+1}(\Omega)$, $\boldsymbol{\varphi}\cdot\boldsymbol{n}\in\boldsymbol{H}^{s+1/2}(\Gamma)$ and $\nabla\cdot\boldsymbol{\varphi}\in H^s(\Omega)$. Let $(\boldsymbol{\varphi}_h,u_h)$ be the least squares approximation of $(\boldsymbol{\varphi},u)$. Furthermore, let $e^u=u-u_h$ and $\boldsymbol{e}^{\boldsymbol{\varphi}}=\boldsymbol{\varphi}-\boldsymbol{\varphi}_h$. Then, for the lowest order case $p_v=1$,*

$$\|e^u\|_{0,\Omega}\lesssim h^{\min\{s+1,2\}}\,\|f\|_{H^s(\Omega)}\,.$$

For $p_v>1$ there holds

$\mathbf{V}_{p_v}(\mathcal{T}_h)=\mathbf{RT}_{p_v-1}(\mathcal{T}_h)$	$\mathbf{V}_{p_v}(\mathcal{T}_h)=\mathbf{BDM}_{p_v}(\mathcal{T}_h)$
$\|e^u\|_{0,\Omega}\lesssim\left(\frac{h}{p}\right)^{\min\{s+1,p_s,p_v+1/2\}+1}C_{f,g}$	$\|e^u\|_{0,\Omega}\lesssim\left(\frac{h}{p}\right)^{\min\{s+1,p_s,p_v+1\}+1}C_{f,g}$

Furthermore, the estimates

$\mathbf{V}_{p_v}(\mathcal{T}_h)=\mathbf{RT}_{p_v-1}(\mathcal{T}_h)$	$\mathbf{V}_{p_v}(\mathcal{T}_h)=\mathbf{BDM}_{p_v}(\mathcal{T}_h)$
$\|\nabla e^u\|_{0,\Omega}\lesssim\left(\frac{h}{p}\right)^{\min\{s+1,p_s,p_v+1/2\}}C_{f,g}$	$\|\nabla e^u\|_{0,\Omega}\lesssim\left(\frac{h}{p}\right)^{\min\{s+1,p_s,p_v+1\}}C_{f,g}$

and

$\mathbf{V}_{p_v}(\mathcal{T}_h)=\mathbf{RT}_{p_v-1}(\mathcal{T}_h)$	$\mathbf{V}_{p_v}(\mathcal{T}_h)=\mathbf{BDM}_{p_v}(\mathcal{T}_h)$
$\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}\lesssim\left(\frac{h}{p}\right)^{\min\{s+1/2,p_s+1/2,p_v\}}C_{f,g}$	$\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}\lesssim\left(\frac{h}{p}\right)^{\min\{s+1/2,p_s+1/2,p_v+1\}}C_{f,g}$

hold. Finally, we have

$\mathbf{V}_{p_v}(\mathcal{T}_h) = \mathbf{RT}_{p_v-1}(\mathcal{T}_h)$	$\mathbf{V}_{p_v}(\mathcal{T}_h) = \mathbf{BDM}_{p_v}(\mathcal{T}_h)$
$\Vert \boldsymbol{e}^{\boldsymbol{\varphi}} \cdot \boldsymbol{n} \Vert_{0,\Gamma} \lesssim \left(\frac{h}{p}\right)^{\min\{s+1/2, p_s+1/2, p_v\}} C_{f,g}$	$\Vert \boldsymbol{e}^{\boldsymbol{\varphi}} \cdot \boldsymbol{n} \Vert_{0,\Gamma} \lesssim \left(\frac{h}{p}\right)^{\min\{s+1/2, p_s+1/2, p_v+1\}} C_{f,g}$

Proof. The regularity follows by the standard shift theorem with the fact that $\boldsymbol{\varphi} = -\nabla u$. We now analyze the quantities in the estimates of the Theorems 5.3.11 , 5.3.13 , 5.3.15 and 5.3.17:

$$\begin{aligned}
\Vert u - \tilde{u}_h \Vert_{0,\Omega} &\lesssim (h/p)^{\min\{s+1,p_s\}+1} \Vert u \Vert_{H^{s+2}(\Omega)} \lesssim (h/p)^{\min\{s+1,p_s\}+1} C_{f,g}, \\
\Vert u - \tilde{u}_h \Vert_{1,\Omega} &\lesssim (h/p)^{\min\{s+1,p_s\}} \Vert u \Vert_{H^{s+2}(\Omega)} \lesssim (h/p)^{\min\{s+1,p_s\}} C_{f,g}, \\
\Vert u - \tilde{u}_h \Vert_{0,\Gamma} &\lesssim (h/p)^{\min\{s+1,p_s\}+1/2} \Vert u \Vert_{H^{s+2}(\Omega)} \lesssim (h/p)^{\min\{s+1,p_s\}+1/2} C_{f,g}.
\end{aligned}$$

Furthermore, the following estimates hold for the choices $\mathbf{V}_{p_v}(\mathcal{T}_h) = \mathbf{RT}_{p_v-1}(\mathcal{T}_h)$ and $\mathbf{V}_{p_v}(\mathcal{T}_h) = \mathbf{BDM}_{p_v}(\mathcal{T}_h)$, respectively

$$\begin{aligned}
\Vert \boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h \Vert_{0,\Omega} &\lesssim \begin{cases} (h/p)^{\min\{s+1,p_v\}} \Vert \boldsymbol{\varphi} \Vert_{H^{s+1}(\Omega)} \lesssim (h/p)^{\min\{s+1,p_v\}} C_{f,g}, \\ (h/p)^{\min\{s+1,p_v+1\}} \Vert \boldsymbol{\varphi} \Vert_{H^{s+1}(\Omega)} \lesssim (h/p)^{\min\{s+1,p_v+1\}} C_{f,g}, \end{cases} \\
\Vert (\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h) \cdot \boldsymbol{n} \Vert_{0,\Gamma} &\lesssim \begin{cases} (h/p)^{\min\{s+1/2,p_v\}} \Vert \boldsymbol{\varphi} \Vert_{H^{s+1}(\Omega)} \lesssim (h/p)^{\min\{s+1/2,p_v\}} C_{f,g}, \\ (h/p)^{\min\{s+1/2,p_v+1\}} \Vert \boldsymbol{\varphi} \Vert_{H^{s+1}(\Omega)} \lesssim (h/p)^{\min\{s+1/2,p_v+1\}} C_{f,g}, \end{cases} \\
\Vert \nabla \cdot (\boldsymbol{\varphi} - \tilde{\boldsymbol{\varphi}}_h) \Vert_{0,\Omega} &\lesssim (h/p)^{\min\{s,p_v\}} \Vert \nabla \cdot \boldsymbol{\varphi} \Vert_{H^{s}(\Omega)} \lesssim (h/p)^{\min\{s,p_v\}} C_{f,g}.
\end{aligned}$$

The estimates of the Theorems 5.3.11 , 5.3.13 , 5.3.15 and 5.3.17 together with the above estimates give, after straightforward calculations, the asserted rates. □

Remark 5.3.19. Note that the rates predicted by Corollary 5.3.18 for the error of the vector valued variable $\Vert \boldsymbol{e}^{\boldsymbol{\varphi}} \Vert_{0,\Omega}$ and the normal trace of the vector valued variable $\Vert \boldsymbol{e}^{\boldsymbol{\varphi}} \cdot \boldsymbol{n} \Vert_{0,\Gamma}$ are the same. This again suggests the suboptimality of the estimate for $\Vert \boldsymbol{e}^{\boldsymbol{\varphi}} \Vert_{0,\Omega}$. ■

5.4. Numerical examples

For the presentation of the numerical results we employ the same conventions as in Section 4.4. An additional quantity of interest is now the error of the normal trace $\Vert \boldsymbol{e}^{\boldsymbol{\varphi}} \cdot \boldsymbol{n} \Vert_{0,\Gamma}$.

Example 5.4.1. We consider as the domain Ω the unit sphere in $\mathbb{R}^2$. The exact solution is given by $u(x, y) = \sin(x + y)$ and therefore smooth. The right-hand sides f and g are calculated according to the choice $\alpha = 1$ and $\gamma = 3$. The numerical results are plotted in Figure 5.1 and B.1 for $\Vert e^u \Vert_{0,\Omega}$, in Figure B.2 and B.3 for $\Vert \nabla e^u \Vert_{0,\Omega}$, in Figure 5.2 and 5.3 for $\Vert \boldsymbol{e}^{\boldsymbol{\varphi}} \Vert_{0,\Omega}$, in Figure 5.4 and 5.5 for $\Vert \boldsymbol{e}^{\boldsymbol{\varphi}} \cdot \boldsymbol{n} \Vert_{0,\Gamma}$.

Example 5.4.2. We again consider as the domain Ω the unit sphere in $\mathbb{R}^2$. The exact solution $u(x, y)$ is calculated corresponding to the right-hand side $f(x, y) = \mathbb{1}_{[0,1/2]}(\sqrt{x^2 + y^2})$, $\gamma = 2$ and satisfying $\partial_n u = 0$. The right-hand side g is calculated according to the choice $\alpha = 1$. The numerical results are plotted in Figure 5.6 and B.4 for $\Vert e^u \Vert_{0,\Omega}$, in Figure 5.7 and B.5 for $\Vert \nabla e^u \Vert_{0,\Omega}$, in Figure 5.8 and B.6 for $\Vert \boldsymbol{e}^{\boldsymbol{\varphi}} \Vert_{0,\Omega}$, in Figure 5.9 and B.7 for $\Vert \boldsymbol{e}^{\boldsymbol{\varphi}} \cdot \boldsymbol{n} \Vert_{0,\Gamma}$.

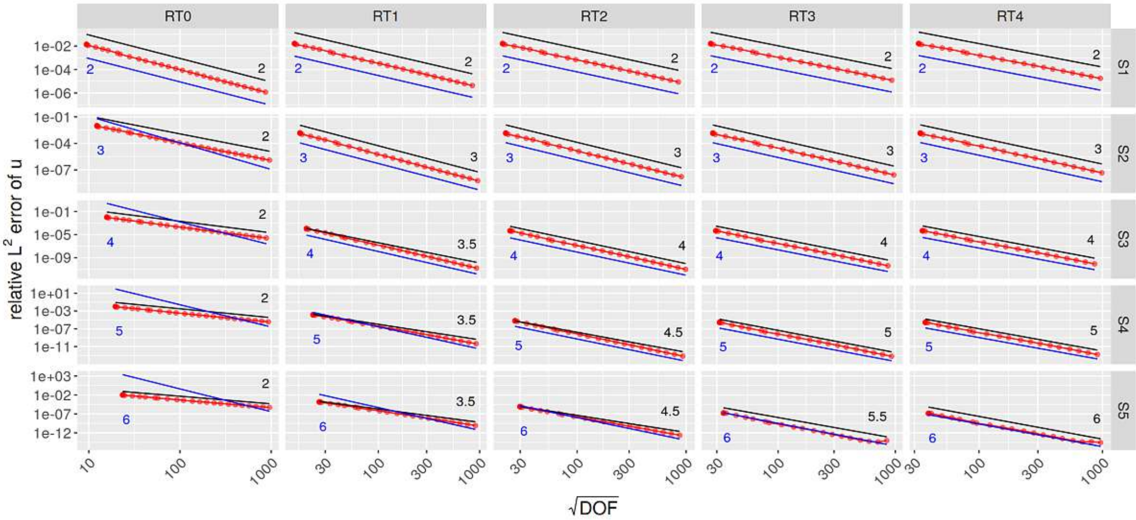

Figure 5.1.: (cf. Example 5.4.1) Convergence of $\|e^u\|_{0,\Omega}$ vs. $\sqrt{\mathrm{DOF}} \sim 1/h$ employing $\mathbf{V}_{p_v}(\mathcal{T}_h) = \mathbf{RT}_{p_v-1}(\mathcal{T}_h)$.

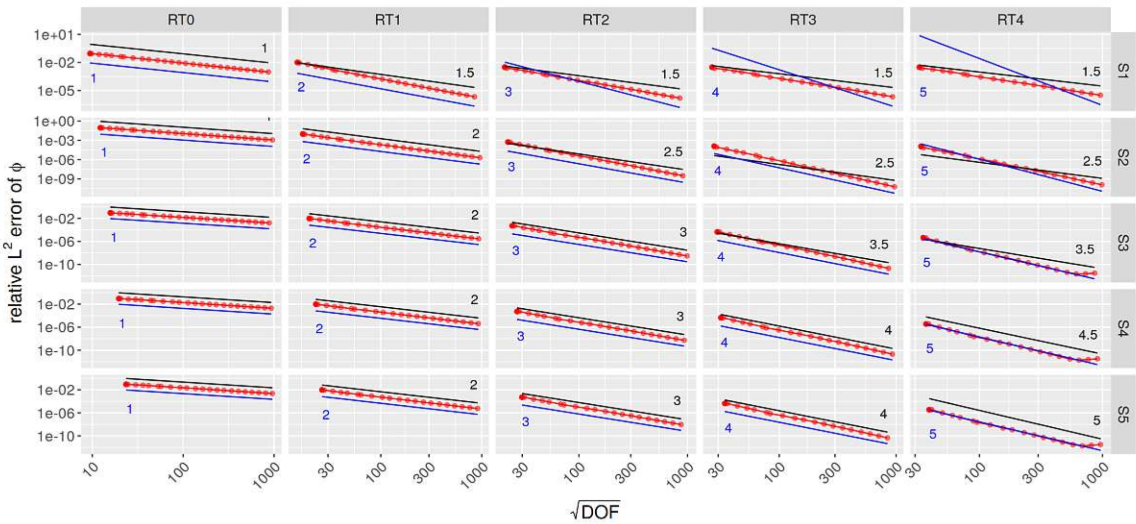

Figure 5.2.: (cf. Example 5.4.1) Convergence of $\|\boldsymbol{e}^{\varphi}\|_{0,\Omega}$ vs. $\sqrt{\mathrm{DOF}} \sim 1/h$ employing $\mathbf{V}_{p_v}(\mathcal{T}_h) = \mathbf{RT}_{p_v-1}(\mathcal{T}_h)$.

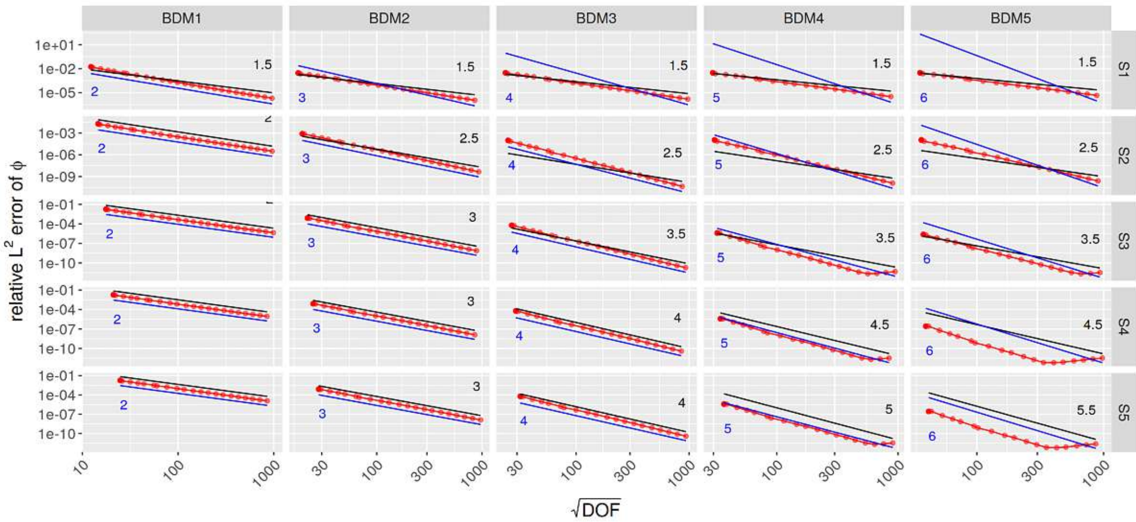

Figure 5.3.: (cf. Example 5.4.1) Convergence of $\|\boldsymbol{e}^{\varphi}\|_{0,\Omega}$ vs. $\sqrt{\mathrm{DOF}} \sim 1/h$ employing $\mathbf{V}_{p_v}(\mathcal{T}_h) = \mathbf{BDM}_{p_v}(\mathcal{T}_h)$.

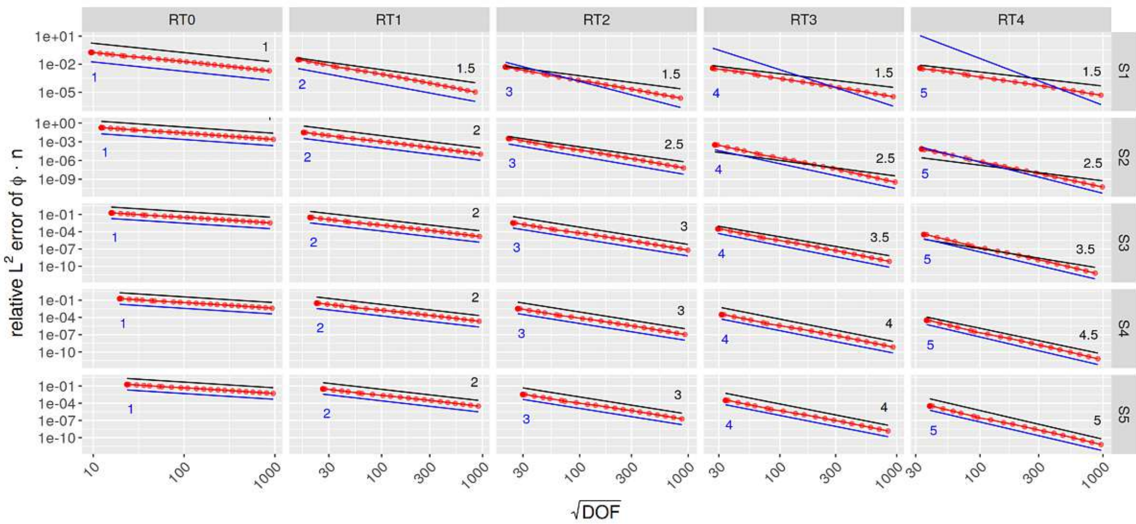

Figure 5.4.: (cf. Example 5.4.1) Convergence of $\|\boldsymbol{e}^{\varphi} \cdot \boldsymbol{n}\|_{0,\Gamma}$ vs. $\sqrt{\mathrm{DOF}} \sim 1/h$ employing $\mathbf{V}_{p_v}(\mathcal{T}_h) = \mathbf{RT}_{p_v-1}(\mathcal{T}_h)$.

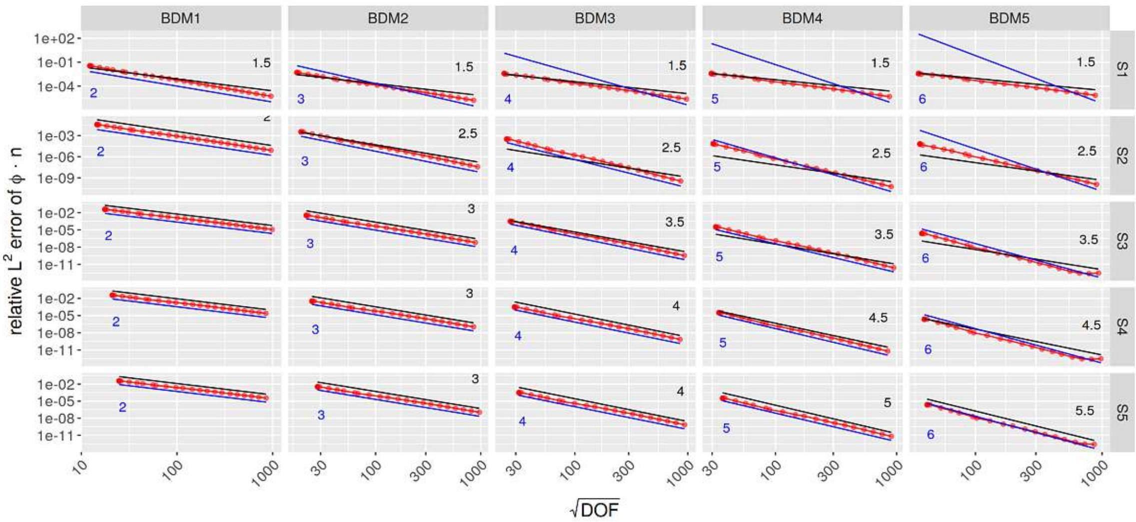

Figure 5.5.: (cf. Example 5.4.1) Convergence of $\|\boldsymbol{e}^{\varphi} \cdot \boldsymbol{n}\|_{0,\Gamma}$ vs. $\sqrt{\text{DOF}} \sim 1/h$ employing $\mathbf{V}_{p_v}(\mathcal{T}_h) = \mathbf{BDM}_{p_v}(\mathcal{T}_h)$.

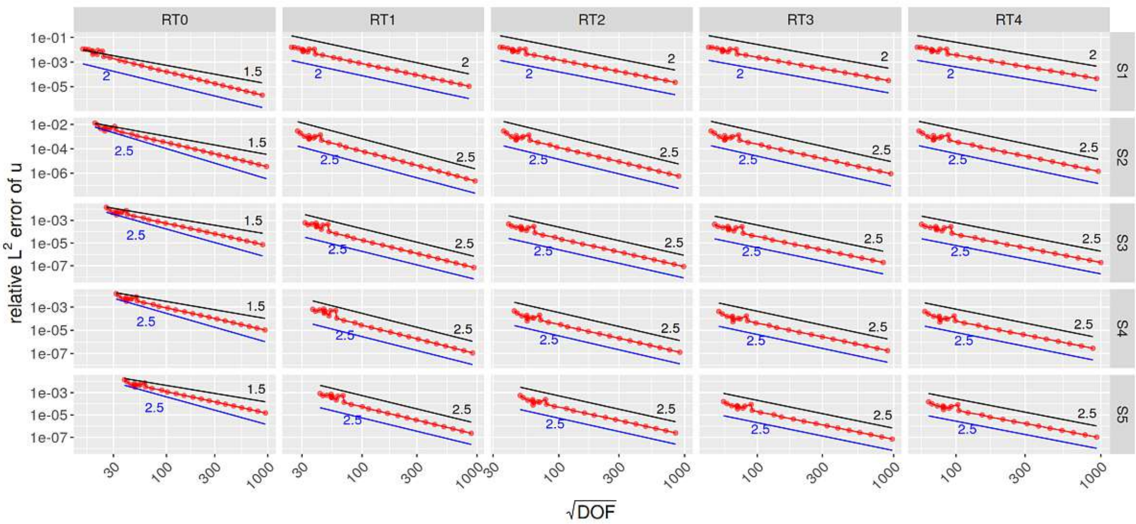

Figure 5.6.: (cf. Example 5.4.2) Convergence of $\|e^{u}\|_{0,\Omega}$ vs. $\sqrt{\text{DOF}} \sim 1/h$ employing $\mathbf{V}_{p_v}(\mathcal{T}_h) = \mathbf{RT}_{p_v-1}(\mathcal{T}_h)$.

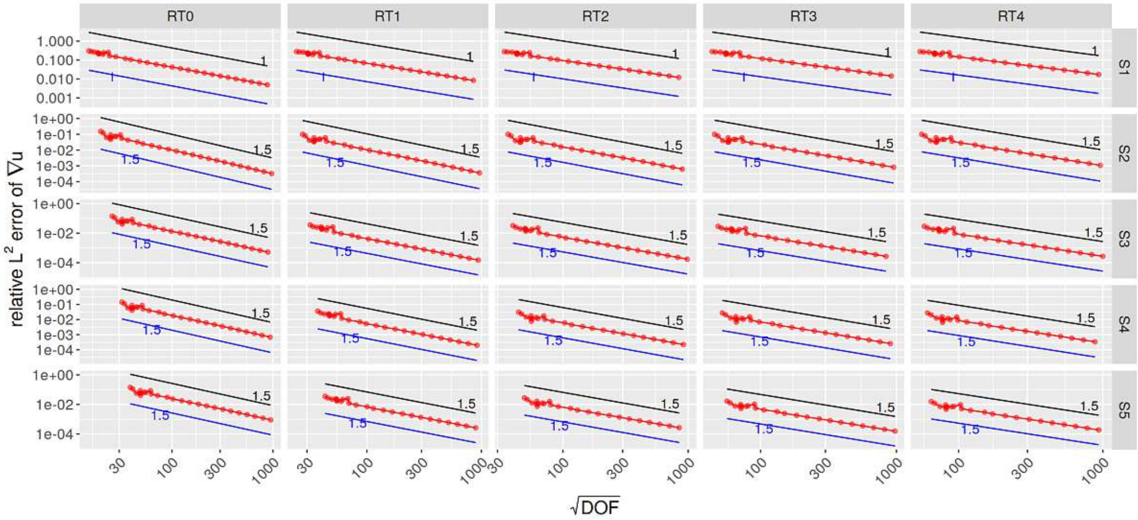

Figure 5.7.: (cf. Example 5.4.2) Convergence of $\|\nabla e^u\|_{0,\Omega}$ vs. $\sqrt{\mathrm{DOF}} \sim 1/h$ employing $\mathbf{V}_{p_v}(\mathcal{T}_h) = \mathbf{RT}_{p_v-1}(\mathcal{T}_h)$.

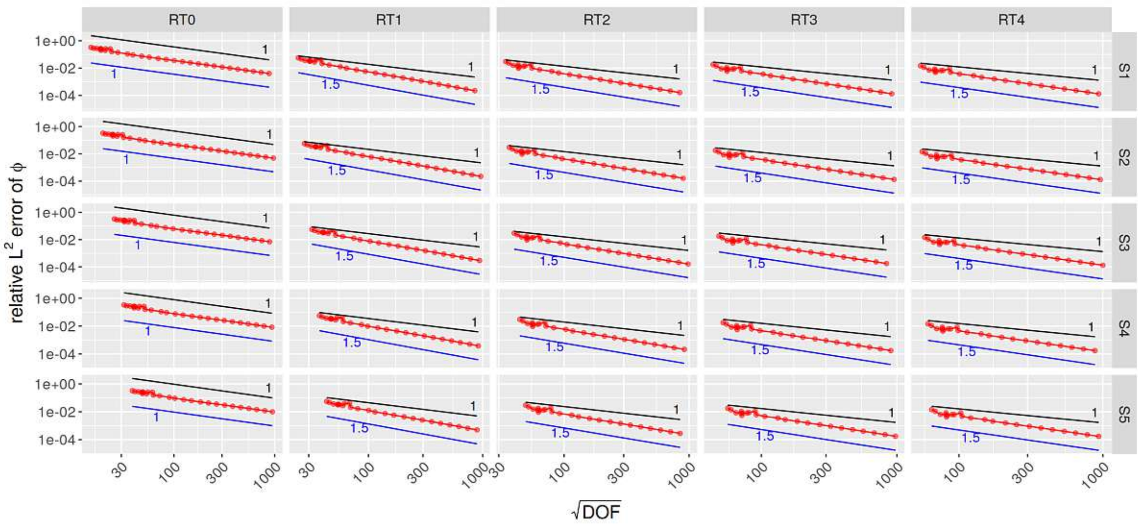

Figure 5.8.: (cf. Example 5.4.2) Convergence of $\|e^{\varphi}\|_{0,\Omega}$ vs. $\sqrt{\mathrm{DOF}} \sim 1/h$ employing $\mathbf{V}_{p_v}(\mathcal{T}_h) = \mathbf{RT}_{p_v-1}(\mathcal{T}_h)$.

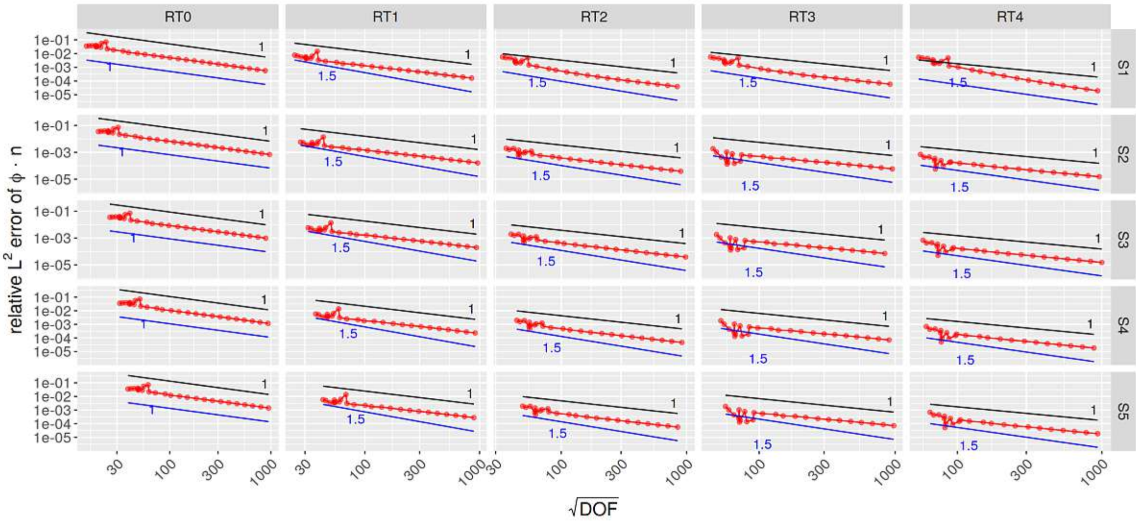

Figure 5.9.: (cf. Example 5.4.2) Convergence of $\|\boldsymbol{e}^{\boldsymbol{\varphi}} \cdot \boldsymbol{n}\|_{0,\Gamma}$ vs. $\sqrt{\text{DOF}} \sim 1/h$ employing $\mathbf{V}_{p_v}(\mathcal{T}_h) = \mathbf{RT}_{p_v-1}(\mathcal{T}_h)$.

6. Galerkin discretizations of Heterogeneous Helmholtz problems

In the present chapter we perform wavenumber-explicit regularity theory as well as convergence analysis of the Galerkin discretization for a class of time-harmonic wave propagation problems in piecewise smooth media. Our model problems include heterogeneous Helmholtz problems with piecewise analytic coefficients endowed with Robin, Dirichlet-to-Neumann and second order absorbing boundary conditions. Furthermore, our theory covers perfectly matched layers as well as volume damping terms. The results presented in the current chapter are part of the work [BCFM21].
The outline of this chapter is as follows. In Section 6.1 we give an informal road map of our results. To that end, we consider a prototypical heterogeneous Helmholtz model problem. Section 6.2 introduces problem specific notation and lists the assumptions of our theory. In Section 6.3 we perform an abstract contraction argument and prove a wavenumber-explicit regularity splitting of the solution, see Theorem 6.3.10. In Theorem 6.3.11 we perform this splitting for higher order Sobolev data. Section 6.4 reviews the adjoint problem, which arises naturally in the duality argument when analyzing a Galerkin discretization. Section 6.5 verifies the assumptions for the wavenumber-explicit regularity splitting of Section 6.3 for a variety of problems. Next, we perform an abstract Galerkin analysis in Section 6.6. We conclude this section with an application to the *hp*-FEM. In Section 6.7 we present numerical results which support our findings. In Section 6.8 we prove analytic regularity of a model problem with second order boundary conditions. Finally, in Section 6.9 we analyze the Dirichlet-to-Neumann map for the exterior Helmholtz equations and its relation to the Dirichlet-to-Neumann map for the Laplacian. We prove a splitting of the difference of the two, see Lemma 6.5.12. Closing this chapter, we propose a splitting of the Dirichlet-to-Neumann map for linear elasticity in Section 6.10.

6.1. Whetting the appetite

Consider the following heterogeneous Helmholtz problem with Robin boundary conditions:

$$\begin{aligned} -\Delta u - k^2 n^2 u &= f \quad \text{in } \Omega, \\ \partial_n u - iku &= g \quad \text{on } \Gamma, \end{aligned} \tag{6.1}$$

where $k \geq k_0 > 0$ is real. The boundary Γ of the bounded Lipschitz domain Ω as well as the spatial-dependent index of refraction $n = n(x)$ are assumed to be analytic and uniformly bounded away from zero and from above, i.e., there exist constants $n_{\min}$, $n_{\max} > 0$ such that $0 < n_{\min} < n(x) < n_{\max}$ for all $x \in \Omega$. Furthermore, let $f \in L^2(\Omega)$ and $g \in H^{1/2}(\Gamma)$. In one of our main results, Theorem 6.3.10, which is applicable to the above problem, see

Section 6.5, we prove a wavenumber-explicit splitting of the weak solution $u \in H^1(\Omega)$ to problem (6.1). Theorem 6.3.10 allows to write $u = u_F + u_A$, where $u_F \in H^2(\Omega)$ and u_A is analytic. The function u_F satisfies

$$\|u_F\|_{2,\Omega} + k\|u_F\|_{1,\Omega} + k^2\|u_F\|_{0,\Omega} \lesssim \|f\|_{0,\Omega} + \|g\|_{1/2,\Gamma},$$

which expresses favorable wavenumber-dependence. The analytic part u_A is oscillatory. Note that for $n \equiv 1$, we recover the results of [MS11, Thm. 4.10] for the homogeneous Helmholtz equation.

The weak formulation of (6.1) reads: Find $u \in H^1(\Omega)$ such that

$$(\nabla u, \nabla v) - k^2(n^2 u, v) - ik\langle u, v\rangle = (f, v) + \langle g, v\rangle \qquad \forall v \in H^1(\Omega). \tag{6.2}$$

We denote by S_k^- the solution operator to problem (6.1), i.e., $S_k^-(f,g) := u$, where u solves (6.2). Let us also introduce the sesquilinear form b_k^- given by

$$b_k^-(u,v) := (\nabla u, \nabla v) - k^2(n^2 u, v) - ik\langle u, v\rangle \qquad \forall u, v \in H^1(\Omega),$$

as well as the associated differential operator $L_k^- u := -\Delta u - k^2 n^2 u$ and the boundary operator $T_{k,\Gamma}^- u := iku$. Problem (6.1) can therefore be written as

$$\begin{aligned} L_k^- u &= f \quad \text{in } \Omega,\\ \partial_n u - T_{k,\Gamma}^- u &= g \quad \text{on } \Gamma. \end{aligned} \tag{6.3}$$

We additionally introduce the solution operator $S_k^+(f,g) := w$ of the auxiliary problem

$$\begin{aligned} -\Delta w + k^2 w &= f \quad \text{in } \Omega,\\ \partial_n w &= g \quad \text{on } \Gamma, \end{aligned} \tag{6.4}$$

as well as the corresponding differential operator $L_k^+ w := -\Delta w + k^2 w$ and the sesquilinear form b_k^+. If we additionally introduce the (in this case) trivial boundary operator $T_{k,\Gamma}^+ = 0$ we find that w solves

$$\begin{aligned} L_k^+ w &= f \quad \text{in } \Omega,\\ \partial_n w - T_{k,\Gamma}^+ w &= g \quad \text{on } \Gamma. \end{aligned} \tag{6.5}$$

By construction problem (6.5) is coercive and the solution features favorable k-explicit *a priori* bounds as well as estimates in $H^2(\Omega)$, since a shift theorem is applicable. The difference between the differential operators and the boundary operators only consists of lower order terms. More generally, we allow for a diffusion matrix A and lower order operators $T_{k,\Omega}^-$ and $T_{k,\Omega}^+$, such that the differential operators L_k^- and L_k^+ take the form

$$\begin{aligned} L_k^- u &= -\nabla \cdot (A\nabla u) - T_{k,\Omega}^- u,\\ L_k^+ u &= -\nabla \cdot (A\nabla u) - T_{k,\Omega}^+ u. \end{aligned}$$

The above fits into this framework, with the choice $-T_{k,\Omega}^- u = -k^2 n^2 u$ and $-T_{k,\Omega}^+ u = +k^2 u$.

We will later see that S_k^- and S_k^+ act very similar on high-frequency data. For the readers' convenience, we summarize this notation in Table 6.1.

Minus	Plus
$u = S_k^-(f,g)$	$u = S_k^+(f,g)$
$b_k^-(u,v)$	$b_k^+(u,v)$
$L_k^- u = f$	$L_k^+ u = f$
$L_k^- u = -\nabla\cdot(A\nabla u) - T_{k,\Omega}^- u$	$L_k^+ u = -\nabla\cdot(A\nabla u) - T_{k,\Omega}^+ u$
$\partial_n u - T_{k,\Gamma}^- u = g$	$\partial_n u - T_{k,\Gamma}^+ u = g$

Table 6.1.: Notational overview.

6.2. Assumptions and problem specific notation

The goal of this section is to depict the abstract settings for which the proposed analysis is valid. A large part of our analysis relies on smoothness properties of the domain. For ease of reference we introduce the following

Assumption 6.2.1 (Assumptions on the domain Ω, the boundary Γ and the interface Γ_i)**.** In spatial dimension $d = 2, 3$ the bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$ has an analytic boundary $\Gamma := \partial\Omega$. The interface $\Gamma_i \subset \Omega$ is an analytic $d-1$ dimensional manifold, possibly consisting of a finite number of connected components. Furthermore, it is nonintersecting and bounded away from Γ.

In view of Assumption 6.2.1 we introduce the piecewise Sobolev spaces of order $s \geq 0$. For Ω and Γ_i as in Assumption 6.2.1 let $H^s(\Omega \setminus \Gamma_i)$ denote the space of functions u, such that $u\big|_\omega \in H^s(\omega)$ for all components $\omega \subset \Omega\setminus\Gamma_i$. The corresponding norm for $u \in H^s(\Omega\setminus\Gamma_i)$ is given by $\|u\|^2_{s,\Omega\setminus\Gamma_i} = \sum_\omega \|u\|^2_{s,\omega}$.

Assumption 6.2.2 (Assumptions on the wavenumber k and the diffusion matrix A)**.** The wavenumber k is real and bounded away from zero, i.e., $k \geq k_0 > 0$. Furthermore, the matrix-valued variable heterogeneity A is analytic on $\overline{\Omega}$ or piecewise analytic with an analytic interface Γ_i. We assume the existence of constants C_A, $\gamma_A \geq 0$ such that

$$\|\nabla^p A\|_{L^\infty(\Omega\setminus\Gamma_i)} \leq C_A \gamma_A^p p! \qquad \forall p \in \mathbb{N}_0.$$

Furthermore, we assume A to be homogeneous at the boundary, i.e., $A \equiv I$ on Γ.

For our presentation, the energy space will be the space $H^{1,t}(\Omega,\Gamma)$. For $t \geq 0$ let

$$H^{1,t}(\Omega,\Gamma) := \{u \in H^1(\Omega) \colon u \in H^t(\Gamma)\} \subset H^1(\Omega).$$

with k-dependent norm

$$\|u\|^2_{1,t,k} := \|\nabla u\|^2_{0,\Omega} + k^2\|u\|^2_{0,\Omega} + k^{-2t+1}|u|^2_{t,\Gamma}.$$

Furthermore, we denote by $\|\cdot\|_{1,k,\Omega}$ the k-dependent norm

$$\|u\|_{1,k,\Omega}^2 := \|\nabla u\|_{0,\Omega}^2 + k^2\|u\|_{0,\Omega}^2.$$

It is easy to show

Lemma 6.2.3. *For $0 \le t \le 1/2$ the space $H^{1,t}(\Omega,\Gamma)$ coincides with $H^1(\Omega)$ and $\|\cdot\|_{1,t,k}$ is equivalent to $\|\cdot\|_{1,k,\Omega}$.*

We stress at this point that our theory also covers vector valued problems. For notational convenience we stick to the space $H^{1,t}(\Omega,\Gamma)$ in order to cover heterogeneous Helmholtz problems with different boundary conditions, including the Robin boundary condition, the full space problem employing the Dirichlet-to-Neumann operator DtN_k on a coupling interface Γ as well as second order absorbing boundary conditions, second order ABCs for short.
For the readers' convenience we present an overview table of our covered problems in Table 6.2. We further discuss these problems in Section 6.5.

<table>
<tr><th>Model</th><th>Energy Space</th><th>L_k^-</th><th>L_k^+</th><th>$T_{k,\Omega}^- - T_{k,\Omega}^+$</th><th>$T_{k,\Gamma}^-$</th><th>$T_{k,\Gamma}^+$</th><th>$T_{k,\Gamma}^- - T_{k,\Gamma}^+$</th></tr>
<tr><td>HH + RBC</td><td rowspan="5">$H^1(\Omega)$</td><td rowspan="3">$-\nabla\cdot(A\nabla u) - k^2n^2u$</td><td rowspan="4">$-\nabla\cdot(A\nabla u) + k^2u$</td><td rowspan="3">$-k^2(n^2+1)u$</td><td>$iku$</td><td>$0$</td><td>$iku$</td></tr>
<tr><td>HH + DtN Sphere</td><td>DtN_k</td><td>DtN_0</td><td>kR_Γ</td></tr>
<tr><td>HH + DtN Γ</td><td>DtN_k</td><td>DtN_0</td><td>$kR_\Gamma + A_\Gamma$</td></tr>
<tr><td>HH + Damping + RBC</td><td>$-\nabla\cdot(A\nabla u) - k^2n^2u + ikmu$</td><td>$-k^2(n^2+1)u + ikmu$</td><td>$iku$</td><td>$0$</td><td>$iku$</td></tr>
<tr><td>HH + PML + RBC</td><td>$-\nabla(A_k^{PML}\nabla u) - k^2n^2u$</td><td>$-\nabla(A_k^{PML}\nabla u) + k^2u$</td><td rowspan="2">$-k^2(n^2+1)u$</td><td>$iku$</td><td>$0$</td><td>$iku$</td></tr>
<tr><td>HH + second order ABC</td><td>$H^{1,1}(\Omega,\Gamma)$</td><td>$-\nabla\cdot(A\nabla u) - k^2n^2u$</td><td>$-\nabla\cdot(A\nabla u) + k^2u$</td><td>$\alpha\Delta_\Gamma u + kR_\Gamma u$</td><td>$\alpha\Delta_\Gamma u$</td><td>$kR_\Gamma$</td></tr>
</table>

Table 6.2.: Overview of covered problems.

The primal problem reads: For $f \in L^2(\Omega)$ and $g \in L^2(\Gamma)$ find $S_k^-(f,g) := u \in H^{1,t}(\Omega,\Gamma)$ such that

$$b_k^-(u,v) := (A\nabla u, \nabla v) - (T_{k,\Omega}^- u, v) - \langle T_{k,\Gamma}^- u, v\rangle = (f,v) + \langle g, v\rangle \qquad \forall v \in H^{1,t}(\Omega,\Gamma). \tag{6.6}$$

The auxiliary problem problem reads: For $f \in L^2(\Omega)$ and $g \in L^2(\Gamma)$ find $S_k^+(f,g) := u \in H^{1,t}(\Omega,\Gamma)$ such that

$$b_k^+(u,v) := (A\nabla u, \nabla v) - (T_{k,\Omega}^+ u, v) - \langle T_{k,\Gamma}^+ u, v\rangle = (f,v) + \langle g, v\rangle \qquad \forall v \in H^{1,t}(\Omega,\Gamma). \tag{6.7}$$

We specify the operators $T_{k,\Omega}^-$, $T_{k,\Gamma}^-$, $T_{k,\Omega}^+$ and $T_{k,\Gamma}^+$ in the Assumptions 6.2.4 and 6.2.5 as well as 6.2.6 below. Furthermore, for $C_1, \gamma_1 \ge 0$ we introduce the analyticity class in the volume

$$\mathcal{A}(C_1,\gamma_1,\Omega\setminus\Gamma_i) := \{v \in L^2(\Omega) \colon \|\nabla^n v\|_{0,\Omega\setminus\Gamma_i} \le C_1\gamma_1^n \max\{n,k\}^n \quad n \in \mathbb{N}_0\},$$

as well as the analyticity class on the boundary. For $g \in L^2(\Gamma)$ we write $g \in \mathcal{A}(C_1,\gamma_1,\Gamma)$, if there exists a one-sided tubular neighborhood T of the boundary Γ, such that g is the restriction of an analytic function G which satisfies

$$\|\nabla^n G\|_{0,T} \le C_1\gamma_1^n \max\{n,k\}^n \qquad \forall n \ge 0.$$

We start with assumptions on the primal problem.

Assumption 6.2.4 (Assumptions regarding L_k^-, b_k^-, $T_{k,\Omega}^-$, $T_{k,\Gamma}^-$ and S_k^-)**.** The following is satisfied:

M.1 The sesquilinear form $b_k^-\colon H^{1,t}(\Omega,\Gamma)\times H^{1,t}(\Omega,\Gamma)\to\mathbb{C}$ is continuous, i.e., there exists a constant $C_{\text{cont},k}^-\geq 0$ such that

$$|b_k^-(u,v)|\leq C_{\text{cont},k}^-\|u\|_{1,t,k}\|v\|_{1,t,k}\qquad\forall u,v\in H^{1,t}(\Omega,\Gamma).$$

M.2 The linear operators $T_{k,\Omega}^-\colon H^{1,t}(\Omega,\Gamma)\to H^{1,t}(\Omega,\Gamma)'$ and $T_{k,\Gamma}^-\colon H^t(\Gamma)\to H^{-t}(\Gamma)$ admit splittings into linear operators

$$T_{k,\Omega}^-=D_\Omega^-+A_\Omega^-,\qquad T_{k,\Gamma}^-=D_\Gamma^-+A_\Gamma^-,$$

such that

$$|(D_\Omega^- u,v)|+|\langle D_\Gamma^- u,v\rangle|\lesssim\|u\|_{1,t,k}\|v\|_{1,t,k}\qquad\forall u,v\in H^{1,t}(\Omega,\Gamma).$$

Furthermore, the operators A_Ω^- and A_Γ^- have the mapping properties

$$A_\Omega^- u\in\mathcal{A}(C_{A,\Omega,k}^-\|u\|_{1,t,k},\gamma_{A,\Omega}^-,\Omega\setminus\Gamma_i),\qquad A_\Gamma^- v\in\mathcal{A}(C_{A,\Gamma,k}^-\|u\|_{1,t,k},\gamma_{A,\Gamma}^-,\Gamma)$$

for all $u\in H^{1,t}(\Omega,\Gamma)$. Let $C_{\text{ana},k}^-:=\max(C_{A,\Omega,k}^-,C_{A,\Gamma,k}^-)$.

M.3 Problem (6.6) is well-posed, i.e., for every $f\in L^2(\Omega)$ and $g\in L^2(\Gamma)$ it admits a unique weak solution $S_k^-(f,g)=u\in H^{1,t}(\Omega,\Gamma)$. The *a priori* energy bound

$$\|u\|_{1,t,k}\leq C_{\text{sol},k}^-(\|f\|_{0,\Omega}+\|g\|_{0,\Gamma})\tag{6.8}$$

holds with a constant $C_{\text{sol},k}^-\gtrsim 1$ independent of f and g.

M.4 For (piecewise) analytic data $f\in\mathcal{A}(C_f,\gamma_f,\Omega\setminus\Gamma_i)$ and $g\in\mathcal{A}(C_g,\gamma_g,\Gamma)$ the solution $u=S_k^-(f,g)$ to Problem (6.6) is again (piecewise) analytic and satisfies

$$\begin{aligned}\|u\|_{1,t,k}&\leq CC_{\text{sol},k}^-(C_f+C_g),\\ \|\nabla^n u\|_{0,\Omega\setminus\Gamma_i}&\leq CC_{\text{sol},k}^-k^{-1}\gamma^p\max\{k,n\}^n(C_f+C_g)\qquad\forall n\geq 2,\end{aligned}$$

with constants C, $\gamma\geq 0$ independent of k.

Assumption 6.2.5 (Assumptions regarding L_k^+, b_k^+, $T_{k,\Omega}^+$, $T_{k,\Gamma}^+$ and S_k^+)**.** The following is satisfied:

P.1 The sesquilinear form $b_k^+\colon H^{1,t}(\Omega,\Gamma)\times H^{1,t}(\Omega,\Gamma)\to\mathbb{C}$ is continuous, i.e., there exists a constant $C_{\text{cont},k}^+\geq 0$ such that

$$|b_k^+(u,v)|\leq C_{\text{cont},k}^+\|u\|_{1,t,k}\|v\|_{1,t,k}\qquad\forall u,v\in H^{1,t}(\Omega,\Gamma).$$

P.2 There exists $\sigma \in \mathbb{C}$ with $|\sigma| = 1$, such that

$$\operatorname{Re}(\sigma b_k^+(u,u)) \gtrsim \|u\|_{1,t,k}^2 \qquad \forall u \in H^{1,t}(\Omega,\Gamma)$$

independently of k.

P.3 The linear operators $T_{k,\Omega}^+ \colon H^{1,t}(\Omega,\Gamma) \to H^{1,t}(\Omega,\Gamma)'$ and $T_{k,\Gamma}^+ \colon H^t(\Gamma) \to H^{-t}(\Gamma)$ are such that

$$T_{k,\Omega}^- - T_{k,\Omega}^+ = R_\Omega + A_\Omega, \qquad T_{k,\Gamma}^- - T_{k,\Gamma}^+ = R_\Gamma + A_\Gamma,$$

with linear operators R_Ω, A_Ω, R_Γ and A_Γ.

P.4 The linear operator R_Ω admits a splitting $R_\Omega = \sum_{i=1}^n k^{2-s_i} R_{\Omega,s_i}$. The linear operators $R_{\Omega,s_i} \colon H^{s_i}(\Omega) \to L^2(\Omega)$ are bounded linear operators and satisfy the estimate $\|R_{\Omega,s_i} u\|_{0,\Omega} \lesssim \|u\|_{s_i,\Omega}$, $0 \leq s_i \leq 1$.

P.5 The linear operator R_Γ admits a splitting $R_\Gamma = \sum_{i=1}^n k^{3/2-s_i} R_{\Gamma,s_i}$. The linear operators $R_{\Gamma,s_i} \colon H^{s_i}(\Gamma) \to H^{1/2}(\Gamma)$ are bounded linear operators and satisfy the estimate $\|R_{\Gamma,s_i} u\|_{1/2,\Gamma} \lesssim \|u\|_{s_i,\Gamma}$, $0 \leq s_i \leq 1/2$.

P.6 The operators A_Ω and A_Γ have the mapping properties

$$A_\Omega u \in \mathcal{A}(C_{A,\Omega,k}\|u\|_{1,t,k}, \gamma_{A,\Omega}, \Omega \setminus \Gamma_i), \qquad A_\Gamma v \in \mathcal{A}(C_{A,\Gamma,k}\|u\|_{1,t,k}, \gamma_{A,\Gamma}, \Gamma)$$

for all $u \in H^{1,t}(\Omega,\Gamma)$. Let $C_{\mathrm{ana},k} \coloneqq \max(C_{A,\Omega,k}, C_{A,\Gamma,k})$.

P.7 For $f \in L^2(\Omega)$ and $g \in H^{1/2}(\Gamma)$ the solution $S_k^+(f,g) = w \in H^2(\Omega \backslash \Gamma_i)$. Furthermore, the regularity shift estimate

$$\|w\|_{2,\Omega\backslash\Gamma_i} \lesssim \|f\|_{0,\Omega} + \|g\|_{1/2,\Gamma} + k^{1/2}\|g\|_{0,\Gamma} \tag{6.9}$$

holds.

Assumption 6.2.6 (Assumptions regarding L_k^+, $T_{k,\Omega}^+$, $T_{k,\Gamma}^+$ and S_k^+ for higher Sobolev regularity)**.** Let $s > 0$ denote the regularity of the data $f \in H^s(\Omega \setminus \Gamma_i)$ and $g \in H^{s+1/2}(\Gamma)$. With the notation of Assumption 6.2.5 we additionally assume the following to be satisfied:

PS.1 The linear operator R_Ω admits a splitting $R_\Omega = \sum_{i=1}^n k^{s+2-s_i} R_{\Omega,s_i}$. The linear operators $R_{\Omega,s_i} \colon H^{s_i}(\Omega \setminus \Gamma_i) \to H^s(\Omega \setminus \Gamma_i)$ are bounded linear operators and satisfy the estimate $\|R_{\Omega,s_i} u\|_{s,\Omega\backslash\Gamma_i} \lesssim \|u\|_{s_i,\Omega\backslash\Gamma_i}$, $0 \leq s_i < s+2$.

PS.2 The linear operator R_Γ admits a splitting $R_\Gamma = \sum_{i=1}^n k^{s+3/2-s_i} R_{\Gamma,s_i}$. The linear operators $R_{\Gamma,s_i} \colon H^{s_i}(\Gamma) \to H^{s+1/2}(\Gamma)$ are bounded linear operators and satisfy the estimate $\|R_{\Gamma,s_i} u\|_{s+1/2,\Gamma} \lesssim \|u\|_{s_i,\Gamma}$, $0 \leq s_i < s + 3/2$.

PS.3 For $f \in H^s(\Omega \setminus \Gamma_i)$ and $g \in H^{s+1/2}(\Gamma)$ the solution $S_k^+(f,g) = w \in H^{s+2}(\Omega \setminus \Gamma_i)$. Furthermore, the regularity shift estimate

$$\|w\|_{s+2,\Omega\backslash\Gamma_i} \lesssim \|f\|_{s,\Omega\backslash\Gamma_i} + k^s\|f\|_{0,\Omega} + \|g\|_{s+1/2,\Gamma} + k^{s+1/2}\|g\|_{0,\Gamma} \tag{6.10}$$

holds.

Remark 6.2.7. For the analysis of an abstract Galerkin discretization, see Section 6.6, we will make additional assumptions. In fact, we will assume that the operators $T^-_{k,\Omega}$ and $T^-_{k,\Gamma}$ are quasi-selfadjoint, see Assumption 6.4.1. Furthermore, in the application to the *hp*-FEM, we will make the additional assumption, that the problem is polynomially well-posed, see Assumption 6.6.6. We do not state these assumptions in the above, since they are not necessary for the wavenumber-explicit regularity theory developed in Section 6.3. ∎

Remark 6.2.8. We stress that the well-posedness in Assumption M.3 as well as the later assumed polynomial bounds in Assumption 6.6.6 are injected into our theory. ∎

Remark 6.2.9 (Conceptual explanation of Assumptions 6.2.4, 6.2.5 and 6.2.6)**.** We quickly discuss the relevance of the Assumptions 6.2.4, 6.2.5 and 6.2.6. Assumption M.1 in Assumption 6.2.4 expresses the continuity of the sesquilinear form b^-_k with possibly wavenumber-dependent continuity constant $C^-_{\text{cont},k}$. In conjunction with M.2 in Assumption 6.2.4 this expresses the fact that unfavorable wavenumber-dependence in the continuity constant $C^-_{\text{cont},k}$ is only caused by operators, which map into a class of sufficiently smooth functions. Assumption M.3 expresses well-posedness of the problem. Assumption M.4 states, that analytic data are mapped to analytic solutions.

We now turn to Assumption 6.2.5. P.1 and P.2 express continuity and coercivity of b^+_k, respectively. This allows for wavenumber-explicit energy estimates of the solution operator S^+_k. P.3 with P.4, P.5 and P.6 expresses that the differential operators L^-_k and L^+_k agree on the leading order operator. Furthermore, the difference admits a splitting into a finite regularity part with favorable k-dependence and an analytic part with possibly unfavorable k-dependence. Finally, P.7 states a shift property associated to S^+_k. Last but not least, Assumption 6.2.6 is the natural assumption for higher order Sobolev data. ∎

Remark 6.2.10. The results of Section 6.3, i.e., the splitting $u = u_F + u_A$ does *not* rely on M.2, nor does it rely on the later Assumptions of quasi-selfadjointness and polynomially well-posedness, see Assumptions 6.4.1 and 6.6.6. These assumptions only enter in the error analysis of an abstract Galerkin discretization. Here an additional adjoint problem arises. In fact, under Assumption 6.4.1 primal and adjoint problem are essentially the same, up to some complex conjugations. This has the advantage that after analyzing the primal problem, the corresponding results also hold for the dual problem. For our model problems in mind, all these assumptions are satisfied. Hence, we are able to perform a complete analysis of the Galerkin discretization of these problems. ∎

Remark 6.2.11 (On implicit and explicit constants in Assumptions 6.2.4, 6.2.5 and 6.2.6)**.** Explicit constants appearing in the Assumptions 6.2.4, 6.2.5 and 6.2.6, which may depend on the wavenumber k, are denoted with an additional subscript k, e.g., $C^-_{\text{cont},k}$ and $C_{\text{ana},k}$. Implicit constants hidden inside $\lesssim$ are independent of k. ∎

6.3. Abstract contraction argument

The main result of the present section is Lemma 6.3.6 and the resulting Theorem 6.3.10, which establishes a wavenumber-explicit regularity theory for our model class.

Lemma 6.3.1 (Stability estimates for S_k^+). *Let P.1 and P.2 in Assumption 6.2.5 be satisfied and let Ω be a bounded Lipschitz domain. Then, for every $f \in \widetilde{H}^{-1}(\Omega)$ and $g \in H^{-t}(\Gamma)$ the problem*

$$\begin{aligned} L_k^+ w &= f \quad \text{in } \Omega, \\ \partial_n w - T_{k,\Gamma}^+ w &= g \quad \text{on } \Gamma \end{aligned}$$

admits a unique weak solution $S_k^+(f,g) := w \in H^{1,t}(\Omega,\Gamma)$. Furthermore, for $f \in L^2(\Omega)$, $g \in L^2(\Gamma)$ and any $0 \le \varepsilon \le 1$ the estimate

$$\|w\|_{1,t,k} \lesssim k^{\varepsilon-1}\|f\|_{\widetilde{H}^{-\varepsilon}(\Omega)} + k^{-1/2}\|g\|_{0,\Gamma} \tag{6.11}$$

holds.

Proof. The weak formulation reads: Find $w \in H^{1,t}(\Omega,\Gamma)$ such that

$$b_k^+(w,v) = (f,v) + \langle g,v\rangle \qquad \forall v \in H^{1,t}(\Omega,\Gamma). \tag{6.12}$$

Unique solvability follows by the Lax-Milgram theorem since P.1 and P.2 in Assumption 6.2.5 are satisfied. For $f \in L^2(\Omega)$, $g \in L^2(\Gamma)$, choosing $v = \overline{\sigma} w$ in (6.12), with σ as in P.2 in Assumption 6.2.5, passing to the real part we find for any $0 \le \varepsilon \le 1$

$$\begin{aligned} \|w\|_{1,t,k}^2 &\lesssim \operatorname{Re}(\sigma b_k^+(w,w)) \le |(f,w)| + |\langle g,w\rangle| \lesssim \|f\|_{\widetilde{H}^{-\varepsilon}(\Omega)}\|w\|_{\varepsilon,\Omega} + \|g\|_{0,\Gamma}\|w\|_{0,\Gamma} \\ &\lesssim (k^{\varepsilon-1}\|f\|_{\widetilde{H}^{-\varepsilon}(\Omega)} + k^{-1/2}\|g\|_{0,\Gamma})\|w\|_{1,k,\Omega}, \end{aligned}$$

where the last estimate follows by interpolation. Estimating $\|w\|_{1,k,\Omega}$ by $\|w\|_{1,t,k}$, canceling one power of $\|w\|_{1,t,k}$ proves (6.11) and concludes the proof. $\square$

We remind the reader of the high and low pass filters introduced in [MS11, Sec. 4.1.1].

Proposition 6.3.2 (Volume high and low pass filters $H_{\eta k}^\Omega$ and $L_{\eta k}^\Omega$). *Let Ω be bounded Lipschitz domain and $\eta > 1$. There exist operators $H_{\eta k}^\Omega \colon L^2(\Omega) \to L^2(\Omega)$ and $L_{\eta k}^\Omega \colon L^2(\Omega) \to L^2(\Omega)$ such that $H_{\eta k}^\Omega f + L_{\eta k}^\Omega f = f$ for every $f \in L^2(\Omega)$. Furthermore, for $0 \le s' \le s$ the operator $H_{\eta k}^\Omega$ satisfies*

$$\|H_{\eta k}^\Omega f\|_{s',\Omega} \lesssim (\eta k)^{s'-s}\|f\|_{s,\Omega} \qquad \forall f \in H^s(\Omega),$$

with hidden constant independent of η and k. Additionally, for $0 \le \varepsilon < 1/2$ the operator $H_{\eta k}^\Omega$ satisfies

$$\|H_{\eta k}^\Omega f\|_{\widetilde{H}^{-\varepsilon}(\Omega)} \lesssim (\eta k)^{-\varepsilon}\|f\|_{0,\Omega} \qquad \forall f \in L^2(\Omega). \tag{6.13}$$

Finally, the low pass filter $L_{\eta k}^\Omega$ satisfies

$$\|\nabla^p L_{\eta k}^\Omega f\|_{0,\Omega} \lesssim (\eta k)^{p-s}\|f\|_{s,\Omega} \qquad \forall f \in H^s(\Omega),\ \forall p \in \mathbb{N}_0,\ p \ge s.$$

Proof. Apart from the estimate (6.13) the results can be found in [MS11, Lemmas 4.2 and 4.3]. Regarding (6.13), note that for $0 \le \varepsilon < 1/2$ and a bounded Lipschitz domain Ω, the space of compactly supported smooth functions $C_0^\infty(\Omega)$ is dense in $H^\varepsilon(\Omega)$. Also note that $H_{\eta k}^\Omega$ is given by

$$H_{\eta k}^\Omega f := H_{\eta k}^{\mathbb{R}^d}(E_\Omega f)\Big|_\Omega,$$

where E_Ω denotes the Stein extension operator. The high pass filter $H_{\eta k}^{\mathbb{R}^d}$ is given by a Fourier procedure, see [MS11, Sec. 4.1.1]. It is easy to check that

$$\|H_{\eta k}^{\mathbb{R}^d} f\|_{-s,\mathbb{R}^d} \lesssim (\eta k)^{-s}\|f\|_{0,\mathbb{R}^d}$$

for all $-1 \le s \le 0$. For $v \in C_0^\infty(\Omega)$ let $\tilde{v}$ denote the trivial extension by zero on all of $\mathbb{R}^d$ of v. These observations now give

$$\begin{aligned}\|H_{\eta k}^\Omega f\|_{\widetilde{H}^{-\varepsilon}(\Omega)} &= \sup_{v\in C_0^\infty(\Omega)} \frac{(H_{\eta k}^\Omega f, v)_\Omega}{\|v\|_{\varepsilon,\Omega}} = \sup_{v\in C_0^\infty(\Omega)} \frac{(H_{\eta k}^{\mathbb{R}^d} f, \tilde{v})_{\mathbb{R}^d}}{\|v\|_{\varepsilon,\Omega}} \\ &\lesssim \sup_{v\in C_0^\infty(\Omega)} \frac{\|H_{\eta k}^{\mathbb{R}^d} E_\Omega f\|_{-\varepsilon,\mathbb{R}^d}\|\tilde{v}\|_{\varepsilon,\mathbb{R}^d}}{\|v\|_{\varepsilon,\Omega}} \\ &\lesssim (\eta k)^{-\varepsilon}\|E_\Omega f\|_{0,\mathbb{R}^d} \lesssim (\eta k)^{-\varepsilon}\|f\|_{0,\Omega},\end{aligned}$$

which concludes the proof. □

Proposition 6.3.3 (Boundary high and low pass filters $H_{\eta k}^\Gamma$ and $L_{\eta k}^\Gamma$)**.** *Let Γ be analytic and $\eta > 1$. There exist operators $H_{\eta k}^\Gamma\colon L^2(\Gamma) \to L^2(\Gamma)$ and $L_{\eta k}^\Gamma\colon L^2(\Gamma) \to L^2(\Gamma)$ such that $H_{\eta k}^\Gamma g + L_{\eta k}^\Gamma g = g$ for every $g \in L^2(\Gamma)$. Furthermore, for $0 \le s' \le s$ the operator $H_{\eta k}^\Gamma$ satisfies*

$$\|H_{\eta k}^\Gamma g\|_{s',\Gamma} \lesssim (\eta k)^{s'-s}\|g\|_{s,\Gamma} \qquad \forall g \in H^s(\Gamma),$$

with hidden constant independent of η and k. Finally, for $s > 0$ and $g \in H^s(\Gamma)$ the function $L_{\eta k}^\Gamma g$ can be obtained as the normal trace of an analytic function, i.e., there exists an analytic function G_g such that $L_{\eta k}^\Gamma g = \boldsymbol{n}\cdot\nabla G_g$, which satisfies

$$\begin{aligned}&\|G_g\|_{3/2+s,\Omega} \lesssim \|g\|_{s,\Gamma}, \\ &\|\nabla^p G_g\|_{0,\Omega} \lesssim (\eta k)^{p-3/2-s}\|g\|_{s,\Gamma} \qquad \forall p \in \mathbb{N}_0, \quad p \ge s + 3/2.\end{aligned}$$

Proof. See [MS11, Lemmas 4.2 and 4.3]. □

Remark 6.3.4 (Other constructions of high and low pass filters)**.** The presented high and low pass filters are by no means the only possible constructions. It is for example also possible to construct them on bounded domains by Fourier series expansion, associated with an eigenvalue problem for the Laplacian. See also [Mel12, Sec 6.1] for these kinds of considerations. ∎

Proposition 6.3.5 (Piecewise volume high and low pass filters $H_{\eta k}^{\Omega,\mathrm{pw}}$ and $L_{\eta k}^{\Omega,\mathrm{pw}}$)**.** *Let Ω be a bounded Lipschitz domain, let the interface Γ_i be as in Assumption 6.2.1 and $\eta > 1$. Then there exist operators $H_{\eta k}^{\Omega,\mathrm{pw}}\colon L^2(\Omega) \to L^2(\Omega)$ and $L_{\eta k}^{\Omega,\mathrm{pw}}\colon L^2(\Omega) \to L^2(\Omega)$ such that*

$H^{\Omega,\mathrm{pw}}_{\eta k} f + L^{\Omega,\mathrm{pw}}_{\eta k} f = f$ *for every* $f \in L^2(\Omega)$. *Furthermore, for* $0 \le s' \le s$ *the operator* $H^{\Omega,\mathrm{pw}}_{\eta k}$ *satisfies*

$$\|H^{\Omega,\mathrm{pw}}_{\eta k} f\|_{s',\Omega\setminus\Gamma_i} \lesssim (\eta k)^{s'-s} \|f\|_{s,\Omega\setminus\Gamma_i} \qquad \forall f \in H^s(\Omega \setminus \Gamma_i),$$

with hidden constant independent of η *and* k*. Finally, the low pass filter* $L^{\Omega,\mathrm{pw}}_{\eta k}$ *satisfies*

$$\|\nabla^p L^{\Omega,\mathrm{pw}}_{\eta k} f\|_{0,\Omega\setminus\Gamma_i} \lesssim (\eta k)^{p-s} \|f\|_{s,\Omega\setminus\Gamma_i} \qquad \forall f \in H^s(\Omega \setminus \Gamma_i),\, \forall p \in \mathbb{N}_0,\, p \ge s.$$

Proof. The filters are constructed piecewise using the filters given in Proposition 6.3.2. In fact, it is easy to see that the operator given by $(H^{\Omega,\mathrm{pw}}_{\eta k} f)\Big|_\omega := H^\omega_{\eta k}(f\chi_\omega)$, where χ_ω denotes the indicator function on each component $\omega \subset \Omega$, satisfies the asserted properties, due to the properties of the high and low pass filters given in Proposition 6.3.2. □

Lemma 6.3.6 (Unified Contraction Argument)**.** *Let the Assumptions 6.2.1 (smoothness of the* Ω*,* Γ *and* Γ_i*) , 6.2.4 (assumptions on the minus problem) and 6.2.5 (assumptions on the plus problem) be satisfied. Let* $q \in (0,1)$ *be given. Then for every* $f \in L^2(\Omega)$ *and* $g \in H^{1/2}(\Gamma)$ *the function* $u = S^-_k(f,g)$ *can be written as* $u = u_F + u_A + S^-_k(\tilde f, \tilde g)$*, where*

$$\|u_F\|_{2,\Omega\setminus\Gamma_i} + k\|u_F\|_{1,t,k} \lesssim \|f\|_{0,\Omega} + \|g\|_{1/2,\Gamma}, \tag{6.14}$$

$$\|u_A\|_{1,t,k} \lesssim C^-_{\mathrm{sol},k}(1 + C_{\mathrm{ana},k}k^{-1})(\|f\|_{0,\Omega} + \|g\|_{1/2,\Gamma}). \tag{6.15}$$

Furthermore, the function u_A *is given by the solution operator* S^-_k *applied to analytic data via*

$$u_A = S^-_k(L^\Omega_{\eta k} f, L^\Gamma_{\eta k} g) + S^-_k(A_\Omega u_F, A_\Gamma u_F)$$

and there exists a constant $\gamma > 0$ *independent of* k *such that*

$$\|\nabla^n u_A\|_{0,\Omega\setminus\Gamma_i} \lesssim C^-_{\mathrm{sol},k} k^{-1}(1+C_{\mathrm{ana},k}k^{-1})\gamma^n \max\{k,n\}^n(\|f\|_{0,\Omega}+\|g\|_{1/2,\Gamma}) \quad \forall n \ge 2. \tag{6.16}$$

Finally, the data $\tilde f$*,* $\tilde g$ *contracts, i.e.,*

$$\|\tilde f\|_{0,\Omega} + \|\tilde g\|_{1/2,\Gamma} \le q(\|f\|_{0,\Omega} + \|g\|_{1/2,\Gamma}). \tag{6.17}$$

Proof. The solution $u = S^-_k(f,g)$ of

$$\begin{aligned} L^-_k u &= f \quad \text{in } \Omega,\\ \partial_n u - T^-_{k,\Gamma} u &= g \quad \text{on } \Gamma \end{aligned}$$

is split as follows

$$u = \underbrace{S^+_k(H^\Omega_{\eta k} f, H^\Gamma_{\eta k} g)}_{=:u_F} + \underbrace{S^-_k(L^\Omega_{\eta k} f, L^\Gamma_{\eta k} g)}_{=:u_{A,\mathrm{I}}} + r_{\mathrm{I}}.$$

Due to P.3 the remainder r_{I} satisfies

$$\begin{aligned} L^-_k r_{\mathrm{I}} &= -(L^-_k - L^+_k)u_F = R_\Omega u_F + A_\Omega u_F \quad \text{in } \Omega,\\ \partial_n r_{\mathrm{I}} - T^-_{k,\Gamma} r_{\mathrm{I}} &= (T^-_{k,\Gamma} - T^+_{k,\Gamma})u_F = R_\Gamma u_F + A_\Gamma u_F \quad \text{on } \Gamma. \end{aligned}$$

We again split r_{I} as follows

$$r_{\mathrm{I}} = \underbrace{S_k^-(A_\Omega u_F, A_\Gamma u_F)}_{=:u_{A,\mathrm{II}}} + r,$$

where r satisfies

$$\begin{aligned} L_k^- r &= R_\Omega u_F =: \tilde{f} \quad \text{in } \Omega, \\ \partial_n r - T_{k,\Gamma}^- r &= R_\Gamma u_F =: \tilde{g} \quad \text{on } \Gamma. \end{aligned}$$

The final splitting now reads

$$\begin{aligned} u &= u_F + u_A + S_k^-(\tilde{f}, \tilde{g}), \\ u_F &= S_k^+(H_{\eta k}^\Omega f, H_{\eta k}^\Gamma g), \\ u_A &= S_k^-(L_{\eta k}^\Omega f, L_{\eta k}^\Gamma g) + S_k^-(A_\Omega u_F, A_\Gamma u_F), \\ \tilde{f} &= R_\Omega u_F, \\ \tilde{g} &= R_\Gamma u_F. \end{aligned}$$

We first prove the estimates for u_F. Since $u_F = S_k^+(H_{\eta k}^\Omega f, H_{\eta k}^\Gamma g)$ with estimate (6.11) in Lemma 6.3.1 we find for any $0 < \varepsilon < 1/2$

$$\begin{aligned} k\|u_F\|_{1,t,k} &\lesssim k(k^{\varepsilon-1}\|H_{\eta k}^\Omega f\|_{\widetilde{H}^{-\varepsilon}(\Omega)} + k^{-1/2}\|H_{\eta k}^\Gamma g\|_{0,\Gamma}) \\ &\lesssim k^\varepsilon \|H_{\eta k}^\Omega f\|_{\widetilde{H}^{-\varepsilon}(\Omega)} + k^{1/2}\|H_{\eta k}^\Gamma g\|_{0,\Gamma} \\ &\lesssim \eta^{-\varepsilon}\|f\|_{0,\Omega} + \eta^{-1/2}\|g\|_{1/2,\Gamma} \\ &\leq \eta^{-\varepsilon}(\|f\|_{0,\Omega} + \|g\|_{1/2,\Gamma}). \end{aligned} \tag{6.18}$$

Furthermore, with (6.9) in P.7 we find

$$\|u_F\|_{2,\Omega\setminus\Gamma_i} \lesssim \|H_{\eta k}^\Omega f\|_{0,\Omega} + \|H_{\eta k}^\Gamma g\|_{1/2,\Gamma} + k^{1/2}\|H_{\eta k}^\Gamma g\|_{0,\Gamma} \lesssim \|f\|_{0,\Omega} + \|g\|_{1/2,\Gamma},$$

which proves the regularity estimate (6.14). We proceed with the proof of the contraction estimates. We prove the remaining estimates for the case $R_\Omega = k^{2-s}R_{\Omega,s}$ for some $0 \leq s \leq 1$ and $R_\Gamma = k^{3/2-r}R_{\Gamma,r}$ for some $0 \leq r \leq 1/2$. The general case $R_\Omega = \sum_{i=1}^n k^{2-s_i}R_{\Omega,s_i}$ and $R_\Gamma = \sum_{i=1}^n k^{3/2-s_i}R_{\Gamma,s_i}$ follows by the triangle inequality. We use the estimates in P.4 and P.5 as well as (6.18). We have

$$\begin{aligned} \|\tilde{f}\|_{0,\Omega} &= \|k^{2-s}R_{\Omega,s}u_F\|_{0,\Omega} \lesssim k^{2-s}\|u_F\|_{s,\Omega} \lesssim k^{2-s}k^{s-1}\|u_F\|_{1,k,\Omega} \\ &= k\|u_F\|_{1,k,\Omega} \lesssim \eta^{-\varepsilon}(\|f\|_{0,\Omega} + \|g\|_{1/2,\Gamma}). \end{aligned}$$

For the data $\tilde{g}$ we have

$$\begin{aligned} \|\tilde{g}\|_{1/2,\Gamma} &= \|k^{3/2-r}R_{\Gamma,r}u_F\|_{1/2,\Gamma} \lesssim k^{3/2-r}\|u_F\|_{r,\Gamma} \lesssim k^{3/2-r}\|u_F\|_{r+1/2,\Omega} \\ &\lesssim k^{3/2-r}k^{r+1/2-1}\|u_F\|_{1,k,\Omega} \lesssim k\|u_F\|_{1,k,\Omega} \lesssim \eta^{-\varepsilon}(\|f\|_{0,\Omega} + \|g\|_{1/2,\Gamma}). \end{aligned}$$

Since the parameter $\eta > 1$ is still at our disposal, we can choose it such that the contraction estimate (6.17) is satisfied. Finally, the estimate for u_A follows by M.3 as well as the Lemmas 6.3.2 and 6.3.3. We find

$$\|S_k^-(L_{\eta k}^\Omega f, L_{\eta k}^\Gamma g)\|_{1,t,k} \lesssim C_{\mathrm{sol},k}^-(\|L_{\eta k}^\Omega f\|_{0,\Omega} + \|L_{\eta k}^\Gamma g\|_{0,\Gamma}) \lesssim C_{\mathrm{sol},k}^-(\|f\|_{0,\Omega} + \|g\|_{1/2,\Gamma}),$$

$$\|S_k^-(A_\Omega u_F, A_\Gamma u_F)\|_{1,t,k} \lesssim C^-_{\mathrm{sol},k}(\|A_\Omega u_F\|_{0,\Omega}+\|A_\Gamma u_F\|_{0,\Gamma}) \lesssim C^-_{\mathrm{sol},k}(C_{A,\Omega,k}+C_{A,\Gamma,k})\|u_F\|_{1,t,k},$$

which together with (6.18) yields (6.15). Finally, by M.4 we find

$$\|\nabla^n S_k^-(L^\Omega_{\eta k} f, L^\Gamma_{\eta k} g)\|_{0,\Omega\setminus\Gamma_i} \lesssim C^-_{\mathrm{sol},k} k^{-1}\gamma^n \max\{k,n\}^n(\|f\|_{0,\Omega}+\|g\|_{1/2,\Gamma}),$$

$$\|\nabla^n S_k^-(A_\Omega u_F, A_\Gamma u_F)\|_{0,\Omega\setminus\Gamma_i} \lesssim C^-_{\mathrm{sol},k} k^{-1}\gamma^n \max\{k,n\}^n(C_{A,\Omega,k}+C_{A,\Gamma,k})\|u_F\|_{1,t,k}$$

for all $n \geq 2$, which again together with (6.18) yields (6.16). □

Remark 6.3.7 (Contraction argument as parametrix)**.** Assume for simplicity $A_\Omega u = 0$ and $A_\Gamma u = 0$. Then the splitting performed in the proof of Lemma 6.3.6 reads

$$S_k^-(f,g) = S_k^+(H^\Omega_{\eta k} f, H^\Gamma_{\eta k} g) + S_k^-(L^\Omega_{\eta k} f, L^\Gamma_{\eta k} g) + r.$$

Consequently, we find

$$S_k^-(H^\Omega_{\eta k} f, H^\Gamma_{\eta k} g) = S_k^+(H^\Omega_{\eta k} f, H^\Gamma_{\eta k} g) + r, \tag{6.19}$$

which expresses, that S_k^+ is an approximate solution operator for L_k^- on high-frequency data. ■

We now perform a similar contraction argument for data in higher order Sobolev spaces:

Lemma 6.3.8 (Unified Contraction Argument, higher Sobolev regularity)**.** *Assume the hypothesis of Lemma 6.3.6. Let $q \in (0,1)$ and $s > 0$ be given. Additionally, let Assumption 6.2.6 be satisfied. Then for every $f \in H^s(\Omega\setminus\Gamma_i)$ and $g \in H^{s+1/2}(\Gamma)$ the function $u = S_k^-(f,g)$ can be written as $u = u_F + u_A + S_k^-(\tilde f, \tilde g)$, where*

$$\|u_F\|_{s+2,\Omega\setminus\Gamma_i} + k^{s+1}\|u_F\|_{1,t,k} \lesssim \|f\|_{s,\Omega\setminus\Gamma_i} + \|g\|_{s+1/2,\Gamma}, \tag{6.20}$$

$$\|u_A\|_{1,t,k} \lesssim C^-_{\mathrm{sol},k}(1 + C_{\mathrm{ana},k}k^{-1})(\|f\|_{0,\Omega} + \|g\|_{1/2,\Gamma}). \tag{6.21}$$

Furthermore, the function u_A is given by the solution operator S_k^- applied to analytic data via

$$u_A = S_k^-(L^{\Omega,\mathrm{pw}}_{\eta k} f, L^\Gamma_{\eta k} g) + S_k^-(A_\Omega u_F, A_\Gamma u_F)$$

and there exists a constant $\gamma > 0$ independent of k such that

$$\|\nabla^n u_A\|_{0,\Omega\setminus\Gamma_i} \lesssim C^-_{\mathrm{sol},k} k^{-1}(1+C_{\mathrm{ana},k}k^{-1})\gamma^n \max\{k,n\}^n(\|f\|_{0,\Omega}+\|g\|_{1/2,\Gamma}) \quad \forall n \geq 2. \tag{6.22}$$

Finally, the data $\tilde f$, $\tilde g$ contracts, i.e.,

$$\|\tilde f\|_{s,\Omega\setminus\Gamma_i} + \|\tilde g\|_{s+1/2,\Gamma} \leq q(\|f\|_{s,\Omega\setminus\Gamma_i} + \|g\|_{s+1/2,\Gamma}). \tag{6.23}$$

Proof. The main difference compared to the proof of Lemma 6.3.6 is the application of the piecewise high and low pass filters. We only highlight the crucial differences. The solution $u = S_k^-(f,g)$ is split as follows

$$u = \underbrace{S_k^+(H^{\Omega,\mathrm{pw}}_{\eta k} f, H^\Gamma_{\eta k} g)}_{=:u_F} + \underbrace{S_k^-(L^{\Omega,\mathrm{pw}}_{\eta k} f, L^\Gamma_{\eta k} g)}_{=:u_{A,\mathrm{I}}} + r_{\mathrm{I}}.$$

With the same notation and lines of proof as in Lemma 6.3.6 the final splitting now reads

$$\begin{aligned} u &= u_F + u_A + S_k^-(\tilde f,\tilde g),\\ u_F &= S_k^+(H_{\eta k}^{\Omega,\mathrm{pw}} f, H_{\eta k}^{\Gamma} g),\\ u_A &= S_k^-(L_{\eta k}^{\Omega,\mathrm{pw}} f, L_{\eta k}^{\Gamma} g) + S_k^-(A_\Omega u_F, A_\Gamma u_F),\\ \tilde f &= R_\Omega u_F,\\ \tilde g &= R_\Gamma u_F. \end{aligned}$$

We first prove the estimates for u_F. Since $u_F = S_k^+(H_{\eta k}^{\Omega,\mathrm{pw}} f, H_{\eta k}^{\Gamma} g)$ with estimate (6.11) in Lemma 6.3.1 we find with the choice $\varepsilon = 0$

$$\begin{aligned} k\|u_F\|_{1,t,k} &\lesssim k(k^{-1}\|H_{\eta k}^{\Omega,\mathrm{pw}} f\|_{0,\Omega} + k^{-1/2}\|H_{\eta k}^{\Gamma} g\|_{0,\Gamma})\\ &\lesssim \|H_{\eta k}^{\Omega,\mathrm{pw}} f\|_{0,\Omega} + k^{1/2}\|H_{\eta k}^{\Gamma} g\|_{0,\Gamma}\\ &\lesssim (\eta k)^{-s}\|f\|_{s,\Omega\setminus\Gamma_i} + \eta^{-s-1/2}k^{-s}\|g\|_{s+1/2,\Gamma}\\ &\lesssim (\eta k)^{-s}(\|f\|_{s,\Omega\setminus\Gamma_i} + \|g\|_{s+1/2,\Gamma}) \end{aligned}$$

and with PS.3 we find

$$\begin{aligned} \|u_F\|_{s+2,\Omega\setminus\Gamma_i} &\lesssim \|H_{\eta k}^{\Omega,\mathrm{pw}} f\|_{s,\Omega\setminus\Gamma_i} + k^s\|H_{\eta k}^{\Omega,\mathrm{pw}} f\|_{0,\Omega} + \|H_{\eta k}^{\Gamma} g\|_{s+1/2,\Gamma} + k^{s+1/2}\|H_{\eta k}^{\Gamma} g\|_{0,\Gamma}\\ &\lesssim \|f\|_{s,\Omega\setminus\Gamma_i} + \|g\|_{s+1/2,\Gamma}, \end{aligned}$$

which proves the regularity estimate (6.20). We proceed with the proof of the contraction estimates. Note that by interpolation we have for any $t \in [0, s+2]$ the estimate

$$\|u_F\|_{t,\Omega\setminus\Gamma_i} \lesssim \|u_F\|_{0,\Omega}^{\frac{s+2-t}{s+2}}\|u_F\|_{s+2,\Omega\setminus\Gamma_i}^{\frac{t}{s+2}} \lesssim \eta^{-s\frac{s+2-t}{s+2}}k^{t-s-2}(\|f\|_{s,\Omega\setminus\Gamma_i} + \|g\|_{s+1/2,\Gamma}).$$

Again, we prove the remaining estimates for the case $R_\Omega = k^{s+2-t}R_{\Omega,t}$ for some $0 \le t < s+2$ and $R_\Gamma = k^{s+3/2-r}R_{\Gamma,r}$ for some $0 \le r < s+3/2$. The general case $R_\Omega = \sum_{i=1}^n k^{s+2-s_i}R_{\Omega,s_i}$ and $R_\Gamma = \sum_{i=1}^n k^{s+3/2-s_i}R_{\Gamma,s_i}$ follows by the triangle inequality. We have

$$\|\tilde f\|_{s,\Omega\setminus\Gamma_i} = \|k^{s+2-t}R_{\Omega,t}u_F\|_{s,\Omega\setminus\Gamma_i} \lesssim k^{s+2-t}\|u_F\|_{t,\Omega\setminus\Gamma_i} \lesssim \eta^{-s\frac{s+2-t}{s+2}}(\|f\|_{s,\Omega\setminus\Gamma_i} + \|g\|_{s+1/2,\Gamma}).$$

Note that since $0 \le t < s+2$ the exponent of η is negative. For the data $\tilde g$ we have

$$\begin{aligned} \|\tilde g\|_{s+1/2,\Gamma} &= \|k^{s+3/2-r}R_{\Gamma,r}u_F\|_{s+1/2,\Gamma} \lesssim k^{s+3/2-r}\|u_F\|_{r,\Gamma} \lesssim k^{s+3/2-r}\|u_F\|_{r+1/2,\Omega\setminus\Gamma_i}\\ &\lesssim \eta^{-s\frac{s+2-r-1/2}{s+2}}(\|f\|_{s,\Omega\setminus\Gamma_i} + \|g\|_{s+1/2,\Gamma}). \end{aligned}$$

Note that since $0 \le r < s+3/2$ the exponent of η is negative. Since the parameter $\eta > 1$ is still at our disposal, we can choose it such that the contraction estimate (6.23) is satisfied. The remaining estimates are proven in the same manner as in Lemma 6.3.6 and are therefore omitted. □

Remark 6.3.9. The contraction argument presented in Lemma 6.3.8 allows for the operator R_Ω to act on broken Sobolev spaces with a wider range than compared to the result in Lemma 6.3.6. Comparison of the proofs of Lemma 6.3.8 and Lemma 6.3.6 shows that one can allow in Assumption 6.2.5 the following:

- $R_\Omega = \sum_{i=1}^n k^{2-s_i} R_{\Omega,s_i}$ with $R_{\Omega,s_i}\colon H^{s_i}(\Omega\setminus\Gamma_i) \to L^2(\Omega)$ being bounded linear operators with $\|R_{\Omega,s_i} u\|_{0,\Omega} \lesssim \|u\|_{s_i,\Omega\setminus\Gamma_i}$, $0 \le s_i < 2$.
- $R_\Gamma = \sum_{i=1}^n k^{3/2-s_i} R_{\Gamma,s_i}$ with $R_{\Gamma,s_i}\colon H^{s_i}(\Gamma) \to H^{1/2}(\Gamma)$ being bounded linear operators with $\|R_{\Gamma,s_i} u\|_{1/2,\Gamma} \lesssim \|u\|_{s_i,\Gamma}$, $0 \le s_i < 3/2$. ■

Theorem 6.3.10. *(Unified Iteration Argument) Assume the hypothesis of Lemma 6.3.6. Then there exists a constant $\gamma > 0$ independent of k such that for every $f \in L^2(\Omega)$ and $g \in H^{1/2}(\Gamma)$ the function $u = S_k^-(f,g)$ can be written as $u = u_F + u_A$, where*

$$\|u_F\|_{2,\Omega\setminus\Gamma_i} + k\|u_F\|_{1,t,k} \lesssim \|f\|_{0,\Omega} + \|g\|_{1/2,\Gamma}, \tag{6.24}$$

$$\|u_A\|_{1,t,k} \lesssim C^-_{\text{sol},k}(1 + C_{\text{ana},k}k^{-1})(\|f\|_{0,\Omega} + \|g\|_{1/2,\Gamma}), \tag{6.25}$$

$$\|\nabla^n u_A\|_{0,\Omega\setminus\Gamma_i} \lesssim C^-_{\text{sol},k} k^{-1}(1 + C_{\text{ana},k}k^{-1})\gamma^n \max\{k,n\}^n(\|f\|_{0,\Omega} + \|g\|_{1/2,\Gamma}) \tag{6.26}$$

for all $n \ge 2$.

Proof. The proof iterates the contraction argument in Lemma 6.3.6. To that end, let $f^{(0)} := f$ and $g^{(0)} := g$. By Lemma 6.3.6 we can split $S_k^-(f^{(0)}, g^{(0)})$ as follows

$$S_k^-(f^{(0)}, g^{(0)}) = u_F^{(0)} + u_A^{(0)} + S_k^-(f^{(1)}, g^{(1)})$$

with

$$\|f^{(1)}\|_{0,\Omega} + \|g^{(1)}\|_{1/2,\Gamma} \le q(\|f^{(0)}\|_{0,\Omega} + \|g^{(0)}\|_{1/2,\Gamma}).$$

We iteratively define the function sequences $u_F^{(i)}$, $u_A^{(i)}$ and $f^{(i)}$, $g^{(i)}$ for $i \in \mathbb{N}$. Due to contraction we find

$$\|f^{(i)}\|_{0,\Omega} + \|g^{(i)}\|_{1/2,\Gamma} \le q^i(\|f^{(0)}\|_{0,\Omega} + \|g^{(0)}\|_{1/2,\Gamma})$$

and therefore

$$u = \underbrace{\sum_{i\in\mathbb{N}_0} u_F^{(i)}}_{=:u_F} + \underbrace{\sum_{i\in\mathbb{N}_0} u_A^{(i)}}_{=:u_A}$$

with u_F and u_A being well-defined and satisfying the asserted estimates due to a geometric series argument. We showcase this argument for the $H^2(\Omega\setminus\Gamma_i)$ norm of u_F:

$$\|u_F\|_{2,\Omega\setminus\Gamma_i} \le \sum_{i\in\mathbb{N}_0} \|u_F^{(i)}\|_{2,\Omega\setminus\Gamma_i} \lesssim \sum_{i\in\mathbb{N}_0} \|f^{(i)}\|_{0,\Omega} \le \sum_{i\in\mathbb{N}_0} q^i \|f\|_{0,\Omega} \lesssim \|f\|_{0,\Omega}.$$

The other estimates follow analogously. □

Theorem 6.3.11. *(Unified Iteration Argument, higher Sobolev regularity) Assume the hypothesis of Lemma 6.3.8. Then there exists a constant $\gamma \ge 0$ independent of k such that for every $f \in L^2(\Omega)$ and $g \in H^{1/2}(\Gamma)$ the function $u = S_k^-(f,g)$ can be written as $u = u_F + u_A$, where*

$$\|u_F\|_{s+2,\Omega\setminus\Gamma_i} + k^{s+1}\|u_F\|_{1,t,k} \lesssim \|f\|_{s,\Omega\setminus\Gamma_i} + \|g\|_{s+1/2,\Gamma}. \tag{6.27}$$

Furthermore, u_A satisfies the same estimates as in Theorem 6.3.10.

Proof. The proof is completely analogous to the one of Theorem 6.3.10 with the application of Lemma 6.3.8 instead of Lemma 6.3.6. □

Remark 6.3.12 (On implicit and explicit constants in Theorems 6.3.10 and 6.3.11)**.** The implicit constants in the estimates of u_F and u_A in the Theorems 6.3.10 and 6.3.11 are independent of k. Note however, that the implicit constants depend on the hidden (k independent) constants in Assumptions 6.2.4, 6.2.5 and 6.2.6. In particular, as in Section 6.5, these implicit constants depend on the coefficients of the differential equation. For the heterogeneous Helmholtz problems in heterogeneous media, as considered in Section 6.5, these implicit constants then depend for example on the index of refraction n. ▪

6.4. Adjoint problem

The proof of quasi-optimality of an abstract Galerkin discretization of problem (6.6) hinges on the approximability of the adjoint problem, i.e., one performs a duality argument. Below, Lemma 6.4.2 will characterize the solution of the adjoint problem. The primal problem reads:

$$\text{Find } u =: S_k^-(f,g) \in H^{1,t}(\Omega,\Gamma) \quad \text{such that} \quad b_k^-(u,v) = (f,v) + \langle g, v\rangle \quad \forall v \in H^{1,t}(\Omega,\Gamma).$$

The corresponding adjoint problem reads:

$$\text{Find } u^\star =: S_k^{-,\star}(f,g) \in H^{1,t}(\Omega,\Gamma) \quad \text{such that} \quad b_k^-(v,u^\star) = (v,f) + \langle v, g\rangle \quad \forall v \in H^{1,t}(\Omega,\Gamma).$$

Assumption 6.4.1 (Quasi-selfadjointness of $T^-_{k,\Omega}$ and $T^-_{k,\Gamma}$)**.** The linear operators $T^-_{k,\Omega}$: $H^{1,t}(\Omega,\Gamma) \to H^{1,t}(\Omega,\Gamma)'$ and $T^-_{k,\Gamma}\colon H^t(\Gamma)\colon \to H^{-t}(\Gamma)$ are quasi-selfadjoint, i.e.,

$$(T^-_{k,\Omega}u,\overline{v}) = (T^-_{k,\Omega}v,\overline{u}) \quad \text{as well as} \quad \langle T^-_{k,\Gamma}u,\overline{v}\rangle = \langle T^-_{k,\Gamma}v,\overline{u}\rangle$$

for all $u, v \in H^{1,t}(\Omega,\Gamma)$.

For the adjoint problem to be again of the same *character* as the primal problem one may assume Assumption 6.4.1.

Lemma 6.4.2 (Adjoint Problems)**.** *Let Assumption 6.4.1 be satisfied. Let* $S_k^-(f,g;A)$ *denote the solution operator specifying the diffusion matrix* A*. Then the adjoint solution operator satisfies* $S_k^{-,\star}(f,g) = \overline{S_k^-(\overline{f},\overline{g};A^T)}$.

Proof. Let $u^\star = S_k^{-,\star}(f,g)$. Therefore, we have

$$(A\nabla v, \nabla u^\star) - (T^-_{k,\Omega}v, u^*) - \langle T^-_{k,\Gamma}v, u^\star\rangle = (v,f) + \langle v, g\rangle \quad \forall v \in H^{1,t}(\Omega,\Gamma).$$

Due to the assumed quasi-selfadjointness in Assumption 6.4.1 we find

$$(A^T\overline{\nabla u^\star}, \overline{\nabla v}) - (T^-_{k,\Omega}\overline{u^\star}, \overline{v}) - \langle T^-_{k,\Gamma}\overline{u^\star}, \overline{v}\rangle = (\overline{f},\overline{v}) + \langle \overline{f}, \overline{v}\rangle \quad \forall v \in H^{1,t}(\Omega,\Gamma).$$

Replacing $\overline{v}$ by v we find $\overline{u^\star} = S_k^-(\overline{f},\overline{g};A^T)$, which yields the result. □

Remark 6.4.3. Loosely speaking Lemma 6.4.2 states that under the assumption of quasi-selfadjointness, see Assumption 6.4.1, the adjoint problem is again of similar character to the primal problem. This in turn is convenient, since the same regularity splitting of Section 6.3 can be performed for both the primal and the adjoint problem. See also [MS11, Lemma 3.1] for similar results in the case of homogeneous media. ▪

6.5. Covered problems

Before turning to the analysis of an abstract Galerkin discretization we first verify that the Assumptions 6.2.4, 6.2.5 and 6.2.6 in Section 6.2 are in fact satisfied for a variety of time-harmonic wave propagation problems. At this point, we stress again that the well-posedness in Assumption M.3 as well as the later assumed polynomial bounds in Assumption 6.6.6 are injected into our theory. Consider the setting of the standard heterogeneous Helmholtz problem, i.e., where the differential operator L_k^- is given by

$$L_k^- u = -\nabla \cdot (A(x)\nabla u) - k^2 n^2(x) u.$$

There is rich literature studying on how the stability constant $C^-_{\mathrm{sol},k}$ depends on the geometry of the domain Ω, the boundary conditions, and the coefficient n and A in the Helmholtz problem: For homogeneous media we refer the reader to [Spe14]. For nontrapping heterogeneous media see [BCFG17, GPS19, GSW20]. The general one-dimensional case is analyzed in [CF15, GS20]. For weak trapping see [CWSGS20]. Finally we mention the weak effect of strong trapping analyzed in [LSW], justifying Assumptions M.3 and 6.6.6.
We present an overview table of the covered problems in Table 6.3.

<table>
<tr><th>Model</th><th>Energy Space</th><th>L_k^-</th><th>L_k^+</th><th>$T^-_{k,\Omega} - T^+_{k,\Omega}$</th><th>$T^-_{k,\Gamma}$</th><th>$T^+_{k,\Gamma}$</th><th>$T^-_{k,\Gamma} - T^+_{k,\Gamma}$</th></tr>
<tr><td>HH + RBC</td><td rowspan="5">$H^1(\Omega)$</td><td rowspan="3">$-\nabla \cdot (A\nabla u) - k^2 n^2 u$</td><td rowspan="4">$-\nabla \cdot (A\nabla u) + k^2 u$</td><td rowspan="3">$-k^2(n^2+1)u$</td><td>$iku$</td><td>0</td><td>$iku$</td></tr>
<tr><td>HH + DtN Sphere</td><td>DtN_k</td><td>DtN_0</td><td>kR_Γ</td></tr>
<tr><td>HH + DtN Γ</td><td>DtN_k</td><td>DtN_0</td><td>$kR_\Gamma + A_\Gamma$</td></tr>
<tr><td>HH + Damping + RBC</td><td>$-\nabla \cdot (A\nabla u) - k^2 n^2 u + ikmu$</td><td>$-k^2(n^2+1)u + ikmu$</td><td>$iku$</td><td>0</td><td>$iku$</td></tr>
<tr><td>HH + PML + RBC</td><td>$-\nabla(A_k^{PML}\nabla u) - k^2 n^2 u$</td><td>$-\nabla(A_k^{PML}\nabla u) + k^2 u$</td><td rowspan="2">$-k^2(n^2+1)u$</td><td>$iku$</td><td>0</td><td>$iku$</td></tr>
<tr><td>HH + second order ABC</td><td>$H^{1,1}(\Omega,\Gamma)$</td><td>$-\nabla \cdot (A\nabla u) - k^2 n^2 u$</td><td>$-\nabla \cdot (A\nabla u) + k^2 u$</td><td>$\alpha\Delta_\Gamma u + kR_\Gamma u$</td><td>$\alpha\Delta_\Gamma u$</td><td>$kR_\Gamma$</td></tr>
</table>

Table 6.3.: Overview of covered problems.

For the problems listed below the choice

$$L_k^+ u = -\nabla \cdot (A\nabla u) + k^2 u$$

suffices. We now turn to the covered problems in detail.

6.5.1. Overview of covered problems

Heterogeneous Helmholtz problem with Robin boundary conditions

As a first example we consider the heterogeneous Helmholtz problem with Robin boundary condition. The problem reads

$$\begin{aligned} -\nabla \cdot (A\nabla u) - k^2 n^2 u &= f \quad \text{in } \Omega, \\ \partial_n u - iknu &= g \quad \text{on } \Gamma. \end{aligned}$$

The variable coefficient A is assumed to map into the class of symmetric positive definite real matrices. Let A be uniformly positive and bounded, i.e.,

$$0 < a_{\min} I \leq A(x) \leq a_{\max} I \qquad \forall x \in \Omega,$$

in the sense of SPD matrices and be homogeneous at the boundary, i.e., $A \equiv I$ on Γ. Furthermore, we assume A to satisfy

$$\|\nabla^p A\|_{L^\infty(\Omega\setminus\Gamma_i)} \leq C_A \gamma_A^p p! \qquad \forall p \in \mathbb{N}_0.$$

Finally, the index of refraction n is uniformly bounded and satisfies the analyticity estimate

$$\|\nabla^p n\|_{L^\infty(\Omega\setminus\Gamma_i)} \leq C_n \gamma_n^p p! \qquad \forall p \in \mathbb{N}_0.$$

For the boundary operator $T^+_{k,\Gamma}$ we may choose $T^+_{k,\Gamma} = 0$. Note that we may assume n to be complex valued. Therefore, volume damping, classically written in the form $-\nabla \cdot (A\nabla u) - k^2 n^2 u + ikmu$ is also included in this setting.

Heterogeneous Helmholtz problem in the full space

The heterogeneous Helmholtz problem with Sommerfeld radiation condition in full space $\mathbb{R}^d$ is to find $U \in H^1_{\mathrm{loc}}(\mathbb{R}^d)$ such that

$$\begin{aligned} -\nabla \cdot (A\nabla U) - k^2 n^2 U &= f && \text{in } \mathbb{R}^d, \\ |\partial_r U - ikU| &= o\left(\|x\|^{\frac{1-d}{2}}\right) && \text{for } \|x\| \to \infty \end{aligned} \tag{6.28}$$

is satisfied in the weak sense. Here, ∂_r denotes the derivative in the radial direction. We assume f, A and n to be local in the sense that there exists a bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$, such that $\operatorname{supp} f \subset \Omega$, $\operatorname{supp}(A - I) \subset \Omega$ and $\operatorname{supp}(n - 1) \subset \Omega$. Problem (6.28) can be reformulated using the Dirichlet-to-Neumann operator $\mathrm{DtN}_k \colon H^{1/2}(\Gamma) \to H^{-1/2}(\Gamma)$, which is given by $g \mapsto \mathrm{DtN}_k g := \partial_n w$, where $w \in H^1_{\mathrm{loc}}(\mathbb{R}^d \setminus \Omega)$ is the unique weak solution to

$$\begin{aligned} -\Delta w - k^2 w &= 0 && \text{in } \mathbb{R}^d \setminus \Omega, \\ w &= g && \text{on } \Gamma, \\ |\partial_r w - ikw| &= o\left(\|x\|^{\frac{1-d}{2}}\right) && \text{for } \|x\| \to \infty. \end{aligned}$$

The model problem is to find $u \in H^1(\Omega)$ such that

$$\begin{aligned} -\nabla \cdot (A\nabla u) - k^2 n^2 u &= f && \text{in } \Omega, \\ \partial_n u - \mathrm{DtN}_k u &= g && \text{on } \Gamma \end{aligned} \tag{6.29}$$

for $f \in L^2(\Omega)$ and $g \in H^{1/2}(\Gamma)$. We make the same assumptions on A and n as in the heterogeneous Helmholtz problem with Robin boundary conditions. The coupling interface Γ between the interior and exterior domain may be chosen to be arbitrary as long as the exterior domain is nontrapping [BSW16, Def. 1.1] and Γ itself is analytic. The operator $T^+_{k,\Gamma}$ can be chosen to be the Dirichlet-to-Neumann operator corresponding to the Laplacian, i.e., $T^+_{k,\Gamma} = \mathrm{DtN}_0$. The splitting of $T^-_{k,\Gamma} - T^+_{k,\Gamma}$ is not obvious, see Lemma 6.5.12. The Dirichlet-to-Neumann map DtN_0 is again given by the Neumann trace of the solution to an exterior problem, i.e., the operator $\mathrm{DtN}_0 \colon H^{1/2}(\Gamma) \to H^{-1/2}(\Gamma)$ is given by $g \mapsto \mathrm{DtN}_0 g := \partial_n w$, where $w \in H^1_{\mathrm{loc}}(\mathbb{R}^d \setminus \Omega)$ is the unique weak solution to

$$\begin{aligned}
-\Delta w &= 0 && \text{in } \mathbb{R}^d \setminus \Omega,\\
w &= g && \text{on } \Gamma,\\
w(x) &= a_\infty + b_\infty \log \|x\| + o(1) && \text{for } \|x\| \to \infty \text{ in spatial dimension 2},\\
w(x) &= O\left(\|x\|^{-1}\right) && \text{for } \|x\| \to \infty \text{ in spatial dimension 3},
\end{aligned}$$

where we can fix either a_∞ or b_∞. We may choose $a_\infty = 0$. The constant b_∞ is then determined as part of the problem such that the exterior Calderón identities (6.69) hold. For further discussion see [Era12] as well as [McL00, Thm. 8.9]. In the following we will refer to the asymptotic condition at infinity simply as *radiation condition.*

Heterogeneous Helmholtz problem with second order ABCs

On a sphere in spatial dimension $d = 2$ there are formulas available for second order absorbing boundary conditions. Our theory covers those in the style of Bayliss-Gunzberger-Turkel [BGT82], Enquist-Majda [EM77, EM79] and Feng [Fen84], see [Ihl98, Sec. 3.3.3, Table 3.2] for a comprehensive comparison. We denote by Δ_Γ and ∇_Γ the surface Laplacian and the surface gradient, respectively. The model problem for a heterogeneous Helmholtz equation with second order absorbing boundary conditions is

$$\begin{aligned}
-\nabla \cdot (A\nabla u) - k^2 n^2 u &= f \quad \text{in } \Omega,\\
\partial_n u - T^-_{k,\Gamma} u &= g \quad \text{on } \Gamma,
\end{aligned} \tag{6.30}$$

with $T^-_{k,\Gamma}$ realizing second order absorbing boundary conditions. In fact all of the mentioned ABCs can be cast into the form

$$T^-_{k,\Gamma} u = \beta u + \alpha \Delta_\Gamma u,$$

with the following specifications for α:

BGT	EM	F
$\alpha = -\frac{1+ik}{2(1+k^2)}$	$\alpha = \frac{1+ik}{2k^2}$	$\alpha = -\frac{i}{2k}$

The parameter β is k-dependent and satisfies $|\beta| \sim k$, for the precise k-dependence see again [Ihl98, Sec.3.3.3, Table 3.2]. In the above models the coefficient α is such that

$$\mathrm{Im}\alpha \neq 0 \quad \text{and} \quad \mathrm{Im}\alpha \sim \frac{1}{k} \quad \text{and} \quad |\mathrm{Re}\alpha| \lesssim \frac{1}{k^2} \tag{6.31}$$

for $k \geq k_0 > 0$. We make the same assumptions on A and n as in the heterogeneous Helmholtz problem with Robin boundary conditions. We will see that (6.31) allows for application of our framework. At this point, it is worth noting that our theory covers general problems of this form, however, it is not clear how to formulate second order ABCs on arbitrary domains. The operator $T^+_{k,\Gamma}$ can be chosen to be the leading term in the seconder order ABCs, i.e., $T^+_{k,\Gamma} = \alpha \Delta_\Gamma$.

Perfectly Matched Layers

Consider again the setup of the full space problem. Enclosing the heterogeneities in a sufficiently large ball and using the method of perfectly matched layers one again arrives at a heterogeneous Helmholtz problem of the form

$$\begin{aligned} -\nabla \cdot (A_{\mathrm{PML},k} \nabla u) - k^2 n^2_{\mathrm{PML},k} u &= f \quad \text{in } \Omega, \\ \partial_n u - iku &= g \quad \text{on } \Gamma, \end{aligned}$$

see [CM98, Sec. 3]. In this case, the index of refraction $n^2_{\mathrm{PML},k}$ is homogeneous at the boundary and again piecewise analytic. The matrix-valued function $A_{\mathrm{PML},k}$ is also piecewise analytic, but not positive definite. However, inspection of the proof of [CM98, Thm. 2] shows that in fact

$$\mathrm{Re}(A_{\mathrm{PML},k} \nabla u, \nabla u) \gtrsim \|\nabla u\|^2_{0,\Omega},$$

which in turn makes our theory applicable.

6.5.2. Verification of assumptions

We now turn to the verification of the nontrivial assumptions made in Section 6.2 for the problems discussed in Subsection 6.5.1. We will see that for the model problems presented above the operator $T^+_{k,\Gamma}$ actually satisfies a stronger assumption. For ease of reference we introduce the following

Assumption 6.5.1 (Coercivity of $T^+_{k,\Gamma}$)**.** Let $t \geq 0$. There exists $\sigma \in \mathbb{C}$ such that $\mathrm{Re}\,\sigma > 0$ and $|\sigma| = 1$ such that

$$-\mathrm{Re}(\sigma \langle T^+_{k,\Gamma} u, u \rangle) \geq 0 \qquad \forall u \in H^{1,t}(\Omega, \Gamma)$$

if $t > 1/2$, additionally

$$-\mathrm{Re}(\sigma \langle T^+_{k,\Gamma} u, u \rangle) \gtrsim k^{-2t+1} |u|^2_{t,\Gamma} \qquad \forall u \in H^{1,t}(\Omega, \Gamma).$$

Lemma 6.5.2. *Let A map into the class of symmetric positive definite real matrices. Let A be uniformly positive, i.e.,*

$$0 < a_{\min} I \leq A(x) \qquad \forall x \in \Omega,$$

in the sense of SPD matrices. Then under Assumption 6.5.1 (coercivity of $T^+_{k,\Gamma}$) also P.2 (coercivity of b^+_k) in Assumption 6.2.5 holds with the choice $L^+_k u = -\nabla \cdot (A \nabla u) + k^2 u$.

Proof. Trivial. □

Lemma 6.5.3 (Continuity and coercivity of b^+_k)**.** *Let A map into the class of symmetric positive definite real matrices. Let A be uniformly positive and bounded, i.e.,*

$$0 < a_{\min} I \leq A(x) \leq a_{\max} I \qquad \forall x \in \Omega,$$

in the sense of SPD matrices. Then with

$$L^+_k u = -\nabla \cdot (A \nabla u) + k^2 u,$$

and any boundary operators $T^+_{k,\Gamma}$ in Table 6.3 the conditions P.1 (continuity of b^+_k) and P.2 (coercivity of b^+_k) in Assumption 6.2.5 are satisfied. In the case of second order ABCs let α be as in (6.31)*. The same holds true, if A is only bounded and satisfies*

$$\operatorname{Re}(A\nabla u, \nabla u) \gtrsim \|\nabla u\|^2_{0,\Omega}$$

for all $u \in H^1(\Omega)$ with the choice $T^+_{k,\Gamma} = 0$.

Proof. **Robin boundary conditions**: The Robin case corresponds to the choice $t = 1/2$ and $T^+_{k,\Gamma}u = 0$. Obviously, $T^+_{k,\Gamma} \colon H^{1/2}(\Gamma) \to H^{-1/2}(\Gamma)$ is a bounded linear operator. Furthermore, $T^+_{k,\Gamma}$ trivially satisfies Assumption 6.5.1 and therefore, together with Lemma 6.5.2 we find P.2 with $\sigma = 1$ to be satisfied. Finally, P.1 follows since A is uniformly bounded.
Full space: The full space problem corresponds to the choice $t = 1/2$ and $T^+_{k,\Gamma}u = \mathrm{DtN}_0 u$. The Dirichlet-to-Neumann map $\mathrm{DtN}_0 \colon H^{1/2}(\Gamma) \to H^{-1/2}(\Gamma)$ is a bounded linear operator. Furthermore, we have $-\langle \mathrm{DtN}_0 u, u\rangle \geq 0$, see Item (i) in Lemma 6.5.12. Hence, Assumption 6.5.1 is satisfied. Again, together with Lemma 6.5.2 we find P.2 with $\sigma = 1$ to be satisfied. Finally, P.1 follows readily since A is uniformly bounded.
Second order ABCs: The second order absorbing boundary conditions correspond to the choice $t = 1$ and $T^+_{k,\Gamma}u = \alpha\Delta_\Gamma u$. The Laplace-Beltrami operator $\Delta_\Gamma \colon H^1(\Gamma) \to H^{-1}(\Gamma)$ is a bounded linear operator since Γ has no boundary. Furthermore, we have $-\langle \alpha\Delta_\Gamma u, u\rangle = \alpha\langle \nabla_\Gamma u, \nabla_\Gamma u\rangle$. We verify that Assumption 6.5.1 is satisfied. To that end, note, that

$$-\operatorname{Re}(\sigma\langle T^+_{k,\Gamma}u, u\rangle) = \operatorname{Re}(\sigma\alpha)\langle \nabla_\Gamma u, \nabla_\Gamma u\rangle = \operatorname{Re}(\sigma\alpha)|u|^2_{1,\Gamma} \overset{?}{\gtrsim} k^{-1}|u|^2_{1,\Gamma}.$$

Since $\operatorname{Im}\alpha \neq 0$ for all $k \geq k_0 > 0$, $\operatorname{Im}\alpha \sim k^{-1}$ and $|\operatorname{Re}\alpha| \lesssim k^{-2}$, it is easy to see that a σ as in 6.5.1 exists. Again, together with Lemma 6.5.2 we find P.2 to be satisfied. Finally, P.1 follows readily since A is uniformly bounded.
PML: The PML corresponds again to the choice $t = 1/2$ and $T^+_{k,\Gamma}u = 0$, with a complex matrix-valued function satisfying

$$\operatorname{Re}(A\nabla u, \nabla u) \gtrsim \|\nabla u\|^2_{0,\Omega}$$

for all $u \in H^1(\Omega)$. It is easy to see that again P.1 and P.2 in Assumption 6.2.5 are satisfied. □

A crucial ingredient in our analysis is the H^2 shift of the solution operator S^+_k. In the following lemma we verify this H^2 shift for the presented model problems. Especially the shift property for problems involving second order ABCs are of independent interest. For simplicity we assume in Lemma 6.5.4 that Assumption 6.2.1 (smoothness assumption on Ω) is satisfied. However, it is worth noting that the results also hold for example for C^2 boundary Γ and interface Γ_i.

Lemma 6.5.4 (H^2 regularity shift of S^+_k)**.** *Let Assumption 6.2.1 be satisfied. Let A map into the class of symmetric positive definite real matrices. Let A be uniformly positive and bounded, i.e.,*

$$0 < a_{\min}I \leq A(x) \leq a_{\max}I \qquad \forall x \in \Omega,$$

in the sense of SPD matrices and $C^1(\Omega \setminus \Gamma_i)$, with one-sided continuous extensions. Then, with

$$L_k^+ u = -\nabla \cdot (A\nabla u) + k^2 u,$$

and any boundary operators $T^+_{k,\Gamma}$ in Table 6.3 the condition P.7 in Assumption 6.2.5 is satisfied, i.e., for $f \in L^2(\Omega)$ and $g \in H^{1/2}(\Gamma)$ the solution $S_k^+(f,g) = w \in H^2(\Omega \setminus \Gamma_i)$ and the estimate

$$\|w\|_{2,\Omega\setminus\Gamma_i} \lesssim \|f\|_{0,\Omega} + \|g\|_{1/2,\Gamma} + k^{1/2}\|g\|_{0,\Gamma} \tag{6.32}$$

holds, in the case of second order ABCs let α be as in (6.31)*. The same holds true, if A is only bounded, in $C^1(\Omega \setminus \Gamma_i)$, with one-sided continuous extensions and satisfies*

$$\mathrm{Re}(A\nabla u, \nabla u) \gtrsim \|\nabla u\|^2_{0,\Omega}$$

for all $u \in H^1(\Omega)$ with the choice $T^+_{k,\Gamma} = 0$.

Proof. **Robin boundary conditions**: The Robin boundary condition corresponds to the choice $t = 1/2$ and $T^+_{k,\Gamma} u = 0$. The function $w = S_k^+(f,g)$ satisfies

$$\begin{aligned} -\nabla \cdot (A\nabla w) &= f - k^2 w && \text{in } \Omega, \\ \partial_n w &= g && \text{on } \Gamma, \end{aligned}$$

therefore, standard regularity results apply and we find $u \in H^2(\Omega \setminus \Gamma_i)$ with the estimate

$$\begin{aligned} \|w\|_{2,\Omega\setminus\Gamma_i} &\lesssim \|f\|_{0,\Omega} + k^2\|w\|_{0,\Omega} + \|g\|_{1/2,\Gamma} \\ &\lesssim \|f\|_{0,\Omega} + k(k^{-1}\|f\|_{0,\Omega} + k^{-1/2}\|g\|_{0,\Gamma}) + \|g\|_{1/2,\Gamma} \\ &\lesssim \|f\|_{0,\Omega} + \|g\|_{1/2,\Gamma} + k^{1/2}\|g\|_{0,\Gamma}, \end{aligned}$$

where we applied the a priori estimate of Lemma 6.3.1, which is applicable since the conditions P.1 and P.2 in Assumption 6.2.5 are satisfied by Lemma 6.5.3.
Full space: The full space problem corresponds to the choice $t = 1/2$ and $T^+_{k,\Gamma} u = \mathrm{DtN}_0 u$. The function $w = S_k^+(f,g)$ satisfies

$$\begin{aligned} -\nabla \cdot (A\nabla w) &= f - k^2 w && \text{in } \Omega, \\ \partial_n w - \mathrm{DtN}_0 w &= g && \text{on } \Gamma. \end{aligned}$$

We reformulate the equations for w as a full space interface problem via the solution of the exterior Dirichlet problem which gives rise to the Dirichlet-to-Neumann operator. Consequently w satisfies, with $\Omega^+ := \overline{\Omega}^c$:

$$\begin{aligned} -\nabla \cdot (A\nabla w) + k^2 w &= f && \text{in } \Omega, \\ \partial_n w &= g + \partial_n u_w && \text{on } \Gamma, \\ -\Delta u_w &= 0 && \text{in } \Omega^+, \\ w &= u_w && \text{on } \Gamma, \\ u_w & \text{ satisfies radiation condition.} \end{aligned} \tag{6.33}$$

Upon defining z as w in Ω and u_w in Ω^+ we find that z satisfies

$$\begin{aligned} -\nabla\cdot(A\nabla z)+k^2 z &= f && \text{in } \Omega,\\ -\Delta z &= 0 && \text{in } \Omega^+,\\ [z] &= 0 && \text{on } \Gamma,\\ [\partial_n z] &= g && \text{on } \Gamma,\\ &z \text{ satisfies radiation condition.} \end{aligned} \tag{6.34}$$

Let $\widetilde{\Omega}$ be a ball with boundary $\widetilde{\Gamma} := \partial\widetilde{\Omega}$ such that $\Omega \subsetneq \widetilde{\Omega}$ and let χ denote a smooth cut-off function such that $\chi \equiv 1$ in Ω and $\chi \equiv 0$ in $\widetilde{\Omega}^c$. The function $\hat{z} := z\chi \in H^1_0(\widetilde{\Omega})$ satisfies, with $\hat{A}$ being the extension of A by the identity matrix,

$$\begin{aligned} -\nabla(\hat{A}\nabla\hat{z})+k^2\mathbb{1}_\Omega\hat{z} &= \hat{f} && \text{in } \widetilde{\Omega},\\ [\hat{z}] &= 0 && \text{on } \Gamma,\\ [\partial_n\hat{z}] &= g && \text{on } \Gamma,\\ \hat{z} &= 0 && \text{on } \widetilde{\Gamma}, \end{aligned} \tag{6.35}$$

with $\hat{f} = f$ in Ω and $\hat{f} = 2\nabla z\nabla\chi + z\Delta\chi$ in $\widetilde{\Omega}\setminus\Omega$. Since $f \in L^2(\Omega)$, $\hat{f} \in L^2(\widetilde{\Omega})$ and $g \in H^{1/2}(\Gamma)$ and due to the shift theorem for transmission problems, see e.g., [Mel02, Prop. 5.4.8], we find $\hat{z} \in H^2(\tilde{\Omega}\setminus(\Gamma_i\cup\Gamma))$ and consequently $w = \hat{z}\big|_\Omega \in H^2(\Omega\setminus\Gamma_i)$. Finally, we have

$$\begin{aligned} \|w\|_{2,\Omega\setminus\Gamma_i} = \|w\chi\|_{2,\Omega\setminus\Gamma_i} &= \|\hat{z}\|_{2,\Omega\setminus\Gamma_i} \le \|\hat{z}\|_{2,\widetilde{\Omega}\setminus(\Gamma_i\cup\Gamma)}\\ &\lesssim \|\hat{f}\|_{0,\widetilde{\Omega}} + \|k^2\mathbb{1}_\Omega\hat{z}\|_{0,\widetilde{\Omega}} + \|g\|_{1/2,\Gamma}\\ &\lesssim \|f\|_{0,\Omega} + \|2\nabla z\nabla\chi + z\Delta\chi\|_{0,\widetilde{\Omega}\setminus\Omega} + k^2\|w\|_{0,\Omega} + \|g\|_{1/2,\Gamma}\\ &\lesssim \|z\|_{1,\widetilde{\Omega}\setminus\Omega} + \|f\|_{0,\Omega} + k^2\|w\|_{0,\Omega} + \|g\|_{1/2,\Gamma}\\ &\lesssim \|u_w\|_{1,\widetilde{\Omega}\setminus\Omega} + \|f\|_{0,\Omega} + k^2\|w\|_{0,\Omega} + \|g\|_{1/2,\Gamma}. \end{aligned}$$

Since $\|u_w\|_{1,\widetilde{\Omega}\setminus\Omega} \lesssim \|w\|_{1/2,\Gamma} \lesssim \|w\|_{1,\Omega}$, due to the fact that u_w is the solution to the exterior Dirichlet problem as well as $k^2\|w\|_{0,\Omega} \lesssim k(k^{-1}\|f\|_{0,\Omega}+k^{-1/2}\|g\|_{0,\Gamma})$, by the a priori estimate of Lemma 6.3.1, which is applicable since the conditions P.1 and P.2 in Assumption 6.2.5 are satisfied by Lemma 6.5.3, we find

$$\|w\|_{2,\Omega\setminus\Gamma_i} \lesssim \|w\|_{1,\Omega} + \|f\|_{0,\Omega} + k^{1/2}\|g\|_{0,\Gamma} + \|g\|_{1/2,\Gamma}.$$

The asserted estimate follows by the a priori estimate for $\|w\|_{1,\Omega}$ in Lemma 6.3.1.
Second order ABCs: The second order absorbing boundary conditions correspond to the choice $t = 1$ and $T^+_{k,\Gamma}u = \alpha\Delta_\Gamma u$. The function $w = S^+_k(f,g)$ satisfies

$$\begin{aligned} -\nabla\cdot(A\nabla w) &= f - k^2 w && \text{in } \Omega,\\ \partial_n w - \alpha\Delta_\Gamma w &= g && \text{on } \Gamma. \end{aligned}$$

Note that for $f \in L^2(\Omega)$ and $g \in L^2(\Gamma)$ we have by the a priori estimate (6.11) of Lemma 6.3.1

$$\|\nabla w\|_{0,\Omega} + k\|w\|_{0,\Omega} + k^{-1/2}|w|_{1,\Gamma} \lesssim k^{-1}\|f\|_{0,\Omega} + k^{-1/2}\|g\|_{0,\Gamma}.$$

Especially, we have

$$\|\nabla_\Gamma w\|_{0,\Gamma} \lesssim k^{-1/2}\|f\|_{0,\Omega} + \|g\|_{0,\Gamma}.$$

Note that the following surface PDE is satisfied

$$\alpha\Delta_\Gamma w = -g + \partial_n w \quad \text{on } \Gamma,$$

which gives

$$\|w\|_{2,\Gamma} \lesssim k\|g\|_{0,\Gamma} + k\|\partial_n w\|_{0,\Gamma} + \|w\|_{1,\Gamma},$$

where we used the properties of α given in (6.31). Since $\|w\|_{1,\Gamma} \lesssim k^{-1/2}\|f\|_{0,\Omega} + \|g\|_{0,\Gamma}$ we have

$$\|w\|_{2,\Gamma} \lesssim k^{-1/2}\|f\|_{0,\Omega} + k\|g\|_{0,\Gamma} + k\|\partial_n w\|_{0,\Gamma}.$$

In order to estimate $\|\partial_n w\|_{0,\Gamma}$ we introduce the auxiliary problem for $\tilde{w} \in H^1(\Omega)$ which realizes the interior Dirichlet-to-Neumann map, i.e.,

$$\begin{aligned} -\nabla \cdot (A\nabla\tilde{w}) &= 0 \quad \text{in } \Omega, \\ \tilde{w} &= w \quad \text{on } \Gamma. \end{aligned} \tag{6.36}$$

Note that $w - \tilde{w} \in H_0^1(\Omega)$ solves

$$\begin{aligned} -\nabla \cdot (A\nabla(w - \tilde{w})) &= -k^2n^2 w \quad \text{in } \Omega, \\ w - \tilde{w} &= 0 \qquad\qquad \text{on } \Gamma. \end{aligned}$$

We have the following estimates

$$\begin{aligned} \|\tilde{w}\|_{1,\Omega} &\lesssim \|w\|_{1/2,\Gamma} \lesssim \|w\|_{1,\Omega} \lesssim k^{-1}\|f\|_{0,\Omega} + k^{-1/2}\|g\|_{0,\Gamma}, \\ \|w - \tilde{w}\|_{1,\Omega} &\leq \|w\|_{1,\Omega} + \|\tilde{w}\|_{1,\Omega} \lesssim \|w\|_{1,\Omega} \lesssim k^{-1}\|f\|_{0,\Omega} + k^{-1/2}\|g\|_{0,\Gamma}, \\ \|w - \tilde{w}\|_{2,\Omega} &\lesssim k^2\|w\|_{0,\Omega} \lesssim \|f\|_{0,\Omega} + k^{1/2}\|g\|_{0,\Gamma}, \\ \|\partial_n\tilde{w}\|_{0,\Gamma} &\lesssim \|\tilde{w}\|_{3/2,\Omega} \lesssim \|w\|_{1,\Gamma} \lesssim k^{-1/2}\|f\|_{0,\Omega} + \|g\|_{0,\Gamma}, \end{aligned}$$

where the last estimate is the crucial one. We postpone the verification of this estimate to the end of the proof. We now estimate using a multiplicative trace inequality and the above estimates

$$\begin{aligned} \|w\|_{2,\Gamma} &\lesssim k^{-1/2}\|f\|_{0,\Omega} + k\|g\|_{0,\Gamma} + k\|\partial_n w\|_{0,\Gamma}, \\ &\lesssim k^{-1/2}\|f\|_{0,\Omega} + k\|g\|_{0,\Gamma} + k\|\partial_n\tilde{w}\|_{0,\Gamma} + k\|\partial_n(w - \tilde{w})\|_{0,\Gamma}, \\ &\lesssim k^{1/2}\|f\|_{0,\Omega} + k\|g\|_{0,\Gamma} + k\|w - \tilde{w}\|_{1,\Omega}^{1/2}\|w - \tilde{w}\|_{2,\Omega}^{1/2}, \\ &\lesssim k^{1/2}\|f\|_{0,\Omega} + k\|g\|_{0,\Gamma}. \end{aligned}$$

We therefore have

$$\|w\|_{2,\Gamma} \lesssim k^{1/2}\|f\|_{0,\Omega} + k\|g\|_{0,\Gamma}. \tag{6.37}$$

We proceed to estimate $\|w\|_{2,\Omega}$. To that end, we will interpolate $H^1(\Gamma)$ and $H^2(\Gamma)$ to get an estimate for $H^{3/2}(\Gamma)$. Since w trivially satisfies the Dirichlet problem

$$\begin{aligned} -\nabla \cdot (A\nabla w) + k^2 w &= f \quad \text{in } \Omega, \\ w &= w \quad \text{on } \Gamma, \end{aligned}$$

we find

$$\begin{aligned} \|w\|_{2,\Omega} &\lesssim \|f\|_{0,\Omega} + k^2\|w\|_{0,\Omega} + \|w\|_{3/2,\Gamma} \\ &\lesssim \|f\|_{0,\Omega} + k^{1/2}\|g\|_{0,\Gamma} + \|w\|_{1,\Gamma}^{1/2}\|w\|_{2,\Gamma}^{1/2} \\ &\lesssim \|f\|_{0,\Omega} + k^{1/2}\|g\|_{0,\Gamma}. \end{aligned}$$

Hence, we have

$$\|w\|_{2,\Omega} \lesssim \|f\|_{0,\Omega} + k^{1/2}\|g\|_{0,\Gamma}.$$

We turn to the proof of the crucial estimate for $\|\partial_n \tilde{w}\|_{0,\Gamma}$. First, we find $\tilde{w} \in H^{3/2}(\Omega)$ and consequently $\nabla\tilde{w} \in H^{1/2}(\Omega)$, since $\tilde{w}$ satisfies (6.36). Therefore, we find $A\nabla\tilde{w} \in H^{1/2}(\Omega \setminus \Gamma_i)$. Again, using (6.36) gives $A\nabla\tilde{w} \in H^{1/2}(\mathrm{div}, \Omega)$. Its normal trace is therefore in $L^2(\Gamma)$ and since A is homogeneous near the boundary we have $\partial_n\tilde{w} \in L^2(\Gamma)$ with the estimate

$$\|\partial_n\tilde{w}\|_{0,\Gamma} = \|A\nabla\tilde{w}\cdot\boldsymbol{n}\|_{0,\Gamma} \lesssim \|A\nabla\tilde{w}\|_{1/2,\mathrm{div},\Omega} = \|A\nabla\tilde{w}\|_{1/2,\Omega} \lesssim \|\tilde{w}\|_{3/2,\Omega}, \tag{6.38}$$

which yields the result for second order ABCs.
PML: For the setting of the PML note that since the complex matrix-valued function A satisfies

$$\mathrm{Re}(A\nabla u, \nabla u) \gtrsim \|\nabla u\|_{0,\Omega}^2.$$

Therefore, the usual procedure of the difference quotient method works out, giving rise to a shift theorem as used in the Robin boundary case. □

Remark 6.5.5 ($H^2(\Gamma)$ estimate for second order ABCs)**.** In the case of second order ABCs with α as in (6.31) the solution $S_k^+(f,g) = w \in H^2(\Omega\setminus\Gamma_i)$ also satisfies the estimate (6.37):

$$k^{-1/2}\|w\|_{2,\Gamma} \lesssim \|f\|_{0,\Omega} + k^{1/2}\|g\|_{0,\Gamma}.$$

■

Lemma 6.5.6 (H^s regularity shift of S_k^+)**.** *Let the assumptions of Lemma 6.5.4 hold. Assume additionally that A is $C^{s+1}(\Omega\setminus\Gamma_i)$. Then, for $f \in H^s(\Omega\setminus\Gamma_i)$ and $g \in H^{s+1/2}(\Gamma)$ the solution $S_k^+(f,g) = w \in H^{s+2}(\Omega\setminus\Gamma_i)$ and the estimate*

$$\|w\|_{s+2,\Omega\setminus\Gamma_i} \lesssim \|f\|_{s,\Omega\setminus\Gamma_i} + k^s\|f\|_{0,\Omega} + \|g\|_{s+1/2,\Gamma} + k^{s+1/2}\|g\|_{0,\Gamma}$$

holds.

Proof. The proof is done by induction over $s \in \mathbb{N}$. Corresponding results for noninteger s follow readily by interpolation. The case $s = 0$ is covered in Lemma 6.5.4. As in the proof of Lemma 6.5.4 the regularity shift $w \in H^{s+2}(\Omega \setminus \Gamma_i)$ follows. Differentiating the differential equation we find

$$\|w\|_{s+2,\Omega\setminus\Gamma_i} \lesssim \|f\|_{s,\Omega\setminus\Gamma_i} + k^2\|w\|_{s,\Omega\setminus\Gamma_i} + \|g\|_{s+1/2,\Gamma}.$$

For $0 \leq s \leq 2$ we interpolate between the a priori estimate (6.11) in Lemma 6.3.1 and the estimate (6.32) in Lemma 6.5.4, in order to estimate $k^2\|w\|_{s,\Omega\setminus\Gamma_i}$. For $s \geq 2$ we use the induction hypothesis and find

$$\begin{aligned}\|w\|_{s+2,\Omega\setminus\Gamma_i} \lesssim \|f\|_{s,\Omega\setminus\Gamma_i} + \|g\|_{s+1/2,\Gamma} + k^2\|f\|_{s-2,\Omega\setminus\Gamma_i}\\ + k^s\|f\|_{0,\Omega} + k^2\|g\|_{s-2+1/2,\Gamma} + k^{s+1/2}\|g\|_{0,\Gamma}.\end{aligned}$$

Finally, interpolation between $H^s(\Omega \setminus \Gamma_i)$ and $L^2(\Omega)$ as well as $H^{s+1/2}(\Gamma)$ and $L^2(\Gamma)$ with appropriate use of Young's inequality yields the result. □

Remark 6.5.7 ($H^{s+2}(\Gamma)$ estimate for second order ABCs)**.** Analogous considerations as in Lemma 6.5.6 show in the case of second order ABCs with α as in (6.31) with $f \in H^s(\Omega\setminus\Gamma_i)$ and $g \in H^{s+1/2}(\Gamma)$ for $s > 0$ that the solution $S_k^+(f,g) = w \in H^{s+2}(\Omega \setminus \Gamma_i)$ also satisfies the estimate

$$k^{-1/2}\|w\|_{s+2,\Gamma} \lesssim \|f\|_{s,\Omega\setminus\Gamma_i} + k^s\|f\|_{0,\Omega} + \|g\|_{s+1/2,\Gamma} + k^{s+1/2}\|g\|_{0,\Gamma}.$$

■

Lemma 6.5.8 (Analytic regularity of S_k^-)**.** *Let Assumption 6.2.1 and M.3 be satisfied. Furthermore, let the hypothesis of Lemma 6.5.4 be satisfied. Finally, let A and n be such that*

$$\begin{aligned}\|\nabla^p A\|_{L^\infty(\Omega\setminus\Gamma_i)} \leq C_A\gamma_A^p p!,\\ \|\nabla^p n\|_{L^\infty(\Omega\setminus\Gamma_i)} \leq C_n\gamma_n^p p!\end{aligned}$$

for all $p \geq 0$. Let $f \in L^2(\Omega)$ be piecewise analytic and satisfy

$$\|\nabla^p f\|_{0,\Omega\setminus\Gamma_i} \leq \widetilde{C}_f\gamma_f^p \max\{p,k\}^p \qquad \forall p \geq 0.$$

Let g be the restriction of an analytic functions G in a one-sided tubular neighborhood T of the boundary Γ and satisfy

$$\|\nabla^p G\|_{0,T} \leq \widetilde{C}_g\gamma_g^p \max\{p,k\}^p \qquad \forall p \geq 0.$$

Then, for any of the problems considered in Subsection 6.5.1 there exist constants C, $\gamma \geq 0$ independent of k, such that the function $u := S_k^-(f,g)$ is piecewise analytic and satisfies

$$\|u\|_{1,t,k} \leq CC^-_{\mathrm{sol},k}(\widetilde{C}_f + \widetilde{C}_g), \tag{6.39}$$

$$\|\nabla^p u\|_{0,\Omega\setminus\Gamma_i} \leq CC^-_{\mathrm{sol},k}k^{-1}\gamma^p \max\{k,p\}^p(\widetilde{C}_f + \widetilde{C}_g) \qquad \forall p \geq 2. \tag{6.40}$$

Proof. Estimate (6.39) is just a restatement of the energy estimate (6.8) in M.3. The proof of analytic regularity proceeds as usual. The domain Ω is covered with balls and an analytic change of variables is performed to flatten the boundary and the interface. It is important to note that membership in the analyticity classes is invariant under analytic change of variables and multiplication by analytic functions, see [MS21, Lemma 2.6]. This in turn makes the results of [Mel02, Sec. 5.5] as well as Section 6.8 (in the case of second order ABCs) applicable.
Robin boundary conditions: The Robin case corresponds to the choice $T^-_{k,\Gamma}u = iknu$. Note that, upon setting $\varepsilon = 1/k$, the function $u = S^-_k(f,g)$ satisfies

$$\begin{aligned} -\varepsilon^2\nabla\cdot(A\nabla u) - n^2u &= \varepsilon^2 f && \text{in } \Omega,\\ \varepsilon^2\partial_n u &= \varepsilon(\varepsilon g + inu) && \text{on } \Gamma. \end{aligned} \tag{6.41}$$

We apply [Mel02, Prop. 5.5.1] (interior analytic regularity) , [Mel02, Prop. 5.5.3] (boundary analytic regularity for Neumann problems) as well as [Mel02, Prop. 5.5.4] (interface analytic regularity) if $\Gamma_i \neq \emptyset$, to problem (6.41). We find

$$\begin{aligned} \|\nabla^p u\|_{0,\Omega\setminus\Gamma_i} &\lesssim C\gamma^p \max\{k,p\}^p(k^{-2}\widetilde{C}_f + k^{-1}\widetilde{C}_g + k^{-1}\|\nabla u\|_{L^2(\Omega)} + \|u\|_{L^2(\Omega)})\\ &\lesssim C\gamma^p \max\{k,p\}^p(k^{-2}\widetilde{C}_f + k^{-1}\widetilde{C}_g + C^-_{\mathrm{sol},k}k^{-1}(\widetilde{C}_f + \widetilde{C}_g))\\ &\lesssim C\gamma^p \max\{k,p\}^p C^-_{\mathrm{sol},k}k^{-1}(\widetilde{C}_f + \widetilde{C}_g) \qquad \forall p \geq 2, \end{aligned}$$

where we applied the estimate (6.39) as well as $C^-_{\mathrm{sol},k} \gtrsim 1$. See also [MS11, Lemma 4.12] for similar arguments in the homogeneous case.
Full space: The full space problem corresponds to the choice $T^-_{k,\Gamma}u = \mathrm{DtN}_k u$. As in the proof of Lemma 6.5.4 we can extend u to satisfy

$$\begin{aligned} -\nabla\cdot(A\nabla u) - k^2n^2u &= f && \text{in } \Omega\cup\Omega^+,\\ [u] &= 0 && \text{on } \Gamma,\\ [\partial_n u] &= g && \text{on } \Gamma,\\ &u \text{ satisfies radiation condition.} \end{aligned}$$

From now on the proof is completely analogous to the Robin case above.
Second order ABCs: For second order absorbing boundary conditions we have $T^-_{k,\Gamma}u = \beta u + \alpha\Delta_\Gamma u$. We proceed similar to the case of Robin boundary conditions. We apply Theorem 6.8.5 instead of [Mel02, Prop. 5.5.3]. □

Lemma 6.5.9 (Quasi-selfadjointness)**.** *The operators $T^-_{k,\Omega}$ and $T^-_{k,\Gamma}$ considered in Subsection 6.5.1 are quasi-selfadjoint.*

Proof. The mapping $T^-_{k,\Omega}\colon u \mapsto k^2n^2u$ is trivially quasi-selfadjoint since

$$(T^-_{k,\Omega}u, \overline{v}) = \int_\Omega n^2uv = \int_\Omega n^2vu = (T^-_{k,\Omega}v, \overline{u}).$$

Analogously we find the mapping $T^-_{k,\Omega} : u \mapsto k^2n^2u + ikmu$ to be quasi-selfadjoint, as well as $T^-_{k,\Gamma} : u \mapsto iknu$. In the case of second order ABCs, $T^-_{k,\Gamma}$ is also trivially quasi-selfadjoint. For the $T^-_{k,\Gamma} = \mathrm{DtN}_k$, see [CWGLS12, Sec. 2.7, Eq. (2.84)] as well as [MS10, Lemma 3.10], in case of a sphere. □

Lemma 6.5.10 (Splitting of $T^-_{k,\Omega}$, $T^-_{k,\Gamma}$, $T^-_{k,\Omega}-T^+_{k,\Omega}$ and $T^-_{k,\Gamma}-T^+_{k,\Gamma}$)**.** *Let Assumption 6.2.1 be satisfied. Let m, $n^2 \in L^\infty(\Omega)$ and $n \in L^\infty(\Gamma)$. Then the operators $T^-_{k,\Omega}$, $T^-_{k,\Gamma}$, $T^+_{k,\Omega}$ and $T^+_{k,\Gamma}$ considered in Subsection 6.5.1 and specified in Table 6.3 satisfy M.2, P.3, P.4, P.5 and P.6. Furthermore, for $s > 0$ let additionally m, $n^2 \in W^{s,\infty}(\Omega \setminus \Gamma_i)$. Then the operators $T^-_{k,\Omega}$, $T^-_{k,\Gamma}$, $T^+_{k,\Omega}$ and $T^+_{k,\Gamma}$ considered in Subsection 6.5.1 and specified in Table 6.3 additionally satisfy PS.1 and PS.2*

Proof. The proof is trivial except the case of the Dirichlet-to-Neumann operator. In case of a sphere the improved splitting holds true by Item (iii) in Lemma 6.5.12. The result for the Dirichlet-to-Neumann in the general nontrapping case follows by application of Item (ii) in Lemma 6.5.12 with $s = 1/2$ for the splitting of $T^-_{k,\Gamma} - T^+_{k,\Gamma}$ and with $s = 0$ for the splitting of $T^-_{k,\Gamma}$ itself. □

Remark 6.5.11 (Stronger estimate for u_F for second order ABCs)**.** In the case of second order absorbing boundary conditions as considered in Subsection 6.5.1, inspection of the proof of Lemma 6.3.6 and Lemma 6.3.8 together with the Remarks 6.5.5 and 6.5.7 yields the following stronger result. The function u_F in the splitting $u = u_F + u_A$ additionally satisfies

$$k^{-1/2}\|u_F\|_{s+2,\Gamma} \lesssim \|f\|_{s,\Omega\setminus\Gamma_i} + \|g\|_{s+1/2,\Gamma}.$$

It is important to note that this improved regularity estimate is not necessary for establishing quasi-optimality as in Corollary 6.6.9. However, it is crucial in order to extract optimal convergence rates as in Corollary 6.6.11. ■

Lemma 6.5.12. *Let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$ be a bounded Lipschitz domain with boundary Γ. Then the following holds:*

(i) $-\langle \mathrm{DtN}_0 u, u\rangle \geq 0$ *for all* $u \in H^{1/2}(\Gamma)$.

(ii) With the notation of Section 6.9 let Ω^+ be nontrapping with analytic boundary Γ. Let $s \geq 0$ be given. Then

$$\mathrm{DtN}_k - \mathrm{DtN}_0 = kB + [\![\partial_n \tilde{A}]\!],$$

where the linear operators $B\colon H^s(\Gamma) \to H^s(\Gamma)$ and $\tilde{A}\colon H^s(\Gamma) \to C^\infty(\Omega_R)$ satisfying for all $u \in H^s(\Gamma)$

$$\|Bu\|_{s,\Gamma} \lesssim \|u\|_{s,\Gamma}, \qquad \tilde{A}u \in \mathcal{A}(Ck^\beta\|u\|_{s,\Gamma}, \gamma, \Omega_R)$$

with $\beta = 7/2 + d/2$, and constants C, $\gamma \geq 0$ independent of k.

(iii) If Γ is the unit ball in dimension $d = 3$ then the symbol of the operator $\mathrm{DtN}_k - \mathrm{DtN}_0$ is given by $z_l(k) + l + 1$, where $z_l(k)$ denotes the symbol of DtN_k. Furthermore, the estimate

$$|z_l(k) + l + 1| \leq 2k \qquad \forall l \geq 0$$

holds. Finally, $\mathrm{DtN}_k - \mathrm{DtN}_0\colon H^s(\Gamma) \to H^s(\Gamma)$ satisfies

$$\|\mathrm{DtN}_k u - \mathrm{DtN}_0 u\|_{s,\Gamma} \lesssim k\|u\|_{s,\Gamma} \qquad \forall u \in H^s(\Gamma)$$

for every $s \geq 0$.

Proof. For $-\langle \mathrm{DtN}_0 u, u\rangle \geq 0$ see [AMP15, Lemma 2.2, (vi)]. We proceed with the proof of Item (iii). If Γ is the unit sphere in spatial dimension $d = 3$ the operator DtN_k has an explicit series representation in terms of spherical harmonics. Let Y_l^m denote the standard spherical harmonics. On the unit sphere Γ in spatial dimension $d = 3$ we can expand u as

$$u(x) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} u_l^m Y_l^m(\theta, \varphi)$$

in spherical coordinates $(\theta, \varphi) \in \Gamma$. The operator DtN_k as well as the operator DtN_0 can be written as

$$\mathrm{DtN}_0 u = -\sum_{l=0}^{\infty} \sum_{m=-l}^{l} (l+1) u_l^m Y_l^m,$$

$$\mathrm{DtN}_k u = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} z_l(k) u_l^m Y_l^m,$$

with explicit estimates for the symbol $z_l(k)$. The formulas for DtN_k and DtN_0 immediately give

$$\mathrm{DtN}_k u - \mathrm{DtN}_0 u = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} (z_l(k) + l + 1) u_l^m Y_l^m.$$

From [DI01, Lemma 3.2, Eq. (3.28)], where the operator DtN_k has opposite sign, we have

$$l + 1 - k \leq -\mathrm{Re}\, z_l(k) \leq l + 1 + k.$$

Consequently, we immediately have

$$|\mathrm{Re}\, z_l(k) + l + 1| \leq k.$$

From [Né01, Thm. 2.6.1, Eq. (2.6.24)] we have

$$0 \leq \mathrm{Im}\, z_l(k) \leq k,$$

which together with the previous estimate gives

$$|z_l(k) + l + 1| \leq 2k.$$

For $u \in H^s(\Gamma)$ and with the previous estimate we have

$$\begin{aligned} \|\mathrm{DtN}_k u - \mathrm{DtN}_0 u\|_{s,\Gamma}^2 &= \sum_{l=0}^{\infty} \sum_{m=-l}^{l} (l+1)^{2s} |z_l(k) + l + 1|^2 |u_l^m|^2 Y_l^m \\ &\leq (2k)^2 \sum_{l=0}^{\infty} \sum_{m=-l}^{l} (l+1)^{2s} |u_l^m|^2 Y_l^m = (2k)^2 \|u\|_{s,\Gamma}^2, \end{aligned}$$

which yields the result. The proof of Item (ii) is given in Section 6.9. □

Collecting the above results we have therefore proven

Theorem 6.5.13 (Regularity theory for heterogeneous Helmholtz problems)**.** *The regularity theory of Section 6.3 is applicable to the problems considered in Subsection 6.5.1 and specified in Table 6.3.*

6.6. Stability and convergence of abstract Galerkin discretizations

The present section establishes quasi-optimality of an abstract Galerkin discretization of (6.6) under the condition that the ansatz space satisfies certain approximation properties. We apply our results to the *hp* Finite Element Method discretizations of the model problems considered in Subsection 6.5.1. Furthermore, a complete convergence analysis is performed for higher order Sobolev data.

6.6.1. Quasi-optimality of abstract Galerkin discretizations

We consider a Galerkin discretization of (6.6) with a subspace $V_h \subset H^{1,t}(\Omega,\Gamma)$. In the analysis a variety of approximability quantities arise. We employ the following notation. The approximability quantities will all be denoted by η with different sub- and superscripts. The superscripts exp in η^{exp} indicate that this quantity gets exponentially small (when working with *hp*-FEM spaces) since it quantifies the approximability of smooth functions. We introduce

$$\eta_1^{\mathrm{exp}} := \sup_{v\in H^{1,t}(\Omega,\Gamma)} \inf_{s_h\in V_h} \frac{\|S_k^{-,\star}(A_\Omega^- v, A_\Gamma^- v) - s_h\|_{1,t,k}}{\|v\|_{1,t,k}}, \tag{6.42}$$

$$\eta_2^{\mathrm{exp}} := \sup_{v\in H^{1,t}(\Omega,\Gamma)} \inf_{s_h\in V_h} \frac{\|S_k^{-,\star}(A_\Omega v, A_\Gamma v) - s_h\|_{1,t,k}}{\|v\|_{1,t,k}}, \tag{6.43}$$

as well as

$$\eta^\star := \sup_{\substack{f\in L^2(\Omega),\\ g\in H^{1/2}(\Gamma)}} \inf_{s_h\in V_h} \frac{\|S_k^{-,*}(f,g) - s_h\|_{1,t,k}}{\|f\|_{0,\Omega} + \|g\|_{1/2,\Gamma}}. \tag{6.44}$$

We consider a fixed right-hand side $f \in L^2(\Omega)$ and $g \in H^{1/2}(\Gamma)$ and the corresponding solution $u = S_k^-(f,g) \in H^{1,t}(\Omega,\Gamma)$ of

$$b_k^-(u,v) = (f,v) + \langle g,v\rangle \quad \forall v \in H^{1,t}(\Omega,\Gamma).$$

Then, in the following we let u_h denote any element of V_h such that

$$b_k^-(u-u_h, v_h) = 0 \quad \forall v_h \in V_h. \tag{6.45}$$

Lemma 6.6.1. *Let Assumptions M.1 , M.2 and 6.4.1 be satisfied. Furthermore, let A be uniformly bounded. Then the estimate*

$$|b_k^-(u-u_h,v)| \lesssim \left(1 + C_{\mathrm{cont},k}^- \eta_1^{\mathrm{exp}}\right) \|u-u_h\|_{1,t,k} \|v\|_{1,t,k} \tag{6.46}$$

holds true for all $v \in H^{1,t}(\Omega,\Gamma)$.

Proof. Let $v \in H^{1,t}(\Omega,\Gamma)$ be arbitrary. For sake of shortness we write $e_h \coloneqq u - u_h$. We have

$$\begin{aligned}
|b_k^-(e_h,v)| &= |(A\nabla e_h,\nabla v) - (T_{k,\Omega}^- e_h, v) - \langle T_{k,\Gamma}^- e_h, v\rangle| \\
&\overset{6.4.1}{=} |(A\nabla e_h,\nabla v) - (T_{k,\Omega}^- \overline{v}, \overline{e_h}) - \langle T_{k,\Gamma}^- \overline{v}, \overline{e_h}\rangle| \\
&\overset{M.2}{\leq} |(A\nabla e_h,\nabla v)| + |(D_\Omega^- \overline{v}, \overline{e_h})| + |\langle D_\Gamma^- \overline{v}, \overline{e_h}\rangle| + |(A_\Omega^- \overline{v}, \overline{e_h}) + \langle A_\Gamma^- \overline{v}, \overline{e_h}\rangle| \\
&\overset{M.2}{\lesssim} \|e_h\|_{1,t,k}\|v\|_{1,t,k} + |(e_h, \overline{A_\Omega^- \overline{v}}) + \langle e_h, \overline{A_\Gamma^- \overline{v}}\rangle|
\end{aligned}$$

The analytic part is now treated using a duality argument. Using the Galerkin orthogonality (6.45) , M.1 and the definition of $\eta_1^{\exp}$ in (6.42), we can estimate

$$\begin{aligned}
|(e_h, \overline{A_\Omega^- \overline{v}}) + \langle e_h, \overline{A_\Gamma^- \overline{v}}\rangle| &= |b_k^-(e_h, S_k^{-,\star}(A_\Omega^- \overline{v}, A_\Gamma^- \overline{v}))| \\
&\overset{(6.45)}{=} |b_k^-(e_h, S_k^{-,\star}(A_\Omega^- \overline{v}, A_\Gamma^- \overline{v}) - v_h)| \\
&\overset{M.1}{\lesssim} C_{\mathrm{cont},k}^- \|e_h\|_{1,t,k} \|S_k^{-,\star}(A_\Omega^- \overline{v}, A_\Gamma^- \overline{v}) - v_h\|_{1,t,k} \\
&\overset{(6.42)}{\lesssim} C_{\mathrm{cont},k}^- \eta_1^{\exp} \|e_h\|_{1,t,k}\|v\|_{1,t,k},
\end{aligned}$$

which concludes the proof. □

Remark 6.6.2. In Lemma 6.6.1 we assume quasi-selfadjointness of the operators $T_{k,\Omega}^-$ and $T_{k,\Gamma}^-$. We note however, that one could also assume a splitting of the adjoint operators as in M.1 and M.2 and derive the same result. ■

Theorem 6.6.3. *Assume the hypothesis of Lemma 6.6.1. Let Assumptions P.2 , P.3 , P.4 and P.5 be satisfied. Assume $C_{\mathrm{cont},k}^- \eta_1^{\exp}$, $k\eta^\star$ and $\eta_2^{\exp}$ to be sufficiently small. Then the Galerkin solution $u_h \in V_h$ of* (6.45) *exists, is unique and satisfies*

$$\|u - u_h\|_{1,t,k} \lesssim \inf_{v_h \in V_h} \|u - v_h\|_{1,t,k}, \tag{6.47}$$

which hidden constant independent of k.

Proof. For the sake of simplicity, we write $e_h \coloneqq u - u_h$, and pick an arbitrary element $v_h \in V_h$. By P.2, and using Galerkin orthogonality (6.45), the first step of the proof consists in estimating

$$\begin{aligned}
\|e_h\|_{1,t,k}^2 &\overset{P.2}{\lesssim} \mathrm{Re}(\sigma b_k^+(e_h,e_h)) \leq |b_k^+(e_h,e_h)| \\
&\leq |b_k^-(e_h,e_h) - b_k^+(e_h,e_h)| + |b_k^-(e_h,e_h)| \\
&\overset{(6.45)}{=} |b_k^-(e_h,e_h) - b_k^+(e_h,e_h)| + |b_k^-(e_h, u - v_h)|.
\end{aligned}$$

Then, using P.3, the Galerkin orthogonality (6.45), the refined continuity estimate (6.46) in Lemma 6.6.1, the definitions of $\eta_2^{\exp}$ and $\eta^\star$ in (6.43) and (6.44), respectively and finally P.4

and P.5 to derive that

$$
\begin{aligned}
|b_k^-(e_h,e_h) - b_k^+(e_h,e_h)| &\overset{P.3}{\leq} |(R_\Omega e_h, e_h) + \langle R_\Gamma e_h, e_h\rangle| + |(A_\Omega e_h, e_h) + \langle A_\Gamma e_h, e_h\rangle| \\
&= |(e_h, R_\Omega e_h) + \langle e_h, R_\Gamma e_h\rangle| + |(e_h, A_\Omega e_h) + \langle e_h, A_\Gamma e_h\rangle| \\
&= |b_k^-(e_h, S_k^{-,\star}(R_\Omega e_h, R_\Gamma e_h))| + |b_k^-(e_h, S_k^{-,\star}(A_\Omega e_h, A_\Gamma e_h))| \\
&\overset{(6.45)}{=} |b_k^-(e_h, S_k^{-,\star}(R_\Omega e_h, R_\Gamma e_h) - s_h^R)| \\
&\qquad + |b_k^-(e_h, S_k^{-,\star}(A_\Omega e_h, A_\Gamma e_h) - s_h^A)| \\
&\overset{(6.46)}{\lesssim} \left(1 + C^-_{\mathrm{cont},k}\eta_1^{\mathrm{exp}}\right) \|e_h\|_{1,t,k} \Big[\|S_k^{-,\star}(R_\Omega e_h, R_\Gamma e_h) - s_h^R\|_{1,t,k} \\
&\qquad + \|S_k^{-,\star}(A_\Omega e_h, A_\Gamma e_h) - s_h^A)\|_{1,t,k} \Big] \\
&\overset{(6.43),(6.44)}{\lesssim} \left(1 + C^-_{\mathrm{cont},k}\eta_1^{\mathrm{exp}}\right) \|e_h\|_{1,t,k} \\
&\qquad \cdot \left[\eta^\star \left(\|R_\Omega e_h\|_{0,\Omega} + \|R_\Gamma e_h\|_{1/2,\Gamma}\right) + \eta_2^{\mathrm{exp}} \|e_h\|_{1,t,k}\right] \\
&\overset{P.4,\ P.5}{\lesssim} \left(1 + C^-_{\mathrm{cont},k}\eta_1^{\mathrm{exp}}\right) (k\eta^\star + \eta_2^{\mathrm{exp}}) \|e_h\|^2_{1,t,k}
\end{aligned}
$$

We therefore find

$$
\|e_h\|^2_{1,t,k} \lesssim \left(1 + C^-_{\mathrm{cont},k}\eta_1^{\mathrm{exp}}\right) (k\eta^\star + \eta_2^{\mathrm{exp}}) \|e_h\|^2_{1,t,k} + |b_k^-(e_h, u - v_h)|.
$$

The assumed smallness of $C^-_{\mathrm{cont},k}\eta_1^{\mathrm{exp}}$, $k\eta^*$ and η_2^{exp} allows to absorb $\|e_h\|^2_{1,t,k}$ on the left-hand side, which yields

$$
\|e_h\|^2_{1,t,k} \lesssim |b_k^-(e_h, u - v_h)|
$$

for any $v_h \in V_h$. Finally, application of Lemma 6.6.1 concludes the proof. □

Remark 6.6.4. Theorem 6.6.3 shows that quasi-optimality of an abstract Galerkin method holds if the quantifies $C^-_{\mathrm{cont},k}\eta_1^{\mathrm{exp}}$, $k\eta^\star$ and η_2^{exp} are sufficiently small. All of these depend on the approximability of the solution to the adjoint problem. The primal problem only enters these quantities in terms of the continuity constant $C^-_{\mathrm{cont},k}$ of the sesquilinear form b_k^- and the operators A_Ω^- and A_Γ^-. Hence, in the application to the hp-FEM, if the adjoint problem is such that the regularity theory of Section 6.3 is applicable one can further estimate $C^-_{\mathrm{cont},k}\eta_1^{\mathrm{exp}}$, $k\eta^\star$ and η_2^{exp} via the splitting stated in Theorem 6.3.10. ■

6.6.2. Application to hp-FEM

We start with assumptions on the triangulation.

Assumption 6.6.5 (quasi-uniform regular fitted meshes)**.** Let Assumption 2.0.1 be satisfied. Furthermore, if $\Gamma_i \neq \emptyset$ we assume the mesh to resolve the interface Γ_i, i.e., each element K lies on one side of the interface Γ_i and at most two mapped vertices of $\widehat{K}$ lie on Γ_i.

Note that for all $t \leq 1$ the space $S_p(\mathcal{T}_h)$ is a conforming subspace of $H^{1,t}(\Omega, \Gamma)$, with the limiting case $t = 1$ included to cover second order ABCs.

Assumption 6.6.6 (Polyomial well-posedness)**.** There exist constants α_1, α_2, α_3, $\alpha_4 \geq 0$ such that

$$C^-_{\mathrm{sol},k} \lesssim k^{\alpha_1}, \quad C^-_{\mathrm{cont},k} \lesssim k^{\alpha_2}, \quad C_{\mathrm{ana},k} \lesssim k^{\alpha_3}, \quad C^-_{\mathrm{ana},k} \lesssim k^{\alpha_4},$$

holds true.

Theorem 6.6.7 (discrete stability of hp-FEM)**.** *Let the Assumptions 6.2.1 (smoothness of the Ω, Γ and Γ_i) , 6.2.4 (assumptions on the minus problem) , 6.2.5 (assumptions on the plus problem) , 6.4.1 , 6.6.5 and 6.6.6 be satisfied. Assume additionally $t \leq 1$. Then there exist constants c_1, $c_2 > 0$ independent of h, p, and k, such that under the scale resolution condition*

$$\frac{kh}{p} \leq c_1 \qquad \text{and} \qquad p \geq c_2(\log k + 1) \tag{6.48}$$

the Galerkin solution $u_{hp} \in S_p(\mathcal{T}_h)$ exists, is unique and satisfies

$$\|u - u_{hp}\|_{1,t,k} \lesssim \inf_{v_{hp} \in S_p(\mathcal{T}_h)} \|u - v_{hp}\|_{1,t,k},$$

with hidden constant independent of k, h and p.

Proof. The proof is standard as in [MS10, MS11]. Due to the assumed polynomial well-posedness (Assumption 6.6.6) and applying the regularity splitting in Theorem 6.3.10, we find that the quantities $C^-_{\mathrm{cont},k}\eta_1^{\mathrm{exp}}$, $k\eta^\star$ and η_2^{exp} get arbitrarily small for appropriate choice of c_1, $c_2 > 0$. There are two simple adjustments to be made. First, regarding the piecewise regularity of u_F and u_A. Since the mesh is fitted to the interface Γ_i, see Assumption 6.6.5, the approximation properties of $S_p(\mathcal{T}_h)$ stay the same. Second, the additional boundary term in the energy norm $\|\cdot\|_{1,t,k}$, which is treated with a trace inequality. The abstract quasi-optimality result in Theorem 6.6.3 then yields the result. □

Remark 6.6.8 (On the boundary term in $\|\cdot\|_{1,t,k}$.)**.** It is worth mentioning that in the proof of Theorem 6.6.7 the boundary term $k^{-t+1/2}\|\cdot\|_{t,\Gamma}$, which appears in the case $1/2 < t \leq 1$, is treated with a trace inequality: When estimating $k\eta^\star$ and after applying the regularity splitting of Theorem 6.3.10 a term of the form

$$\inf_{s_h \in S_p(\mathcal{T}_h)} k k^{-t+1/2}\|u_F - s_h\|_{t,\Gamma}$$

arises. This term is treated by applying a trace inequality and using standard approximation properties of $S_p(\mathcal{T}_h)$:

$$\begin{aligned}
\inf_{s_h \in S_p(\mathcal{T}_h)} k k^{-t+1/2}\|u_F - s_h\|_{t,\Gamma} &\lesssim \inf_{s_h \in S_p(\mathcal{T}_h)} k^{3/2-t}\|u_F - s_h\|_{t+1/2,\Omega\setminus\Gamma_i} \\
&\lesssim k^{3/2-t}\left(\frac{h}{p}\right)^{2-t-1/2}\|u_F\|_{2,\Omega\setminus\Gamma_i} \\
&= \left(\frac{kh}{p}\right)^{3/2-t}\|u_F\|_{2,\Omega\setminus\Gamma_i}.
\end{aligned}$$

Therefore, the condition that kh/p be sufficiently small also ensures $(kh/p)^{3/2-t}$ to be small, and vice versa. However, we will see below, that in order to derive optimal rates, the additional regularity on the boundary, see Remark 6.5.11, needs to be exploited. For that we need an approximation operator featuring simultaneous approximation properties in $H^{1,1}(\Omega, \Gamma)$, see Proposition 6.6.10 ■

As a simple corollary of Theorem 6.6.3 we have

Corollary 6.6.9 (Application of *hp*-FEM to the problems in Subsection 6.5.1)**.** *Consider any of the model problems of Subsection 6.5.1. Let this problem be polynomially well-posed, i.e., let Assumptions M.3 and 6.6.6 be satisfied. Additionally let Assumptions 6.2.1 and 6.6.5 be satisfied. Then there exist constants c_1, $c_2 > 0$ independent of h, p, and k, such that under the scale resolution condition*

$$\frac{kh}{p} \leq c_1 \qquad \text{and} \qquad p \geq c_2(\log k + 1) \tag{6.49}$$

the Galerkin solution $u_{hp} \in S_p(\mathcal{T}_h)$ exists, is unique and satisfies

$$\|u - u_{hp}\|_{1,t,k} \lesssim \inf_{v_{hp} \in S_p(\mathcal{T}_h)} \|u - v_{hp}\|_{1,t,k},$$

with hidden constant independent of k, h and p.

Proposition 6.6.10 (Simultaneous Approximation in $H^{1,1}(\Omega, \Gamma)$)**.** *Let $\widehat{K}$ denote the reference triangle in spatial dimension two. Let $s \geq 1$. Then for every p there exists a linear operator $\widehat{\Pi}_p^{\mathrm{grad}} \colon H^{3/2}(\widehat{K}) \to \mathcal{P}_p(\widehat{K})$, which satisfies*

$$p\|u - \widehat{\Pi}_p^{\mathrm{grad}} u\|_{0,\widehat{K}} + \|u - \widehat{\Pi}_p^{\mathrm{grad}} u\|_{1,\widehat{K}} \lesssim p^{-s}\|u\|_{s+1,\widehat{K}} \tag{6.50}$$

for $p \geq s - 1$. Additionally, there holds

$$p\|u - \widehat{\Pi}_p^{\mathrm{grad}} u\|_{0,\partial\widehat{K}} + \|u - \widehat{\Pi}_p^{\mathrm{grad}} u\|_{1,\partial\widehat{K}} \lesssim p^{-s}\|u\|_{s+1,\partial\widehat{K}} \tag{6.51}$$

for $p \geq s - 1$.

Proof. The desired operator is defined in [MR20, Def. 2.5]. The estimate (6.50) is given in [MR20, Cor. 2.14]. The estimate (6.51) is a combination of [MR20, Def. 2.5 and Lemma 4.1]. □

Corollary 6.6.11 (Convergence rates of *hp*-FEM for the problems in Subsection 6.5.1)**.** *Assume the hypothesis of Corollary 6.6.9. Let $s \geq 0$, $f \in H^s(\Omega \setminus \Gamma_i)$ and $g \in H^{s+1/2}(\Gamma)$ be given. Let $\theta = \max\{\alpha_1, \alpha_1 + \alpha_3 - 1\}$, with α_1 and α_3 given in Assumption 6.6.6. Then there exist constants c_1, c_2, $\sigma > 0$ independent of h, p, and k, such that under the scale resolution condition*

$$\frac{kh}{p} \leq c_1 \qquad \text{and} \qquad p \geq s + c_2(\log k + 1)$$

the Galerkin solution $u_{hp} \in S_p(\mathcal{T}_h)$ satisfies the estimate

$$\|u - u_{hp}\|_{1,t,k} \lesssim \left[\left(\frac{h}{p}\right)^{s+1} + k^{\theta-1} \left\{ \left(\frac{h}{h+\sigma}\right)^p + k\left(\frac{kh}{\sigma p}\right)^p \right\} \right] (\|f\|_{s,\Omega\setminus\Gamma_i} + \|g\|_{s+1/2,\Gamma}),$$

with hidden constant independent of k, h and p.

Proof. Assumption 6.2.6 is satisfied for the problems considered in Subsection 6.5.1. The infimum in Corollary 6.6.9 is quantified by applying the splitting given in Theorem 6.3.11. Note that due to Theorem 6.3.11 as well as Remark 6.5.11 we can split $u = u_F + u_A$ with

$$\begin{aligned}
\|u_F\|_{s+2,\Omega\setminus\Gamma_i} + k^{s+1}\|u_F\|_{1,t,k} &\lesssim \|f\|_{s,\Omega\setminus\Gamma_i} + \|g\|_{s+1/2,\Gamma}, \\
k^{-1/2}\|u_F\|_{s+2,\Gamma} &\lesssim \|f\|_{s,\Omega\setminus\Gamma_i} + \|g\|_{s+1/2,\Gamma}, \\
\|u_A\|_{1,t,k} &\lesssim k^{\theta}(\|f\|_{0,\Omega} + \|g\|_{1/2,\Gamma}), \\
\|\nabla^n u_A\|_{0,\Omega} &\lesssim k^{\theta-1}\gamma^n \max\{k,n\}^n(\|f\|_{0,\Omega} + \|g\|_{1/2,\Gamma}) \qquad \forall n \geq 2,
\end{aligned}$$

with $\theta = \max\{\alpha_1, \alpha_1+\alpha_3-1\}$, see the estimate (6.25). Applying this splitting and employing the approximation properties of the finite element spaces yields the result, see [MPS13, Sec. 4] for these kinds of arguments. In the case of second order absorbing boundary conditions, one additionally applies the approximation operator given in Proposition 6.6.10. □

6.7. Numerical examples

All our calculations are performed with the *hp*-FEM code NETGEN / NGSOLVE by J. Schöberl, [Sch, Sch97]. The curved boundary and interface are implemented using second order rational splines. We plot different errors against N_λ, the number of degrees of freedom per wavelength,

$$N_\lambda = \frac{2\pi \sqrt[d]{\mathrm{DOF}}}{k\sqrt[d]{|\Omega|}},$$

where the wavelength λ and the wavenumber k are related via $k = 2\pi/\lambda$ and DOF denotes the size of the linear system to be solved.

Example 6.7.1. Let Ω be the unit circle in $\mathbb{R}^2$ and consider the problem

$$\begin{aligned}
-\Delta u - k^2 n^2 u &= 1 \qquad \text{in } \Omega, \\
\partial_n u - iku &= 0 \qquad \text{on } \Gamma.
\end{aligned}$$

The index of refraction n is given in polar coordinates by $n \equiv n_1 \equiv 1$ for $r \leq 1/2$ and $n \equiv n_2 \equiv 2$ for $1/2 < r \leq 1$. The exact solution can be derived by elementary calculations. In fact, the solution can be derived by separation of variables, in polar coordinates, and is given by

$$u(r) = \begin{cases} c_1 J_0(kn_1 r) - \frac{1}{(kn_1)^2} & r \leq 1/2, \\ c_2 J_0(kn_2 r) + c_3 Y_0(kn_2 r) - \frac{1}{(kn_2)^2} & r > 1/2, \end{cases}$$

where J_0 and Y_0 are the Bessel functions of order zero and the constants c_1, c_2 and c_3 can be determined using the Robin boundary conditions as well as the interface conditions. For the numerical studies, we solve this problem using h-FEM with polynomial degrees $p = 1$, 2, 3 and 4. It is important to note that the interface Γ_i is resolved by the mesh. Therefore, the observed rates are optimal with respect to the employed finite element space, since the solution is piecewise smooth. The results are visualized in Figure 6.1. Note that as in the homogeneous case higher order versions are less prone to the pollution effect, see also [EM12, Sec. 4.3].

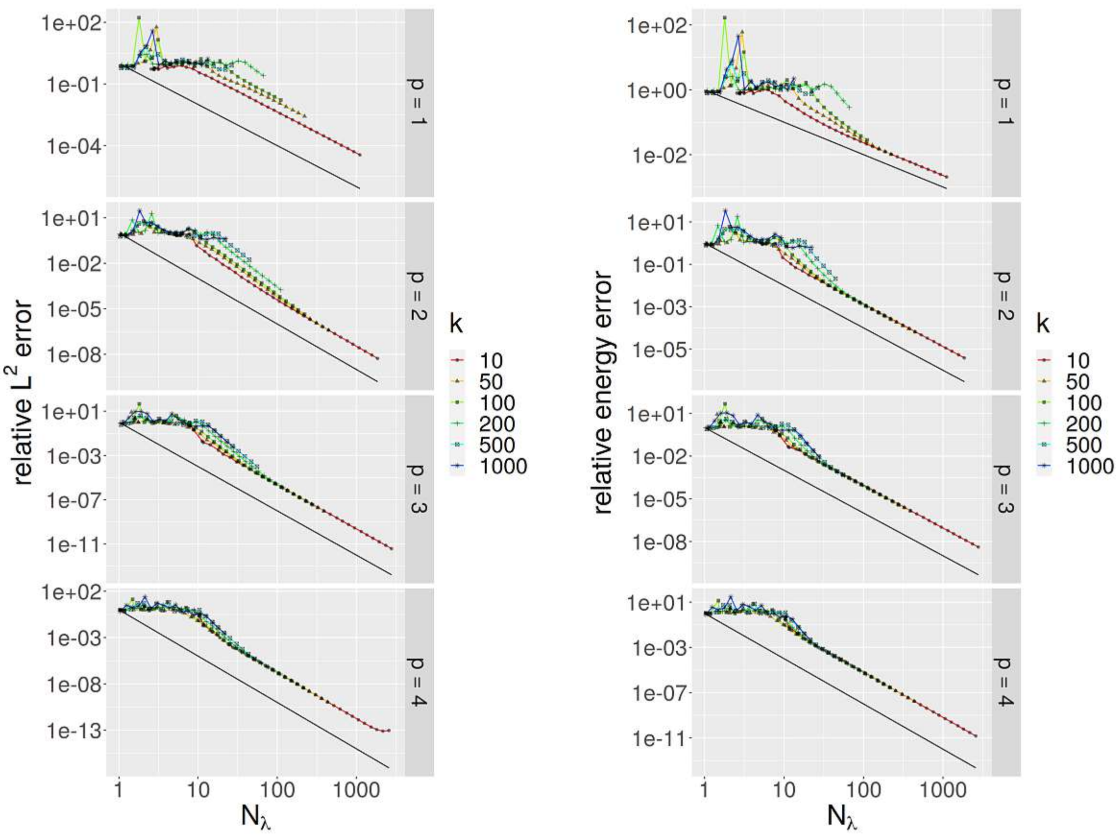

Figure 6.1.: Numerical results of the h-FEM for $p = 1, 2, 3, 4$ as described in Example 6.7.1. Relative $L^2(\Omega)$ error (left) with reference line in black corresponding to h^{p+1}. Relative energy error (right) with reference line in black corresponding to h^p.

Example 6.7.2. Let Ω be the unit circle in $\mathbb{R}^2$ and Γ_i the quadrilateral with corners $(-1/2, -1/2)$, $(1/2, -1/2)$, $(-1/2, 1/2)$ and $(1/2, 1/2)$. The index of refraction n is given by $n \equiv n_1 \equiv 1$ inside of Γ_i and $n \equiv n_2 \equiv 2$ otherwise. We chose $u(x, y) = \mathrm{e}^{i(k_1 x + k_2 y)}$ with $k_1 = -k_2 = \frac{1}{\sqrt{2}} k$ to be the exact solution and calculate the data f and g such that

$$\begin{aligned} -\Delta u - k^2 n^2 u &= f && \text{in } \Omega, \\ \partial_n u - iku &= g && \text{on } \Gamma. \end{aligned}$$

For the numerical studies, this problem will be solved using h-FEM with polynomial degrees $p = 1, 2, 3$ and 4. Again the interface Γ_i is resolved by the mesh. The results are visualized in Figure 6.2.

Example 6.7.3. Let again Ω be the unit circle in $\mathbb{R}^2$ and consider the problem

$$\begin{aligned} -\Delta u - k^2 n^2 u &= f && \text{in } \Omega, \\ \partial_n u - \alpha \Delta_\Gamma u - \beta u &= g && \text{on } \Gamma. \end{aligned}$$

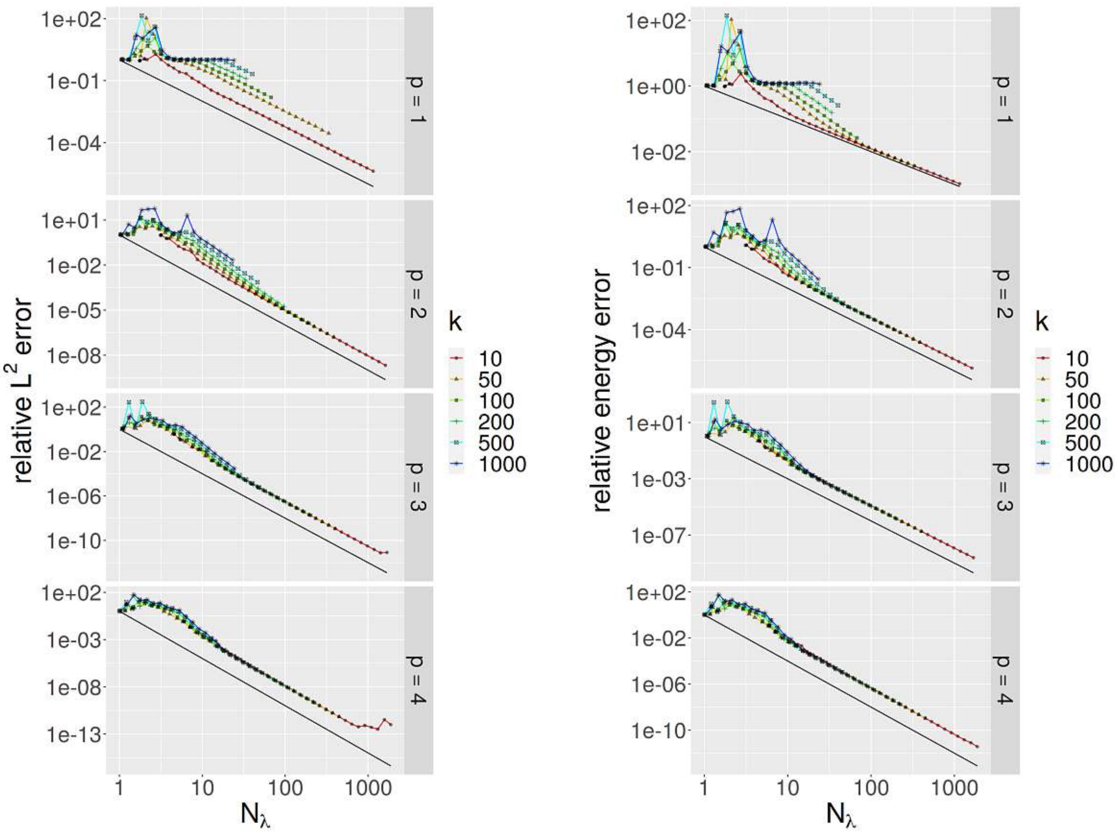

Figure 6.2.: Numerical results of the h-FEM for $p = 1, 2, 3, 4$ as described in Example 6.7.2. Relative $L^2(\Omega)$ error (left) with reference line in black corresponding to h^{p+1}. Relative energy error (right) with reference line in black corresponding to h^p.

The index of refraction n is given in polar coordinates by $n \equiv n_1 \equiv 1$ for $r \leq 1/2$ and $n \equiv n_2 \equiv 2$ for $1/2 < r \leq 1$. The parameters α and β are chosen according to [Fen84]:

$$\alpha = -\frac{i}{2k} \quad \text{and} \quad \beta = ik - \frac{1}{2} - \frac{i}{8k}.$$

The exact solution is chosen to be $u(x, y) = \sin(k(x + y))$. The right-hand sides f and g are calculated accordingly. Note that on the unit sphere the surface gradient can be expressed for sufficiently smooth functions as the trace of $(-y, x)^T \cdot \nabla u$, which allows for straightforward numerical discretization of the problem in question. For the numerical studies, we solve this problem using h-FEM with polynomial degrees $p = 1$, 2, 3 and 4. It is important to note that the interface Γ_i is resolved by the mesh. The observed rates are optimal with respect to the employed finite element space, since the solution is piecewise smooth. The results are visualized in Figure 6.3.

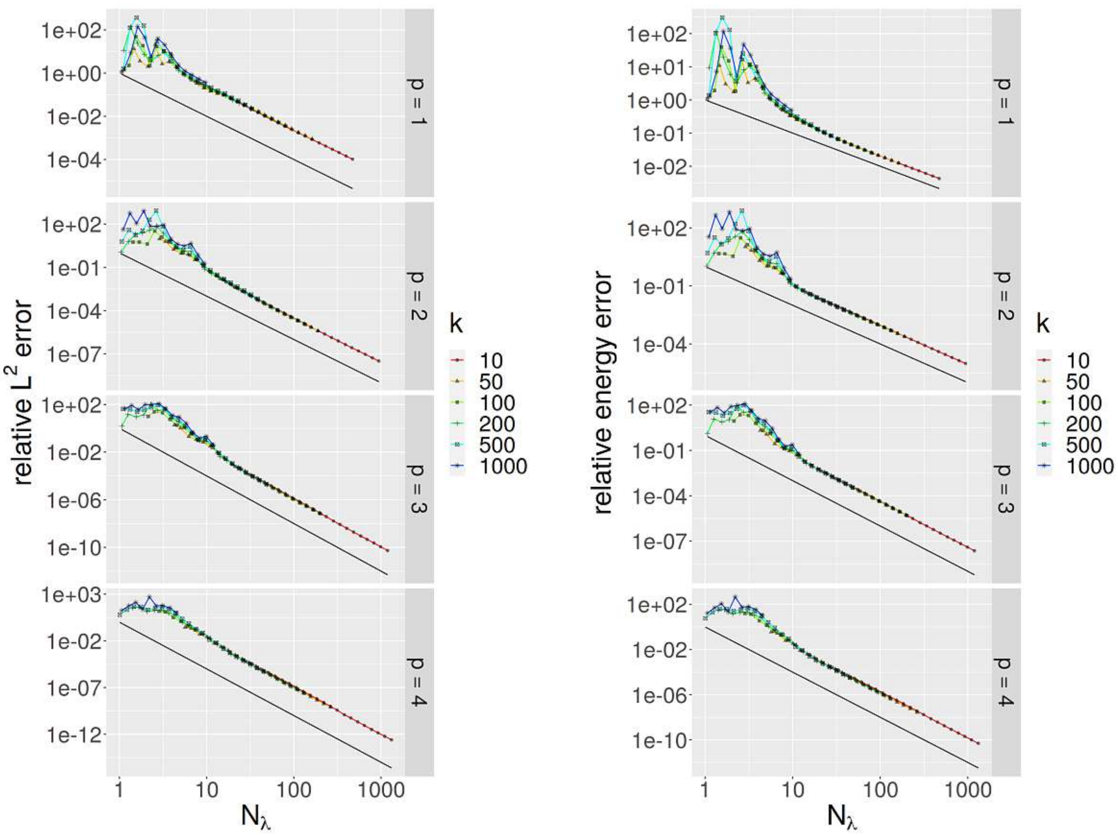

Figure 6.3.: Numerical results of the h-FEM for $p = 1, 2, 3, 4$ as described in Example 6.7.3. Relative $L^2(\Omega)$ error (left) with reference line in black corresponding to h^{p+1}. Relative energy error (right) with reference line in black corresponding to h^p.

6.8. Analytic regularity for second order absorbing boundary conditions

The present section develops similar results as [Mel02, Sec. 5.5] for a model problem with second order boundary conditions. We start by introducing general notation: For $d \geq 2$ and $R > 0$ let $B_R \subset \mathbb{R}^d$ denote the ball of radius R with center in the origin. Let $B_R^+ \subset \mathbb{R}^d$ be a half ball with radius R, i.e., $B_R^+ = \{x \in B_R \colon x_d > 0\}$. Furthermore, let $\Gamma_R := \{x \in B_R \colon x_d = 0\}$. We consider functions u that satisfy

$$\begin{aligned} -\nabla \cdot (A\nabla u) &= f && \text{in } B_R^+, \\ \partial_{n_A} u + \alpha \nabla_\Gamma \cdot (A_\Gamma \nabla_\Gamma u) &= \alpha^{1/2} g + G && \text{on } \Gamma_R, \end{aligned} \tag{6.52}$$

with the conormal derivative $\partial_{n_A} u = n \cdot (A\nabla u)$. We assume that solutions u of (6.52) satisfy

$$a(u,v) := \int_{B_R^+} (A\nabla u)\cdot\nabla\overline{v} + \alpha\int_{\Gamma_R}(A_\Gamma\nabla_\Gamma u)\cdot\nabla_\Gamma\overline{v} \tag{6.53}$$
$$= \int_{B_R^+} f\overline{v} + \int_{\Gamma_R}(\alpha^{1/2}g+G)\overline{v} \qquad \forall v\in C_0^\infty(B_R^+).$$

To describe the parameter α, it is convenient to introduce, for fixed $\theta>0$, the sector

$$\mathcal{S}_\theta := \{\alpha\in\mathbb{C} \,|\, |\arg\alpha| \le \pi-\theta\}.$$

6.8.1. The shift theorem in tangential direction

The proof of Lemma 6.8.2 uses the well-established difference quotients method of Nirenberg that can be found, e.g., in [Eva10, Sec. 6.3]. For $j\in\{1,\dots,d-1\}$, the j-th unit vector $e_j\in\mathbb{R}^d$, and $h\in\mathbb{R}\setminus\{0\}$ we introduce the translation operator τ_j^h by $(\tau_j^h v)(x) = v(x+he_j)$ and the difference quotient $(D_j^h v)(x) = h^{-1}(v(x+he_j)-v(x))$. Inspection of the proof of [Eva10, Sec. 5.8.2, Thm. 3] shows that for fixed $0<r_1<r_2$

$$\partial_{x_j} v\in L^2(B_{r_2}) \quad\Longrightarrow\quad \|D_j^h v\|_{L^2(B_{r_1})} \le \|\partial_{x_j} v\|_{L^2(B_{r_2})} \quad \forall |h|<r_2-r_1, \tag{6.54}$$

as well as

$$\|D_j^h v\|_{L^2(B_{r_1})} \le C_v \quad \forall |h|\le r_2-r_1 \quad\Longrightarrow\quad \|\partial_{x_j} v\|_{L^2(B_{r_1})} \le C_v. \tag{6.55}$$

Lemma 6.8.1. *Let G, $u\in H^1(B_R^+)$. Let r, $\delta>0$ with $r+\delta<R$ be given. Let $\chi\in C_0^\infty(\mathbb{R}^d;\mathbb{R})$ be a cut-off function with $\operatorname{supp}\chi\subset B_{r+\delta/2}$ and $\chi\equiv 1$ on B_r. Then there exists a constant $C>0$ depending only on the spatial dimension, such that*

$$\left|\int_{\Gamma_R} G D_{x_j}^{-h}(\chi D_{x_j}^h u)\right| \le$$
$$C\left(\delta^{-1}\|G\|_{L^2(B_{r+\delta}^+)} + \|\nabla G\|_{L^2(B_{r+\delta}^+)}\right)\left(\delta^{-1}\|\nabla u\|_{L^2(B_{r+\delta}^+)} + \|\chi\nabla D_{x_j}^h u\|_{L^2(B_{r+\delta}^+)}\right).$$

Proof. Let χ' be another cut-off function with $\operatorname{supp}\chi'\subset B_{r+\delta}$, $\chi'\equiv 1$ on $\operatorname{supp}\chi$ and $\|\nabla\chi'\|_{L^\infty}\le C\delta^{-1}$. Then, for h sufficiently small (depending only on r, δ)

$$\int_{\Gamma_R} G D_j^{-h}(\chi D_j^h u) = \int_{\Gamma_R} \chi' G D_j^{-h}(\chi D_j^h u).$$

Let $v = D_j^{-h}(\chi D_j^h u)$. Scaling B_R^+ to a half ball of radius 1, we denote by $\widehat{G}$ the scaled version of $\chi' G$ and by $\widehat{v}$ the scaled version of v. It is also convenient to define ψ as the solution of the Neumann problem $-\Delta\psi = 0$ in B_1^+ and $\partial_n\psi = \widehat{v}$ on ∂B_1^+. Note that $\widehat{v}\equiv 0$ on $\partial B_1^+\setminus\Gamma_1$. Furthermore, $\int_{\Gamma_1}\widehat{v} = 0$ and therefore the solvability condition for the above Neumann problem is satisfied. We then have

$$R^{-(d-1)}\int_{\Gamma_R}\chi' G v = \int_{\partial B_1^+}\widehat{G}\widehat{v} = \int_{\partial B_1^+}\widehat{G}\partial_n\psi = \int_{B_1^+}\nabla\widehat{G}\cdot\nabla\psi$$
$$\le \|\nabla\widehat{G}\|_{L^2(B_1^+)}\|\nabla\psi\|_{L^2(B_1^+)}$$
$$\le R^{(1-d)/2}\|\nabla(\chi' G)\|_{L^2(B_R^+)}\|\widehat{v}\|_{H^{-1/2}(\partial B_1^+)}.$$

Since $\widehat{v} \equiv 0$ on $\partial B_1^+ \setminus \Gamma_1$, we estimate $\|\widehat{v}\|_{H^{-1/2}(\partial B_1^+)} \leq C\|\widehat{v}\|_{H^{-1/2}(\Gamma_1)}$. To estimate this last norm, we write $w = \chi D_j^h u$ and denote by $\widehat{w}$ the corresponding scaled function. We observe for $\varphi \in C_0^\infty(\Gamma_1)$ and h sufficiently small (depending only on δ, r, R)

$$\begin{aligned}\left|\int_{\Gamma_1} (D_j^{-h/R}\widehat{w})\varphi\right| &= \left|\int_{\Gamma_1} \widehat{w} D_j^{h/R}\varphi\right| \\ &\leq \|\widehat{w}\|_{L^2(\Gamma_1)}\|D_j^{h/R}\varphi\|_{L^2(\Gamma_2)} \\ &\overset{(6.54)}{\leq} \|\widehat{w}\|_{L^2(\Gamma_1)}\|\partial_{x_j}\varphi\|_{L^2(\Gamma_1)} \\ &\leq \|\widehat{w}\|_{L^2(\Gamma_1)}\|\varphi\|_{H^1(\Gamma_1)}.\end{aligned}$$

Hence, $\|D_j^{-h/R}\widehat{w}\|_{H^{-1}(\Gamma_1)} \leq C\|\widehat{w}\|_{L^2(\Gamma_1)}$ uniformly in h. Similarly, (6.54) shows

$$\|D_{x_j}^{-h/R}\widehat{w}\|_{L^2(\Gamma_1)} \leq C\|\partial_{x_j}\widehat{w}\|_{L^2(\Gamma_1)} \leq C\|\widehat{w}\|_{H^1(\Gamma_1)}$$

uniformly in h. By interpolation, we arrive at

$$\|D_{x_j}^{-h/R}\widehat{w}\|_{H^{-1/2}(\Gamma_1)} \leq C\|\widehat{w}\|_{H^{1/2}(\Gamma_1)} \leq C\|\widehat{w}\|_{H^1(B_1^+)} \leq C\|\nabla\widehat{w}\|_{L^2(B_1^+)},$$

where the penultimate estimate follows from the trace inequality and the last one by a Poincaré inequality, which is applicable due to the support properties of $\widehat{w}$. In conclusion, we arrive at

$$\begin{aligned}&\left|\int_{\Gamma_R} G D_{x_j}^{-h}(\chi D_{x_j}^h u)\right| \leq C\|\nabla(\chi' G)\|_{L^2(B_R^+)}\|\nabla(\chi D_{x_j}^h u)\|_{L^2(B_R^+)} \\ &\quad\leq C\left(\delta^{-1}\|G\|_{L^2(B_{r+\delta}^+)} + \|\nabla G\|_{L^2(B_{r+\delta}^+)}\right)\left(\delta^{-1}\|\nabla u\|_{L^2(B_{r+\delta}^+)} + \|\chi\nabla D_{x_j}^h u\|_{L^2(B_{r+\delta}^+)}\right).\end{aligned}$$

□

Lemma 6.8.2. *Let $A \in C^1(\overline{B}_R^+)$, $A_\Gamma \in C^1(\overline{\Gamma}_R^+)$ be matrix-valued functions that are pointwise symmetric positive definite with lower bound on the eigenvalues $\lambda_{\min} > 0$. Let $\alpha \in \mathcal{S}_\theta$. Let $f \in L^2(B_R^+)$, $g \in L^2(\Gamma_R)$ and $G \in H^1(B_R^+)$. Then there exists $C_{\text{stab}} > 0$ depending only on θ, a lower bound on $\lambda_{\min}$, and an upper bound on $\|A\|_{L^\infty} + R\|\nabla A\|_{L^\infty} + \|A_\Gamma\|_{L^\infty} + R\|\nabla A_\Gamma\|_{L^\infty}$ such that any solution u of (6.52) satisfies for all r, $\delta > 0$ with $r + \delta < R$*

$$\begin{aligned}&\|\nabla^2 u\|_{L^2(B_r)} + |\alpha|^{1/2}\|\nabla_\Gamma^2 u\|_{L^2(\Gamma_r)} \leq \\ &\quad C_{\text{stab}}\Big(\|f\|_{L^2(B_{r+\delta}^+)} + \|g\|_{L^2(\Gamma_{r+\delta})} + \delta^{-1}\|G\|_{L^2(B_{r+\delta}^+)} + \|\nabla G\|_{L^2(B_{r+\delta}^+)} \\ &\qquad + \delta^{-1}\|\nabla u\|_{L^2(B_{r+\delta}^+)} + |\alpha|^{1/2}\delta^{-1}\|\nabla_\Gamma u\|_{L^2(\Gamma_{r+\delta})}\Big).\end{aligned} \tag{6.56}$$

Proof. Step 1: A calculation reveals that $\alpha \in S_\theta$ implies the existence of $c_{\text{coer}} > 0$ such that

$$|y_1 + \alpha y_2| \geq c_{\text{coer}}(y_1 + |\alpha| y_2) \qquad \forall y_1, y_2 \geq 0. \tag{6.57}$$

Step 2: Let $\chi \in C_0^\infty(\mathbb{R}^d;\mathbb{R})$ be a cut-off function with $\operatorname{supp}\chi \subset B_{r+\delta/2}$ and $\chi \equiv 1$ on B_r. We assume furthermore $\|\nabla\chi\|_{L^\infty} \leq C\delta^{-1}$, and $\partial_{n_A}\chi = 0$ on Γ_R, see [Mel02, Lemma 5.5.21] for a similar construction. For h sufficiently small, we select the test function $v = -D_j^{-h}\chi^2 D_j^h u$ in (6.53) and get

$$a(u, -D_j^{-h}\chi^2 D_j^h u) = -\int_{B_R^+} f D_j^{-h}\chi^2 D_j^h \overline{u} - \int_{\Gamma_R} (\alpha^{1/2} g + G) D_j^{-h}\chi^2 D_j^h \overline{u}. \tag{6.58}$$

We treat the left-hand and the right-hand side separately. We proceed as in the proof of [Eva10, Sec. 6.3, Thm. 1]. We have

$$\begin{aligned}
-\int_{B_R^+} (A\nabla u)\cdot\nabla(D_j^{-h}\chi^2 D_j^h\overline{u}) &= \int_{B_R^+} D_j^h(A\nabla u)\cdot\nabla(\chi^2 D_j^h\overline{u}) \\
&= \int_{B_R^+} ((\tau_j^h A)(D_j^h\nabla u) + (D_j^h A)\nabla u)\cdot\nabla(\chi^2 D_j^h\overline{u}) \\
&= \int_{B_R^+} ((\tau_j^h A)(D_j^h\nabla u) + (D_j^h A)\nabla u)\cdot\left(\chi^2\nabla D_j^h\overline{u} + 2\chi\nabla\chi D_j^h\overline{u}\right) \\
&= \int_{B_R^+} \chi^2((\tau_j^h A)\nabla D_j^h u)\cdot\nabla D_j^h\overline{u} + R_{\mathrm{vol}}
\end{aligned}$$

with R_{vol} given by

$$R_{\mathrm{vol}} = \int_{B_R^+} (\tau_j^h A)(D_j^h\nabla u)2\chi\nabla\chi D_j^h\overline{u} + (D_j^h A)\nabla u\cdot\left(\chi^2\nabla D_j^h\overline{u} + 2\chi\nabla\chi D_j^h\overline{u}\right).$$

Hence, we can estimate

$$\begin{aligned}
|R_{\mathrm{vol}}| \leq C\delta^{-1}\|A\|_{L^\infty}\|\chi\nabla D_j^h u\|_{L^2(B_R^+)}\|D_j^h u\|_{L^2(B_{r+\delta/2}^+)} \\
+ C\|\nabla A\|_{L^\infty}\|\nabla u\|_{L^2(B_{r+\delta/2}^+)}\|\chi\nabla D_j^h u\|_{L^2(B_R^+)} \\
+ C\delta^{-1}\|\nabla A\|_{L^\infty}\|\nabla u\|_{L^2(B_{r+\delta/2}^+)}\|D_j^h u\|_{L^2(B_{r+\delta/2}^+)}.
\end{aligned}$$

Analogously, we get

$$-\int_{\Gamma_R} (A_\Gamma\nabla_\Gamma)\cdot\nabla_\Gamma D_j^{-h}\chi^2 D_j^h\overline{u} = \int_{\Gamma_R} \chi^2((\tau_j^h A_\Gamma)\nabla_\Gamma D_j^h u)\cdot\nabla_\Gamma D_j^h u + R_{\mathrm{bnd}}$$

with

$$\begin{aligned}
|R_{\mathrm{bnd}}| \leq C\delta^{-1}\|A_\Gamma\|_{L^\infty}\|\chi\nabla_\Gamma D_j^h u\|_{L^2(\Gamma_R)}\|D_j^h u\|_{L^2(\Gamma_{r+\delta/2})} \\
+ C\|\nabla A_\Gamma\|_{L^\infty}\|\nabla_\Gamma u\|_{L^2(\Gamma_{r+\delta})}\|\chi\nabla_\Gamma D_j^h u\|_{L^2(\Gamma_R)} \\
+ C\delta^{-1}\|\nabla A_\Gamma\|_{L^\infty}\|\nabla_\Gamma u\|_{L^2(\Gamma_{r+\delta/2})}\|D_j^h u\|_{L^2(\Gamma_{r+\delta/2})}.
\end{aligned}$$

For the right-hand side of (6.58), we get with (6.54) applied to $v = -D_j^{-h}\chi^2 D_j^h u$ and Lemma 6.8.1

$$\begin{aligned}
&\left|\int_{B_R^+} f D_j^{-h}\chi^2 D_j^h \overline{u}\right| \leq C\|f\|_{L^2(B^+_{r+\delta/2})}\left(\delta^{-1}\|D_j^h u\|_{L^2(B^+_{r+\delta/2})} + \|\chi\nabla D_j^h u\|_{L^2(B_R^+)}\right),\\
&\left|\int_{\Gamma_R} g D_j^{-h}\chi^2 D_j^h \overline{u}\right| \leq C\|g\|_{L^2(\Gamma_{r+\delta/2})}\left(\delta^{-1}\|D_j^h u\|_{L^2(\Gamma_{r+\delta/2})} + \|\chi\nabla_\Gamma D_j^h u\|_{L^2(\Gamma_R)}\right),\\
&\left|\int_{\Gamma_R} G D_j^{-h}\chi^2 D_j^h \overline{u}\right| = \left|\int_{\Gamma_R} D_j^h G\chi^2 D_j^h \overline{u}\right|\\
&\qquad \leq C\left(\delta^{-1}\|G\|_{L^2(B^+_{r+\delta})} + \|\nabla G\|_{L^2(B^+_{r+\delta})}\right)\left(\delta^{-1}\|\nabla u\|_{L^2(B^+_{r+\delta})} + \|\chi\nabla D_j^h u\|_{L^2(B^+_{r+\delta})}\right).
\end{aligned}$$

Step 3: Using (6.57) we get

$$\begin{aligned}
&c_{\text{coer}}\left(\int_{B_R^+}\chi^2((\tau_j^h A)\nabla D_j^h u)\cdot\nabla D_j^h\overline{u} + |\alpha|\int_{\Gamma_R}\chi^2((\tau_j^h A)\nabla_\Gamma D_j^h u)\cdot\nabla_\Gamma D_j^h\overline{u}\right)\\
&\quad\leq |a(u, -D_j^{-h}\chi^2 D_j^h u) - R_{\text{vol}} - \alpha R_{\text{bnd}}|\\
&\quad\leq \left|\int_{B_R^+} f D_j^{-h}\chi^2 D_j^h\overline{u}\right| + \left|\int_{\Gamma_R}(\alpha^{1/2}g + G)D_j^{-h}\chi^2 D_j^h\overline{u}\right| + |R_{\text{vol}}| + |\alpha||R_{\text{bnd}}|.
\end{aligned}$$

Using the lower bound for A and A_Γ and the above estimates together with (6.54) we find

$$\begin{aligned}
&c_{\text{coer}}\lambda_{\min}\left(\|\chi\nabla D_j^h u\|^2_{L^2(B_R^+)} + |\alpha|\|\chi\nabla_\Gamma D_j^h u\|^2_{L^2(\Gamma_R)}\right) \leq\\
&\quad C\Big(\|f\|_{L^2(B^+_{r+\delta})}\delta^{-1}\|\partial_{x_j} u\|_{L^2(B_{r+\delta})} + \|f\|_{L^2(B^+_{r+\delta})}\|\chi\nabla D_j^h u\|_{L^2(B_R^+)}\\
&\qquad + \|g\|_{L^2(B^+_{r+\delta})}\delta^{-1}|\alpha|^{1/2}\|\partial_{x_j} u\|_{L^2(\Gamma_{r+\delta})} + \|g\|_{L^2(\Gamma_{r+\delta})}|\alpha|^{1/2}\|\chi\nabla_\Gamma D_j^h u\|_{L^2(\Gamma_R)}\\
&\qquad + \left(\delta^{-1}\|G\|_{L^2(B^+_{r+\delta})} + \|\nabla G\|_{L^2(B^+_{r+\delta})}\right)\left(\delta^{-1}\|\nabla u\|_{L^2(B^+_{r+\delta})} + \|\chi\nabla D_j^h u\|_{L^2(B^+_{r+\delta})}\right)\\
&\qquad + (\|A\|_{L^\infty} + \delta\|\nabla A\|_{L^\infty})\delta^{-1}\|\nabla u\|_{L^2(B^+_{r+\delta})}\|\chi\nabla D_j^h u\|_{L^2(B_R^+)}\\
&\qquad + \delta\|\nabla A\|_{L^\infty}\delta^{-2}\|\nabla u\|^2_{L^2(B^+_{r+\delta})}\\
&\qquad + (\|A_\Gamma\|_{L^\infty} + \delta\|\nabla A_\Gamma\|_{L^\infty})|\alpha|\delta^{-1}\|\nabla_\Gamma u\|_{L^2(\Gamma_{r+\delta})}\|\chi\nabla_\Gamma D_j^h u\|_{L^2(\Gamma_R)}\\
&\qquad + \delta\|\nabla A_\Gamma\|_{L^\infty}|\alpha|\delta^{-2}\|\nabla_\Gamma u\|^2_{L^2(\Gamma_{r+\delta})}\Big).
\end{aligned}$$

The Cauchy-Schwarz inequality with epsilon allows us to absorb the terms $\|\chi\nabla D_j^h u\|_{L^2(B_R^+)}$

and $\|\chi\nabla_\Gamma D_j^h u\|_{L^2(\Gamma_R)}$ from the right-hand side into the left-hand side and we get

$$\begin{aligned}
&\frac{1}{2}c_{\mathrm{coer}}\lambda_{\min}\left(\|\chi\nabla D_j^h u\|^2_{L^2(B_R^+)} + |\alpha|\|\chi\nabla_\Gamma D_j^h u\|^2_{L^2(\Gamma_R)}\right) \leq \\
&\quad C\Big(\|f\|_{L^2(B^+_{r+\delta})}\delta^{-1}\|\partial_{x_j}u\|_{L^2(B_{r+\delta})} + \lambda_{\min}^{-1}\|f\|^2_{L^2(B^+_{r+\delta})} \\
&\qquad + \|g\|_{L^2(B^+_{r+\delta})}\delta^{-1}|\alpha|^{1/2}\|\partial_{x_j}u\|_{L^2(\Gamma_{r+\delta})} + \lambda_{\min}^{-1}\|g\|^2_{L^2(\Gamma_{r+\delta})} \\
&\qquad + \left(\delta^{-1}\|G\|_{L^2(B^+_{r+\delta})} + \|\nabla G\|_{L^2(B^+_{r+\delta})}\right)\delta^{-1}\|\nabla u\|_{L^2(B^+_{r+\delta})} \\
&\qquad + R^2\lambda_{\min}^{-1}\left(\delta^{-1}\|G\|_{L^2(B^+_{r+\delta})} + \|\nabla G\|_{L^2(B^+_{r+\delta})}\right)^2 \\
&\qquad + \lambda_{\min}^{-1}(\|A\|_{L^\infty} + \delta\|\nabla A\|_{L^\infty})^2\delta^{-2}\|\nabla u\|^2_{L^2(B^+_{r+\delta})} \\
&\qquad + \delta\|\nabla A\|_{L^\infty}\delta^{-2}\|\nabla u\|^2_{L^2(B^+_{r+\delta})} \\
&\qquad + \lambda_{\min}^{-1}(\|A_\Gamma\|_{L^\infty} + \delta\|\nabla A_\Gamma\|_{L^\infty})^2|\alpha|\delta^{-2}\|\nabla_\Gamma u\|^2_{L^2(\Gamma_{r+\delta})} \\
&\qquad + \delta\|\nabla A_\Gamma\|_{L^\infty}|\alpha|\delta^{-2}\|\nabla_\Gamma u\|^2_{L^2(\Gamma_{r+\delta})}\Big).
\end{aligned}$$

From (6.55), we get in the limit $h \to 0$ that $\nabla\partial_{x_j}u \in L^2(B_r^+)$ and $\nabla_\Gamma\partial_{x_j}u \in L^2(\Gamma_r^+)$ together with

$$\begin{aligned}
&\|\nabla\nabla_{x'}u\|_{L^2(B_r)} + |\alpha|^{1/2}\|\nabla_\Gamma\nabla_{x'}u\|_{L^2(\Gamma_r)} \leq \\
&\quad C_{\mathrm{stab}}\Big(\|f\|_{L^2(B^+_{r+\delta})} + \|g\|_{L^2(\Gamma_{r+\delta})} + \delta^{-1}\|G\|_{L^2(B^+_{r+\delta})} + \|\nabla G\|_{L^2(B^+_{r+\delta})} \\
&\qquad + \delta^{-1}\|\nabla u\|_{L^2(B^+_{r+\delta})} + |\alpha|^{1/2}\delta^{-1}\|\nabla_\Gamma u\|_{L^2(\Gamma_{r+\delta})}\Big)
\end{aligned} \tag{6.59}$$

with $C_{\mathrm{stab}} > 0$ depending only on θ, a lower bound on $\lambda_{\min}$, and an upper bound on $\|A\|_{L^\infty} + R\|\nabla A\|_{L^\infty} + \|A_\Gamma\|_{L^\infty} + R\|\nabla A_\Gamma\|_{L^\infty}$.

Step 4: We complete the proof by controlling $\|\partial^2_{x_d}u\|_{L^2(B_r^+)}$. This follows from the differential equation

$$-A_{dd}\partial^2_{x_d}u = f + \sum_{i,j=1}^{d}(\partial_{x_i}A_{ij})\partial_{x_j}u - \sum_{(i,j)\neq(d,d)}A_{ij}\partial_{x_i}\partial_{x_j}u.$$

We have $A_{dd} = e_d^\top A e_d \geq \lambda_{\min}$ and therefore

$$\lambda_{\min}\|\partial^2_{x_d}u\|_{L^2(B_r^+)} \leq \|f\|_{L^2(B_r^+)} + C\delta\|\nabla A\|_{L^\infty}\delta^{-1}\|\nabla u\|_{L^2(B_r^+)} + \|A\|_{L^\infty}\|\nabla\nabla_{x'}u\|_{L^2(B_r^+)}.$$

Noting $\frac{\|A\|_{L^\infty}}{\lambda_{\min}} \geq 1$ together with (6.59) concludes the proof.

□

6.8.2. Control of the tangential derivatives

For functions $v = (v_i)_{i\in\mathbf{I}}$, ($\mathbf{I}$ some finite index set) defined on $\mathbb{R}^d$ we introduce the notation

$$|\nabla^p v|^2 = \sum_{i\in\mathbf{I}}\sum_{\boldsymbol{\alpha}\in\mathbb{N}_0^d:\,|\boldsymbol{\alpha}|=p}\frac{|\boldsymbol{\alpha}|!}{\boldsymbol{\alpha}!}|D^{\boldsymbol{\alpha}}v_i|^2. \tag{6.60}$$

Analogously, the notation $|\nabla^p_{x'} v|$ for $x' \in \mathbb{R}^{d-1}$ indicates that $\boldsymbol{\alpha} \in \mathbb{N}^{d-1}$ in the sum. We proceed as described, e.g., in [Mel02, Sec. 5.5.3 and 5.5.4]. We write $\nabla_{x'}$ for the tangential derivatives in contrast to [Mel02, Sec. 5.5.3], where ∇_x denotes the tangential derivatives and introduce

$$
\begin{aligned}
M'_{R,p}(f) &:= \frac{1}{p!} \sup_{R/2 \le r < R} (R-r)^{p+2} \|\nabla^p_{x'} f\|_{L^2(B^+_r)}, \\
M'_{R,p,\Gamma}(g) &:= \frac{1}{p!} \sup_{R/2 \le r < R} (R-r)^{p+2} \|\nabla^p_{x'} g\|_{L^2(\Gamma_r)}, \\
N'_{R,p}(v) &:= \begin{cases} \frac{1}{p!} \sup_{R/2 \le r < R} (R-r)^{p+2} \|\nabla^2 \nabla^p_{x'} v\|_{L^2(B^+_r)} & \text{if } p \ge 0, \\ \sup_{R/2 \le r < R} (R-r)^{p+2} \|\nabla^{2+p} v\|_{L^2(B^+_r)} & \text{if } p = -1,\, -2, \end{cases} \\
N'_{R,p,\Gamma}(v) &:= \begin{cases} \frac{1}{p!} \sup_{R/2 \le r < R} (R-r)^{p+2} \|\nabla^{p+2}_{x'} v\|_{L^2(\Gamma_r)} & \text{for } p \ge 0, \\ \sup_{R/2 \le r < R} (R-r)^{p+2} \|\nabla^{p+2}_{x'} v\|_{L^2(\Gamma_r)} & \text{for } p = -2,\, -1, \end{cases} \\
H_{R,p}(v) &:= \frac{1}{[p-1]!} \sup_{R/2 \le r < R} (R-r)^{p+1} \left[\|\nabla^p_{x'} v\|_{L^2(B^+_r)} + \frac{R-r}{[p]} \|\nabla^p_{x'} \nabla v\|_{L^2(B^+_r)} \right],
\end{aligned}
$$

where

$$[p]! := \max\{p, 1\}!.$$

Analogous to [Mel02, Lemmas 5.5.15 and 5.5.23] is

Lemma 6.8.3. *Let $p \in \mathbb{N}_0$. Let $A \in C^{p+1}(\overline{B}^+_R)$, $A_\Gamma \in C^{p+1}(\overline{\Gamma}^+_R)$ be matrix-valued functions that are pointwise symmetric positive definite with lower bound on the eigenvalues $\lambda_{\min} > 0$. Let $f \in H^p(B^+_R)$, $g \in H^p(\Gamma^+_R)$, $G \in H^{p+1}(B^+_R)$ and $\alpha \in \mathcal{S}_\theta$. There exists a constant $C_B > 0$ depending only on the same quantities as the constant C_{stab} in Lemma 6.8.2 such that*

$$
\begin{aligned}
N'_{R,p}(u) + |\alpha|^{1/2} N'_{R,p,\Gamma}(u) \le C_B \Bigg\{ & M'_{R,p}(f) + M'_{R,p,\Gamma}(g) + H_{R,p}(G) \\
&+ \sum_{q=1}^{p+1} \binom{p+1}{q} \left(\left(\frac{R}{2}\right)^q \|\nabla^q A\|_{L^\infty(B^+_R)} + \left(\frac{R}{2}\right)^{q-1} q \|\nabla^{q-1} A\|_{L^\infty(B^+_R)} \right) \frac{[p-q]!}{p!} N'_{R,p-q}(u) \\
&+ N'_{R,p-1}(u) + N'_{R,p-2}(u) \\
&+ |\alpha|^{1/2} \left(\sum_{q=1}^{p+1} \binom{p+1}{q} \left(\frac{R}{2}\right)^q \|\nabla^q_{x'} A_\Gamma\|_{L^\infty(\Gamma_R)} \frac{[p-q]!}{p!} N'_{R,p-q,\Gamma}(u) \right. \\
&\left. + N'_{R,p-1,\Gamma}(u) + N'_{R,p-2,\Gamma}(u) \right) \Bigg\}.
\end{aligned}
$$

Proof. The proof follows the lines of the proofs of [Mel02, Lemmas 5.5.12, 5.5.15 and 5.5.23]. We abbreviate $\mathbf{a} = A\mathbf{n}$, where $\mathbf{n}$ is the outer normal vector on Γ_R. The derivative $D^{\boldsymbol{\alpha}'}_{x'} u$

with $\boldsymbol{\alpha}' \in \mathbb{N}_0^{d-1}$ satisfies

$$\begin{aligned}
-\nabla \cdot (A \nabla D_{x'}^{\boldsymbol{\alpha}'} u) &= D_{x'}^{\boldsymbol{\alpha}'} f - \{\nabla \cdot (A \nabla D_{x'}^{\boldsymbol{\alpha}'} u) - D_{x'}^{\boldsymbol{\alpha}'}(\nabla \cdot (A \nabla u))\},\\
\partial_{n_A} D_{x'}^{\boldsymbol{\alpha}'} u + \alpha \nabla_\Gamma \cdot (A_\Gamma \nabla_\Gamma D_{x'}^{\boldsymbol{\alpha}'} u) &= D_{x'}^{\boldsymbol{\alpha}'}(\alpha^{1/2} g_1 + g_2)\\
&\quad + \alpha\{\nabla_{x'} \cdot (A_\Gamma \nabla_{x'} D_{x'}^{\boldsymbol{\alpha}'} u) - D_{x'}^{\boldsymbol{\alpha}'}(\nabla_{x'} \cdot (A_\Gamma \nabla_{x'} u))\}\\
&\quad + \left\{(\mathbf{a} \cdot \nabla D_{x'}^{\boldsymbol{\alpha}'} u) - D_{x'}^{\boldsymbol{\alpha}'}(\mathbf{a} \cdot \nabla u)\right\}.
\end{aligned}$$

The remainder of the proof proceeds analogously to the proof of [Mel02, Lemmas 5.5.12 and 5.5.23] and uses Lemma 6.8.2 with $\delta = \frac{R-r}{p+2}$. Specifically, the terms arising from A, f, G are obtained directly as in the proofs of [Mel02, Lemmas 5.5.12 and 5.5.23]. The terms arising from the Laplace-Beltrami are treated with the same arguments as in the proof of [Mel02, Lemma 5.5.12]. ☐

6.8.3. Tangential control for k-dependent problem

We consider the problem

$$\begin{aligned}
-\nabla \cdot (A \nabla u) - c k^2 u &= f \qquad \text{in } B_R^+,\\
\partial_{n_A} u + \alpha \nabla_\Gamma \cdot (A_\Gamma \nabla_\Gamma u) &= \alpha^{1/2} g + G + bku \qquad \text{on } \Gamma.
\end{aligned} \tag{6.61}$$

For the data, we assume

$$\|\nabla^p f\|_{L^2(B_R^+)} \leq C_f \gamma_f^p R^{-p} \max\{p+1, Rk\}^p \qquad \forall p \in \mathbb{N}_0, \tag{6.62a}$$
$$\|\nabla_{x'}^p g\|_{L^2(\Gamma_R)} \leq C_g \gamma_g^p R^{-p} \max\{p+1, Rk\}^p \qquad \forall p \in \mathbb{N}_0, \tag{6.62b}$$
$$\|\nabla^p G\|_{L^2(B_R^+)} \leq C_G \gamma_G^p R^{-p} \max\{p+1, Rk\}^p \qquad \forall p \in \mathbb{N}_0, \tag{6.62c}$$
$$\|\nabla^p A\|_{L^\infty(B_R^+)} \leq C_A \gamma_A^p p! \qquad \forall p \in \mathbb{N}_0, \tag{6.62d}$$
$$\|\nabla_{x'}^p A_\Gamma\|_{L^\infty(\Gamma_R)} \leq C_{A_\Gamma} \gamma_{A_\Gamma}^p p! \qquad \forall p \in \mathbb{N}_0, \tag{6.62e}$$
$$\|\nabla^p b\|_{L^\infty(B_R^+)} \leq C_b \gamma_b^p p! \qquad \forall p \in \mathbb{N}_0, \tag{6.62f}$$
$$\|\nabla^p c\|_{L^\infty(B_R^+)} \leq C_c \gamma_c^p p! \qquad \forall p \in \mathbb{N}_0. \tag{6.62g}$$

Additionally, we assume that

$$|\alpha|^{1/2} \leq C_\alpha k^{-1/2} \tag{6.63}$$

as well as $\alpha \in S_\theta$ for some $\theta > 0$ uniformly in k.

Theorem 6.8.4. *Let $R \leq 1$. Let f, g, G, A, A_Γ, b and α satisfy* (6.62) *and* (6.63), *respectively. Let u solve (6.61). Then there exists a $K \geq 1$ depending only on the coefficients A, A_Γ of the differential operator, θ, and the constants in* (6.62) *and* (6.63) *such that for all $p \geq -1$*

$$N'_{R,p}(u) + k^{-1/2} N'_{R,p,\Gamma}(u) \leq C_u K^{p+2} \frac{\max\{p+3, Rk\}^{p+2}}{[p]!} \tag{6.64}$$

where

$$\begin{aligned} C_u = &\min\{1, Rk\}^2 (k^{-2} C_f + k^{-2} C_g + C_c \|u\|_{L^2(B_R^+)}) \\ &+ \min\{1, Rk\} k^{-1} C_G (1 + \gamma_G R) \\ &+ \min\{1, Rk\}(1 + \gamma_b R) C_b \|u\|_{L^2(B_R^+)} \\ &+ \min\{1, Rk\}(1 + C_b \min\{1, Rk\}) k^{-1} \|\nabla u\|_{L^2(B_R^+)} \\ &+ \min\{1, Rk\} k^{-3/2} C_\alpha \|\nabla_\Gamma u\|_{L^2(\Gamma_R)}). \end{aligned} \tag{6.65}$$

Proof. We will frequently use the elementary property

$$\min\{a, Rk\} \max\{a, Rk\} = aRk \qquad a > 0.$$

We first verify that (6.64) is correct for $p = -1$, if $K \geq (1 + C_\alpha^{-1})/2$. To that end, note that

$$\begin{aligned} N'_{R,-1}(u) &\leq \frac{R}{2} \|\nabla u\|_{L^2(B_R^+)} = \frac{1}{2} kRk^{-1} \|\nabla u\|_{L^2(B_R^+)} \\ &= \frac{1}{2} \max\{1, Rk\} \min\{1, Rk\} k^{-1} \|\nabla u\|_{L^2(B_R^+)} \\ &\leq \frac{1}{2} \max\{1, Rk\} C_u, \\ k^{-1/2} N'_{R,-1,\Gamma}(u) &\leq k^{-1/2} \frac{R}{2} \|\nabla_\Gamma u\|_{L^2(\Gamma_R)} \\ &\leq \frac{C_\alpha^{-1}}{2} \max\{1, Rk\} \min\{1, Rk\} k^{-1} k^{-1/2} C_\alpha \|\nabla_\Gamma u\|_{L^2(\Gamma_R)} \\ &\leq \frac{C_\alpha^{-1}}{2} \max\{1, Rk\} C_u. \end{aligned}$$

Hence, we have

$$N'_{R,-1}(u) + k^{-1/2} N'_{R,-1,\Gamma}(u) \leq \frac{1 + C_\alpha^{-1}}{2} \max\{1, Rk\} C_u,$$

which concludes the case $p = -1$. We next show that it is correct for $p = 0$ and then by induction for all $p \geq 1$. To that end, we rewrite the equation as

$$\begin{aligned} -\nabla \cdot (A \nabla u) &= f + ck^2 u && \text{in } B_R^+, \\ \partial_{n_A} u + \alpha \nabla_\Gamma \cdot (A_\Gamma \nabla_\Gamma u) &= \alpha^{1/2} g + G + bku && \text{on } \Gamma. \end{aligned}$$

For $p = 0$, an application of Lemma 6.8.2 (used with $\delta \sim R - r$) gives

$$\begin{aligned}
&N'_{R,p}(u) + k^{-1/2} N'_{R,p,\Gamma}(u) \\
&\quad \le C_{\text{stab}} \Big(\frac{R^2}{4} (\|f + ck^2 u\|_{L^2(B_R^+)} + \|g\|_{L^2(\Gamma_R)}) \\
&\qquad + \frac{R}{2} \|G + bku\|_{L^2(B_R^+)} + \frac{R^2}{4} \|\nabla(G + bku)\|_{L^2(B_R^+)} \\
&\qquad + \frac{R}{2} \|\nabla u\|_{L^2(B_R^+)} + |\alpha|^{1/2} \frac{R}{2} \|\nabla_\Gamma u\|_{L^2(\Gamma_R)} \Big) \\
&\quad \le C_{\text{stab}} \Big(R^2 (\|f\|_{L^2(B_R^+)} + C_c k^2 \|u\|_{L^2(B_R^+)} + \|g\|_{L^2(\Gamma_R)}) \\
&\qquad + R\|G\|_{L^2(B_R^+)} + R^2 \|\nabla G\|_{L^2(B_R^+)} \\
&\qquad + Rk(1 + \gamma_b R) C_b \|u\|_{L^2(B_R^+)} + R^2 k C_b \|\nabla u\|_{L^2(B_R^+)} \\
&\qquad + R\|\nabla u\|_{L^2(B_R^+)} + C_\alpha k^{-1/2} R \|\nabla_\Gamma u\|_{L^2(\Gamma_R)} \Big) \\
&\quad \le C_{\text{stab}} \Big(R^2 (C_f + C_g + C_c k^2 \|u\|_{L^2(B_R^+)}) \\
&\qquad + RC_G + R^2 \gamma_G C_G \\
&\qquad + Rk(1 + \gamma_b R) C_b \|u\|_{L^2(B_R^+)} + R^2 k C_b \|\nabla u\|_{L^2(B_R^+)} \\
&\qquad + R\|\nabla u\|_{L^2(B_R^+)} + Rk^{-1/2} C_\alpha \|\nabla_\Gamma u\|_{L^2(\Gamma_R)} \Big) \\
&\quad \le C_{\text{stab}} C_u \max\{3, Rk\}^2,
\end{aligned}$$

where the last estimate follows by similar estimates as in the case $p = -1$. Let us assume that (6.64) is correct for all $-1 \le p' \le p-1$ for some $p \ge 1$. We show that it is correct for p if K is chosen sufficiently large. In fact, we will implicitly assume $K > \max\{\gamma_f, \gamma_g, \gamma_G, \gamma_c, \gamma_b\}$, so that the various geometric series below converge.

It is convenient to abbreviate

$$m(p) := \max\{p + 3, Rk\}^{p+2}. \tag{6.66}$$

In order to apply Lemma 6.8.3, we have to estimate $M'_{R,p}(f - k^2 cu)$, $M'_{R,p,\Gamma}(g)$ and $H'(kbu)$. Since $M'_{R,p}(f - k^2 cu) \le M'_{R,p}(f) + M'_{R,p}(k^2 cu)$, we first estimate $M'_{R,p}(f)$.

$$\begin{aligned}
M'_{R,p}(f) &\le \frac{1}{p!} \left(\frac{R}{2}\right)^{p+2} C_f \gamma_f^p R^{-p} \max\{p + 1, Rk\}^p \\
&\le \frac{C_f}{4} \left(\frac{\gamma_f}{2}\right)^p R^2 \frac{1}{p!} \max\{p + 3, Rk\}^p \\
&\le \frac{C_f}{4} \left(\frac{\gamma_f}{2}\right)^p k^{-2} \min\{1, Rk\}^2 \frac{m(p)}{p!}.
\end{aligned}$$

From [Mel02, Lemma 5.5.13] and the induction hypothesis, we have

$$\begin{aligned}
M'_{R,p}(k^2cu) &\leq k^2 C_c \sum_{q=0}^{p} \left(\gamma_c \frac{R}{2}\right)^q \left(\frac{R}{2}\right)^2 \frac{[p-q-2]!}{(p-q)!} N'_{R,p-q-2}(u) \\
&= \frac{C_c}{4} \min\{1, Rk\}^2 \max\{1, Rk\}^2 \sum_{q=0}^{p} \left(\frac{\gamma_c R}{2}\right)^q \frac{[p-q-2]!}{(p-q)!} N'_{R,p-q-2}(u) \\
&\leq \frac{C_c}{4} \min\{1, Rk\}^2 \max\{1, Rk\}^2 \sum_{q=0}^{p} \left(\frac{\gamma_c R}{2}\right)^q \frac{1}{(p-q)!} K^{p-q} \max\{p-q+1, Rk\}^{p-q} C_u \\
&= \frac{C_c}{4} \min\{1, Rk\}^2 K^p \max\{1, Rk\}^2 \sum_{q=0}^{p} \left(\frac{\gamma_c R}{2K}\right)^q \frac{1}{(p-q)!} \max\{p-q+1, Rk\}^{p-q} C_u \\
&\leq \frac{C_c}{4} \min\{1, Rk\}^2 K^p \max\{1, Rk\}^2 \sum_{q=0}^{p} \left(\frac{\gamma_c R}{2K}\right)^q \frac{1}{p!} p^q \max\{p-q+1, Rk\}^{p-q} C_u \\
&\leq \frac{C_c}{4} \min\{1, Rk\}^2 K^{p+2} \frac{m(p)}{p!} K^{-2} \sum_{q=0}^{p} \left(\frac{\gamma_c R}{2K}\right)^q C_u \\
&\leq \frac{C_c}{4} \min\{1, Rk\}^2 K^{p+2} \frac{m(p)}{p!} \frac{K^{-2}}{1-\gamma_c R/(2K)} C_u,
\end{aligned}$$

where the last step follows from a geometric series argument and assumes $K > \gamma_c R/2$. Similarly, we find

$$M'_{R,p,\Gamma}(g) \leq \frac{C_g}{4} \left(\frac{\gamma_g}{2}\right)^p k^{-2} \min\{1, Rk\}^2 \frac{m(p)}{p!}.$$

From [Mel02, Lemma 5.5.24], similar estimates to the above and the induction hypothesis, we have

$$\begin{aligned}
H_{R,p}(kbu) &\leq C_b \frac{1}{[p]!} \frac{1}{2} \min\{1, Rk\} \times \\
&\left\{ \max\{p+1, Rk\}^2 \sum_{q=0}^{p} \binom{p}{q} \left[\left(\frac{\gamma_b R}{2}\right)^q + \left(\frac{\gamma_b R}{2}\right)^{q+1} \right] q! [p-q-2]! N'_{R,p-q-2}(u) \right. \\
&\quad \left. + \max\{p+1, Rk\} \sum_{q=0}^{p} \binom{p}{q} \left(\frac{\gamma_b R}{2}\right)^q q! [p-q-1]! N'_{R,p-q-1}(u) \right\} \\
&\leq \frac{C_b}{2} \min\{1, Rk\} K^{p+2} \frac{m(p)}{p!} \left(\frac{K^{-2}}{1-\gamma_b R/(2K)} + \frac{K^{-2}\gamma_b R/2}{1-\gamma_b R/(2K)} + \frac{K^{-1}}{1-\gamma_b R/(2K)} \right) C_u.
\end{aligned}$$

Finally, we estimate $H_{R,p}(G)$ as at the end of the proof of [Mel02, Prop. 5.5.25]

$$H_{R,p}(G) \leq C_G \gamma_G^p k^{-1} \min\{1, Rk\}(1+\gamma_G R) \frac{m(p)}{p!}. \tag{6.67}$$

We are now in the place to perform the induction argument. Lemma 6.8.3 gives

$$
\begin{aligned}
&N'_{R,p}(u) + k^{-1/2} N'_{R,p,\Gamma}(u) \\
&\leq C_B \Big[M'_{R,p}(f - k^2 cu) + M'_{R,p,\Gamma}(g) + H_{R,p}(G + bku) \\
&\quad + C_A \sum_{q=1}^{p+1} \binom{p+1}{q} q! \left(\left(\frac{R\gamma_A}{2}\right)^q + \left(\frac{R\gamma_A}{2}\right)^{q-1} \right) \frac{[p-q]!}{p!} N'_{R,p-q}(u) \\
&\quad + N'_{R,p-1}(u) + N'_{R,p-2}(u) \\
&\quad + |\alpha|^{1/2} C_{A_\Gamma} \sum_{q=1}^{p+1} \binom{p+1}{q} q! \left(\frac{R\gamma_{A_\Gamma}}{2}\right)^q \frac{[p-q]!}{p!} N'_{R,p-q,\Gamma}(u) \\
&\quad + |\alpha|^{1/2} (N'_{R,p-1,\Gamma}(u) + N'_{R,p-2,\Gamma}(u)) \Big] \\
&\leq C_B \Big[C_f k^{-2} \min\{1, Rk\}^2 \gamma_f^p \frac{m(p)}{p!} + C_g k^{-2} \min\{1, Rk\}^2 \gamma_g^p \frac{m(p)}{p!} \\
&\quad + \frac{C_c}{4} \min\{1, Rk\}^2 K^{p+2} \frac{m(p)}{p!} \frac{K^{-2}}{1 - \gamma_c R/(2K)} C_u \\
&\quad + C_G \gamma_G^p k^{-1} \min\{1, Rk\} (1 + \gamma_G R) \frac{m(p)}{p!} \\
&\quad + \frac{C_b}{2} \min\{1, Rk\} K^{p+2} \frac{m(p)}{p!} \left(\frac{K^{-2}}{1 - \gamma_b R/(2K)} + \frac{K^{-2}\gamma_b R/2}{1 - \gamma_b R/(2K)} + \frac{K^{-1}}{1 - \gamma_b R/(2K)} \right) C_u \\
&\quad + C_A K^{p+2} \frac{m(p)}{p!} \left(\frac{K^{-2}}{1 - \gamma_A R/(2K)} + \frac{K^{-2}\gamma_A R/2}{1 - \gamma_A R/(2K)} \right) C_u \\
&\quad + K^{p+2} \frac{m(p)}{p!} \left(K^{-2} + K^{-1}\right) C_u \\
&\quad + C_\alpha C_{A_\Gamma} K^{p+2} \frac{m(p)}{p!} \frac{K^{-2}}{1 - \gamma_{A_\Gamma} R/(2K)} C_u \\
&\quad + C_\alpha K^{p+2} \frac{m(p)}{p!} \left(K^{-2} + K^{-1}\right) C_u \Big] \\
&\leq C_B K^{p+2} \frac{m(p)}{p!} \Big[C_f k^{-2} \min\{1, Rk\}^2 K^{-2} \left(\frac{\gamma_f}{K}\right)^p + C_g k^{-2} \min\{1, Rk\}^2 K^{-2} \left(\frac{\gamma_g}{K}\right)^p \\
&\quad + \frac{C_c}{4} \min\{1, Rk\}^2 \frac{K^{-2}}{1 - \gamma_c R/(2K)} C_u + C_G k^{-1} \min\{1, Rk\} (1 + \gamma_G R) K^{-2} \left(\frac{\gamma_G}{K}\right)^p \\
&\quad + \frac{C_b}{2} \min\{1, Rk\} \left(\frac{K^{-2}}{1 - \gamma_b R/(2K)} + \frac{K^{-2}\gamma_b R/2}{1 - \gamma_b R/(2K)} + \frac{K^{-1}}{1 - \gamma_b R/(2K)} \right) C_u \\
&\quad + C_A \left(\frac{K^{-2}}{1 - \gamma_A R/(2K)} + \frac{K^{-2}\gamma_A R/2}{1 - \gamma_A R/(2K)} \right) C_u + \left(K^{-2} + K^{-1}\right) C_u \\
&\quad + C_\alpha C_{A_\Gamma} \frac{K^{-2}}{1 - \gamma_{A_\Gamma} R/(2K)} C_u + C_\alpha \left(K^{-2} + K^{-1}\right) C_u \Big] \\
&\leq C_u K^{p+2} \frac{m(p)}{p!} C_B \Big[\cdots \Big].
\end{aligned}
$$

The induction argument is complete, once we established that there exists a sufficiently large K, which can be chosen uniformly in p such that $C_B\Big[\cdots\Big] \leq 1$. This can be easily seen, since

$$
\begin{aligned}
C_B\Big[\cdots\Big] = C_B\Big[& K^{-2}\left(\frac{\gamma_f}{K}\right)^p + K^{-2}\left(\frac{\gamma_g}{K}\right)^p + \frac{C_c}{4}\frac{K^{-2}}{1-\gamma_c R/(2K)} + K^{-2}\left(\frac{\gamma_G}{K}\right)^p \\
&+ \frac{C_b}{2}\left(\frac{K^{-2}}{1-\gamma_b R/(2K)} + \frac{K^{-2}\gamma_b R/2}{1-\gamma_b R/(2K)} + \frac{K^{-1}}{1-\gamma_b R/(2K)}\right) \\
&+ C_A\left(\frac{K^{-2}}{1-\gamma_A R/(2K)} + \frac{K^{-2}\gamma_A R/2}{1-\gamma_A R/(2K)}\right) + \left(K^{-2}+K^{-1}\right) \\
&+ C_\alpha C_{A_\Gamma}\frac{K^{-2}}{1-\gamma_{A_\Gamma} R/(2K)} + C_\alpha\left(K^{-2}+K^{-1}\right)\Big],
\end{aligned}
$$

which concludes the proof. □

6.8.4. Control of all derivatives for k-dependent problems

Control of the derivatives $\partial_{x_d}^q$ is achieved using the differential equation. For that purpose, we introduce

$$
N'_{p,q,R}(u) := \frac{1}{[p+q]!}\sup_{R/2\leq r<R}(R-r)^{p+q+2}\|\nabla_{x'}^p\partial_{x_d}^{q+2}u\|_{L^2(B_r^+)}.
$$

Theorem 6.8.5. *Assume the hypotheses of Theorem 6.8.4. Then there exist constants K_1, K_2 depending only on the coefficients A, A_Γ of the differential operator, θ, and the constants in* (6.62) *and* (6.63) *such that for all p, $q \in \mathbb{N}_0 \cup \{-1,-2\}$ with $p+q \neq -2$*

$$
N'_{R,p,q}(u) \leq C_u K_1^{p+2} K_2^{q+2}\frac{\max\{p+q+3, Rk\}^{p+q+2}}{[p+q]!}, \tag{6.68}
$$

where C_u is given in Theorem 6.8.4.

Proof. The proof follows directly from the proof of [Mel02, Prop. 5.5.2]. Namely, [Mel02, Prop. 5.5.2] proceeds by induction on q. The induction step relies on a) the fact that an elliptic equation in B_R^+ of the form studied here is considered and b) control of the tangential derivatives, which is provided in Theorem 6.8.4. Hence, the proof of [Mel02, Prop. 5.5.2] applies. □

6.9. Dirichlet-to-Neumann maps via Boundary Integral Operators

The main goal of the present section is prove Item (ii) in Lemma 6.5.12. To that end, we rewrite the Dirichlet-to-Neumann operators DtN_k and DtN_0 in terms of boundary integral operators.

6.9.1. Preliminaries

Let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, be a bounded Lipschitz domain with analytic boundary $\Gamma := \partial\Omega$. We denote by Ω^+ the exterior domain, i.e., $\Omega^+ := \mathbb{R}^d \backslash \overline{\Omega}$. Throughout this section we assume Ω^+ to be nontrapping, see [BSW16, Def. 1.1]. Furthermore, we assume that the open ball B_R of radius R around the origin contains $\overline{\Omega}$, i.e., $\overline{\Omega} \subset B_R$. We set $\Omega_R := (\Omega \cup \Omega^+) \cap B_R = B_R \backslash \Gamma$. Following standard notation we introduce the interior and exterior trace operators γ_0^{int}, γ_1^{int}, γ_0^{ext} and γ_1^{ext}. Furthermore, we denote by V_k, K_k, K_k' and D_k the single layer, double layer, adjoint double layer and hypersingular boundary integral operators, see [Ste08, Sec. 6.9 and 7.9]. The corresponding potentials are denoted with an additional tilde $(\tilde{\cdot})$. Finally, given a coupling parameter $\eta \in \mathbb{R} \setminus \{0\}$ we introduce the combined field operator $A'_{k,\eta}$ given by

$$A'_{k,\eta} := \frac{1}{2} + K_k' + i\eta V_k.$$

We remind the reader of the exterior Calderón identities

$$\begin{pmatrix} \gamma_0^{ext} u \\ \gamma_1^{ext} u \end{pmatrix} = \begin{pmatrix} \frac{1}{2} + K_k & -V_k \\ -D_k & \frac{1}{2} - K_k' \end{pmatrix} \cdot \begin{pmatrix} \gamma_0^{ext} u \\ \gamma_1^{ext} u \end{pmatrix}. \tag{6.69}$$

Given Dirichlet data u we can now express the Dirichlet-to-Neumann operator DtN_k for any $k \geq 0$ by a complex linear combination of the two equations in the Calderón identity: For any $\eta \in \mathbb{R} \setminus \{0\}$ we have

$$\left(\frac{1}{2} + K_k' + i\eta V_k\right) \mathrm{DtN}_k u = \left(-D_k + i\eta(-\frac{1}{2} + K_k)\right) u, \tag{6.70}$$

or using the combined field operator $A'_{k,\eta}$ we have

$$A'_{k,\eta} \mathrm{DtN}_k u = \left(-D_k + i\eta(-\frac{1}{2} + K_k)\right) u. \tag{6.71}$$

Our analysis relies on invertibility of the combined field operator $A'_{k,\eta}$ as an operator mapping $H^s(\Gamma)$ into itself. In fact, wavenumber-explicit estimates of $\|(A'_{k,\eta})^{-1}\|_{L^2(\Gamma) \leftarrow L^2(\Gamma)}$ are available in the literature. We refer to [BSW16, Sec. 1.4] for detailed discussion and the references therein. For nontrapping $\Omega^+ \subset \mathbb{R}^d$, $d = 2, 3$ it is known that

$$\|(A'_{k,\eta})^{-1}\|_{L^2(\Gamma) \leftarrow L^2(\Gamma)} \lesssim k^{5/4} \left(1 + \frac{k^{3/4}}{|\eta|}\right) \tag{6.72}$$

for all $k \geq k_0$ and $\eta \in \mathbb{R} \setminus \{0\}$, see [Spe14, Thm 1.11]. This bound can be sharpened assuming $|\eta| \sim k$. In fact, for nontrapping $\Omega^+ \subset \mathbb{R}^d$, $d = 2, 3$ and $|\eta| \sim k$ there holds

$$\|(A'_{k,\eta})^{-1}\|_{L^2(\Gamma) \leftarrow L^2(\Gamma)} \lesssim 1 \tag{6.73}$$

for all $k \geq k_0$, see [BSW16, Thm. 1.13]. We now collect certain results concerning mapping properties as well as invertibility of the involved operators.

Proposition 6.9.1. *Let Γ be analytic and $\eta \in \mathbb{R} \setminus \{0\}$ fixed. If $d = 2$, assume additionally* $\operatorname{diam}\Omega < 1$. *Then*

(i) $A_{0,-\eta} = \frac{1}{2} + K_0 - i\eta V_0 \colon H^s(\Gamma) \to H^s(\Gamma)$ *is boundedly invertible for* $s \geq 0$.

(ii) $A'_{0,\eta} = \frac{1}{2} + K'_0 + i\eta V_0 \colon H^s(\Gamma) \to H^s(\Gamma)$ *is boundedly invertible for* $s \geq -1$.

(iii) For $k > 0$ *the combined field operator* $A'_{k,\eta} = \frac{1}{2} + K'_k + i\eta V_k \colon H^s(\Gamma) \to H^s(\Gamma)$ *is boundedly invertible for* $s \geq -1$.

For $k \geq 0$ *the operators*

$$\begin{aligned} V_k &\colon H^{-1/2+s}(\Gamma) \to H^{1/2+s}(\Gamma), \\ K_k &\colon H^{1/2+s}(\Gamma) \to H^{1/2+s}(\Gamma), \\ K'_k &\colon H^{-1/2+s}(\Gamma) \to H^{-1/2+s}(\Gamma), \\ D_k &\colon H^{1/2+s}(\Gamma) \to H^{-1/2+s}(\Gamma) \end{aligned} \tag{6.74}$$

are bounded for $s \geq -1/2$. *Finally, for* $k \geq k_0 > 0$ *the splittings*

$$\begin{aligned} V_k - V_0 &= S_V + \gamma_0^{int} \tilde{A}_V, \\ K_k - K_0 &= S_K + \gamma_0^{int} \tilde{A}_K, \\ K'_k - K'_0 &= S_{K'} + \gamma_1^{int} \tilde{A}_V, \\ D_k - D_0 &= S_D + \gamma_1^{int} \tilde{A}_K \end{aligned} \tag{6.75}$$

with linear maps $\tilde{A}_V \colon H^{-3/2}(\Gamma) \to C^\infty(\overline{\Omega})$ *and* $\tilde{A}_K \colon H^{-1/2}(\Gamma) \to C^\infty(\overline{\Omega})$ *and bounded linear operators* S_V, S_K, $S_{K'}$ *and* S_D *having the following mapping properties for* $s \geq -1$

$$\begin{aligned} \|S_V u\|_{-1/2+s,\Gamma} &\leq C_{s,s'} k^{-(1+s-s')} \|u\|_{-1/2+s',\Gamma}, && 1/2 \leq s' \leq s+3, \\ \|S_K u\|_{1/2+s,\Gamma} &\leq C_{s,s'} k^{-(1+s-s')} \|u\|_{-1/2+s',\Gamma}, && 1/2 \leq s' \leq s+3, \\ \|S_{K'} u\|_{-1/2+s,\Gamma} &\leq C_{s,s'} k^{-(1+s-s')} \|u\|_{-3/2+s',\Gamma}, && 3/2 \leq s' \leq s+3, \\ \|S_D u\|_{1/2+s,\Gamma} &\leq C_{s,s'} k^{-(1+s-s')} \|u\|_{-3/2+s',\Gamma}, && 3/2 \leq s' \leq s+3 \end{aligned} \tag{6.76}$$

hold true. Furthermore, the operator $\tilde{A}_K$ *has the mapping property*

$$\tilde{A}_K f \in \mathcal{A}(C_K \|f\|_{-1/2,\Gamma}, \gamma_K, \Omega) \qquad \forall f \in H^{-1/2}(\Gamma), \tag{6.77}$$

with constants C_K, γ_K *independent of* $k \geq k_0$. *Finally, for* $t \geq 0$ *the following mapping properties*

$$\begin{aligned} \|S_V u\|_{t,\Gamma} &\leq C_t k^{-1} \|u\|_{t,\Gamma}, \\ \|S_K u\|_{t,\Gamma} &\leq C_t \|u\|_{t,\Gamma}, \\ \|S_{K'} u\|_{t,\Gamma} &\leq C_t \|u\|_{t,\Gamma}, \\ \|S_D u\|_{t,\Gamma} &\leq C_t k \|u\|_{t,\Gamma} \end{aligned} \tag{6.78}$$

hold true.

Proof. For Item (i) see [Mel12, Lemma 3.5, (ii)]. For Item (ii) in the case $s \geq 0$ see [Mel12, Lemma 3.5, (iv)]. We turn to the case $s \in [-1, 0]$. Note that the adjoint of $A_{0,-\eta}$ is precisely the operator $A'_{0,\eta}$. Furthermore, by Item (i) the operator $A_{0,-\eta}\colon H^t(\Gamma) \to H^t(\Gamma)$ is boundedly invertible in particular for $t \in [0, 1]$. Hence, due to the adjoint of $A_{0,-\eta}$ being $A'_{0,\eta}$, we find that $A'_{0,\eta}\colon H^s(\Gamma) \to H^s(\Gamma)$ is also boundedly invertible for $s \in [-1, 0]$. For Item (iii) see [CWGLS12, Thm. 2.27] in the case $s \in [-1, 0]$ as well as [BSW16, Sec. 6.1]. Consequently, by [Mel12, Lemma 2.14] invertibility holds for any $s \geq 0$. The mapping properties (6.74) are standard. For (6.75) and (6.76) see [MMPR20, Lemma A.1] for the $1/2 < s'$ and $3/2 < s'$, respectively. The limiting cases follow by inspection of the proof, the therein used estimates for the potentials, and applying a multiplicative trace estimate. For (6.77) see also [MMPR20, Lemma A.1]. (6.78) is just a simplification of (6.76). □

6.9.2. Decomposition of the Dirichlet-to-Neumann map

Before proceeding with the proof of Item (ii) in Lemma 6.5.12 let us introduce the jumps of the trace operators:

$$[\![u]\!] := \gamma_0^{ext} u - \gamma_0^{int} u, \qquad [\![\partial_n u]\!] := \gamma_1^{ext} u - \gamma_1^{int} u.$$

For linear operators $\tilde{A}$ mapping into spaces of piecewise defined functions we define the operator $[\![\tilde{A}]\!]$ and $[\![\partial_n \tilde{A}]\!]$ analogously, e.g., $[\![\tilde{A}]\!] u := [\![\tilde{A} u]\!]$.

We now collect further technical results of [Mel12]. We closely follow the notation and results of [Mel12]. As in [Mel12] we assume

$$C_\eta^{-1} k \leq |\eta| \leq C_\eta k \tag{6.79}$$

for some $C_\eta > 0$ independent of k.

In Proposition 6.9.2 below we extend the results of [Mel12, Lemma 6.3] to a wider range of Sobolev spaces.

Proposition 6.9.2 ([Mel12, Lemma 6.3])**.** *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain with an analytic boundary Γ. Let $q \in (0, 1)$. Then one can construct operators $L^{neg}_{\Gamma,q}$, $H^{neg}_{\Gamma,q}$ on $H^{-1}(\Gamma)$ with the following properties:*

(i) $L^{neg}_{\Gamma,q} f + H^{neg}_{\Gamma,q} f = f$ *for all* $f \in H^{-1}(\Gamma)$.

(ii) *For* $-1 \leq s' \leq s$ *there holds* $\|H^{neg}_{\Gamma,q} f\|_{s',\Gamma} \leq C_{s,s'} (q/k)^{s-s'} \|f\|_{s,\Gamma}$.

(iii) $L^{neg}_{\Gamma,q} f$ *is the restriction to* Γ *of a function that is analytic on a tubular neighborhood* T *of* Γ *and satisfies*

$$\begin{aligned}
\|\nabla^n L^{neg}_{\Gamma,q} f\|_{0,T} &\leq C_q k^{d/2} \gamma_q^n \max\{k, n\}^n \|f\|_{-1/2,\Gamma} \qquad \forall n \in \mathbb{N}_0, \\
\|\nabla^n L^{neg}_{\Gamma,q} f\|_{0,T} &\leq C_q k^{d/2+1} \gamma_q^n \max\{k, n\}^n \|f\|_{-1,\Gamma} \qquad \forall n \in \mathbb{N}_0.
\end{aligned}$$

Here, $C_{s,s'}$ is independent of q and k; the constants C_q, $\gamma_q > 0$ are independent of k.

Proof. For Item (i), Item (ii) in the case $-1 \leq s' \leq s \leq 1$ and Item (iii) see [Mel12, Lemma 6.3 and Remark 6.4]. The crucial extension is the estimate stated in Item (ii) in the case $-1 \leq s' \leq s$ for $s \geq 1$. In the proof of [Mel12, Lemma 6.3] the operators $H^{neg}_{\Gamma,q}$ and $L^{neg}_{\Gamma,q}$ are explicitly constructed. We collect the important ingredients of the proof of [Mel12, Lemma 6.3] in the following. On the compact manifold Γ consider the eigenvalue problem for the Laplace-Beltrami operator

$$-\Delta_\Gamma \varphi = \lambda^2 \varphi \quad \text{on } \Gamma. \tag{6.80}$$

There are countably many eigenfunctions φ_m, $m \in \mathbb{N}_0$, with corresponding eigenvalues $\lambda_m^2 \geq 0$, which we assume to be sorted in ascending order. Without loss of generality, these eigenfunctions are normalized in $L^2(\Gamma)$. The functions $(\varphi_m)_{m\in\mathbb{N}_0}$ are an orthonormal basis of $L^2(\Gamma)$ and an orthogonal basis of $H^1(\Gamma)$. With $u_m := (u, \varphi_m)$ we have

$$\|u\|^2_{0,\Gamma} = \sum_{m=0}^{\infty} |u_m|^2, \quad \text{and} \quad \|u\|^2_{1,\Gamma} = \sum_{m=0}^{\infty} (1+\lambda_m^2)|u_m|^2. \tag{6.81}$$

For $s \in \mathbb{R}$ we introduce the sequence space h^s by

$$h^s := \left\{ (u_m)_{m\in\mathbb{N}} \colon \sum_{m=0}^{\infty} (1+\lambda_m^2)^s |u_m|^2 < \infty \right\}.$$

The mapping $\iota\colon u \mapsto (\,(u,\varphi_m)\,)_{m\in\mathbb{N}_m}$ then provides an isomorphism between the Sobolev space $H^s(\Gamma)$ and the sequence space h^s for $s \in [-1,1]$, with corresponding norm equivalence, see [Mel12, Lemma C.3]. However, as we will see below ι is in fact an isomorphism for all $s \geq -1$. Inspection of the proof of [Mel12, Lemma 6.3], in particular the proof of the estimate for $H^{neg}_{\Gamma,q}$, reveals that

$$\|H^{neg}_{\Gamma,q} f\|_{s',\Gamma} \leq C_{s,s'} (q/k)^{s-s'} \|f\|_{s,\Gamma}$$

holds for all $-1 \leq s' \leq s$, for which $\iota\colon H^s(\Gamma) \to h^s$ and $\iota\colon H^{s'}(\Gamma) \to h^{s'}$ are isomorphisms. Hence, the proof is complete once we establish that $\iota\colon H^s(\Gamma) \to h^s$ is an isomorphism for all $s > 1$. We show the case $s = 2$.

The inclusion $h^2 \hookrightarrow H^2(\Gamma)$: Let $u = \sum_{m=0}^{\infty} u_m \varphi_m$ be such that $\sum_{m=0}^{\infty}(1+\lambda_m^2)^2|u_m|^2 < \infty$. Let $u^N = \sum_{m=0}^{N} u_m\varphi_m$. By the above construction, $u^N \to u$ in $H^1(\Gamma)$ and $\|u\|_{1,\Gamma} = \|(u_m)_{m\in\mathbb{N}}\|_{h^1}$. Furthermore, we have

$$\begin{aligned}
\|\Delta_\Gamma u^N - \Delta_\Gamma u^{M-1}\|^2_{0,\Gamma} &= \left\| \Delta_\Gamma \sum_{m=1}^{N} u_m\varphi_m - \Delta_\Gamma \sum_{m=1}^{M-1} u_m\varphi_m \right\|^2_{0,\Gamma} \\
&= \left\| \sum_{m=M}^{N} u_m \Delta_\Gamma \varphi_m \right\|^2_{0,\Gamma} \\
&= \left\| \sum_{i=M}^{N} u_m \lambda_m^2 \varphi_m \right\|^2_{0,\Gamma} \\
&= \sum_{i=M}^{N} |u_m|^2 \lambda_m^4 \to 0,
\end{aligned}$$

where we used (6.80), the fact that the eigenfunctions are an orthonormal basis of $L^2(\Gamma)$, and the assumed convergence $\sum_{m=0}^{\infty}(1+\lambda_m^2)^2|u_m|^2 < \infty$. Therefore, u^N is a Cauchy sequence in $H^1(\Gamma, \Delta_\Gamma) = \{u \in H^1(\Gamma)\colon \Delta_\Gamma u \in L^2(\Gamma)\}$, with corresponding graph norm. Consequently u^N converges in $H^1(\Gamma, \Delta_\Gamma)$. Since $\Delta_\Gamma : H^1(\Gamma, \Delta_\Gamma) \to L^2(\Gamma)$ is continuous, we conclude $\Delta_\Gamma u = \sum_{m\in\mathbb{N}_0} u_m \Delta_\Gamma \varphi_m = -\sum_{m\in\mathbb{N}_0} u_m \lambda_m^2 \varphi_m$. Finally, by elliptic regularity we can now estimate

$$|u|_{2,\Gamma}^2 \lesssim \|\Delta_\Gamma u\|_{0,\Gamma}^2 = \sum_{m\in\mathbb{N}_0} |f_m|^2 \lambda_m^4 = |(u_m)_{m\in\mathbb{N}_0}|_{h^2}^2. \tag{6.82}$$

Finally,

$$\|u\|_{2,\Gamma}^2 \lesssim \|(u_m)_{m\in\mathbb{N}_0}\|_{h^2}^2$$

follows by (6.82) together with (6.81).

The inclusion $H^2(\Gamma) \hookrightarrow h^2$: Let $u \in H^2(\Gamma)$ be be given with the representation $u = \sum_{m=0}^{\infty} u_m \varphi_m$, where the sum converges in $H^1(\Gamma)$. Since $u \in H^2(\Gamma)$ we have $-\Delta_\Gamma u =: f \in L^2(\Gamma)$. In the following we express the coefficient u_m in terms of f_m. Note that

$$\lambda_m^2 u_m = \lambda_m^2 (u, \varphi_m) = (\nabla_\Gamma u, \nabla_\Gamma \varphi_m) = (f, \varphi_m) = f_m.$$

Hence, we have $\lambda_m^2 u_m = f_m$ and consequently

$$\sum_{m=0}^{\infty} \lambda_m^4 |u_m|^2 = \sum_{m=0}^{\infty} |f_m|^2 < \infty.$$

Finally, using 6.82 as well as the fact that $\|u\|_{1,\Gamma} = \|(u_m)_{m\in\mathbb{N}}\|_{h^1}$, we find

$$\|(u_m)_{m\in\mathbb{N}_0}\|_{h^2}^2 = \|(u_m)_{m\in\mathbb{N}_0}\|_{h^1}^2 + |(u_m)_{m\in\mathbb{N}_0}|_{h^2}^2 = \|u\|_{1,\Gamma}^2 + \|\Delta_\Gamma u\|_{0,\Gamma}^2 \leq \|u\|_{2,\Gamma}^2.$$

This concludes the proof for $s = 2$. Interpolation between $s = 1$ and $s = 2$ yields the result for $s \in (1, 2)$, see [Mel12, Lemma C.3]. Inductively one proceeds for the space $H^{2n}(\Gamma)$ by similar arguments. Instead of Δ_Γ one performs the same arguments for Δ_Γ^n. □

Remark 6.9.3. A natural question arising from the proof of Proposition 6.9.2 is whether or not a similar construction allows for high and low pass filters in the volume Ω. The volume filters in Proposition 6.3.2 only allow for estimates in negative Sobolev norms for $-1/2 \leq s'$. In fact similar arguments as in the proof of Proposition 6.9.2 allow to construct high and low pass filters via the eigenvalue problem

$$\begin{aligned} -\Delta\varphi &= \lambda^2 \varphi && \text{in } \Omega, \\ \partial_n u &= 0 && \text{on } \Gamma. \end{aligned}$$

However, the corresponding high pass filter only allows for estimates in the range $-1 \leq s' \leq s \leq 1$, because of the additional boundary terms. ■

In the following we will prove an extension of the following

Proposition 6.9.4 ([Mel12, Thm. 2.9]). *Let Γ be analytic and let $-1/2 \le s \le 0$. Fix $q \in (0,1)$. Then the operator $A'_{k,\eta}$ can be written in the form*

$$A'_{k,\eta} = \frac{1}{2} + K'_0 + R_{A'} + k[\![\tilde{\mathcal{A}}_1]\!] + [\![\partial_n \tilde{\mathcal{A}}_2]\!],$$

where $R_{A'}\colon H^s(\Gamma) \to H^{s+1}(\Gamma)$ and $\tilde{\mathcal{A}}_1$, $\tilde{\mathcal{A}}_2\colon H^{-1}(\Gamma) \to C^\infty(T)$ satisfy

$$\|R_{A'}u\|_{s+1,\Gamma} \le Ck\|u\|_{s,\Gamma}, \qquad \|R_{A'}u\|_{s,\Gamma} \le q\|u\|_{s,\Gamma},$$
$$\tilde{\mathcal{A}}_1\varphi \in \mathcal{A}(C_q C_{1,\varphi}, \gamma_q, T), \qquad C_{1,\varphi} = k\|\varphi\|_{-3/2,\Gamma} + k^{d/2}\|\varphi\|_{-1,\Gamma},$$
$$\tilde{\mathcal{A}}_2\varphi \in \mathcal{A}(C_q C_{2,\varphi}, \gamma_q, T), \qquad C_{2,\varphi} = k\|\varphi\|_{-3/2,\Gamma}.$$

The constant C and the tubular neighborhood T of Γ are independent of $k \ge k_0$ and q; the constants C_q, $\gamma_q > 0$ are independent of $k \ge k_0$ (but may depend of q).

The proof of Proposition 6.9.4 relies on a decomposition of the volume potential $\tilde{V}_k$, which we present below for the readers' convenience.

Proposition 6.9.5 ([Mel12, Thm. 5.3]). *Let Γ be analytic and $q \in (0,1)$. Then*

$$\tilde{V}_k = \tilde{V}_0 + \tilde{S}_{V,pw} + \tilde{\mathcal{A}}_{V,pw},$$

where the linear operators $\tilde{S}_{V,pw}$ and $\tilde{\mathcal{A}}_{V,pw}$ satisfy the following for every $s \ge -1$:

(i) $\tilde{S}_{V,pw}\colon H^{-1/2+s}(\Gamma) \to H^{3+s}(\Omega_R)$ *with*

$$\|\tilde{S}_{V,pw}\varphi\|_{s',\Omega_R} \le C_{s',s} q^2 (qk^{-1})^{1+s-s'} \|\varphi\|_{-1/2+s,\Gamma}, \quad 0 \le s' \le s+3.$$

Here, the constant $C_{s',s} > 0$ is independent of q and $k \ge k_0$.

(ii) $\tilde{\mathcal{A}}_{V,pw}\colon H^{-1/2+s}(\Gamma) \to C^\infty(\Omega)$ *with*

$$\|\nabla^n \tilde{\mathcal{A}}_{V,pw}\varphi\|_{0,\Omega_R} \le C_q\gamma_q \max\{n+1, k\}^{n+1} \|\varphi\|_{-3/2,\Gamma} \quad \forall n \in \mathbb{N}_0.$$

Here, the constants C_q, $\gamma_q > 0$ are independent of $k \ge k_0$ but may depend on q.

Theorem 6.9.6 (Extension of [Mel12, Thm. 2.9]). *Let Γ be analytic and let $s \ge 0$. Fix $q \in (0,1)$. Then the operator $A'_{k,\eta}$ can be written in the form*

$$A'_{k,\eta} = A'_{0,1} + R_{A'} + k[\![\tilde{\mathcal{A}}_1]\!] + [\![\partial_n \tilde{\mathcal{A}}_2]\!],$$

where the linear operator $R_{A'}$ satisfies

$$\|R_{A'}u\|_{s+1,\Gamma} \le Ck\|u\|_{s,\Gamma}, \tag{6.83a}$$
$$\|R_{A'}u\|_{s,\Gamma} \le Ck\|u\|_{s-1,\Gamma}, \tag{6.83b}$$
$$\|R_{A'}u\|_{s,\Gamma} \le q\|u\|_{s,\Gamma}, \tag{6.83c}$$
$$\|R_{A'}u\|_{s-1,\Gamma} \le q\|u\|_{s-1,\Gamma}, \tag{6.83d}$$

and the linear operators $\tilde{\mathcal{A}}_1$, $\tilde{\mathcal{A}}_2\colon H^{-1}(\Gamma) \to C^\infty(T)$ *satisfy*

$$\tilde{\mathcal{A}}_1\varphi \in \mathcal{A}(C_q C_{1,\varphi}, \gamma_q, T), \qquad C_{1,\varphi} = k\|\varphi\|_{-3/2,\Gamma} + k^{d/2}\|\varphi\|_{-1,\Gamma}, \tag{6.84a}$$

$$\tilde{\mathcal{A}}_2\varphi \in \mathcal{A}(C_q C_{2,\varphi}, \gamma_q, T), \qquad C_{2,\varphi} = k\|\varphi\|_{-3/2,\Gamma}. \tag{6.84b}$$

The constant C and the tubular neighborhood T of Γ are independent of $k \ge k_0$ and q; the constants C_q, $\gamma_q > 0$ are independent of $k \ge k_0$ (but may depend of q).

Proof. We perform a similar splitting as in the proof of [Mel12, Thm. 2.9]. The starting point of our analysis is the decomposition

$$A'_{k,\eta} = \frac{1}{2} + K'_0 + \gamma_1^{int}(\tilde{S}_{V,pw} + \tilde{\mathcal{A}}_{V,pw}) + i\eta\gamma_0^{int}(\tilde{V}_0 + \tilde{S}_{V,pw} + \tilde{\mathcal{A}}_{V,pw}),$$

with $\tilde{S}_{V,pw}$ and $\tilde{\mathcal{A}}_{V,pw}$ as in Proposition 6.9.5, see [Mel12, Eq. (6.4)]. Adding and subtracting iV_0 and noting $V_0 = \gamma_0^{int}\tilde{V}_0$ we find

$$\begin{aligned} A'_{k,\eta} &= \frac{1}{2} + K'_0 + iV_0 + \gamma_1^{int}(\tilde{S}_{V,pw} + \tilde{\mathcal{A}}_{V,pw}) + i(\eta-1)\gamma_0^{int}\tilde{V}_0 + i\eta\gamma_0^{int}(\tilde{S}_{V,pw} + \tilde{\mathcal{A}}_{V,pw}) \\ &= A'_{0,1} + \gamma_1^{int}(\tilde{S}_{V,pw} + \tilde{\mathcal{A}}_{V,pw}) + i(\eta-1)\gamma_0^{int}\tilde{V}_0 + i\eta\gamma_0^{int}(\tilde{S}_{V,pw} + \tilde{\mathcal{A}}_{V,pw}). \end{aligned} \tag{6.85}$$

Using the filters $H^{neg}_{\Gamma,q}$ and $L^{neg}_{\Gamma,q}$ in Proposition 6.9.2 we define

$$R_{A'} = H^{neg}_{\Gamma,q}\left(\gamma_1^{int}\tilde{S}_{V,pw} + i\eta\gamma_0^{int}\tilde{S}_{V,pw} + i(\eta-1)V_0\right), \tag{6.86a}$$

$$\tilde{\mathcal{A}}_1 = -k^{-1}\chi_\Omega\left(i\eta\tilde{\mathcal{A}}_{V,pw} + L^{neg}_{\Gamma,q}\left(\gamma_1^{int}\tilde{S}_{V,pw} + i\eta\gamma_0^{int}\tilde{S}_{V,pw} + i(\eta-1)V_0\right)\right),$$

$$\tilde{\mathcal{A}}_2 = -\chi_\Omega\tilde{\mathcal{A}}_{V,pw}.$$

The mapping properties of $\tilde{\mathcal{A}}_1$ and $\tilde{\mathcal{A}}_2$ stay the same as in Proposition 6.9.4. We are left with the mapping properties of $R_{A'}$. In the following the parameter q appearing in Proposition 6.9.2 and 6.9.5 is still at our disposal[1]. We fix it at the end of the proof to guarantee the estimates (6.83c) and (6.83d).

Step 1: We estimate the term $i(\eta-1)H^{neg}_{\Gamma,q}V_0$ in various norms. We heavily use the estimates for $H^{neg}_{\Gamma,q}$ and V_0 given in Proposition 6.9.2 and (6.75) in Proposition 6.9.1, respectively. First estimating η, then using the properties of $H^{neg}_{\Gamma,q}$ in Proposition 6.9.2 and finally the mapping properties of V_0 we find

$$\begin{aligned} \|i(\eta-1)H^{neg}_{\Gamma,q}V_0u\|_{s+1,\Gamma} &\le Ck\|H^{neg}_{\Gamma,q}V_0u\|_{s+1,\Gamma} \le Ck\|V_0u\|_{s+1,\Gamma} &&\le Ck\|u\|_{s,\Gamma}, \\ \|i(\eta-1)H^{neg}_{\Gamma,q}V_0u\|_{s,\Gamma} &\le Ck\|H^{neg}_{\Gamma,q}V_0u\|_{s,\Gamma} \le Ck\|V_0u\|_{s,\Gamma} &&\le Ck\|u\|_{s-1,\Gamma}, \\ \|i(\eta-1)H^{neg}_{\Gamma,q}V_0u\|_{s,\Gamma} &\le Ck\|H^{neg}_{\Gamma,q}V_0u\|_{s,\Gamma} \le Ck(q/k)\|V_0u\|_{s+1,\Gamma} &&\le Cq\|u\|_{s,\Gamma}, \\ \|i(\eta-1)H^{neg}_{\Gamma,q}V_0u\|_{s-1,\Gamma} &\le Ck\|H^{neg}_{\Gamma,q}V_0u\|_{s-1,\Gamma} \le Ck(q/k)\|V_0u\|_{s,\Gamma} &&\le Cq\|u\|_{s-1,\Gamma}. \end{aligned}$$

In the Steps 2 and 3 below we again heavily use the properties of $H^{neg}_{\Gamma,q}$ given in Proposition 6.9.2. Furthermore, we often apply the results of Proposition 6.9.5, especially Item (i).

[1] Do not confuse this q with the one appearing in the statement of the present theorem.

Below, we will write certain exponents nonsimplified in order to indicate the corresponding choices of Sobolev exponents when applying Proposition 6.9.5.

Step 2: We estimate the term $H^{neg}_{\Gamma,q}\gamma_1^{int}\tilde{S}_{V,pw}$ in various norms. We have

$$\begin{aligned}\|H^{neg}_{\Gamma,q}\gamma_1^{int}\tilde{S}_{V,pw}u\|_{s+1,\Gamma} &\leq C\|\gamma_1^{int}\tilde{S}_{V,pw}u\|_{s+1,\Gamma} \leq C\|\tilde{S}_{V,pw}u\|_{s+5/2,\Omega} \\ &\leq Cq^2(qk^{-1})^{1+(s+1/2)-(s+5/2)}\|u\|_{s,\Gamma} = Cqk\|u\|_{s,\Gamma}.\end{aligned}$$

By the previous estimate we also find

$$\|H^{neg}_{\Gamma,q}\gamma_1^{int}\tilde{S}_{V,pw}u\|_{s,\Gamma} \leq Cq/k\|\gamma_1^{int}\tilde{S}_{V,pw}u\|_{s+1,\Gamma} \leq Cq^2\|u\|_{s,\Gamma}.$$

In the case $s \in [0, 1/2)$, we perform a multiplicative trace inequality and find

$$\begin{aligned}\|H^{neg}_{\Gamma,q}\gamma_1^{int}\tilde{S}_{V,pw}u\|_{s-1,\Gamma} &\leq C(q/k)^{-s+1}\|\gamma_1^{int}\tilde{S}_{V,pw}u\|_{0,\Gamma} \\ &\leq C(q/k)^{-s+1}\|\tilde{S}_{V,pw}u\|_{1,\Omega}^{1/2}\|\tilde{S}_{V,pw}u\|_{2,\Omega}^{1/2} \\ &\leq C(q/k)^{-s+1}\left[q^2(qk^{-1})^{1+(s-1/2)-1}\right]^{1/2}\left[q^2(qk^{-1})^{1+(s-1/2)-2}\right]^{1/2}\|u\|_{s-1,\Gamma} \\ &= Cq^2\|u\|_{s-1,\Gamma}.\end{aligned}$$

In the case $s \geq 1/2$ we perform a standard trace estimate and find

$$\begin{aligned}\|H^{neg}_{\Gamma,q}\gamma_1^{int}\tilde{S}_{V,pw}u\|_{s-1,\Gamma} &\leq Cq/k\|\gamma_1^{int}\tilde{S}_{V,pw}u\|_{s,\Gamma} \leq Cq/k\|\tilde{S}_{V,pw}u\|_{s+3/2,\Omega} \\ &\leq Cq/kq^2(qk^{-1})^{1+(s-1/2)-(s+3/2)}\|u\|_{s-1,\Gamma} \\ &= Cq^2\|u\|_{s-1,\Gamma}.\end{aligned}$$

By the previous two estimate we find

$$\|H^{neg}_{\Gamma,q}\gamma_1^{int}\tilde{S}_{V,pw}u\|_{s,\Gamma} \leq C\|\gamma_1^{int}\tilde{S}_{V,pw}u\|_{s,\Gamma} \leq Cqk\|u\|_{s-1,\Gamma}.$$

Summarizing, we found

$$\begin{aligned}\|H^{neg}_{\Gamma,q}\gamma_1^{int}\tilde{S}_{V,pw}u\|_{s+1,\Gamma} &\leq Cqk\|u\|_{s,\Gamma}, \\ \|H^{neg}_{\Gamma,q}\gamma_1^{int}\tilde{S}_{V,pw}u\|_{s,\Gamma} &\leq Cqk\|u\|_{s-1,\Gamma}, \\ \|H^{neg}_{\Gamma,q}\gamma_1^{int}\tilde{S}_{V,pw}u\|_{s,\Gamma} &\leq Cq^2\|u\|_{s,\Gamma}, \\ \|H^{neg}_{\Gamma,q}\gamma_1^{int}\tilde{S}_{V,pw}u\|_{s-1,\Gamma} &\leq Cq^2\|u\|_{s-1,\Gamma}.\end{aligned}$$

Step 3: We estimate the term $\eta H^{neg}_{\Gamma,q}\gamma_0^{int}\tilde{S}_{V,pw}$ in various norms. We have

$$\begin{aligned}\|\eta H^{neg}_{\Gamma,q}\gamma_0^{int}\tilde{S}_{V,pw}u\|_{s+1,\Gamma} &\leq Ck\|\gamma_0^{int}\tilde{S}_{V,pw}u\|_{s+1,\Gamma} \leq Ck\|\tilde{S}_{V,pw}u\|_{s+3/2,\Omega} \\ &\leq Ckq^2(qk^{-1})^{1+(s+1/2)-(s+3/2)}\|u\|_{s,\Gamma} = Cq^2k\|u\|_{s,\Gamma}.\end{aligned}$$

By the previous estimate we also find

$$\|\eta H^{neg}_{\Gamma,q}\gamma_0^{int}\tilde{S}_{V,pw}u\|_{s,\Gamma} \leq Ckq/k\|\gamma_0^{int}\tilde{S}_{V,pw}u\|_{s+1,\Gamma} \leq Cq^3\|u\|_{s,\Gamma}.$$

In the case $s \in [0, 1/2)$, we perform a multiplicative trace inequality and find

$$\begin{aligned}\|\eta H^{neg}_{\Gamma,q}\gamma_0^{int}\tilde{S}_{V,pw}u\|_{s-1,\Gamma} &\leq Ck(q/k)^{-s+1}\|\gamma_0^{int}\tilde{S}_{V,pw}u\|_{0,\Gamma}\\ &\leq C(q/k)^{-s+1}\|\tilde{S}_{V,pw}u\|^{1/2}_{0,\Omega}\|\tilde{S}_{V,pw}u\|^{1/2}_{1,\Omega}\\ &\leq C(q/k)^{-s+1}\left[q^2(qk^{-1})^{1+(s-1/2)-0}\right]^{1/2}\left[q^2(qk^{-1})^{1+(s-1/2)-1}\right]^{1/2}\|u\|_{s-1,\Gamma}\\ &= Cq^3\|u\|_{s-1,\Gamma}.\end{aligned}$$

In the case $s \geq 1/2$ we perform a standard trace estimate and find

$$\begin{aligned}\|\eta H^{neg}_{\Gamma,q}\gamma_0^{int}\tilde{S}_{V,pw}u\|_{s-1,\Gamma} &\leq Ckq/k\|\gamma_0^{int}\tilde{S}_{V,pw}u\|_{s,\Gamma} \leq Ckq/k\|\tilde{S}_{V,pw}u\|_{s+1/2,\Omega}\\ &\leq Ckq/kq^2(qk^{-1})^{1+(s-1/2)-(s+1/2)}\|u\|_{s-1,\Gamma}\\ &= Cq^3\|u\|_{s-1,\Gamma}.\end{aligned}$$

By the previous two estimate we find

$$\|\eta H^{neg}_{\Gamma,q}\gamma_0^{int}\tilde{S}_{V,pw}u\|_{s,\Gamma} \leq k\|\gamma_0^{int}\tilde{S}_{V,pw}u\|_{s,\Gamma} \leq Cq^2k\|u\|_{s-1,\Gamma}.$$

Summarizing, we found

$$\begin{aligned}\|\eta H^{neg}_{\Gamma,q}\gamma_0^{int}\tilde{S}_{V,pw}u\|_{s+1,\Gamma} &\leq Cq^2k\|u\|_{s,\Gamma},\\ \|\eta H^{neg}_{\Gamma,q}\gamma_0^{int}\tilde{S}_{V,pw}u\|_{s,\Gamma} &\leq Cq^2k\|u\|_{s-1,\Gamma},\\ \|\eta H^{neg}_{\Gamma,q}\gamma_0^{int}\tilde{S}_{V,pw}u\|_{s,\Gamma} &\leq Cq^3\|u\|_{s,\Gamma},\\ \|\eta H^{neg}_{\Gamma,q}\gamma_0^{int}\tilde{S}_{V,pw}u\|_{s-1,\Gamma} &\leq Cq^3\|u\|_{s-1,\Gamma}.\end{aligned}$$

Step 4: The definition of the operator $R_{A'}$ in (6.86a), the triangle inequality, and appropriate choice of q yields mapping properties of $R_{A'}$ as stated in (6.83). □

Finally, a simple application of [Mel12, Cor. 7.5] for nontrapping Ω^+ with analytic boundary is the following

Lemma 6.9.7. *Let Ω^+ be nontrapping. Let Γ be analytic, T be a tubular neighborhood of Γ and C_{g_1}, C_{g_2} , $\gamma_g > 0$. Then there exist constants C, $\gamma > 0$ independent of $k \geq k_0$ such that for all $g_1 \in \mathcal{A}(C_{g_1}, \gamma_g, T)$, $g_2 \in \mathcal{A}(C_{g_2}, \gamma_g, T)$ the solution $\varphi \in L^2(\Gamma)$ of*

$$A'_{k,\eta}\varphi = k[\![g_1]\!] + [\![\partial_n g_2]\!]$$

satisfies

$$\varphi = -[\![\partial_n v]\!], \qquad v \in \mathcal{A}(Ck^{5/2}(C_{g_1} + C_{g_2}), \gamma, \Omega_R).$$

Proof. We apply [Mel12, Cor. 7.5] with $s_A = 0$. By Item (iii) in Proposition 6.9.1 the operator $A'_{\eta,k}\colon L^2(\Gamma) \to L^2(\Gamma)$ is boundedly invertible. The result follows immediately from [Mel12, Cor. 7.5] together with the bound (6.73). □

Proof. (Proof of Item (ii) in Lemma 6.5.12)

Step 1: We derive a splitting of $(A'_{k,\eta})^{-1}$, similar to the results of [Mel12, Thm. 2.11]. Fix $\hat{q} \in (0,1)$. Let

$$q := \hat{q} \min \left\{ 1, \frac{1}{\|(A'_{0,1})^{-1}\|_{H^s(\Gamma)\leftarrow H^s(\Gamma)}}, \frac{1}{\|(A'_{0,1})^{-1}\|_{H^{s-1}(\Gamma)\leftarrow H^{s-1}(\Gamma)}} \right\}.$$

Note, that by Proposition 6.9.1 the operator $A'_{0,1} \colon H^t(\Gamma) \to H^t(\Gamma)$ is boundedly invertible for $t \geq -1$ and therefore q is well defined and $q \in (0,1)$. Theorem 6.9.6 applied to q gives the decomposition

$$A'_{k,\eta} = A'_{0,1} + R + [\![\mathcal{A}]\!],$$

with $R = R_{A'}$ and $\mathcal{A} = k\tilde{\mathcal{A}}_1 + \partial_n \tilde{\mathcal{A}}_2$, as in Theorem 6.9.6. Note that by construction

$$\|(A'_{0,1})^{-1}R\|_{H^s(\Gamma)\leftarrow H^s(\Gamma)} \leq \hat{q} \quad \text{and} \quad \|(A'_{0,1})^{-1}R\|_{H^{s-1}(\Gamma)\leftarrow H^{s-1}(\Gamma)} \leq \hat{q}. \tag{6.87}$$

Hence, $A'_{0,1} + R$ is boundedly invertible by a geometric series argument, since

$$(A'_{0,1} + R)^{-1} = (I + (A'_{0,1})^{-1}R)^{-1}(A'_{0,1})^{-1} \tag{6.88}$$

with norm estimates

$$\|(A'_{0,1} + R)^{-1}\|_{H^s(\Gamma)\leftarrow H^s(\Gamma)} \leq (1-\hat{q})^{-1}\|(A'_{0,1})^{-1}\|_{H^s(\Gamma)\leftarrow H^s(\Gamma)}, \tag{6.89a}$$
$$\|(A'_{0,1} + R)^{-1}\|_{H^{s-1}(\Gamma)\leftarrow H^{s-1}(\Gamma)} \leq (1-\hat{q})^{-1}\|(A'_{0,1})^{-1}\|_{H^{s-1}(\Gamma)\leftarrow H^{s-1}(\Gamma)}. \tag{6.89b}$$

By Proposition 6.9.1 the operator $A'_{k,\eta} \colon H^t(\Gamma) \to H^t(\Gamma)$ is boundedly invertible for $t \geq -1$. We may decompose $(A'_{k,\eta})^{-1}$ as follows

$$(A'_{k,\eta})^{-1} = (A'_{0,1} + R)^{-1} + Q. \tag{6.90}$$

The operator Q is in fact given by

$$Q = -(A'_{k,\eta})^{-1}[\![\mathcal{A}]\!](A'_{0,1} + R)^{-1}, \tag{6.91}$$

since

$$\begin{aligned} I = (A'_{k,\eta})(A'_{k,\eta})^{-1} &= (A'_{k,\eta})(A'_{0,1} + R)^{-1} + (A'_{k,\eta})Q \\ &= (A'_{0,1} + R + [\![\mathcal{A}]\!])(A'_{0,1} + R)^{-1} + (A'_{k,\eta})Q \\ &= I + [\![\mathcal{A}]\!](A'_{0,1} + R)^{-1} + (A'_{k,\eta})Q. \end{aligned}$$

Step 2 We rewrite the difference $\mathrm{DtN}_k - \mathrm{DtN}_0$ using the combined field equations. Using the combined field equations (6.71) with for η as in (6.79) (for DtN_k) and $\eta = 1$ (for DtN_0) we find

$$\mathrm{DtN}_k - \mathrm{DtN}_0 = (A'_{k,\eta})^{-1}\left[-D_k + i\eta\left(-\frac{1}{2} + K_k\right)\right] - (A'_{0,1})^{-1}\left[-D_0 + i\left(-\frac{1}{2} + K_0\right)\right]. \tag{6.92}$$

Adding and subtracting D_0 and K_0 in (6.92), employing the splitting of $D_k - D_0$ and $K_k - K_0$ given in (6.75) in Proposition 6.9.1, and applying the splitting of $(A'_{k,1})^{-1}$ in (6.90) we find

$$\begin{aligned}
\mathrm{DtN}_k - \mathrm{DtN}_0 &= -(A'_{k,\eta})^{-1}\left[D_k - D_0\right] - (A'_{k,\eta})^{-1} D_0 \\
&\quad + i\eta (A'_{k,\eta})^{-1}\left[K_k - K_0\right] + i\eta (A'_{k,\eta})^{-1}\left[-1/2 + K_0\right] \\
&\quad + (A'_{0,1})^{-1} D_0 - i(A'_{0,1})^{-1} D_0\left[-1/2 + K_0\right] \\
&= (A'_{0,1})^{-1} D_0 - (A'_{0,1} + R)^{-1} D_0 - Q D_0 \\
&\quad - (A'_{0,1} + R)^{-1} S_D - Q S_D - (A'_{k,\eta})^{-1} \gamma_1^{int} \tilde{A}_K \\
&\quad + i\eta (A'_{0,1} + R)^{-1} S_K + i\eta Q S_K + i\eta (A'_{k,\eta})^{-1} \gamma_0^{int} \tilde{A}_K \\
&\quad + i\eta (A'_{0,1} + R)^{-1}\left[-1/2 + K_0\right] + i\eta Q\left[-1/2 + K_0\right] \\
&\quad - i(A'_{0,1})^{-1}\left[-1/2 + K_0\right] \\
&= \mathrm{FSO} + \mathrm{ASO},
\end{aligned}$$

where the Finite Shift Operators (FSO) and the Analytic Shift Operators (ASO) are given by

$$\begin{aligned}
\mathrm{FSO} &:= (A'_{0,1})^{-1} D_0 - (A'_{0,1} + R)^{-1} D_0 \\
&\quad - (A'_{0,1} + R)^{-1} S_D + i\eta (A'_{0,1} + R)^{-1} S_K \\
&\quad + i\eta (A'_{0,1} + R)^{-1}\left[-1/2 + K_0\right] \\
&\quad - i(A'_{0,1})^{-1}\left[-1/2 + K_0\right], \\
\mathrm{ASO} &:= -Q D_0 - Q S_D - (A'_{k,\eta})^{-1} \gamma_1^{int} \tilde{A}_K \\
&\quad + i\eta Q S_K + i\eta (A'_{k,\eta})^{-1} \gamma_0^{int} \tilde{A}_K + i\eta Q\left[-1/2 + K_0\right].
\end{aligned}$$

Step 3: We analyze the Finite Shift Operators (FSO). We will show that

$$\mathrm{FSO} = kB, \tag{6.93}$$

where the linear operator B maps in fact from $H^s(\Gamma)$ to $H^s(\Gamma)$ and satisfies $\|Bu\|_{s,\Gamma} \lesssim \|u\|_{s,\Gamma}$, as in the assertion of the present lemma. Using the mapping properties of $(A'_{0,1} + R)^{-1}$ in (6.89a) as well as (6.74), (6.78) and Item (ii) in Proposition 6.9.1 we find

$$\begin{aligned}
\|(A'_{0,1} + R)^{-1} S_D u\|_{s,\Gamma} &\lesssim \|S_D u\|_{s,\Gamma} \lesssim k\|u\|_{s,\Gamma}, \\
k\|(A'_{0,1} + R)^{-1} S_K u\|_{s,\Gamma} &\lesssim k\|S_K u\|_{s,\Gamma} \lesssim k\|S_K u\|_{s,\Gamma} \lesssim k\|u\|_{s,\Gamma}, \\
k\|(A'_{0,1} + R)^{-1}\left[-1/2 + K_0\right] u\|_{s,\Gamma} &\lesssim k\|\left[-1/2 + K_0\right] u\|_{s,\Gamma} \lesssim k\|u\|_{s,\Gamma}, \\
\|(A'_{0,1})^{-1}\left[-1/2 + K_0\right] u\|_{s,\Gamma} &\lesssim \|\left[-1/2 + K_0\right] u\|_{s,\Gamma} \lesssim \|u\|_{s,\Gamma}.
\end{aligned}$$

Once we have shown

$$\|(A'_{0,1})^{-1} D_0 - (A'_{0,1} + R)^{-1} D_0 u\|_{s,\Gamma} \lesssim k\|u\|_{s,\Gamma},$$

the assertion in (6.93) follows. Using (6.88) we find

$$\begin{aligned}(A'_{0,1})^{-1}D_0 - (A'_{0,1}+R)^{-1}D_0 &= (A'_{0,1})^{-1}D_0 - (I+(A'_{0,1})^{-1}R)^{-1}(A'_{0,1})^{-1}D_0 \\ &= \left[I - (I+(A'_{0,1})^{-1}R)^{-1}\right](A'_{0,1})^{-1}D_0 \\ &= -\left[\sum_{n=1}^{\infty}(-1)^n((A'_{0,1})^{-1}R)^n\right](A'_{0,1})^{-1}D_0.\end{aligned}$$

Applying the previous calculations, a geometric series argument with (6.87), the mapping properties of $(A'_{0,1})^{-1}$ in Item (ii) in Proposition 6.9.1, the estimate $\|Ru\|_{s,\Gamma} \lesssim k\|u\|_{s-1,\Gamma}$ given by Theorem 6.9.6, again the mapping properties of $(A'_{0,1})^{-1}$ and finally the mapping properties of D_0 given in (6.74) in Proposition 6.9.1 we find

$$\begin{aligned}&\|(A'_{0,1})^{-1}D_0 - (A'_{0,1}+R)^{-1}D_0u\|_{s,\Gamma} \\ &\qquad= \left\|\left[\sum_{n=1}^{\infty}(-1)^n((A'_{0,1})^{-1}R)^{n-1}\right]((A'_{0,1})^{-1}R)(A'_{0,1})^{-1}D_0u\right\|_{s,\Gamma} \\ &\qquad\le \frac{1}{1-\hat{q}}\|(A'_{0,1})^{-1}R(A'_{0,1})^{-1}D_0u\|_{s,\Gamma} \\ &\qquad\lesssim \|R(A'_{0,1})^{-1}D_0u\|_{s,\Gamma} \\ &\qquad\lesssim k\|(A'_{0,1})^{-1}D_0u\|_{s-1,\Gamma} \\ &\qquad\lesssim k\|D_0u\|_{s-1,\Gamma} \\ &\qquad\lesssim k\|u\|_{s,\Gamma}.\end{aligned}$$

Hence, the assertion in (6.93) follows, which concludes the analysis of the finite shift operators FSO. Summarizing, so far we have found that

$$\mathrm{DtN}_k - \mathrm{DtN}_0 = kB + \mathrm{ASO},$$

with B as in the assertions of the present lemma.

Step 4: We analyze the Analytic Shift Operators (ASO). We have

$$\begin{aligned}\mathrm{ASO} = &-QD_0 - Q\left[S_D - i\eta S_K - i\eta\left[-1/2 + K_0\right]\right] \\ &+ (A'_{k,\eta})^{-1}[i\eta\gamma_0^{int}\tilde{A}_K - \gamma_1^{int}\tilde{A}_K].\end{aligned}$$

Step 4a: We analyze the term $-QD_0$. In view of (6.91) we have for $f \in H^s(\Gamma)$

$$\begin{aligned}-QD_0f &= (A'_{k,\eta})^{-1}[\![\mathcal{A}]\!](A'_{0,1}+R)^{-1}D_0f \\ &= (A'_{k,\eta})^{-1}\left[k[\![\mathcal{A}_1]\!](A'_{0,1}+R)^{-1}D_0f + [\![\partial_n\mathcal{A}_2]\!](A'_{0,1}+R)^{-1}D_0f\right].\end{aligned}$$

In order to apply Lemma 6.9.7, we use the mapping properties of $\mathcal{A}_1$ and $\mathcal{A}_2$ given in Theorem 6.9.6 and estimate

$$\begin{aligned}k\|(A'_{0,1}+R)^{-1}D_0f\|_{-3/2,\Gamma} + k^{d/2}\|(A'_{0,1}+R)^{-1}D_0f\|_{-1,\Gamma} &\lesssim k^{d/2}\|(A'_{0,1}+R)^{-1}D_0f\|_{s-1,\Gamma} \\ &\lesssim k^{d/2}\|D_0f\|_{s-1,\Gamma} \\ &\lesssim k^{d/2}\|f\|_{s,\Gamma},\end{aligned}$$

where we used the trivial embedding $H^{s-1}(\Gamma) \subset H^{-1}(\Gamma) \subset H^{-3/2}(\Gamma)$, the fact that $k + k^{d/2} \lesssim k^{d/2}$, the mapping property (6.89b) and finally the mapping properties of D_0 given in (6.74) in Proposition 6.9.1. Hence, for the tubular neighborhood T given in Theorem 6.9.6 we find

$$\mathcal{A}_1(A'_{0,1} + R)^{-1} D_0 f \in \mathcal{A}(C_1 k^{d/2} \|f\|_{s,\Gamma}, \gamma_1, T),$$
$$\mathcal{A}_2(A'_{0,1} + R)^{-1} D_0 f \in \mathcal{A}(C_1 k \|f\|_{s,\Gamma}, \gamma_1, T),$$

for constants C_1, $\gamma_1 > 0$ independent of k. We find Lemma 6.9.7 to be applicable, which yields the representation

$$-QD_0 f = [\![\partial_n v_f^1]\!], \qquad v_f^1 \in \mathcal{A}(\tilde{C}_1 k^{5/2+d/2} \|f\|_{s,\Gamma}, \tilde{\gamma}_1, \Omega_R), \tag{6.94}$$

for constants $\tilde{C}_1$, $\tilde{\gamma}_1 > 0$ independent of k.

Step 4b: We analyze the term

$$-Q\left[S_D - i\eta S_K - i\eta\left[-1/2 + K_0\right]\right].$$

We proceed very similar to Step 4a. We estimate

$$\begin{aligned}
k^{d/2}\|(A'_{0,1} + R)^{-1}\left[S_D - i\eta S_K - i\eta\left[-1/2 + K_0\right]\right] f\|_{-1,\Gamma} \\
&\lesssim k^{d/2}\|(A'_{0,1} + R)^{-1}\left[S_D - i\eta S_K - i\eta\left[-1/2 + K_0\right]\right] f\|_{s,\Gamma} \\
&\lesssim k^{d/2}\|\left[S_D - i\eta S_K - i\eta\left[-1/2 + K_0\right]\right] f\|_{s,\Gamma} \\
&\lesssim k^{d/2+1}\|f\|_{s,\Gamma},
\end{aligned}$$

where we first use the trivial embedding $H^s(\Gamma) \subset H^{-1}(\Gamma)$, the mapping property (6.89a), the mapping properties of S_D, S_K and K_0 given in in (6.78) and (6.74) in Proposition 6.9.1 as well as $|\eta| \lesssim k$. Proceeding as in Step 4a we find the representation

$$-Q\left[S_D - i\eta S_K - i\eta\left[-1/2 + K_0\right]\right] = [\![\partial_n v_f^2]\!], \quad v_f^2 \in \mathcal{A}(\tilde{C}_2 k^{5/2+d/2+1} \|f\|_{s,\Gamma}, \tilde{\gamma}_2, \Omega_R) \tag{6.95}$$

to hold true, for constants $\tilde{C}_2$, $\tilde{\gamma}_2 > 0$ independent of k.

Step 4c: We analyze the term

$$(A'_{k,\eta})^{-1}[i\eta\gamma_0^{int}\tilde{A}_K - \gamma_1^{int}\tilde{A}_K].$$

For $f \in H^s(\Gamma)$ the mapping properties of $\tilde{A}_K$ are such that $\tilde{A}_K f \in \mathcal{A}(C_K\|f\|_{-1/2,\Gamma}, \gamma_K, \Omega)$, see (6.77) in Proposition 6.9.1. Upon extending $\tilde{A}_K f$ with zero outside of Ω, we find Lemma 6.9.7 to be applicable, which yields the representation

$$(A'_{k,\eta})^{-1}[i\eta\gamma_0^{int}\tilde{A}_K - \gamma_1^{int}\tilde{A}_K] f = [\![\partial_n v_f^3]\!], \qquad v_f^3 \in \mathcal{A}(\tilde{C}_3 k^{5/2}\|f\|_{-1/2,\Gamma}, \tilde{\gamma}_3, \Omega_R), \tag{6.96}$$

with constants $\tilde{C}_3$, $\tilde{\gamma}_3 > 0$ independent of k.

Step 5: Collecting the representations derived in (6.94), (6.95) and (6.96), we find

$$\mathrm{ASO} = [\![\partial_n \tilde{A}]\!], \qquad \tilde{A}u \in \mathcal{A}(Ck^{7/2+d/2}\|u\|_{s,\Gamma}, \gamma, \Omega_R)$$

with $\tilde{A}$ as in the assertions of the present lemma. Hence, the splitting

$$\mathrm{DtN}_k - \mathrm{DtN}_0 = kB + [\![\partial_n \tilde{A}]\!]$$

with B and $\tilde{A}$ as asserted, holds true. This concludes the proof. □

6.10. Dirichlet-to-Neumann map for Elasticity

In the present section we prove positivity of the Dirichlet-to-Neumann map DtN_0 and propose a splitting for the difference of the Dirichlet-to-Neumann maps corresponding to linear elasticity. Furthermore, we analyze the symbol of the finite regularity part in Lemma 6.10.5. Following standard notation we denote by λ and μ the Lamé parameters. We assume $\mu > 0$ and $\lambda \geq 0$, which is common for elastic materials, see [McL00, p. 299]. We denote by J_ν and Y_ν the standard Bessel functions. Furthermore, the Hankel function $H_\nu^{(1)}$ is given by $H_\nu^{(1)} = J_\nu + iY_\nu$. The Dirichlet-to-Neumann map is explicitly known on the unit circle in spatial dimension two. Let $\boldsymbol{u} \in \boldsymbol{L}^2(\Gamma)$ be given in terms of a Fourier expansion in polar coordinates

$$\boldsymbol{u}(\theta) = \sum_{n\in\mathbb{Z}} (u_n^r \boldsymbol{e}_r + u_n^\theta \boldsymbol{e}_\theta)\mathrm{e}^{in\theta},$$

where $\boldsymbol{e}_r = (\cos(\theta), \sin(\theta))^T$, $\boldsymbol{e}_\theta = (-\sin(\theta), \cos(\theta))^T$ and u_n^r, u_n^θ are the Fourier coefficients. The Dirichlet-to-Neumann map is given, see [Yua19, Eq. 2.59], by

$$\begin{aligned}\mathrm{DtN}_k\boldsymbol{u} = &\sum_{n\in\mathbb{Z}} \left[\left(-\mu + \frac{\alpha_{2,n}k^2}{\Lambda_n}\right) u_n^r + \left(-in\mu + \frac{ink^2}{\Lambda_n}\right) u_n^\theta\right] \boldsymbol{e}_r \mathrm{e}^{in\theta} \\ &+ \sum_{n\in\mathbb{Z}} \left[\left(-\mu + \frac{\alpha_{1,n}k^2}{\Lambda_n}\right) u_n^\theta + \left(in\mu - \frac{ink^2}{\Lambda_n}\right) u_n^r\right] \boldsymbol{e}_\theta \mathrm{e}^{in\theta},\end{aligned}$$

where Λ_n is given by, see the last equation in the proof of [Yua19, Lemma 2.7.1],

$$\Lambda_n = n^2 - \alpha_{1,n}\alpha_{2,n}$$

with, see second to last equation in the proof of [Yua19, Lemma 2.7.1]

$$\alpha_{1,n} = \kappa_1 \frac{H_n^{(1)\prime}(\kappa_1)}{H_n^{(1)}(\kappa_1)}, \qquad \alpha_{2,n} = \kappa_2 \frac{H_n^{(1)\prime}(\kappa_2)}{H_n^{(1)}(\kappa_2)}$$

with compressional and shear wavenumber κ_1 and κ_2 given by

$$\kappa_1 = \frac{k}{\sqrt{\lambda + 2\mu}}, \qquad \kappa_2 = \frac{k}{\sqrt{\mu}}$$

with λ, μ being the Lamé parameter. The canonical operator DtN_0 is derived by considering the symbol of DtN_k and passing to $k = 0$. We collect some results concerning Hankel functions in the following

Lemma 6.10.1 (Properties of Hankel functions)**.** *There holds:*

(i) $H_{-n}^{(1)} = (-1)^n H_n^{(1)}$ *for all* $n \in \mathbb{N}$*.*

(ii) $H_0^{(1)}(z) \sim \frac{2i}{\pi} \ln z$ *for* $z \to 0$*.*

(iii) $H_n^{(1)}(z) \sim -\frac{i}{\pi}\Gamma(n)\left(\frac{z}{2}\right)^{-n}$ *for* $n > 0$ *fixed and* $z \to 0$*.*

Proof. For Item (i) see [Olv97, Ch. 7, Eq. (4.09)]. Item (ii) and (iii) follow immediately by the asymptotics in [Olv97, Ch. 12, Eq. (1.07),(1.08)]. □

Lemma 6.10.2. *There holds:*

(i) $\alpha_{i,-n} = \alpha_{i,n}$ *for* $i = 1, 2$.

(ii) $\alpha_{i,n} \to -|n|$ *for* $k \to 0$.

(iii) $\frac{k^2}{\Lambda_0} \to 0$ *for* $k \to 0$.

(iv) $\frac{k^2}{\Lambda_n} \to \frac{2\mu(\lambda+2\mu)(|n|-1)}{(\lambda+3\mu)|n|}$ *for* $k \to 0$ *and* $n \geq 1$.

Proof. The result in Item (i) follows immediately from Lemma 6.10.1 Item (i). For Item (ii) note that the Hankel functions satisfy the differential relation

$$H_n^{(1)\prime}(z) = -H_{n+1}^{(1)}(z) + \frac{n}{z} H_n^{(1)}(z), \tag{6.97}$$

see [DLMF, Eq. 10.6.2]. Hence, for $n = 0$ we find together with Lemma 6.10.1

$$z\frac{H_0^{(1)\prime}(z)}{H_0^{(1)}(z)} = -z\frac{H_1^{(1)}(z)}{H_0^{(1)}(z)} \sim z\frac{\frac{i}{\pi}\Gamma(1)\left(\frac{z}{2}\right)^{-1}}{\frac{2i}{\pi}\ln z} = \frac{1}{\ln z} \to 0, \tag{6.98}$$

as $z \to 0$, which proves Item (ii) for $n = 0$. For $n \geq 1$ using (6.97) as well as Lemma 6.10.1 Item (iii) we find for $z \to 0$

$$\begin{aligned} z\frac{H_n^{(1)\prime}(z)}{H_n^{(1)}(z)} &= -z\frac{H_{n+1}^{(1)}(z)}{H_n^{(1)}(z)} + n \sim -z\frac{-\frac{i}{\pi}\Gamma(n+1)\left(\frac{z}{2}\right)^{-(n+1)}}{-\frac{i}{\pi}\Gamma(n)\left(\frac{z}{2}\right)^{-n}} + n \\ &= -zn\left(\frac{z}{2}\right)^{-1} + n = -n, \end{aligned}$$

which proves Item (ii), in view of Item (i) in Lemma 6.10.1. For Item (iii) we use the calculations in (6.98) to conclude

$$\frac{k^2}{\Lambda_0} = -\frac{k^2}{\alpha_{0,1}\alpha_{0,2}} \sim -\frac{k^2}{\frac{1}{\ln\kappa_1}\frac{1}{\ln\kappa_2}} = -k^2 \ln\kappa_1 \ln\kappa_2 \to 0,$$

as $k \to 0$. We turn to Item (iv). For $n = 1$ we use the fact that

$$H_1^{(1)\prime}(z) = H_0^{(1)}(z) - \frac{1}{z}H_1^{(1)}(z), \tag{6.99}$$

see again [DLMF, Eq. 10.6.2]. We find with (6.99)

$$\begin{aligned} z\frac{H_1^{(1)\prime}(z)}{H_1^{(1)}(z)} &= z\frac{H_0^{(1)}(z)}{H_1^{(1)}(z)} - 1 \sim -z\frac{\frac{2i}{\pi}\ln z}{\frac{i}{\pi}\frac{2}{z}} - 1 \\ &= -z^2\ln z - 1, \end{aligned}$$

as $z \to 0$. Hence, we find

$$\frac{k^2}{\Lambda_1} = \frac{k^2}{1-\alpha_{0,1}\alpha_{0,2}} \sim \frac{k^2}{1-(1+\kappa_1^2\ln\kappa_1)(1+\kappa_2^2\ln\kappa_2)}$$
$$= -\frac{k^2}{\kappa_1^2\ln\kappa_1+\kappa_2^2\ln\kappa_2+\kappa_1^2\kappa_2^2\ln\kappa_1\ln\kappa_2} \to 0,$$

as $k \to 0$. We proceed with case $n \geq 2$. We truncate the series expansions in [DLMF, Eq. 10.2.2] and [DLMF, Eq. 10.8.1] for $n \geq 2$ to find

$$H_n^{(1)}(z) \sim -\frac{i}{\pi}(n-1)!\left(\frac{z}{2}\right)^{-n} - \frac{i}{\pi}(n-2)!\left(\frac{z}{2}\right)^{-n+2}, \tag{6.100}$$

as $z \to 0$. Using again (6.97) and inserting (6.100) we find

$$z\frac{H_n^{(1)\prime}(z)}{H_n^{(1)}(z)} = -z\frac{H_{n+1}^{(1)}(z)}{H_n^{(1)}(z)} + n$$
$$\sim -z\frac{-\frac{i}{\pi}n!\left(\frac{z}{2}\right)^{-n-1} - \frac{i}{\pi}(n-1)!\left(\frac{z}{2}\right)^{-n+1}}{-\frac{i}{\pi}(n-1)!\left(\frac{z}{2}\right)^{-n} - \frac{i}{\pi}(n-2)!\left(\frac{z}{2}\right)^{-n+2}} + n$$
$$= -2\frac{n(n-1)+(n-1)\frac{z^2}{4}}{(n-1)+\frac{z^2}{4}} + n.$$

Inserting the above in the definition of Λ_n, using the definition of κ_1 and κ_2, and performing elementary calculations we finally find

$$\frac{k^2}{\Lambda_n} = -\frac{k^2}{n^2-\alpha_{0,1}\alpha_{0,2}} \sim \frac{(k^2+4\mu(n-1))(k^2+4(\lambda+2\mu)(n-1))}{4(n-1)(k^2+2(\lambda+3\mu)n)} \to \frac{2\mu(\lambda+2\mu)(n-1)}{(\lambda+3\mu)n},$$

as $k \to 0$, which concludes the proof. □

We therefore find using Lemma 6.10.2 that

$$\left(-\mu+\frac{\alpha_{2,n}k^2}{\Lambda_n}\right) \to -\mu-\sigma_n, \qquad \left(-in\mu+\frac{ink^2}{\Lambda_n}\right) \to -in\mu+i\operatorname{sgn}(n)\sigma_n,$$
$$\left(-\mu+\frac{\alpha_{1,n}k^2}{\Lambda_n}\right) \to -\mu-\sigma_n, \qquad \left(in\mu-\frac{ink^2}{\Lambda_n}\right) \to in\mu-i\operatorname{sgn}(n)\sigma_n,$$

as $k \to 0$ with

$$\sigma_n := \begin{cases} 0 & n=0,\\ \frac{2\mu(\lambda+2\mu)}{(\lambda+3\mu)}(|n|-1) & \text{else.}\end{cases}$$

Hence, DtN_0 is given by

$$\mathrm{DtN}_0\boldsymbol{u} = \sum_{n\in\mathbb{Z}}\left[(-\mu-\sigma_n)\,u_n^r + (-in\mu+i\operatorname{sgn}(n)\sigma_n)\,u_n^\theta\right]\boldsymbol{e}_r\mathrm{e}^{in\theta}$$
$$+\sum_{n\in\mathbb{Z}}\left[(-\mu-\sigma_n)\,u_n^\theta + (in\mu-i\operatorname{sgn}(n)\sigma_n)\,u_n^r\right]\boldsymbol{e}_\theta\mathrm{e}^{in\theta}.$$

Lemma 6.10.3 (Positivity of $-\mathrm{DtN}_0$)**.** *Let $\mu > 0$ and $\lambda + \frac{5}{3}\mu \geq 0$. Then the operator $-\mathrm{DtN}_0$ is positive, i.e.,*

$$-\langle \mathrm{DtN}_0 u, u\rangle \geq 0$$

for all $u \in H^{1/2}(\Gamma)$.

Proof. The positivity of $-\mathrm{DtN}_0$ follows immediately from the positive definiteness of the matrix corresponding to the symbol. Hence, once we find that for each $n \in \mathbb{Z}$ the matrix

$$M_n := -\begin{pmatrix} -\mu - \sigma_n & -in\mu + i\operatorname{sgn}(n)\sigma_n \\ in\mu - i\operatorname{sgn}(n)\sigma_n & -\mu - \sigma_n \end{pmatrix}$$

is positive definite, the proof is complete. For $n = 0$ we have

$$M_0 = \begin{pmatrix} \mu & 0 \\ 0 & \mu \end{pmatrix}$$

which is trivially positive definite, since $\mu > 0$. For $n \neq 0$ we have

$$M_n = \begin{pmatrix} \mu + \frac{2\mu(\lambda+2\mu)}{(\lambda+3\mu)}(|n|-1) & -in\mu + i\operatorname{sgn}(n)\frac{2\mu(\lambda+2\mu)}{(\lambda+3\mu)}(|n|-1) \\ in\mu - i\operatorname{sgn}(n)\frac{2\mu(\lambda+2\mu)}{(\lambda+3\mu)}(|n|-1) & \mu + \frac{2\mu(\lambda+2\mu)}{(\lambda+3\mu)}(|n|-1) \end{pmatrix}.$$

Since $\mu > 0$ and $\lambda + \frac{5}{3}\mu \geq 0$ we find

$$\mu + \frac{2\mu(\lambda + 2\mu)}{(\lambda + 3\mu)}(|n| - 1) > 0.$$

Hence, the top-left entry of M_n is positive for $n \in \mathbb{Z} \setminus \{0\}$. It is easily verified that in fact the determinant is also positive under the assumptions of the present lemma, which concludes the proof. □

The difference $\mathrm{DtN}_k - \mathrm{DtN}_0$ is given by

$$\begin{aligned}\mathrm{DtN}_k\boldsymbol{u} - \mathrm{DtN}_0\boldsymbol{u} &= \sum_{n\in\mathbb{Z}} \left[\left(\frac{\alpha_{2,n}k^2}{\Lambda_n} + \sigma_n\right)u_n^r + \left(\frac{ink^2}{\Lambda_n} - i\operatorname{sgn}(n)\sigma_n\right)u_n^\theta\right]\boldsymbol{e}_r \mathrm{e}^{in\theta} \\ &\quad + \sum_{n\in\mathbb{Z}} \left[\left(\frac{\alpha_{1,n}k^2}{\Lambda_n} + \sigma_n\right)u_n^\theta + \left(-\frac{ink^2}{\Lambda_n} + i\operatorname{sgn}(n)\sigma_n\right)u_n^r\right]\boldsymbol{e}_\theta \mathrm{e}^{in\theta}.\end{aligned}$$

We propose the k-dependent splitting

$$\mathrm{DtN}_k\boldsymbol{u} - \mathrm{DtN}_0\boldsymbol{u} = \sum_{n\in\mathbb{Z}} \cdots = \sum_{|n|>2k} \cdots + \sum_{|n|\leq 2k} \cdots = R_\Gamma + A_\Gamma. \tag{6.101}$$

In the following we verify that the operator R_Γ is in fact an operator of order zero and its symbol is uniformly bounded by k.

Lemma 6.10.4. *Let $n \geq 2x$. With the above notation the uniform asymptotic expansion*

$$x\frac{H_n^{(1)\prime}(x)}{H_n^{(1)}(x)} = -n + \frac{x^2}{2n} + r,$$

with

$$|r| \lesssim \frac{x^2}{n^2} + \frac{x^4}{n^3}$$

holds.

Proof. The asymptotics in [DLMF, Eq. 10.20.6] and [DLMF, Eq. 10.20.9], respectively give

$$H_n^{(1)}(nz) \sim 2\mathrm{e}^{-\pi i/3}\left(\frac{4\zeta}{1-z^2}\right)^{\frac{1}{4}} \times \left(\frac{\mathrm{Ai}\left(\mathrm{e}^{2\pi i/3}\nu^{\frac{2}{3}}\zeta\right)}{\nu^{\frac{1}{3}}}\sum_{k=0}^{\infty}\frac{A_k(\zeta)}{\nu^{2k}} + \frac{\mathrm{e}^{2\pi i/3}\mathrm{Ai}'\left(\mathrm{e}^{2\pi i/3}\nu^{\frac{2}{3}}\zeta\right)}{\nu^{\frac{5}{3}}}\sum_{k=0}^{\infty}\frac{B_k(\zeta)}{\nu^{2k}}\right),$$

$$H_n^{(1)\prime}(nz) \sim \frac{4}{z}\mathrm{e}^{-2\pi i/3}\left(\frac{1-z^2}{4\zeta}\right)^{\frac{1}{4}} \times \left(\frac{\mathrm{e}^{-2\pi i/3}\mathrm{Ai}\left(\mathrm{e}^{2\pi i/3}\nu^{\frac{2}{3}}\zeta\right)}{\nu^{\frac{4}{3}}}\sum_{k=0}^{\infty}\frac{C_k(\zeta)}{\nu^{2k}} + \frac{\mathrm{Ai}'\left(\mathrm{e}^{2\pi i/3}\nu^{\frac{2}{3}}\zeta\right)}{\nu^{\frac{2}{3}}}\sum_{k=0}^{\infty}\frac{D_k(\zeta)}{\nu^{2k}}\right)$$

uniformly in z, with

$$\frac{2}{3}\zeta^{3/2} = \ln\frac{1+\sqrt{1-z^2}}{z} - \sqrt{1-z^2}. \tag{6.102}$$

We are interested in the case $n \geq 2x$, with $z = x/n$, i.e., $z \in (0, 1/2)$. We will use the following abbreviations for now:

$$A \sim \sum_{k=0}^{\infty}\frac{A_k(\zeta)}{\nu^{2k}}, \quad B \sim \sum_{k=0}^{\infty}\frac{B_k(\zeta)}{\nu^{2k}}, \quad C \sim \sum_{k=0}^{\infty}\frac{C_k(\zeta)}{\nu^{2k}}, \quad D \sim \sum_{k=0}^{\infty}\frac{D_k(\zeta)}{\nu^{2k}}.$$

For the definitions of A_k, B_k, C_k and D_k we refer to [DLMF, Eq. 10.20.10-13]. Furthermore, we will use the asymptotics for the Airy functions Ai and Ai′, given in [DLMF, Eq. 9.7.5 and 9.7.6]. To that end, let us introduce

$$\xi = \frac{2}{3}\left(\mathrm{e}^{2\pi i/3}\nu^{\frac{2}{3}}\zeta\right)^{3/2} = -\frac{2}{3}\nu\zeta^{3/2}.$$

We then have

$$\mathrm{Ai}\left(\mathrm{e}^{2\pi i/3}\nu^{\frac{2}{3}}\zeta\right) \sim \frac{\mathrm{e}^{-\xi}}{2\sqrt{\pi}(\mathrm{e}^{2\pi i/3}\nu^{\frac{2}{3}}\zeta)^{1/4}}\sum_{k=0}^{\infty}(-1)^k\frac{u_k}{\xi^k},$$

$$\mathrm{Ai}'\left(\mathrm{e}^{2\pi i/3}\nu^{\frac{2}{3}}\zeta\right) \sim -\frac{(\mathrm{e}^{2\pi i/3}\nu^{\frac{2}{3}}\zeta)^{1/4}\mathrm{e}^{-\xi}}{2\sqrt{\pi}}\sum_{k=0}^{\infty}(-1)^k\frac{v_k}{\xi^k},$$

with u_k, v_k as in [DLMF, Eq. 9.7.2]. Again, we abbreviate

$$U \sim \sum_{k=0}^{\infty}(-1)^k \frac{u_k}{\xi^k}, \quad V \sim \sum_{k=0}^{\infty}(-1)^k \frac{v_k}{\xi^k}.$$

We find after elementary calculations

$$\begin{aligned} x\frac{H_n^{(1)\prime}(x)}{H_n^{(1)}(x)} &= x\frac{H_n^{(1)\prime}(nz)}{H_n^{(1)}(nz)} \\ &\sim x\frac{\frac{4}{z}\mathrm{e}^{-2\pi i/3}\left(\frac{1-z^2}{4\zeta}\right)^{\frac{1}{4}}\left(\frac{\mathrm{e}^{-2\pi i/3}\mathrm{Ai}\left(\mathrm{e}^{2\pi i/3}\nu^{\frac{2}{3}}\zeta\right)}{\nu^{\frac{4}{3}}}C + \frac{\mathrm{Ai}'\left(\mathrm{e}^{2\pi i/3}\nu^{\frac{2}{3}}\zeta\right)}{\nu^{\frac{2}{3}}}D\right)}{2\mathrm{e}^{-\pi i/3}\left(\frac{4\zeta}{1-z^2}\right)^{\frac{1}{4}}\left(\frac{\mathrm{Ai}\left(\mathrm{e}^{2\pi i/3}\nu^{\frac{2}{3}}\zeta\right)}{\nu^{\frac{1}{3}}}A + \frac{\mathrm{e}^{2\pi i/3}\mathrm{Ai}'\left(\mathrm{e}^{2\pi i/3}\nu^{\frac{2}{3}}\zeta\right)}{\nu^{\frac{5}{3}}}B\right)} \\ &= n\left(\frac{1-z^2}{\zeta}\right)^{1/2}\frac{\frac{\mathrm{e}^{-2\pi i/3}\mathrm{Ai}\left(\mathrm{e}^{2\pi i/3}\nu^{\frac{2}{3}}\zeta\right)}{\nu^{\frac{4}{3}}}C + \frac{\mathrm{Ai}'\left(\mathrm{e}^{2\pi i/3}\nu^{\frac{2}{3}}\zeta\right)}{\nu^{\frac{2}{3}}}D}{\frac{\mathrm{Ai}\left(\mathrm{e}^{2\pi i/3}\nu^{\frac{2}{3}}\zeta\right)}{\nu^{\frac{1}{3}}}A + \frac{\mathrm{e}^{2\pi i/3}\mathrm{Ai}'\left(\mathrm{e}^{2\pi i/3}\nu^{\frac{2}{3}}\zeta\right)}{\nu^{\frac{5}{3}}}B} \\ &= -\sqrt{n^2-x^2}\frac{\zeta^{-1/2}UC + nVD}{\zeta^{1/2}VB + nUA}. \end{aligned} \tag{6.103}$$

We now truncate the series expansions of A, B, C, D, U and V with the remainder denoted with a subscript r as follows:

$$\begin{aligned} A &\sim \underbrace{A_0(\zeta) + \frac{A_1(\zeta)}{n^2}}_{=:A_p} + A_r, & A_r &= O\left(\frac{1}{n^4}\right), \\ B &\sim \underbrace{B_0(\zeta) + \frac{B_1(\zeta)}{n^2}}_{=:B_p} + B_r, & B_r &= O\left(\frac{\zeta^{-1/2}}{n^4}\right), \\ C &\sim \underbrace{C_0(\zeta) + \frac{C_1(\zeta)}{n^2}}_{=:C_p} + C_r, & C_r &= O\left(\frac{\zeta^{1/2}}{n^4}\right), \\ D &\sim \underbrace{D_0(\zeta) + \frac{D_1(\zeta)}{n^2}}_{=:D_p} + D_r, & D_r &= O\left(\frac{1}{n^4}\right), \\ U &\sim \underbrace{1 - \frac{u_1}{\xi} + \frac{u_2}{\xi^2} - \frac{u_3}{\xi^3}}_{=:U_p} + U_r, & U_r &= O\left(\frac{1}{\xi^4}\right), \\ V &\sim \underbrace{1 - \frac{v_1}{\xi} + \frac{v_2}{\xi^2} - \frac{v_3}{\xi^3}}_{=:V_p} + V_r, & V_r &= O\left(\frac{1}{\xi^4}\right). \end{aligned} \tag{6.104}$$

Once we have shown

$$\frac{\zeta^{-1/2}UC + nVD}{\zeta^{1/2}VB + nUA} = 1 + O\left(\frac{x^2}{n^3}\right), \tag{6.105}$$

the result follows, since by Taylor expansion and the calculations in (6.103) we have

$$\begin{aligned}
x\frac{H_n^{(1)\prime}(x)}{H_n^{(1)}(x)} &= -\sqrt{n^2-x^2}\frac{\zeta^{-1/2}UC + nVD}{\zeta^{1/2}VB + nUA} \\
&= -\sqrt{n^2-x^2}\left(1 + O\left(\frac{x^2}{n^3}\right)\right) \\
&= -n\left(1 - \frac{x^2}{2n^2} + O\left(\frac{x^4}{n^4}\right)\right)\left(1 + O\left(\frac{x^2}{n^3}\right)\right) \\
&= -n\left(1 - \frac{x^2}{2n^2} + O\left(\frac{x^4}{n^4}\right)\right)\left(1 + O\left(\frac{x^2}{n^3}\right)\right) \\
&= -n + \frac{x^2}{2n} + r
\end{aligned}$$

with

$$|r| \lesssim \frac{x^2}{n^2} + \frac{x^4}{n^3}.$$

We now verify (6.105). To that end, we insert (6.104) into

$$\frac{\zeta^{-1/2}UC + nVD}{\zeta^{1/2}VB + nUA}.$$

We first simplify the numerator in the above. Elementary calculations show, with the use of (6.104),

$$\frac{\zeta^{-1/2}UC + nVD}{\zeta^{1/2}VB + nUA} = \frac{\zeta^{-1/2}U_pC_p + nV_pD_p}{\zeta^{1/2}VB + nUA} + O\left(\frac{x^2}{n^3}\right).$$

The denominator is treated similarly with the aid of Taylor expansion. Again with appropriate use of (6.104) we find

$$\frac{\zeta^{-1/2}UC + nVD}{\zeta^{1/2}VB + nUA} = \frac{\zeta^{-1/2}U_pC_p + nV_pD_p}{\zeta^{1/2}V_pB_p + nU_pA_p} + O\left(\frac{x^2}{n^3}\right).$$

Inserting the definitions of A_p, B_p, C_p, D_p, U_p and V_p, see [DLMF, Eq. 9.7.2 and 10.20.10-13], and Taylor expanding the above yields the result. □

Lemma 6.10.5. *Let* Γ *be the unit circle in dimension* $d = 2$ *then the operator* $\mathrm{DtN}_k - \mathrm{DtN}_0$ *admits a splitting*

$$\mathrm{DtN}_k - \mathrm{DtN}_0 = R_\Gamma + A_\Gamma$$

defined as in (6.101). *The symbol of* R_Γ *is uniformly bounded by* k.

Proof. The estimate for the symbol of R_Γ follows once we have shown

$$\left|\frac{\alpha_{2,n}k^2}{\Lambda_n} + \sigma_n\right| \lesssim k, \tag{6.106}$$

$$\left|\frac{\alpha_{1,n}k^2}{\Lambda_n} + \sigma_n\right| \lesssim k, \tag{6.107}$$

$$\left|\frac{ink^2}{\Lambda_n} - i\operatorname{sgn}(n)\sigma_n\right| \lesssim k \tag{6.108}$$

for $n > 2k$. In view of Lemma 6.10.2 we just perform the analysis for positive n. By Lemma 6.10.4 we have

$$\alpha_{i,n} = -n + \frac{\kappa_i^2}{2n} + r_i$$

with $|r_i| \lesssim \frac{k^2}{n^2} + \frac{k^4}{n^3}$. We first show

$$\left|\frac{nk^2}{\Lambda_n} - \sigma_n\right| \lesssim k.$$

Since $\sigma_n = \frac{2\mu(\lambda+2\mu)}{(\lambda+3\mu)}(|n|-1)$, let us introduce $\widehat{\sigma} = \frac{2\mu(\lambda+2\mu)}{(\lambda+3\mu)}$. Hence,

$$\left|\frac{nk^2}{\Lambda_n} - \sigma_n\right| \lesssim k \qquad \text{is equivalent to} \qquad \left|\frac{k^2}{\Lambda_n} - \widehat{\sigma}\right| \lesssim \frac{k}{n}.$$

We now use the asymptotics for $\alpha_{i,n}$ to find

$$\alpha_{1,n}\alpha_{2,n} = n^2 - \frac{\kappa_1^2+\kappa_2^2}{2} + \frac{\kappa_1^2\kappa_2^2}{4n^2} - n(r_1+r_2) + \frac{\kappa_1^2}{2n}r_2 + \frac{\kappa_2^2}{2n}r_1 + r_1r_2.$$

Note that

$$\frac{\kappa_1^2+\kappa_2^2}{2} = k^2\frac{1}{\widehat{\sigma}}$$

and therefore

$$\alpha_{1,n}\alpha_{2,n} = n^2 - k^2\frac{1}{\widehat{\sigma}} + \frac{\kappa_1^2\kappa_2^2}{4n^2} - n(r_1+r_2) + \frac{\kappa_1^2}{2n}r_2 + \frac{\kappa_2^2}{2n}r_1 + r_1r_2.$$

We calculate

$$\begin{aligned}
\frac{k^2}{\Lambda_n} - \widehat{\sigma} &= \frac{k^2}{n^2 - \alpha_{1,n}\alpha_{2,n}} - \widehat{\sigma} \\
&= \frac{k^2}{n^2 - n^2 + k^2\frac{1}{\widehat{\sigma}} - \frac{\kappa_1^2\kappa_2^2}{4n^2} + n(r_1+r_2) - \frac{\kappa_1^2}{2n}r_2 - \frac{\kappa_2^2}{2n}r_1 - r_1r_2} - \widehat{\sigma} \\
&= \frac{k^2}{\frac{k^2}{\widehat{\sigma}} - \frac{\kappa_1^2\kappa_2^2}{4n^2} + n(r_1+r_2) - \frac{\kappa_1^2}{2n}r_2 - \frac{\kappa_2^2}{2n}r_1 - r_1r_2} - \widehat{\sigma}.
\end{aligned}$$

Factorizing the denominator we find

$$\begin{aligned}\frac{k^2}{\Lambda_n} - \widehat{\sigma} &= \frac{k^2}{\frac{k^2}{\widehat{\sigma}}(1+\rho)} - \widehat{\sigma} \\ &= -\widehat{\sigma}\frac{\rho}{1+\rho}\end{aligned}$$

with ρ given by

$$\rho = \frac{\widehat{\sigma}}{k^2}\left(-\frac{\kappa_1^2\kappa_2^2}{4n^2} + n(r_1+r_2) - \frac{\kappa_1^2}{2n}r_2 - \frac{\kappa_2^2}{2n}r_1 - r_1 r_2\right).$$

One readily finds

$$|\rho| \lesssim \frac{k}{n},$$

due to the estimates for r_i and the fact that $n \geq 2k$. Hence, estimate (6.108) follows. Estimates (6.106) and (6.107) follow analogously. □

Remark 6.10.6. Analysis of the symbol of A_Γ is subject to future work. ■

A. Additional numerical results - FOSLS I

For completeness we present additional convergence plots concerning the numerical experiments corresponding to the Examples 4.4.1 and 4.4.2 considered in Chapter 4. In Figure A.1 we plot $\|e^u\|_{0,\Omega}$ employing Brezzi-Douglas-Marini elements for the problem considered in Example 4.4.1. The Figures A.2 and A.3 depicting $\|\nabla e^u\|_{0,\Omega}$ are essentially the same just one order less than $\|e^u\|_{0,\Omega}$. The numerical results for the finite regularity solution considered in Example 4.4.2 are plotted in Figure A.4 for $\|e^u\|_{0,\Omega}$, in Figure A.5 for $\|\nabla e^u\|_{0,\Omega}$ and in Figure A.6 for $\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}$.

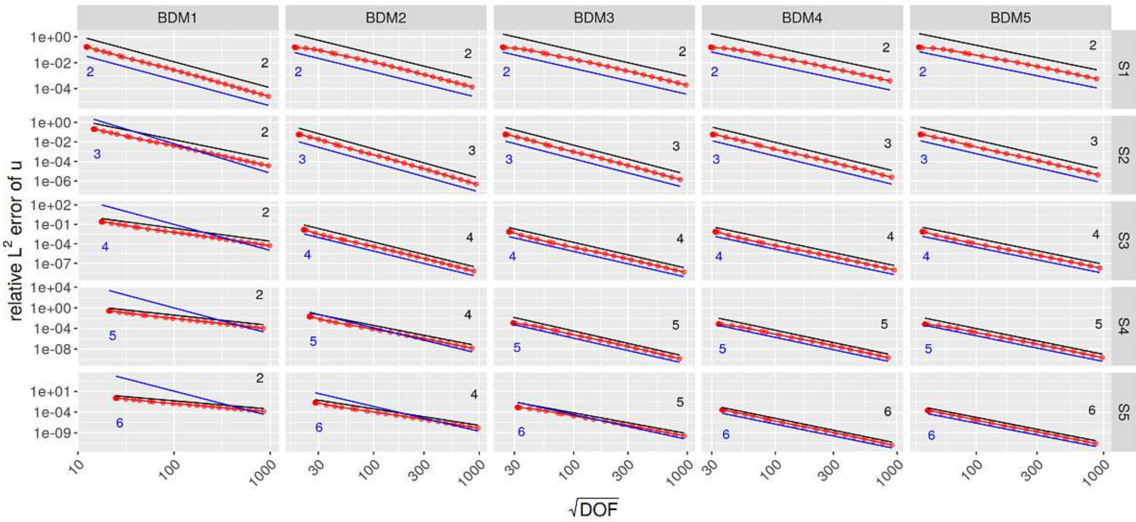

Figure A.1.: (cf. Example 4.4.1) Convergence of $\|e^u\|_{0,\Omega}$ vs. $\sqrt{\text{DOF}} \sim 1/h$ employing $\mathbf{V}^0_{p_v}(\mathcal{T}_h) = \mathbf{BDM}^0_{p_v}(\mathcal{T}_h)$.

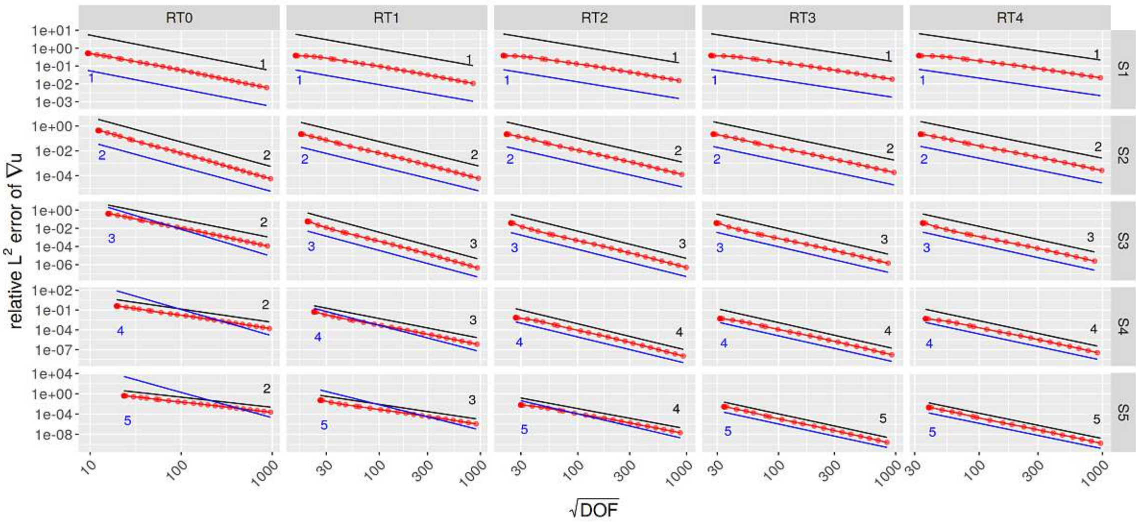

Figure A.2.: (cf. Example 4.4.1) Convergence of $\|\nabla e^u\|_{0,\Omega}$ vs. $\sqrt{\mathrm{DOF}} \sim 1/h$ employing $\mathbf{V}^0_{p_v}(\mathcal{T}_h) = \mathbf{RT}^0_{p_v-1}(\mathcal{T}_h)$.

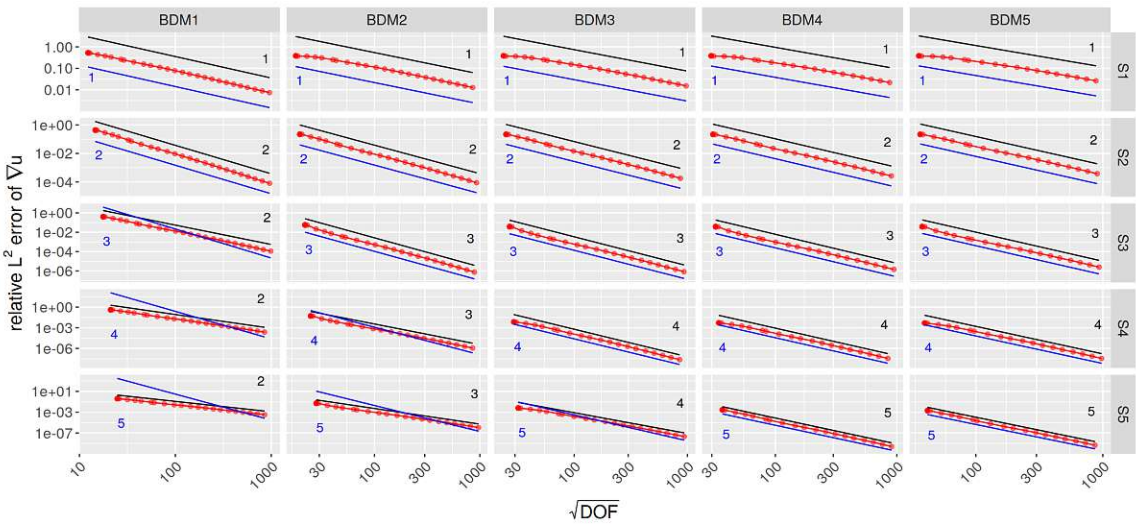

Figure A.3.: (cf. Example 4.4.1) Convergence of $\|\nabla e^u\|_{0,\Omega}$ vs. $\sqrt{\mathrm{DOF}} \sim 1/h$ employing $\mathbf{V}^0_{p_v}(\mathcal{T}_h) = \mathbf{BDM}^0_{p_v}(\mathcal{T}_h)$.

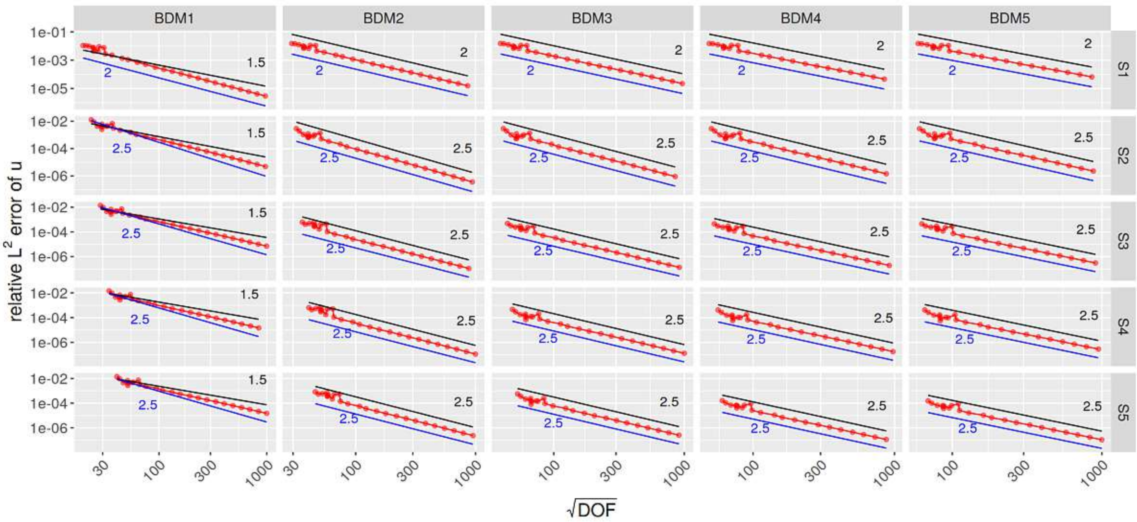

Figure A.4.: (cf. Example 4.4.2) Convergence of $\|e^u\|_{0,\Omega}$ vs. $\sqrt{\mathrm{DOF}} \sim 1/h$ employing $\mathbf{V}^0_{p_v}(\mathcal{T}_h) = \mathbf{BDM}^0_{p_v}(\mathcal{T}_h)$.

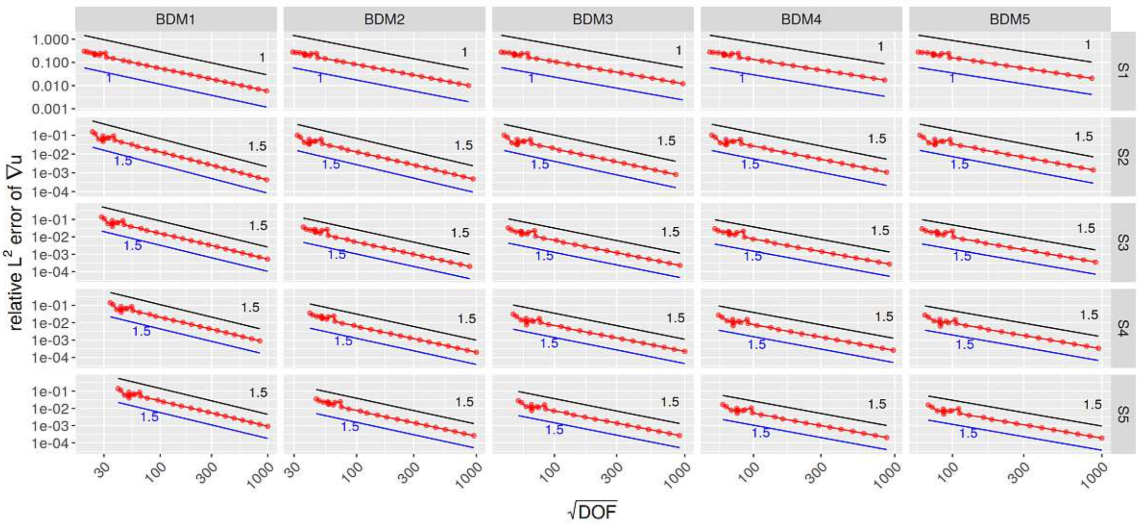

Figure A.5.: (cf. Example 4.4.2) Convergence of $\|\nabla e^u\|_{0,\Omega}$ vs. $\sqrt{\mathrm{DOF}} \sim 1/h$ employing $\mathbf{V}^0_{p_v}(\mathcal{T}_h) = \mathbf{BDM}^0_{p_v}(\mathcal{T}_h)$.

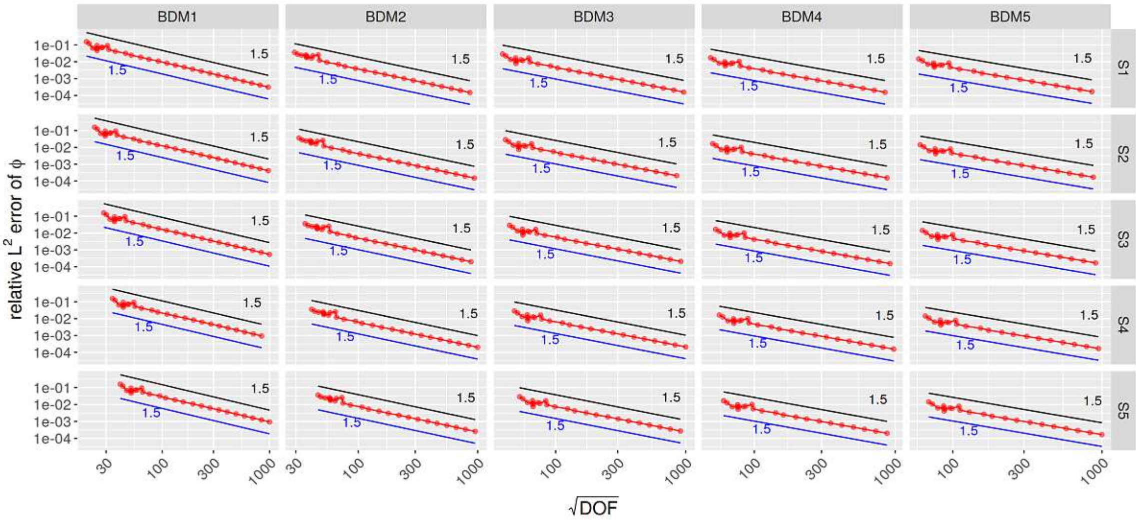

Figure A.6.: (cf. Example 4.4.2) Convergence of $\|\boldsymbol{e}^{\varphi}\|_{0,\Omega}$ vs. $\sqrt{\mathrm{DOF}} \sim 1/h$ employing $\mathbf{V}^0_{p_v}(\mathcal{T}_h) = \mathbf{BDM}^0_{p_v}(\mathcal{T}_h)$.

B. Additional numerical results - FOSLS II

For completeness we present additional convergence plots concerning the numerical experiments corresponding to the Examples 5.4.1 and 5.4.2 considered in Chapter 5.

In Figure B.1 we plot $\|e^u\|_{0,\Omega}$ employing Brezzi-Douglas-Marini elements for the problem considered in Example 5.4.1. The Figures B.2 and B.3 depicting $\|\nabla e^u\|_{0,\Omega}$ are essentially the same just one order less than $\|e^u\|_{0,\Omega}$. The numerical results for the finite regularity solution considered in Example 5.4.2 are plotted in Figure B.4 for $\|e^u\|_{0,\Omega}$, in Figure B.5 for $\|\nabla e^u\|_{0,\Omega}$, in Figure B.6 for $\|\boldsymbol{e}^{\boldsymbol{\varphi}}\|_{0,\Omega}$ and in Figure B.7 $\|\boldsymbol{e}^{\boldsymbol{\varphi}} \cdot \boldsymbol{n}\|_{0,\Gamma}$

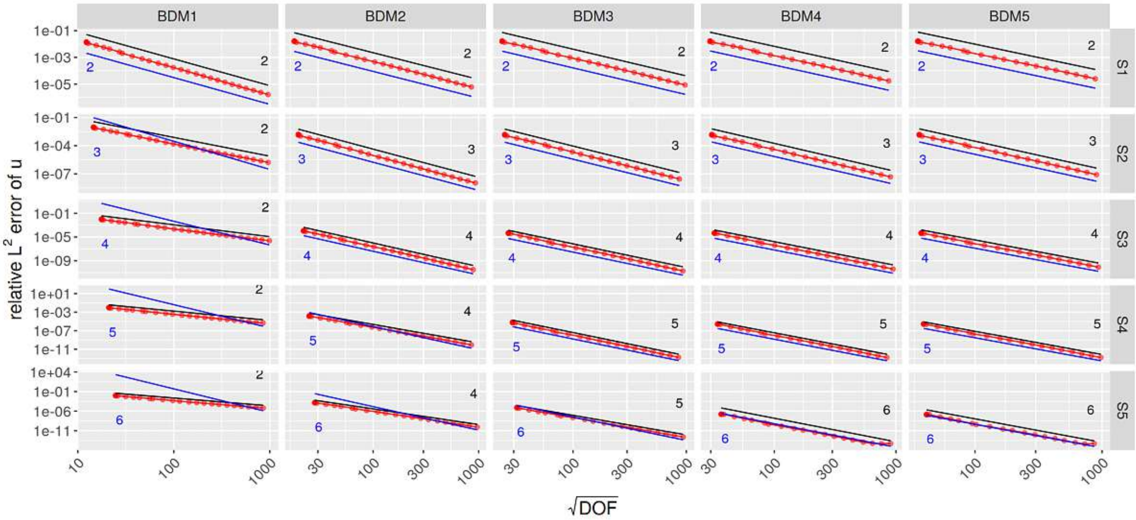

Figure B.1.: (cf. Example 5.4.1) Convergence of $\|e^u\|_{0,\Omega}$ vs. $\sqrt{\text{DOF}} \sim 1/h$ employing $\mathbf{V}_{p_v}(\mathcal{T}_h) = \mathbf{BDM}_{p_v}(\mathcal{T}_h)$.

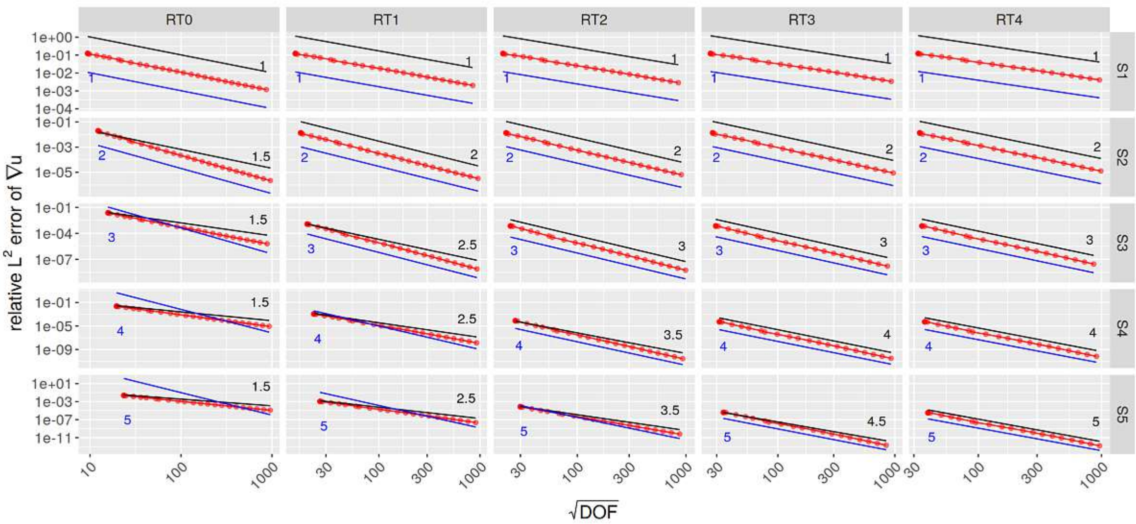

Figure B.2.: (cf. Example 5.4.1) Convergence of $\|\nabla e^u\|_{0,\Omega}$ vs. $\sqrt{\text{DOF}} \sim 1/h$ employing $\mathbf{V}_{p_v}(\mathcal{T}_h) = \mathbf{RT}_{p_v-1}(\mathcal{T}_h)$.

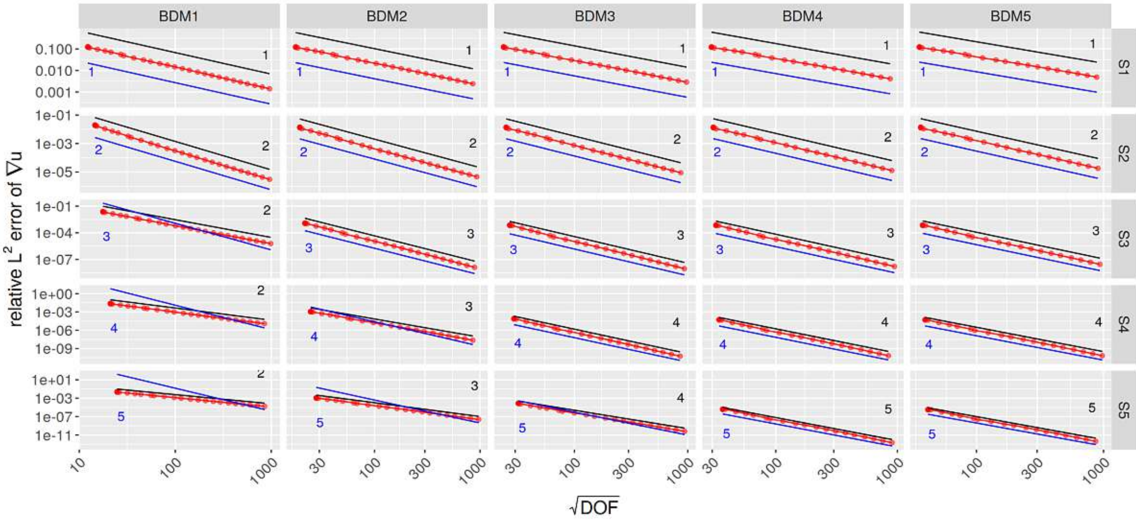

Figure B.3.: (cf. Example 5.4.1) Convergence of $\|\nabla e^u\|_{0,\Omega}$ vs. $\sqrt{\mathrm{DOF}} \sim 1/h$ employing $\mathbf{V}_{p_v}(\mathcal{T}_h) = \mathbf{BDM}_{p_v}(\mathcal{T}_h)$.

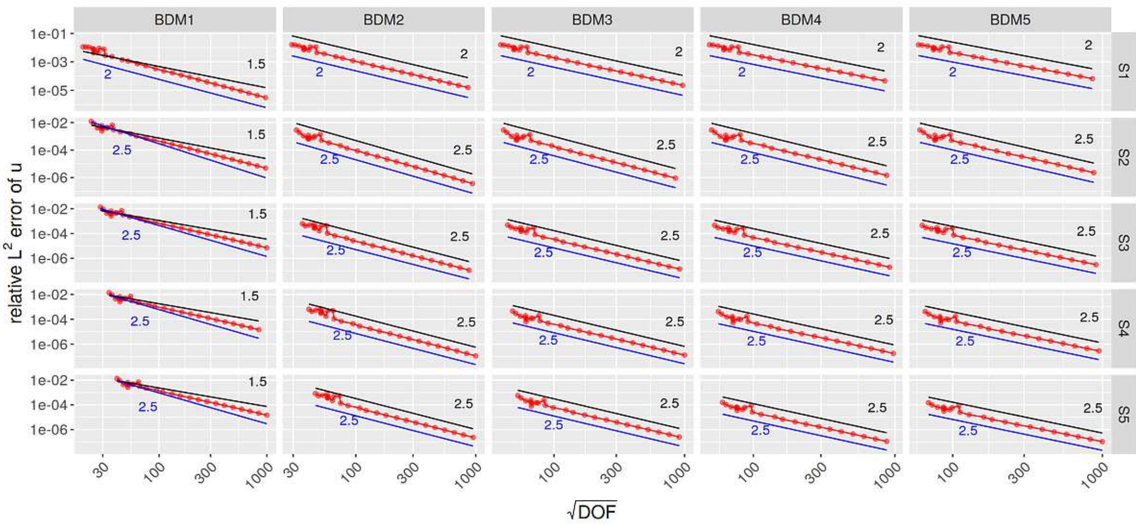

Figure B.4.: (cf. Example 5.4.2) Convergence of $\|e^u\|_{0,\Omega}$ vs. $\sqrt{\mathrm{DOF}} \sim 1/h$ employing $\mathbf{V}_{p_v}(\mathcal{T}_h) = \mathbf{BDM}_{p_v}(\mathcal{T}_h)$.

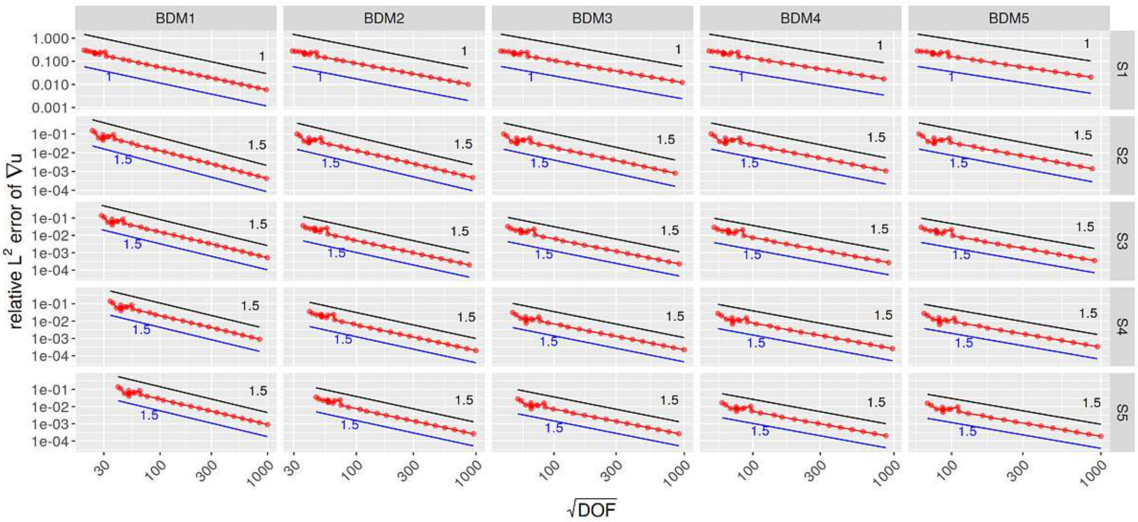

Figure B.5.: (cf. Example 5.4.2) Convergence of $\|\nabla e^u\|_{0,\Omega}$ vs. $\sqrt{\mathrm{DOF}} \sim 1/h$ employing $\mathbf{V}_{p_v}(\mathcal{T}_h) = \mathbf{BDM}_{p_v}(\mathcal{T}_h)$.

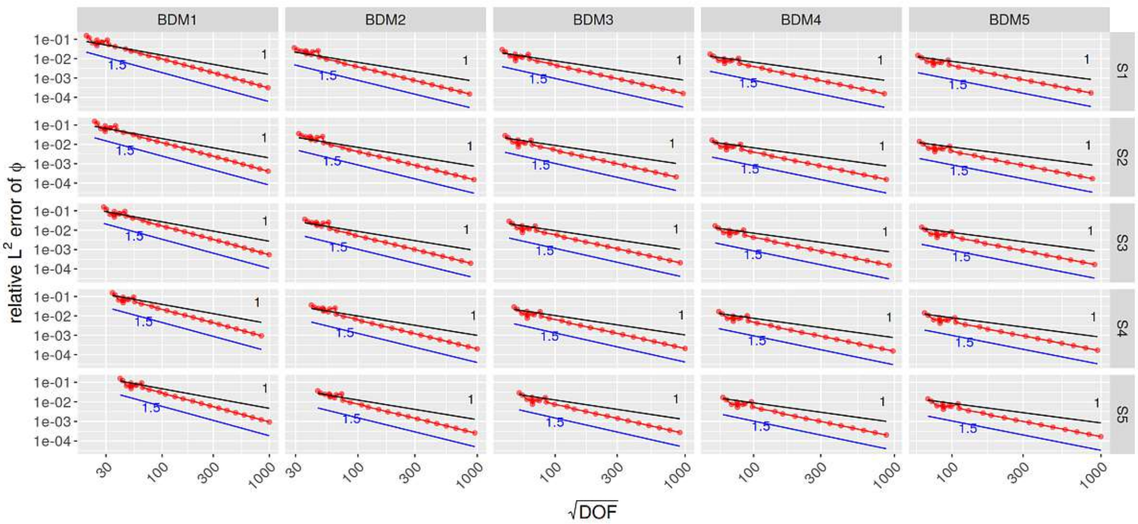

Figure B.6.: (cf. Example 5.4.2) Convergence of $\|\boldsymbol{e}^{\varphi}\|_{0,\Omega}$ vs. $\sqrt{\mathrm{DOF}} \sim 1/h$ employing $\mathbf{V}_{p_v}(\mathcal{T}_h) = \mathbf{BDM}_{p_v}(\mathcal{T}_h)$.

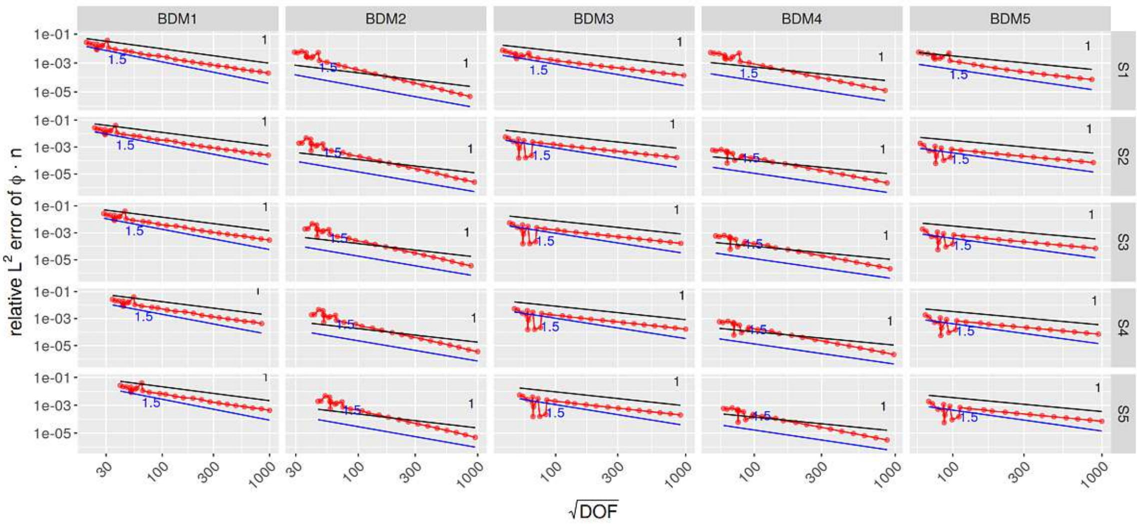

Figure B.7.: (cf. Example 5.4.2) Convergence of $\|\boldsymbol{e}^{\varphi} \cdot \boldsymbol{n}\|_{0,\Gamma}$ vs. $\sqrt{\mathrm{DOF}} \sim 1/h$ employing $\mathbf{V}_{p_v}(\mathcal{T}_h) = \mathbf{BDM}_{p_v}(\mathcal{T}_h)$.

Bibliography

[Ain04] M. Ainsworth. Discrete dispersion relation for *hp*-version finite element approximation at high wave number. *SIAM J. Numer. Anal.*, 42(2):553–575, 2004.

[AMP15] M. Aurada, J. M. Melenk, and D. Praetorius. FEM-BEM coupling for the large-body limit in micromagnetics. *J. Comput. Appl. Math.*, 281:10–31, 2015.

[BBF13] D. Boffi, F. Brezzi, and M. Fortin. *Mixed finite element methods and applications*, volume 44 of *Springer Series in Computational Mathematics*. Springer, Heidelberg, 2013.

[BCFG17] H. Barucq, T. Chaumont-Frelet, and C. Gout. Stability analysis of heterogeneous Helmholtz problems and finite element solution based on propagation media approximation. *Math. Comp.*, 86(307):2129–2157, 2017.

[BCFM21] M. Bernkopf, T. Chaumont-Frelet, and J. M. Melenk. Wavenumber-explicit stability and convergence analysis of *hp* Finite Element discretizations of Helmholtz problems in piecewise smooth media, in preparation, 2021.

[BG05] P. Bochev and M. Gunzburger. On least-squares finite element methods for the Poisson equation and their connection to the Dirichlet and Kelvin principles. *SIAM J. Numer. Anal.*, 43(1):340–362, 2005.

[BG09] P. Bochev and M. Gunzburger. *Least-squares finite element methods*, volume 166 of *Applied Mathematical Sciences*. Springer, New York, 2009.

[BGT82] A. Bayliss, M. Gunzburger, and E. Turkel. Boundary conditions for the numerical solution of elliptic equations in exterior regions. *SIAM Journal on Applied Mathematics*, 42(2):430–451, 1982.

[BM19] M. Bernkopf and J. M. Melenk. Analysis of the *hp*-Version of a First Order System Least Squares Method for the Helmholtz Equation. In Thomas Apel, Ulrich Langer, Arnd Meyer, and Olaf Steinbach, editors, *Advanced Finite Element Methods with Applications: Selected Papers from the 30th Chemnitz Finite Element Symposium 2017*, pages 57–84. Springer International Publishing, Cham, 2019.

[BM20] M. Bernkopf and J. M. Melenk. Optimal convergence rates in L^2 for a first order system least squares finite element method. Part I: homogeneous boundary conditions, submitted. arXiv e-prints arXiv:2012.12919, Dez 2020.

[BM21] M. Bernkopf and J. M. Melenk. Optimal convergence rates in L^2 for a first order system least squares finite element method. Part II: inhomogeneous boundary conditions, in preparation, 2021.

[BSW16] D. Baskin, E. A. Spence, and J. Wunsch. Sharp high-frequency estimates for the Helmholtz equation and applications to boundary integral equations. *SIAM J. Math. Anal.*, 48(1):229–267, 2016.

[CF15] T. Chaumont-Frelet. *Finite element approximation of Helmholtz problems with application to seismic wave propagation.* PhD thesis, INSA Rouen and Inria project-team Magique3D, 2015.

[CLMM94] Z. Cai, R. Lazarov, T. A. Manteuffel, and S. F. McCormick. First-order system least squares for second-order partial differential equations. I. *SIAM J. Numer. Anal.*, 31(6):1785–1799, 1994.

[CLX13] H. Chen, P. Lu, and X. Xu. A hybridizable discontinuous Galerkin method for the Helmholtz equation with high wave number. *SIAM J. Numer. Anal.*, 51(4):2166–2188, 2013.

[CM98] F. Collino and P. Monk. The perfectly matched layer in curvilinear coordinates. *SIAM J. Sci. Comput.*, 19(6):2061–2090, 1998.

[CMM97a] Z. Cai, T. A. Manteuffel, and S. F. McCormick. First-order system least squares for second-order partial differential equations. II. *SIAM J. Numer. Anal.*, 34(2):425–454, 1997.

[CMM97b] Z. Cai, T. A. Manteuffel, and S. F. McCormick. First-order system least squares for the Stokes equations, with application to linear elasticity. *SIAM J. Numer. Anal.*, 34(5):1727–1741, 1997.

[CQ17] H. Chen and W. Qiu. A first order system least squares method for the Helmholtz equation. *Journal of Computational and Applied Mathematics*, 309:145 – 162, 2017.

[CWGLS12] S. N. Chandler-Wilde, I. G. Graham, S. Langdon, and E. A. Spence. Numerical-asymptotic boundary integral methods in high-frequency acoustic scattering. *Acta Numer.*, 21:89–305, 2012.

[CWSGS20] S.N. Chandler-Wilde, E.A. Spence, A. Gibbs, and V.P. Smyshlyaev. High-frequency bounds for the Helmholtz equation under parabolic trapping and applications in numerical analysis. *SIAM J. Math. Anal.*, 52(1):845–893, 2020.

[DGMZ12] L. Demkowicz, J. Gopalakrishnan, I. Muga, and J. Zitelli. Wavenumber explicit analysis of a DPG method for the multidimensional Helmholtz equation. *Comput. Methods Appl. Mech. Engrg.*, 213/216:126–138, 2012.

[DI01] L. Demkowicz and F. Ihlenburg. Analysis of a coupled finite-infinite element method for exterior Helmholtz problems. *Numer. Math.*, 88(1):43–73, 2001.

[DLMF] Nist digital library of mathematical functions. http://dlmf.nist.gov/, Release 1.1.0 of 2020-12-15. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, and M. A. McClain, eds.

[EGSV21] A. Ern, T. Gudi, I. Smears, and M. Vohralík. Equivalence of local- and global-best approximations, a simple stable local commuting projector, and optimal *hp* approximation estimates in $H(\text{div})$. *IMA Journal of Numerical Analysis*, Mar 2021. draa103.

[EM77] B. Engquist and A. Majda. Absorbing boundary conditions for the numerical simulation of waves. *Mathematics of Computation*, 31(139):629–651, Jul 1977.

[EM79] B. Engquist and A. Majda. Radiation boundary conditions for acoustic and elastic wave calculations. *Communications on Pure and Applied Mathematics*, 32(3):313–357, 1979.

[EM12] S. Esterhazy and J. M. Melenk. On stability of discretizations of the Helmholtz equation. In *Numerical analysis of multiscale problems*, volume 83 of *Lect. Notes Comput. Sci. Eng.*, pages 285–324. Springer, Heidelberg, 2012.

[EM14] S. Esterhazy and J. M. Melenk. An analysis of discretizations of the Helmholtz equation in L^2 and in negative norms. *Comput. Math. Appl.*, 67(4):830–853, 2014.

[Era12] Christoph Erath. Coupling of the finite volume element method and the boundary element method: an a priori convergence result. *SIAM J. Numer. Anal.*, 50(2):574–594, 2012.

[Eva10] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[Fen84] K. Feng. Finite element method and natural boundary reduction. In *Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Warsaw, 1983)*, pages 1439–1453. PWN, Warsaw, 1984.

[FW09] X. Feng and H. Wu. Discontinuous Galerkin methods for the Helmholtz equation with large wave number. *SIAM J. Numer. Anal.*, 47(4):2872–2896, 2009.

[FX13a] X. Feng and Y. Xing. Absolutely stable local discontinuous Galerkin methods for the Helmholtz equation with large wave number. *Math. Comp.*, 82(283):1269–1296, 2013.

[FX13b] X. Feng and Y. Xing. Absolutely stable local discontinuous Galerkin methods for the Helmholtz equation with large wave number. *Math. Comp.*, 82(283):1269–1296, 2013.

[GPS19] I.G. Graham, O.R. Pembery, and E.A. Spence. The Helmholtz equation in heterogeneous media: a priori bounds, well-posedness and resonances. *J. Differential Equations*, 266:2869–2923, 2019.

[GS20] I. G. Graham and S. A. Sauter. Stability and finite element error analysis for the Helmholtz equation with variable coefficients. *Math. Comp.*, 89(321):105–138, 2020.

[GSW20] J. Galkowski, E.A. Spence, and J. Wunsch. Optimal constans in non-trapping resolvent estimates and applications in numerical analysis. *Pure Appl. Anal.*, 2(1):157–202, 2020.

[IB95] F. Ihlenburg and I. Babuška. Finite element solution to the Helmholtz equation with high wave number. Part I: The h-version of the FEM. *Comput. Math. Appl.*, 30:9–37, 1995.

[IB97] F. Ihlenburg and I. Babuška. Finite element solution to the Helmholtz equation with high wave number. Part II: The hp-version of the FEM. *SIAM J. Numer. Anal.*, 34:315–358, 1997.

[Ihl98] F. Ihlenburg. *Finite element analysis of acoustic scattering*, volume 132 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1998.

[Jes77] D. C. Jespersen. A least squares decomposition method for solving elliptic equations. *Math. Comp.*, 31(140):873–880, 1977.

[JS92] S. Jensen and M. Suri. On the L^2 Error for the p-version of the Finite Element Method over Polygonal Domains. *Comput. Methods Appl. Mech. Eng.*, 97(2):233–243, Jun 1992.

[LMMR00] B. Lee, T. A. Manteuffel, S. F. McCormick, and J. Ruge. First-order system least-squares for the Helmholtz equation. *SIAM J. Sci. Comput.*, 21(5):1927–1949, 2000. Iterative methods for solving systems of algebraic equations (Copper Mountain, CO, 1998).

[LSW] D. Lafontaine, E. A. Spence, and J. Wunsch. For most frequencies, strong trapping has a weak effect in frequency-domain scattering. *Communications on Pure and Applied Mathematics*, n/a(n/a).

[LSW20] D. Lafontaine, E. A. Spence, and J. Wunsch. Wavenumber-explicit convergence of the hp-FEM for the full-space heterogeneous Helmholtz equation with smooth coefficients. arXiv e-prints arXiv:2010.00585, Oct 2020.

[McL00] W. McLean. *Strongly elliptic systems and boundary integral equations*. Cambridge University Press, Cambridge, 2000.

[Mel02] J. M. Melenk. *hp-Finite Element Methods for Singular Perturbations*, volume 1796 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2002.

[Mel05] J. M. Melenk. On approximation in meshless methods. In *Frontiers of numerical analysis*, Universitext, pages 65–141. Springer, Berlin, 2005.

[Mel12] J. M. Melenk. Mapping properties of combined field Helmholtz boundary integral operators. *SIAM J. Math. Anal.*, 44(4):2599–2636, 2012.

[MMPR20] L. Mascotto, J. M. Melenk, I. Perugia, and A. Rieder. FEM-BEM mortar coupling for the Helmholtz problem in three dimensions. *Comput. Math. Appl.*, 80(11):2351–2378, 2020.

[Mon03] P. Monk. *Finite element methods for Maxwell's equations*. Numerical Mathematics and Scientific Computation. Oxford University Press, New York, 2003.

[MPS13] J. M. Melenk, A. Parsania, and S. A. Sauter. General DG-Methods for Highly Indefinite Helmholtz Problems. *Journal of Scientific Computing*, 57(3):536–581, Dec 2013.

[MR20] J. M. Melenk and C. Rojik. On commuting p-version projection-based interpolation on tetrahedra. *Math. Comp.*, 89(321):45–87, 2020.

[MS10] J. M. Melenk and S. A. Sauter. Convergence analysis for finite element discretizations of the Helmholtz equation with Dirichlet-to-Neumann boundary conditions. *Math. Comp.*, 79(272):1871–1914, 2010.

[MS11] J. M. Melenk and S. A. Sauter. Wavenumber explicit convergence analysis for Galerkin discretizations of the Helmholtz equation. *SIAM J. Numer. Anal.*, 49(3):1210–1243, 2011.

[MS21] J. M. Melenk and S. A. Sauter. Wavenumber-explicit *hp*-FEM analysis for Maxwell's equations with transparent boundary conditions. *Found. Comput. Math.*, 21(1):125–241, 2021.

[NT20] S. Nicaise and J. Tomezyk. Convergence analysis of a *hp*-finite element approximation of the time-harmonic Maxwell equations with impedance boundary conditions in domains with an analytic boundary. *Numerical Methods for Partial Differential Equations*, 36(6):1868–1903, 2020.

[Né01] Jean-Claude Nédélec. *Acoustic and electromagnetic equations*, volume 144 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2001. Integral representations for harmonic problems.

[Olv97] F. W. J. Olver. *Asymptotics and special functions*. AKP Classics. A K Peters, Ltd., Wellesley, MA, 1997. Reprint of the 1974 original [Academic Press, New York; MR0435697 (55 #8655)].

[PD17] S. Petrides and L. F. Demkowicz. An adaptive DPG method for high frequency time-harmonic wave propagation problems. *Comput. Math. Appl.*, 74(8):1999–2017, 2017.

[Roj19] C. Rojik. *p-version projection based interpolation*. PhD thesis, TU Wien, 2019. https://repositum.tuwien.at/handle/20.500.12708/17.

[Sch] J. Schöberl. Finite Element Software NETGEN/NGSolve version 6.2. `https://ngsolve.org/`.

[Sch97] J. Schöberl. NETGEN - An advancing front 2D/3D-mesh generator based on abstract rules. *Computing and Visualization in Science*, 1(1):41–52, Jul 1997.

[Sch98] C. Schwab. *p- and hp- Finite Element Methods: Theory and Applications in Solid and Fluid Mechanics*. G.H.Golub and others. Clarendon Press, 1998.

[Spe14] E. A. Spence. Wavenumber-explicit bounds in time-harmonic acoustic scattering. *SIAM J. Math. Anal.*, 46(4):2987–3024, 2014.

[Ste08] O. Steinbach. *Numerical approximation methods for elliptic boundary value problems*. Springer, New York, 2008. Finite and boundary elements, Translated from the 2003 German original.

[Yua19] X. Yuan. *Direct and inverse scattering problems for elastic waves*. PhD thesis, Purdue University, 2019.

[ZW13] Lingxue Zhu and Haijun Wu. Preasymptotic error analysis of CIP-FEM and FEM for Helmholtz equation with high wave number. Part II: *hp* version. *SIAM J. Numer. Anal.*, 51(3):1828–1852, 2013.

Maximilian Bernkopf

Hafnersteig 10/5
1010 Vienna
Austria
maximilian.bernkopf@tuwien.ac.at
maximilianbernkopf.github.io/math

Education

09.2017–current **PhD candidate**, TU Wien.
- Supervisor: Prof. Jens Markus Melenk, PhD

11.2015–06.2016 **Dipl.-Ing.(equivalent MSc) in Financial and Actuarial Mathematics**, TU Wien.
- Master thesis: "Analysis of the alpha-hypergeometric stochastic volatility model"
- Supervisor: Prof. Dr. Stefan Gerhold

07.2011–11.2015 **BSc in Financial and Actuarial Mathematics**, TU Wien.

10.2002–06.2010 **Matura (High school graduation equivalent)**, Schottengymnasium, Vienna, Austria.

Employment

09.2016–09.2017 **Data Scientist**, IntraBase, Vienna, Austria.
Focus on Statistical Learning

08.2016–06.2017 **Data Scientist**, Mantigma, Vienna, Austria.
Focus on Time Series Analysis

10.2015–06.2016 **Research Assistant**, Research Unit of Financial and Actuarial Mathematics, TU Wien.
Focus on Credit Risk Models

09.2013–12.2013 **Internship**, FMA Finanzmarktaufsicht Österreich, Vienna, Austria.
Focus on Solvency II

12.2010–08.2011 **Community Service**, Arbeiter-Samariter-Bund, Vienna, Austria.

Publications

2021 Bernkopf, M., T. Chaumont-Frelet, and J. M. Melenk (2021). *Wavenumber-explicit stability and convergence analysis of hp Finite Element discretizations of Helmholtz problems in piecewise smooth media, in preparation.*

2021 Bernkopf, M. and J. M. Melenk (2021). *Optimal convergence rates in L^2 for a first order system least squares finite element method. Part II: inhomogeneous boundary conditions, in preparation.*

2020 Bernkopf, M. and J. M. Melenk (2020). *Optimal convergence rates in L^2 for a first order system least squares finite element method. Part I: homogeneous boundary conditions, submitted.* arXiv e-prints arXiv:2012.12919. URL: https://arxiv.org/pdf/2012.12919.

2019 Bernkopf, M. and J. M. Melenk (2019). "Analysis of the hp-Version of a First Order System Least Squares Method for the Helmholtz Equation". In: *Advanced Finite Element Methods with Applications: Selected Papers from the 30th Chemnitz Finite Element Symposium 2017*. Ed. by Thomas Apel et al. Cham: Springer International Publishing, pp. 57–84. ISBN: 978-3-030-14244-5. DOI: 10.1007/978-3-030-14244-5_4.

Teaching

03.2021–current **Tutor**, Institute of Analysis and Scientific Computing, TU Wien.
Analysis 1

03.2019–07.2019 **Seminar Instructor**, Institute of Analysis and Scientific Computing, TU Wien.
Seminar on inverse problems

10.2018–02.2019 **Seminar Instructor**, Institute of Analysis and Scientific Computing, TU Wien.
Seminar on uncertainty quantification and approximation theory of neural networks

10.2017–02.2019 **Tutor**, Institute of Analysis and Scientific Computing, TU Wien.
Analysis 1 - 3

03.2016–07.2016 **Tutor**, Institute of Analysis and Scientific Computing, TU Wien.
Computer Mathematics

03.2015–07.2015 **Tutor**, Institute of Analysis and Scientific Computing, TU Wien.
Computer Mathematics