QUASI-OPTIMAL APPROXIMATION OF SURFACE BASED LAGRANGE MULTIPLIERS IN FINITE ELEMENT METHODS

J.M. MELENK* AND B. WOHLMUTH†

Abstract. We show quasi-optimal *a priori* convergence results in the L^2- and $H^{-1/2}$-norm for the approximation of surface based Lagrange multipliers such as those employed in the mortar finite element method. We improve on the estimates obtained in the standard saddle point theory, where error estimates for both the primal and dual variables are obtained simultaneously and thus only suboptimal *a priori* estimates for the dual variable are reached. For the lowest oder case (i.e., piecewise polynomials of degree $k = 1$ for the primal variable), an additional factor of $\sqrt{h}|\ln h|$ and for higher order cases (i.e., $k > 1$) an additional factor of $\sqrt{h}$ in the *a priori* bound for the dual variable can be recovered. The proof is based on the use of new estimates for the primal variable in strips of width $O(h)$ near these surfaces.

AMS subject classification: 65N30
Key words: anisotropic norms, mortar methods, local FEM error analysis, Lagrange multiplier

1. Introduction. An important goal of many finite element calculations in computational mechanics are accurate and reliable values for the flux across certain interfaces or the boundary of the domain. In non-linear contact problems, for example, the appropriate flux is related to the surface traction in the contact zone and thus plays an important role in various friction models. In numerical methods that are based on a purely primal formulation, the flux can be extracted from the numerical solution in a thin strip adjoining the interface. Hence, it is desirable to understand and quantify the discretization error in such thin strips. Alternative approaches could involve primal-dual formulations that produce the sought fluxes either as the Lagrange multiplier or through a suitable post-processing procedure. Just as in purely primal methods, a sharp *a priori* error analysis of these methods also requires good estimates for the primal variable in a thin strip near the interface. The present note, therefore, provides quasi-optimal estimates for the primal solution in thin tubular neighborhoods of interfaces. As an example of how such estimates for the primal variable can be used in the analysis of the convergence behavior of Lagrange multipliers, we study the mortar method for the Poisson problem and show quasi-optimal convergence in the Lagrange multiplier there as well. While we focus on the Poisson equation as a model problem, the techniques employed may also be used for more general elliptic systems and in other discretization schemes such as DG methods and XFEM; specifically, for the Poisson problem in a mixed setting, we refer to [17].
The results of the present paper improve on standard estimates for the Lagrange multiplier in mortar methods. These methods may be viewed as saddle point problems where the Lagrange multiplier ensures weak continuity of the primal variable on the interfaces. Then, the errors in the primal and dual variables are linked to each other, and the standard saddle point theory [9, 18] leads to *a priori* estimates for the dual variable in the $H^{-1/2}$-norm which are at most of the same order as the error bounds for the primal variable in the H^1-norm. However, the *best approximation* error for the Lagrange multiplier in the $H^{-1/2}$-norm is typically better by a factor $\sqrt{h}$ than

*(melenk@tuwien.ac.at), Institut für Analysis und Scientific Computing, Technische Universität Wien, Wiedner Hauptstrasse 8-10, A–1040 Wien, Austria

†(wohlmuth@ma.tum.de), M2 Zentrum Mathematik, Technische Universität München, Boltzmannstraße 3, D–85748 Garching

the best approximation error for the primal variable. It is this gap in the *a priori* analysis that the present paper removes (up to a logarithmic factor for the lowest order case $k = 1$). Similar observations about the mismatch between best approximation and available *a priori* estimates for the Lagrange multiplier can be made for the L^2-convergence, [8, 24]. Also for this case, our analysis recovers a factor of $\sqrt{h}|\ln h|$ for $k = 1$ and of $\sqrt{h}$ for $k > 1$. Our analysis will also cover the closely related situation of imposing Dirichlet boundary conditions weakly with the aid of a Lagrange multiplier as proposed in [19].

In view of the technical nature of the article, we formulate in Section 2 our model problem and state the two main results. The first result (Theorem 2.1) gives quasi-optimal *a priori* error estimates for the primal solution restricted to a tubular neighborhood of width $\mathcal{O}(h)$ of the domain boundary and the interfaces. The second result (Theorem 2.4) focuses on estimates for the dual variables on the interfaces. The remainder of the paper is devoted to the proofs of these results. In Section 3 we introduce anisotropic norms. Section 4 quantifies the approximation properties of nodal interpolation operators in these new anisotropic norms. Certain dual problems with locally supported data are considered in Section 5. The concluding Section 6 is devoted to the actual proofs of the two main results. Throughout the paper $0 < c$, $C < \infty$ stand for generic constants not depending on the mesh size but possibly depending on the approximation order k of the finite element spaces. For integer k, Sobolev norms on domains ω are denoted by $\|\cdot\|_{H^k(\omega)}$; the seminorm is denoted by $|\cdot|_{H^k(\omega)}$. We will also work with the Besov spaces $B^s_{2,q}(\omega)$, which are defined as interpolation spaces using the "real method" (see [20, 21] for details): for positive $s \notin \mathbb{N}$ and $q \in [1, \infty]$ we set

$$B^s_{2,q}(\omega) := (H^{\lfloor s \rfloor}(\omega), H^{\lceil s \rceil}(\omega))_{s-\lfloor s \rfloor,q} \tag{1.1}$$

2. Model problem and main results. Let $\Omega \subset \mathbb{R}^d$, $d = 2$, 3, be a convex and bounded polyhedral domain and $f \in L^2(\Omega)$. As a model problem, we consider

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

2.1. Subdomain decomposition and notation. The domain Ω is decomposed into M non-overlapping subdomains Ω_i, $i = 1, \ldots, M$, each of which is shape-regular and polyhedral. To obtain a unified notation for the two cases of interest, namely, an approximation of the Neumann values at the outer boundary if $M = 1$ and an approximation of the inner fluxes if $M > 1$, we enrich the interior skeleton $\overline{\Gamma}^{\text{int}} := \cup_{i,j=1,i\neq j}^{M} \partial\Omega_i \cap \partial\Omega_j$ by $\partial\Omega$ and set $\overline{\Gamma} := \overline{\Gamma}^{\text{int}} \cup \partial\Omega$. Moreover, we assume that the skeleton Γ can be written as a finite decomposition of N planar open faces in 3D or straight segments in 2D, i.e., $\Gamma = \cup_{l=1}^{N} \gamma_l$, which we call interfaces. For each $\gamma_l \subset \partial\Omega$ there exists a unique $\Omega_{s(l)}$ such that $\gamma_l \subset \partial\Omega_{s(l)}$. For each interior interface $\gamma_l \subset \Gamma^{\text{int}}$, $l \leq N^{\text{int}} < N$ there are two adjoining subdomains $\Omega_{s(l)}$, $\Omega_{m(l)}$, $s(l) \neq m(l)$. As is standard in the mortar context, the subdomain $\Omega_{s(l)}$ is called slave subdomain and the subdomain $\Omega_{m(l)}$ is called master subdomain. The naming originates from the fact that the discrete Lagrange multiplier will be defined with respect to the mesh on the slave side, and thus the primal solution on the slave side is dominated by the primal solution on the master side. The left part of Fig. 2.1 illustrates a decomposition of Ω into subdomains and an enumeration of the interfaces. In the example of the triangulation shown in the right part of Fig. 2.1 the bold dots indicate degrees of

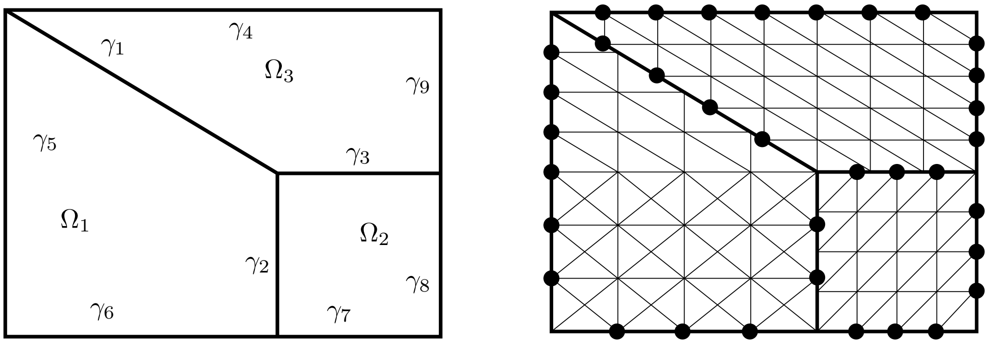

FIG. 2.1. *Left: decomposition of Ω into $M = 3$ subdomains with skeleton consisting of $N = 9$ interfaces; $N^{int} = 3$. Right: Nonmatching meshes and DOF of Lagrange multiplier space (bold dots) for $k = 1$.*

freedom for the Lagrange multiplier. Hence, the slave side for γ_1 is Ω_3, for γ_2 is Ω_1, and for γ_3 is Ω_2; thus, $s(1) = 3$, $s(2) = 1$, $s(3) = 2$, $m(1) = 1$, $m(2) = 2$, $m(3) = 3$. The special case $M = 1$ is included in our analysis. It corresponds to a single subdomain $\Omega_1 = \Omega$ with $N^{\text{int}} = 0$, $\Gamma^{\text{int}} = \emptyset$, and a conforming discretization.

2.2. Mortar finite element spaces and assumptions. For each subdomain Ω_i, let $\mathcal{T}_i$ be a quasi-uniform simplicial[1] triangulation of mesh size h. As is standard in the mortar context, these meshes are not assumed to match at the interfaces. On Ω_i, we define the standard space of order k of conforming finite elements V_i, and on γ_l we denote by $M_{s(l)}$ the Lagrange multiplier space associated with the $(d-1)$-dimensional mesh inherited from the d-dimensional triangulation of the slave side. Associated with γ_l is also the trace space $W_{s(l)} := \{v \in H_0^1(\gamma_l) \colon v = w|_{\gamma_{s(l)}}, w \in V_{s(l)}\}$. Here we restrict ourselves to formulations where $\dim W_{s(l)} = \dim M_{s(l)}$. We assume that our Lagrange multiplier space $M_{s(l)}$ satisfies the following properties:

(A1) *Stability and well-posedness of the mortar projection*: The operator $\Pi_{s(l)}$: $L^2(\gamma_l) \longrightarrow W_{s(l)}$ defined by

$$\int_{\gamma_l} \Pi_{s(l)} v \mu_h ds := \int_{\gamma_l} v \mu_h ds, \quad \forall \mu_h \in M_{s(l)}$$

is uniformly L^2-stable and, if restricted to $H_{00}^{\frac{1}{2}}(\gamma_l)$, also uniformly $H_{00}^{\frac{1}{2}}(\gamma_l)$-stable.

(A2) *Best approximation property*:

$$\inf_{\mu_h \in M_{s(l)}} \|\mu - \mu_h\|_{L^2(\gamma_l)} \le Ch^k |\mu|_{H^k(\gamma_l)}, \quad \forall \mu \in H^k(\gamma_l).$$

We note that in 2D many choices are well established, e.g., standard Lagrange multiplier spaces such as k-th order conforming functions or biorthogonal bases with the cross-point modification satisfy these two conditions, e.g., [4, 15, 25]. For results in 3D, we refer to [7, 14].

From the Assumption (A1) we directly obtain that the pairing $(M_{s(l)}, W_{s(l)})$ is uniformly inf-sup stable with respect to the $(L^2(\gamma_l), L^2(\gamma_l))$ and $(H^{-1/2}(\gamma_l), H_{00}^{1/2}(\gamma_l))$

[1] The restriction to simplicial triangulations is not essential; extensions to triangulations based on quadrilaterals/hexahedra are possible.

norm pairings. Here $H^{-1/2}(\gamma_l)$ stands for the dual norm of $H_{00}^{1/2}(\gamma_l)$. Moreover Assumptions (A1) and (A2) guarantee a best approximation property in the $H^{-1/2}(\gamma_l)$-norm, i.e.,

$$\inf_{\mu_h \in M_{s(l)}} \|\mu - \mu_h\|_{H^{-\frac{1}{2}}(\gamma_l)} \leq C h^{k+\frac{1}{2}} |\mu|_{H^k(\gamma_l)}, \quad \forall \mu \in H^k(\gamma_l).$$

For the Lagrange multiplier on γ_l we work with two different norms, the $H^{-1/2}(\gamma_l)$-norm and the $\sqrt{h}$-weighted $L^2(\gamma_l)$-norm. Correspondingly we work with the $H_{00}^{1/2}(\gamma_l)$-norm and the $\sqrt{h^{-1}}$-weighted $L^2(\gamma_l)$-norm on trace spaces. If it does not matter which one is considered, we use the abbreviated notation $(\|\cdot\|_{M^*(\gamma_l)}, \|\cdot\|_{M(\gamma_l)})$ for the $(H^{-1/2}(\gamma_l), H_{00}^{1/2}(\gamma_l))$-norm and the ($\sqrt{h}$-weighted $L^2(\gamma_l)$, $\sqrt{h^{-1}}$-weighted $L^2(\gamma_l)$)-norm.

The spaces $M_{s(l)}$ and $W_{s(l)}$, $l = 1, \ldots, N$, on the interfaces γ_l form the spaces $M_h := \prod_{l=1}^N M_{s(l)}$ and $W_h := \prod_{l=1}^N W_{s(l)}$, which we view as subspaces of $L^2(\Gamma)$ in the standard way. Then the local mortar projections $\Pi_{s(l)}$ define the global mortar projection $\Pi_h : L^2(\Gamma) \to W_h$ by

$$\Pi_h := \sum_{l=1}^N \Pi_{s(l)}. \tag{2.2}$$

We get from Assumptions (A1) and (A2)

$$\|\mu_h\|_{M^*(\Gamma)} \leq C \sup_{v_h \in W_h} \frac{\int_\Gamma \mu_h v_h ds}{\|v_h\|_{M(\Gamma)}}, \quad \forall \mu_h \in M_h, \tag{2.3a}$$

$$\inf_{\mu_h \in M_h} \|\mu - \mu_h\|_{M^*(\Gamma)} \leq C h^{k+\frac{1}{2}} |\mu|_{H^k(\Gamma)}, \quad \forall \mu \in H^k(\Gamma) := \prod_{l=1}^N H^k(\gamma_l), \tag{2.3b}$$

see, e.g., [3]. Here $\|\cdot\|^2_{M^*(\Gamma)} := \sum_{l=1}^N \|\cdot\|^2_{M^*(\gamma_l)}$ and $\|\cdot\|^2_{M(\Gamma)} := \sum_{l=1}^N \|\cdot\|^2_{M(\gamma_l)}$ stands for the broken norms on the skeleton Γ. Also higher order norms on Γ are always broken norms, e.g., $|\cdot|^2_{H^k(\Gamma)} := \sum_{l=1}^N |\cdot|^2_{H^k(\gamma_l)}$.

Based on these assumptions, we introduce now the finite element spaces of order k on Ω. Let us define the product space V_h^{-1} by

$$V_h^{-1} := \{v \in \prod_{i=1}^M V_i \colon v_{|\partial\gamma_l} = 0, N^{\text{int}} < l \leq N\}, \tag{2.4a}$$

and the constrained space V_h by

$$V_h := \{v \in V_h^{-1} \colon b(v, \mu_h) = 0, \forall \mu_h \in M_h\}, \tag{2.4b}$$

where

$$b(v, \mu) := \sum_{l=1}^N \langle \mu, [v] \rangle_{\gamma_l}.$$

Here $[\cdot]$ denotes the jump, i.e., on γ_l we have $[v] := (v|_{\Omega_{s(l)}})|_{\gamma_l} - (v|_{\Omega_{m(l)}})|_{\gamma_l}$, $1 \leq l \leq N^{\text{int}}$ and $[v] := (v|_{\Omega_{s(l)}})|_{\gamma_l}$, $N^{\text{int}} < l \leq N$; the bracket $\langle \cdot, \cdot \rangle_{\gamma_l}$ stands for the $H^{1/2}$-$(H^{1/2})'$ duality pairing.

We note that if $M = 1$ then V_h is the standard conforming finite element space of order k and if $M > 1$ then V_h is a non-conforming constrained mortar space of order k. Due to the corner/edge constraints in (2.4a), we have that the Dirichlet boundary conditions are strongly satisfied in the definition of V_h.

2.3. Primal formulation and its main result. The weak discrete primal formulation reads: Find $u_h \in V_h$ such that

$$a(u_h, v_h) = l(v_h), \quad \forall v_h \in V_h, \tag{2.5}$$

with

$$a(w, v) = \sum_{i=1}^{M} \int_{\Omega_i} \nabla u \cdot \nabla v dx, \quad l(v) = \int_{\Omega} f v dx, \quad \forall w, v \in \prod_{i=1}^{M} H^1(\Omega_i).$$

It is well-known that, under suitable regularity assumptions, u_h approximates the exact solution u in the broken H^1-norm and the L^2-norm with orders k and $k+1$, respectively, [5, 6]. These *a priori* estimates are based on the best approximation properties of the mortar space V_h and an analysis of the consistency error and are optimal.

The goal of the present section are the quasi-optimal estimates of Theorem 2.1 for the error in the L^2-norm on a strip S_h of width $2h$, which is defined as follows (cf. also the left part of Fig. 3.1):

$$S_h := \cup_{i=1}^{M} S_{h,i}, \tag{2.6a}$$
$$S_{h,i} := \{x \in \Omega_i : \operatorname{dist}(x, \partial\Omega_i) < h\}. \tag{2.6b}$$

The regularity assumption on u in the following Theorem 2.1 is formulated in terms of Besov spaces $B^s_{2,q}(\Omega_i)$, which were defined in (1.1). To help the reader gauge the regularity requirement of Theorem 2.1, we recall the fact that for each $\varepsilon > 0$ and non-integer s we have the embedding $H^{s+\varepsilon}(\Omega_i) \subset B^s_{2,1}(\Omega_i) \subset H^s(\Omega_i)$.

THEOREM 2.1. *Let Ω be convex, let the space M_h satisfy Assumptions (A1) and (A2), and let u_h be given by (2.5). If the solution u of (2.1) satisfies the additional regularity requirement $u \in \prod_{i=1}^{M} B^{k+\frac{3}{2}}_{2,1}(\Omega_i)$, then*

$$\|u - u_h\|_{L^2(S_h)} \le C h^{k+\frac{3}{2}} \left(1 + \delta_{k,1} |\ln h|\right) \|u\|_{B^{k+\frac{3}{2}}_{2,1}},$$

where $\|u\|^2_{B^{k+\frac{3}{2}}_{2,1}} := \sum_{i=1}^{M} \|u\|^2_{B^{k+\frac{3}{2}}_{2,1}(\Omega_i)}$, and $\delta_{k,1}$ denotes the usual Kronecker symbol.

Proof. The proof will be given at the end of Section 6. ◻

REMARK 2.2. Closely related results for general 2D polygons on graded meshes are obtained in [1]. While [1] and the present work are based on similar techniques from the local error analysis in FEM as described in [22, 23], significant differences lie in the regularity theory developed for the analysis. In view of applications in control problems, [1] focuses on elliptic equations with right-hand sides in L^∞ or Hölder spaces; this naturally leads to a regularity theory with solutions in weighted $W^{2,\infty}$-spaces. In contrast, our regularity theory is based on weighted H^2-spaces and the anisotropic spaces introduced in Section 3. ∎

REMARK 2.3. Theorem 2.1 (and analogously Theorem 2.4 below) assume convexity of Ω. This is done to ensure that certain auxiliary problems have H^2-regularity. ∎

2.4. Primal-dual formulation and its main result. Given the primal solution u_h, we can easily define a post-processed Lagrange multiplier $\lambda_h \in M_h$ by

$$b(w_h, \lambda_h) = l(E_h w_h) - a(u_h, E_h w_h), \qquad \forall w_h \in W_h, \tag{2.7}$$

where $E_h : W_h \to V_h^{-1}$ is defined by

$$E_h = \sum_{l=1}^{N} E_{s(l)}, \tag{2.8}$$

and $E_{s(l)} : W_{s(l)} \to V_{s(l)}$ is the extension by zero to all nodal values associated with nodes not in γ_l. We remark that the linear system (2.7) is block diagonal. These blocks are invertible square matrices since we stipulate dim $W_{s(l)} =$ dim $M_{s(l)}$ and Assumption (A1). Hence, λ_h can be computed for each γ_l separately. For special choices of $M_{s(l)}$ we can even guarantee that (2.7) results in a diagonal system.
The pair (u_h, λ_h) satisfies also the saddle-point formulation of a mortar problem and weakly imposed Dirichlet boundary conditions. Due to our definition of V_h there is no difference between strongly and weakly imposed homogeneous Dirichlet boundary conditions. Then the discrete saddle point formulation for (2.1) reads: Find $(u_h, \lambda_h) \in V_h^{-1} \times M_h$ such that

$$a(u_h, v_h) + b(v_h, \lambda_h) = l(v_h), \qquad \forall v_h \in V_h^{-1}, \tag{2.9a}$$

$$b(u_h, \mu_h) = 0, \qquad \forall \mu_h \in M_h. \tag{2.9b}$$

We note that the formulations (2.5), (2.7) on the one hand and (2.9) on the other hand are equivalent. As shown in [9], the abstract theory of saddle point problems yields under suitable regularity assumptions on λ the following *a priori* estimate:

$$\|\lambda - \lambda_h\|_{M^*(\Gamma)} \leq C \left(\left(\sum_{i=1}^{M} \|u - u_h\|^2_{H^1(\Omega_i)} \right)^{1/2} + \inf_{\mu_h \in M_h} \|\lambda - \lambda_h\|_{M^*(\Gamma)} \right), \tag{2.10}$$

where $\lambda|_{\gamma_l} := -\partial_{n_l} u|_{\Omega_{s(l)}}$, and n_l is the outer unit normal of $\partial\Omega_{s(l)} \cap \gamma_l$.
The approximation properties of V_h with respect to the broken $H^1(\Omega)$-norm yield $\mathcal{O}(h^k)$ for the first term in (2.10) whereas the best approximation property of M_h with respect to the $M^*(\Gamma)$-norm yields even $\mathcal{O}(h^{k+1/2})$ by (2.3b). Hence, the *a priori* estimate (2.10) for the dual variable is suboptimal by a factor $\sqrt{h}$. Numerical results [15, 25] show that the upper bound for the Lagrange multiplier provided by (2.10) is not sharp. Up to logarithmic factors in the case of $k = 1$, the following theorem recovers the optimal rate of convergence for the dual variable:
THEOREM 2.4. *Let Ω be convex, let the mortar space M_h satisfy Assumptions (A1) and (A2), and let (u_h, λ_h) be given by (2.9). If the solution u of (2.1) satisfies the additional regularity requirement $u \in \prod_{i=1}^{M} B_{2,1}^{k+\frac{3}{2}}(\Omega_i)$, then*

$$\|\lambda - \lambda_h\|_{L^2(\Gamma)} \leq Ch^k \left(1 + \delta_{k,1} |\ln h|\right) \|u\|_{B_{2,1}^{k+\frac{3}{2}}}.$$

If additionally $\Omega_{s(l)}$ is convex, then

$$\|\lambda - \lambda_h\|_{H^{-\frac{1}{2}}(\gamma_l)} \leq Ch^{k+\frac{1}{2}} \left(1 + \delta_{k,1} |\ln h|\right) \|u\|_{B_{2,1}^{k+\frac{3}{2}}},$$

where $\delta_{k,1}$ is the usual Kronecker symbol.
Proof. The proof will be given at the end of Section 6. □
REMARK 2.5. We recall $H^{k+3/2+\varepsilon}(\Omega_i) \subset B_{2,1}^{k+3/2}(\Omega_i)$ for all $\varepsilon > 0$. Therefore, in the 2D case of a polygon Ω and $k = 1$, the solution u of (2.1) satisfies the regularity assumption $u \in B_{2,1}^{5/2}(\Omega)$ if all interior angles of the polygon Ω are smaller than $2\pi/3$. Then, Theorem 2.4 shows that already piecewise constant approximation of the Lagrange multiplier converges with rate $\mathcal{O}(h|\ln h|)$ in the $L^2(\Gamma)$-norm. ■

2.5. A review of local error estimation in FEM and comments on key aspects of the proofs of Theorems 2.1, 2.4. In order to motivate the procedure of the present paper, we briefly review the classical local L^2-error analysis for $H^1_0(\Omega)$-conforming methods. It is based on the Aubin–Nitsche trick in the following way: For the exact solution $u \in H^1_0(\Omega)$ and its FEM approximation u_h and the bilinear form D (given by $D(u,v) = \int_\Omega \nabla u \cdot \nabla v \, dx$) one defines a dual solution w by requiring

$$(v, \chi(u - u_h))_{L^2(\Omega)} = D(v, w) \qquad \forall v \in H^1_0(\Omega);$$

here, χ is a cut-off function that reflects the interest in local L^2-estimates. Letting w_h be the FEM approximation to w, and letting $I_h u$ be an arbitrary approximation to u from the FEM space (typically, an interpolant), one can exploit Galerkin orthogonality twice:

$$(u - u_h, \chi(u - u_h))_{L^2(\Omega)} = D(u - u_h, w) = D(u - u_h, w - w_h) = D(u - I_h u, w - w_h).$$

To proceed further, a variant of a Cauchy–Schwarz inequality is employed, and the estimate takes the form

$$(u - u_h, \chi(u - u_h))_{L^2(\Omega)} \le C \|\nabla(u - I_h u)\|_X \|\nabla(w - w_h)\|_{X'},$$

with suitable norms $\|\cdot\|_X$ and $\|\cdot\|_{X'}$. They should take into account regularity properties of the solution u and the dual solution w. Well-established choices in the literature include:

1. $\|\cdot\|_X$ and $\|\cdot\|_{X'}$ are both the $L^2(\Omega)$-norm. This is the classical choice discussed in most FEM textbooks. It allows one to exploit regularity of u measured in L^2-based Sobolev norms; likewise, a regularity theory for w in classical L^2-based Sobolev norms ("shift theorems") can be brought to bear. This setting, however, does not provide an easy mechanism to incorporate extra regularity of u (e.g., $W^{k+1,\infty}$-regularity vs. H^{k+1}-regularity) or special properties of w in the case that $\operatorname{supp}\chi$ is small.
2. $\|\cdot\|_X = \|\cdot\|_{L^\infty(\Omega)}$ and $\|\cdot\|_{X'} = \|\cdot\|_{L^1(\Omega)}$. This choice is particularly useful in L^∞-estimates for the FEM and is discussed, for example, in [23, Sec. 5.4]. It allows one to exploit additional regularity of u in the form of the rather strong $W^{k+1,\infty}$-norm as opposed to the weaker H^{k+1}-norm.

In order to keep the regularity requirements for u to a minimum (while still retaining the maximal, achievable rate of convergence) we employ an anisotropic norm $\|\cdot\|_X$ that is between the L^2-norm and the L^∞-norm; we will formally introduce it in Section 3 below. Essentially, we employ fitted coordinates near the interface Γ that single out a variable that measures the distance from Γ. The norm may then be thought of as an L^2-norm in the tangential coordinates and an L^∞-norm in the normal coordinate. Correspondingly, the dual norm $\|\cdot\|_{X'}$ is based on L^2 and L^1. The estimates for $\|\nabla(u - I_h u)\|_X$ are provided in Section 4. In Section 5 we deal with $\|\nabla(w - w_h)\|_{X'}$ by rephrasing it as an estimate for the weighted L^2-norm $\|\sigma\nabla(w - w_h)\|_{L^2(\Omega)}$ (where $\sigma = \sqrt{\delta_\Gamma + h}$ is the square root of the regularized distance from Γ).

3. Anisotropic spaces and norms. A technical tool for the proof of Theorem 2.1 are anisotropic norms that reflect the anisotropic structure of tubular neighborhoods of Γ. Near Γ, one can introduce fitted coordinates that single out a special variable τ that is a measure for the distance from Γ. An integration over the tubular neighborhood can then be performed by integrating over the scalar variable τ and

$(d-1)$-dimensional manifolds that are "parallel" to Γ. In view of this observation, our anisotropic norms are based on L^2-norms over these $(d-1)$-dimensional manifolds and L^p-norms with respect to the τ-variable. The cases $p = 1$ and $p = \infty$ will be of particular interest to us. Our analysis will rest on the simple observations (3.1) and Lemma 3.1.

As is standard in the context of Lipschitz domains, we employ a localization technique to define fitted coordinate systems. As we will discuss in more detail below, the subdomains Ω_i (which are assumed to the Lipschitz) are covered by "cylinders" $\mathcal{C}_{ji} \subset \Omega_i$, $j = 1, \ldots, J_i$, and each cylinder $\mathcal{C}_{ji}$ is a region above a Lipschitz graph φ_{ji}. On each such cylinder $\mathcal{C}_{ji}$ we may then define anisotropic norms $\|\cdot\|_{L^2(\gamma_{ji};L^p)}$. The global anisotropic norm is obtained by combining the local anisotropic norms.

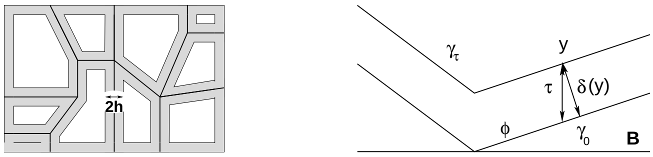

FIG. 3.1. *Left: strip S_h defined with respect to the distance function δ_Γ. Right: local shift γ_τ with respect to ϕ.*

Let us give more details concerning the anisotropic norms. To that end, let B, B' with $B \subset\subset B' \subset \mathbb{R}^{d-1}$ be two $(d-1)$-dimensional balls and ϕ be a Lipschitz continuous function on B'. For $0 < D < D'$, we define the open cylinders $\mathcal{C}$, $\mathcal{C}'$ and the open strip $S(\alpha, \beta)$ by

$$\begin{aligned} \mathcal{C} &:= \{(x, \phi(x) + \tau) \colon x \in B, \quad 0 < \tau < D\}, \\ \mathcal{C}' &:= \{(x, \phi(x) + \tau) \colon x \in B', \quad 0 < \tau < D'\}, \\ S(\alpha, \beta) &:= \{(x, \phi(x) + \tau) \colon x \in B', \quad \alpha < \tau < \beta\} \cap \mathcal{C}', \quad \alpha \leq \beta \end{aligned}$$

and the $(d-1)$-dimensional manifolds

$$\gamma_\tau := \{(x, \phi(x) + \tau) \colon x \in B\}, \qquad \tau \geq 0.$$

The Fubini–Tonelli formula for integration over $\mathcal{C}$ yields

$$\int_{\mathcal{C}} w = \int_{\tau=0}^{D} \int_{x \in B} w(x, \phi(x) + \tau)\, dx\, d\tau.$$

This motivates the definition of a measure μ^τ on γ_τ by defining the integral over γ_τ by

$$\int_{\gamma_\tau} w\, d\mu^\tau := \int_{x \in B} w(x, \phi(x) + \tau)\, dx.$$

The Lipschitz continuity of ϕ guarantees that the measure μ^τ is equivalent to the classical surface measure on the $(d-1)$-dimensional manifold γ_τ: The surface measure on γ_τ is given by $ds = (1 + \|\nabla\phi(x)\|_2^2)^{1/2}\, dx$, where $\|\cdot\|_2$ is the Euclidean norm on $\mathbb{R}^{d-1}$. Hence, the constant in the equivalence depends only on the Lipschitz constant of ϕ.

Let δ_{γ_0} be the distance function to γ_0 with respect to the Euclidean norm. Since ϕ is assumed to be Lipschitz continuous, we have $\delta_{\gamma_0}(y) \sim \tau$ uniformly in $y \in \gamma_\tau$ (see also the right picture in Fig. 3.1).

Now, we introduce anisotropic norms on $\mathcal{C}$ by

$$\|v\|_{L^2(\gamma_0;L^p)} := \left(\int_{\tau=0}^{D}\left(\int_{\gamma_\tau} v^2 d\mu^\tau\right)^{\frac{p}{2}} d\tau\right)^{\frac{1}{p}}, \quad 1 \le p < \infty,$$
$$\|v\|_{L^2(\gamma_0;L^\infty)} := \sup_{\tau\in(0,D)}\left(\int_{\gamma_\tau} v^2 d\mu^\tau\right)^{\frac{1}{2}}$$

and observe that for $p = 2$ we recover the standard $L^2(\mathcal{C})$-norm, i.e., $\|\cdot\|_{L^2(\mathcal{C})} = \|\cdot\|_{L^2(\gamma_0;L^2)}$. We recall that the definition of these norms is based on a decomposition of the d-dimensional cylinder into a one-dimensional and a $(d-1)$-dimensional subset. Roughly speaking the $L^2(\gamma_0; L^p)$-norm has a $(d-1)$-dimensional L^2-component and a one-dimensional L^p-part. We note that for all $\epsilon > 0$ we have:

$$\begin{aligned}\int_{\tau=0}^{D} f(\tau)\,d\tau &= \int_{\tau=0}^{h} f(\tau)\,d\tau + \int_{\tau=h}^{D} f(\tau)\,d\tau \\ &\le \sqrt{h}\|f\|_{L^2(0,h)} + \left(\int_{\tau=h}^{D}\tau^{-(1+2\epsilon)}\,d\tau\int_{\tau=h}^{D}\tau^{1+2\epsilon}f^2(\tau)\,d\tau\right)^{1/2} \\ &\le C\left(\sqrt{h}\|f\|_{L^2(0,h)} + h^{-\epsilon}\|\tau^{\frac{1}{2}+\epsilon}f\|_{L^2(0,D)}\right). \qquad (3.1a)\end{aligned}$$

The limiting case $\epsilon = 0$ yields

$$\int_{\tau=0}^{D} f(\tau)\,d\tau \le C\left(\sqrt{h}\|f\|_{L^2(0,h)} + \sqrt{|\ln h|}\|\sqrt{\tau}f\|_{L^2(0,D)}\right). \qquad (3.1b)$$

The following lemma shows that a Hölder type inequality holds for our newly defined anisotropic norms and that the L^2-norm on a family of strips can be bounded by a weighted L^2-norm.

Lemma 3.1. *For all v, w with well defined norms on $\mathcal{C}'$, we have*

$$\left|\int_{\mathcal{C}} vw\,dx\right| \le \|v\|_{L^2(\gamma_0;L^p)}\|w\|_{L^2(\gamma_0;L^q)}, \qquad \frac{1}{p}+\frac{1}{q} = 1. \qquad (3.2)$$

For $0 < \alpha < \beta$ and $s > 0$, we find

$$\int_{\tau=0}^{D'} \|v\|^2_{L^2(S(\alpha\tau,\beta\tau))}\,d\tau \le C(\alpha,\beta)\|\sqrt{\delta_{\gamma_0}}v\|^2_{L^2(\mathcal{C}')} \qquad (3.3)$$
$$\int_{\tau=0}^{D'} \tau\|v\|^2_{L^2(S(\tau-s,\tau+s))}\,d\tau \le Cs\|\sqrt{(s+\delta_{\gamma_0})}v\|^2_{L^2(\mathcal{C}')}, \qquad (3.4)$$

where C, $C(\alpha,\beta)$ are independent of s and v but depend on the Lipschitz constant of ϕ defining the cylinders $\mathcal{C}$, $\mathcal{C}'$.

Proof. Rewriting $\int_{\mathcal{C}} \ldots$ as $\int_{\tau=0}^{D}\int_{\gamma_\tau} \ldots$, we obtain (3.2) from the standard one-dimensional

Hölder inequality. To show (3.3), we apply Fubini–Tonelli and get

$$\begin{aligned}
\int_{\tau=0}^{D'} \|v\|^2_{L^2(S(\alpha\tau,\beta\tau))}\, d\tau &= \int_{\tau=0}^{D'} \int_{x\in B'} \int_{t=\min\{D',\alpha\tau\}}^{\min\{D',\beta\tau\}} |v(x,\phi(x)+t)|^2 \, dt\, dx\, d\tau \\
&= \int_{t=0}^{D'} \int_{x\in B'} \int_{\tau=\min\{t/\beta,D'\}}^{\min\{t/\alpha,D'\}} |v(x,\phi(x)+t)|^2 \, d\tau\, dx\, dt \\
&\le (\frac{1}{\alpha}-\frac{1}{\beta}) \int_{t=0}^{D'} \int_{x\in B'} t|v(x,\phi(x)+t)|^2 \, dx\, dt \\
&\le C(\alpha,\beta)\|\sqrt{\delta_{\gamma_0}}v\|^2_{L^2(\mathcal{C}')}.
\end{aligned}$$

Finally, for (3.4) we first note that the case $s \ge D'$ is trivial. For $s < D'$, we split the integral $\int_{\tau=0}^{D'} \dots$ into $\int_{\tau=0}^{s} \dots + \int_{\tau=s}^{D'} \dots$ and observe that the first term is straightforward since $\tau \le s \le s + \delta_{\gamma_0}$. The second term can be bounded by

$$\begin{aligned}
\int_{\tau=s}^{D'} \tau\|v\|^2_{L^2(S(\tau-s,\tau+s))}\, d\tau &= \int_{\tau=s}^{D'} \int_{x\in B'} \int_{t=\tau-s}^{\min\{\tau+s,D'\}} \tau|v(x,\phi(x)+t)|^2 \, dt\, dx\, d\tau \\
&= \int_{t=0}^{D'} \int_{x\in B'} \int_{\tau=\max\{t-s,0\}}^{\min\{t+s,D'\}} \tau|v(x,\phi(x)+t)|^2 \, d\tau\, dx\, dt \\
&\le C \int_{t=0}^{D'} \int_{x\in B'} st|v(x,\phi(x)+t)|^2 \, dx\, dt \le Cs\|\sqrt{\delta_{\gamma_0}}v\|^2_{L^2(\mathcal{C}')}.
\end{aligned}$$

□

Since each subdomain Ω_i is a Lipschitz domain, it can be represented by finitely many cylinders $\mathcal{C}_{ji}$, $j = 1,\dots,J_i$, of essentially the above form. More precisely, there exist J_i Cartesian coordinate systems (described by the variables $(\widehat{x}_{ji}, y_{ji}) \in \mathbb{R}^{d-1}\times\mathbb{R}$) and the cylinders $\mathcal{C}_{ji}$ (with corresponding balls B_{ji} and Lipschitz maps ϕ_{ji} and a fixed $0 < D_{ji} < D'_{ji}$) such that $\mathcal{C}_{ji}$ is described in these Cartesian coordinates by $\mathcal{C}_{ji} = \{(\widehat{x}_{ji}, \phi_{ji}(\widehat{x}_{ji}) + \tau) \colon \widehat{x}_{ji} \in B_{ji}, 0 < \tau < D_{ji}\}$. We note that cylinders that cover the "interior" of Ω_i can be described by the function $\phi_{ji} \equiv 0$ and that the ones associated with the "boundary" are given in terms of the Lipschitz boundary functions which are assumed to be piecewise affine. Furthermore, we have $\Omega_i = \cup_{j=1}^{J_i}\mathcal{C}_{ji}$ and note that some $\mathcal{C}_{ji}$ do overlap. Combining the contributions from the cylinders, we can then define broken anisotropic norms on Ω by

$$\|v\|^p_{L^2(\Gamma;L^p)} := \sum_{i=1}^{M}\sum_{j=1}^{J_i} \|v\|^p_{L^2(\gamma_{ji},L^p)}, \qquad 1 \le p < \infty, \tag{3.5a}$$

$$\|v\|_{L^2(\Gamma;L^\infty)} := \max_{i=1,\dots,M}\ \max_{j=1,\dots,J_i} \|v\|_{L^2(\gamma_{ji},L^\infty)}, \tag{3.5b}$$

where we describe, in local coordinates, $\gamma_{ji} := \{(\widehat{x}_{ji}, \phi_{ji}(\widehat{x}_{ji})) \colon \widehat{x}_{ji} \in B_{ji}\}$. We note that in the "interior" of Ω_i the choice of ϕ_{ji} is rather arbitrary, and thus is the anisotropy. However this is in no conflict with our theoretical results since the proofs exploit the anisotropic structure of the norm only close to the boundary of Ω_i.
Following the lines of the proof of Lemma 3.4 and using the definition (3.5), we obtain the global Hölder-type inequality for our anisotropic norms

$$\left|\int_\Omega vw\, dx\right| \le \|v\|_{L^2(\Gamma;L^p)}\|w\|_{L^2(\Gamma;L^q)} \qquad \text{for all } p,q \in [1,\infty] \text{ with } \frac{1}{p}+\frac{1}{q} = 1. \tag{3.6}$$

REMARK 3.2. The Hölder type inequality (3.6) shows $\|u\|_{L^2(\Omega)} \le \|u\|_{L^2(\Gamma;L^2)}$. For the converse estimate, we note that

$$\|u\|^2_{L^2(\Gamma;L^2)} \le C_{overlap}\|u\|^2_{L^2(\Omega)}, \qquad C_{overlap} := \sup_{x\in\Omega} \operatorname{card}\{(j,i)\colon x \in \mathcal{C}_{ji}\}.$$

In other words, $C_{overlap}$ measures the amount of overlap of the cylinders $\mathcal{C}_{ji}$. Using a (non-negative) partition of unity $(\psi_{ji})_{ji}$ subordinate to $\mathcal{C}_{ji}$, one can show that $\|u\|^2_{L^2(\Omega_i)} = \sum_j \|\sqrt{\psi_{ji}}u\|^2_{L^2(\gamma_{ji};L^2)}$, and we can recover the standard L^2-norm. Since we are not interested in the constants in our *a priori* bounds, we do not use a partition of unity. ■

4. The mortar interpolation operator in an anisotropic norm. Our analysis requires approximation results in the previously introduced anisotropic norms, in particular in $\|\cdot\|_{L^2(\Gamma;L^\infty)}$. Since we consider a mortar setting, we study a modification of the subdomain-wise defined operator I_h^k, which we call P_h. Its approximation properties are the topic of Lemma 4.1.

We recall the following approximation and stability results for functions w that are sufficiently smooth on each $T \in \mathcal{T}_h$:

$$\|w - I_h^k w\|_{H^\ell(T)} \le Ch^{m-\ell}\|\nabla^m w\|_{L^2(T)}, \qquad \ell \in \{0,1\}, m \in \{2,\dots,k+1\}. \tag{4.1}$$

If $M > 1$, a crucial step in the proof of the optimal *a priori* estimate in the energy norm is to establish best approximation properties of the constrained space V_h defined by (2.4b). This can be done with the aid of the operator $P_h : C(\overline{\Omega}) \cap H_0^1(\Omega) \to V_h$ given by

$$P_h v := I_h^k v - E_h\Pi_h[I_h^k v], \tag{4.2}$$

where the mortar projection Π_h is defined in (2.2), and E_h is given in (2.8). We note that the operator P_h has best approximation properties of order k in the broken H^1-norm and of order $k+1$ in the L^2-norm on Ω. On Γ, we have for $v \in H_0^1(\Omega)$

$$\|[v - P_h v]\|_{L^2(\gamma_l)} \le ch^{k+1}|v|_{H^{k+1}(\gamma_l)} \quad \text{if } v_{|\gamma_l} \in H^{k+1}(\gamma_l) \quad \forall l \tag{4.3a}$$

Moreover, due to its construction, we get the more local estimate

$$|u - P_h u|_{H^1(S_h^{\mathrm{s}})} + \frac{1}{h}\|u - P_h u\|_{L^2(S_h^{\mathrm{s}})} \le Ch^k|u|_{H^{k+1}(S_h)}, \tag{4.3b}$$

where $S_h^{\mathrm{s}} := \cup_{l=1}^N S_{h;l}^{\mathrm{s}} \subset S_h$. Here $S_{h;l}^{\mathrm{s}}$ is the union of all elements $T \in \mathcal{T}_{s(l)}$ such that $\overline{T} \cap \gamma_l \ne \emptyset$. The estimates (4.3a)–(4.3b) are standard and result from the L^2-stability of the mortar projection and the local definition of the Lagrange interpolation operator. Additionally, we have to establish order $k+1$ approximation properties of P_h with respect to the $L^2(\Gamma;L^\infty)$-norm.

LEMMA 4.1. *Let the space M_h satisfy Assumptions (A1), (A2). Then there exists a constant $C > 0$ independent of the mesh size but depending on the subdomain decomposition and the approximation order k such that*

$$\|\nabla(v - P_h v)\|_{L^2(\Gamma;L^\infty)} \le Ch^k\|v\|_{B_{2,1}^{k+\frac{3}{2}}}, \quad \forall v \in \prod_{i=1}^M B_{2,1}^{k+\frac{3}{2}}(\Omega_i) \cap H_0^1(\Omega).$$

Proof. We restrict ourselves to one single pair of cylinders $\mathcal{C} \subset \mathcal{C}' \subset \Omega_i$. Then the definition of the $L^2(\gamma_0; L^\infty)$-norm shows that we have to consider the L^2-norm on γ_τ with $\tau > 0$ in more detail. The definition of P_h implies

$$\|\nabla(v - P_h v)\|_{L^2(\gamma_\tau)} \leq \|\nabla(v - I_h^k v)\|_{L^2(\gamma_\tau)} + \|\nabla E_h \Pi_h[I_h^k v]\|_{L^2(\gamma_\tau)}. \tag{4.4}$$

We start with the first term in (4.4). Applying the trace inequality for elements T we have

$$\|\nabla(v - I_h^k v)\|^2_{L^2(\gamma_\tau \cap \overline{T})} \leq C \left(\frac{1}{h} \|\nabla(v - I_h^k v)\|^2_{L^2(T)} + h \|\nabla^2(v - I_h^k v)\|^2_{L^2(T)} \right). \tag{4.5}$$

Introducing the subdomain

$$\overline{S}_h(\tau) := \cup_{T \in \mathcal{I}_i} \overline{T}, \quad \mathcal{I}_i := \{T \in \mathcal{T}_i \colon \gamma_\tau \cap \overline{T} \neq \emptyset\} \tag{4.6}$$

and observing $S_h(\tau) \subset S(\tau - h, \tau + h)$ (here, we use that h is sufficiently small) we get in view of (4.5)

$$\begin{aligned}
\|\nabla(v - I_h^k v)\|^2_{L^2(\gamma_\tau)} &= \sum_{T \in \mathcal{I}_i} \|\nabla(v - I_h^k v)\|^2_{L^2(\gamma_\tau \cap \overline{T})} \\
&\leq C \left(\frac{1}{h} \|\nabla(v - I_h^k v)\|^2_{L^2(S_h(\tau))} + h \sum_{T \in \mathcal{I}_i} \|\nabla^2(v - I_h^k v)\|^2_{L^2(T)} \right) \\
&\leq C h^{2k-1} |v|^2_{H^{k+1}(S_h(\tau))} \leq C h^{2k} \|v\|^2_{B^{k+\frac{3}{2}}_{2,1}};
\end{aligned}$$

the last inequality results from a 1D Sobolev embedding of L^∞ in $B^{1/2}_{2,1}$, (see [16, Lemma 2.1] for details), and the fact that by (4.6) the width of $S_h(\tau)$ is $\mathcal{O}(h)$. We now turn to the second term in (4.4). For $\tau \geq h$, this term is equal to zero. For $\tau < h$, we exploit that $E_h \Pi_h[I_h^k v]$ is a discrete function (hence, permitting inverse estimates) and, in fact by definition of E_h, the trivial extenion of $\Pi_h[I_h^k v]$. Thus,

$$\begin{aligned}
\|\nabla E_h \Pi_h[I_h^k v]\|^2_{L^2(\gamma_\tau)} &\leq C \frac{1}{h} \|\nabla E_h \Pi_h[I_h^k v]\|^2_{L^2(S_h(\tau))} \leq C \frac{1}{h^3} \|E_h \Pi_h[I_h^k v]\|^2_{L^2(S_h(\tau))} \\
&\leq C \frac{1}{h^2} \|\Pi_h[I_h^k v]\|^2_{L^2(\Gamma)} \leq C \frac{1}{h^2} \|[I_h^k v]\|^2_{L^2(\Gamma)} \leq C \frac{1}{h^2} \|[v - I_h^k v]\|^2_{L^2(\Gamma)} \\
&\leq C \frac{1}{h^2} h^{2(k+1)} |v|^2_{H^{k+1}(\Gamma)} \leq C h^{2k} \|v\|^2_{B^{k+\frac{3}{2}}_{2,1}},
\end{aligned}$$

where the last bound follows from the fact that the trace map is a continuous operator from $B^{1/2}_{2,1}(\Omega_i)$ onto $L^2(\partial\Omega_i)$, [21, Thm. 2.9.3]. The global result is then obtained from the local result by noting that the number of required cylinders is finite and independent of the mesh size. ∎

5. Bounds for dual problems with locally supported data. As discussed in Section 2.5, a classical tool to obtain L^2-estimates in finite element methods is the Aubin–Nitsche trick, which exploits properties of a dual problem with the finite element error as the right-hand side data. Here, we consider two types of dual problems. The first one, studied in Section 5.1, is associated with the global domain Ω and Dirichlet boundary conditions and ultimately leads to Theorem 2.1 and the $L^2(\Gamma)$-estimate in Theorem 2.4. The second dual problem is concerned with a subdomain $\Omega_{s(l)}$ and Neumann boundary data and is used in the proof of the $H^{-1/2}(\gamma_l)$-estimate of Theorem 2.4. In both cases we are particularly interested in right-hand sides that are supported by strips of width $\mathcal{O}(h)$.

5.1. Global dual problem with Dirichlet data. We consider

$$-\Delta w = v \in L^2(\Omega) \quad \text{in } \Omega, \qquad w = 0 \quad \text{on } \partial\Omega \tag{5.1}$$

with locally supported data, i.e., $\operatorname{supp} v \subset \overline{S_h}$ (see (2.6a) for a definition of S_h). We introduce the solution operator $T^D : H^{-1}(\Omega) \to H_0^1(\Omega)$ and assume that the following shift theorem holds:

$$T^D : H^{-1+s_0}(\Omega) \to H^{1+s_0}(\Omega) \cap H_0^1(\Omega) \text{ is linear, bounded for some } s_0 > 1/2 \tag{5.2}$$

Remark 5.1. For convex domains, it is well-known that $s_0 = 1$ is admissible, [13]. ∎

5.1.1. Regularity. We start with a regularity result which is similar to [2], where the 2D case is studied.

Lemma 5.2. *Let $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$ satisfy (5.2). Then the solution operator T^D for the problem (5.1) maps*

$$\left(B_{2,1}^{1/2}(\Omega)\right)' \to B_{2,\infty}^{3/2}(\Omega) \cap H_0^1(\Omega), \tag{5.3a}$$

and moreover, for $\operatorname{supp} v \subset \overline{S_h}$, we have

$$\|T^D v\|_{B_{2,\infty}^{3/2}(\Omega)} \leq C\sqrt{h}\|v\|_{L^2(\Omega)}. \tag{5.3b}$$

Proof. We start with the proof of (5.3a). By interpolation and assumption (5.2), we have for $0 < \theta < 1$:

$$T^D : (H^{-1+s_0}(\Omega), H^{-1}(\Omega))_{\theta,\infty} \to (H^{1+s_0}(\Omega), H^1(\Omega))_{\theta,\infty} = B_{2,\infty}^{1+s_0(1-\theta)}(\Omega).$$

For $s_0 > 1/2$, we get that $\theta := 1 - (2s_0)^{-1} \in (0,1)$. By [21, Thm. 1.11.2] or [20, Lemma 41.3], we then have

$$B_{2,\infty}^{3/2}(\Omega) = B_{2,\infty}^{1+s_0(1-\theta)}(\Omega) = (H^{1+s_0}(\Omega), H^1(\Omega))_{\theta,\infty}.$$

In view of the continuous embeddings $H_0^1(\Omega) \subset H^1(\Omega)$ and $H_0^{1-s_0}(\Omega) \subset H^{1-s_0}(\Omega)$, we find

$$\begin{aligned}
\left(B_{2,1}^{1/2}(\Omega)\right)' &= \left(B_{2,1}^{1-s_0(1-\theta)}(\Omega)\right)' = ((H^{1-s_0}(\Omega), H^1(\Omega))_{\theta,1})' \\
&\subset ((H_0^{1-s_0}(\Omega), H_0^1(\Omega))_{\theta,1})' = ((H_0^{1-s_0}(\Omega))', (H_0^1(\Omega))')_{\theta,\infty} \\
&= (H^{s_0-1}(\Omega), H^{-1}(\Omega))_{\theta,\infty}.
\end{aligned}$$

This shows (5.3a). To see (5.3b), let $v \in L^2(\Omega)$ with $\operatorname{supp} v \subset \overline{S_h}$. Then (5.3a) in combination with [16, Lemma 2.1] shows

$$\begin{aligned}
\|T^D v\|_{B_{2,\infty}^{3/2}(\Omega)} &\leq C\|v\|_{(B_{2,1}^{1/2}(\Omega))'} = C \sup_{z \in B_{2,1}^{1/2}(\Omega)} \frac{(v,z)_{L^2(\Omega)}}{\|z\|_{B_{2,1}^{1/2}(\Omega)}} \\
&\leq C\|v\|_{L^2(\Omega)} \sup_{z \in B_{2,1}^{1/2}(\Omega)} \frac{\|z\|_{L^2(S_h)}}{\|z\|_{B_{2,1}^{1/2}(\Omega)}} \leq C\sqrt{h}\|v\|_{L^2(\Omega)}.
\end{aligned}$$

□

REMARK 5.3. For Lipschitz domains Ω and $\operatorname{supp} v \subset \overline{S_h}$, we obtain trivially (without assuming (5.2)) in view of [16, Lemma 2.1] the weaker bound

$$\|T^D v\|_{H^1(\Omega)} \le C \sup_{z \in H_0^1(\Omega)} \frac{(v,z)_{L^2(\Omega)}}{\|z\|_{H^1(\Omega)}} \le C \sup_{z \in H_0^1(\Omega)} \frac{\|v\|_{L^2(\Omega)} \|z\|_{L^2(S_h)}}{\|z\|_{H^1(\Omega)}} \le C\sqrt{h}\|v\|_{L^2(\Omega)}.$$

However, this weaker bound is insufficient for the optimal rate of Theorem 2.1. ∎

LEMMA 5.4. *Let the bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$, be a polygon ($d=2$) or a polyhedron ($d=3$). Assume that $w \in B^{3/2}_{2,\infty}(\Omega)$ is the solution of (5.1) and that* $\operatorname{supp} v \subset \overline{S_h}$*, then there exists constants C, $\tilde{c} > 0$ independent of v such that*

$$\|\sqrt{\delta_\Gamma}\nabla^2 w\|_{L^2(\Omega\setminus S_{\tilde{c}h})} \le C\sqrt{|\ln h|}\|w\|_{B^{3/2}_{2,\infty}(\Omega)}, \tag{5.4}$$

where δ_Γ is the distance function to Γ. Additionally, by taking $\tilde{c} > 0$ sufficiently large, there exists $c' > 0$ such that for every $\alpha > 0$

$$\|\delta_\Gamma^\alpha \nabla^3 w\|_{L^2(\Omega\setminus S_{\tilde{c}h})} \le C\|\delta_\Gamma^{\alpha-1}\nabla^2 w\|_{L^2(\Omega\setminus S_{c'h})}. \tag{5.5}$$

Proof. Step 1: Let $\mathcal{C} \subset \mathcal{C}' \subset \Omega_i$ be a pair of cylinders as described in Section 3 such that $\{(x,\phi(x))\colon x \in B'\}$ is a part of Γ. We assume furthermore that on $\mathcal{C}'$ we have

$$C_1 t \le \operatorname{dist}(z,\Gamma) \le C_2 t \qquad \forall z = (x, \phi(x)+t) \in \mathcal{C}'.$$

Let $\mathcal{C}''$ be a second cylinder of the form $\mathcal{C}'' = \{(x,\phi(x)+t)\colon 0 < t < D'', x \in B''\}$ where $B \subset\subset B'' \subset\subset B'$ and $D < D'' < D'$. Let $\chi \in C^\infty(\mathbb{R}^d)$ be such that $\chi|_{\mathcal{C}''} \equiv 1$ and $\chi|_{\Omega_i\setminus\mathcal{C}'} \equiv 0$. To simplify the notation, we assume that the function w is given in a coordinate system commensurate with the coordinate system describing the cylinders $\mathcal{C}$, $\mathcal{C}'$, viz., w evaluated at a point $(x,\phi(x)+t) \in \mathcal{C}'$ is given by $w(x,\phi(x)+t)$. A translation in the last variable defines the function $\widetilde{w}$ by $\widetilde{w}(x,\phi(x)+t) := w(x,\phi(x)+t+2h/C_1)$. We note

$$-\Delta\widetilde{w} = 0 \qquad \text{on } \{(x,\phi(x)+t)\colon x \in B', -h/C_1 < t < D' - 2h/C_1\} \tag{5.6}$$

if h is sufficiently small. In this step, we show

$$\|\widetilde{w}\|_{H^{3/2}(\mathcal{C}'')} \le C\sqrt{|\ln h|}\|w\|_{B^{3/2}_{2,\infty}(\Omega_i)}. \tag{5.7}$$

Using the characterization of $H^{3/2}(\mathcal{C}')$ in terms of the K-functional, we write (cf. also [10, p.193, eqn. (7.4)])

$$\begin{aligned}\|\chi\widetilde{w}\|^2_{H^{3/2}(\mathcal{C}')} &= \int_{t=0}^1 \left(t^{-1/2}K(t,\chi\widetilde{w})\right)^2 \frac{dt}{t} \\ &= \int_{t=0}^\varepsilon \left(t^{-1/2}K(t,\chi\widetilde{w})\right)^2 \frac{dt}{t} + \int_{t=\varepsilon}^1 \left(t^{-1/2}K(t,\chi\widetilde{w})\right)^2 \frac{dt}{t}. \end{aligned} \tag{5.8}$$

The second integral in (5.8) can be estimated by

$$\int_{t=\varepsilon}^1 \left(t^{-1/2}K(t,\chi\widetilde{w})\right)^2 \frac{dt}{t} \le \int_{t=\varepsilon}^1 \frac{dt}{t} \sup_{t>0} \left(t^{-1/2}K(t,\chi\widetilde{w})\right)^2 \le \ln\varepsilon \|\chi\widetilde{w}\|^2_{B^{3/2}_{2,\infty}(\mathcal{C}')}.$$

For the first integral in (5.8) we employ interior regularity estimates for solutions of the (homogeneous) Laplace equation. Specifically, (5.6) and interior regularity (see, e.g., [12, Thm. 8.8]) give

$$\|\chi\widetilde{w}\|_{H^2(\mathcal{C}')} \leq Ch^{-1}\|w\|_{H^1(\mathcal{C}')}.$$

Hence, estimating $K(t,\chi\widetilde{w}) = \inf_{v\in H^2}\|\chi\widetilde{w}-v\|_{H^1(\mathcal{C}')}+t\|v\|_{H^2(\mathcal{C}')} \leq t\|\chi\widetilde{w}\|_{H^2(\mathcal{C}')}$, we obtain

$$\int_{t=0}^{\varepsilon} t^{-2}K^2(t,\chi\widetilde{w})\,dt \leq \varepsilon\|\chi\widetilde{w}\|^2_{H^2(\mathcal{C}')} \leq C\varepsilon h^{-2}\|w\|^2_{H^1(\mathcal{C}')}.$$

We conclude with

$$\begin{aligned}\|\chi\widetilde{w}\|^2_{H^{3/2}(\mathcal{C}')} &\leq C\left[\varepsilon h^{-2}\|w\|^2_{H^1(\mathcal{C}')} + \ln\varepsilon\|\chi\widetilde{w}\|^2_{B^{3/2}_{2,\infty}(\mathcal{C}')}\right]\\ &\leq C\left[\varepsilon h^{-2}+\ln\varepsilon\right]\|w\|^2_{B^{3/2}_{2,\infty}(\mathcal{C}')} \leq C\left[\varepsilon h^{-2}+\ln\varepsilon\right]\|w\|^2_{B^{3/2}_{2,\infty}(\Omega_i)}\end{aligned}$$

where, in the penultimate last step we have employed that multiplication by a smooth function and translation are bounded operations on Sobolev (and therefore also Besov) spaces. Selecting $\varepsilon = h^2$ shows $\|\chi\widetilde{w}\|_{H^{3/2}(\mathcal{C}')} \leq C\sqrt{|\ln h|}\|w\|_{B^{3/2}_{2,\infty}(\Omega_i)}$ from which we get (5.7) in view of the support properties of χ.

Step 2: Let z solve $-\Delta z = 0$ on a ball $B_{1+\rho}$ of radius $1+\rho$ for a fixed $\rho > 0$. Then standard interior regularity (see, e.g., [12, Thm. 8.8]) gives $\|\nabla^2 z\|_{L^2(B_1)} \leq C\|z\|_{H^1(B_{1+\rho})} \leq C\|z\|_{H^{3/2}(B_{1+\rho})}$. Since linear polynomials are harmonic, we even get $\|\nabla^2 z\|_{L^2(B_1)} \leq C|z|_{H^{3/2}(B_{1+\rho})}$ with the $H^{3/2}$-seminorm on the right-hand side. For the remainder of the argument, we use the Aronstein–Slobodeckij characterization of the $H^{3/2}$-seminorm. In view of (5.6) we get with a scaling argument for balls B_r such that $B_{(1+\rho)r} \subset \mathcal{C}'$

$$\|\nabla^2\widetilde{w}\|_{L^2(B_r)} \leq Cr^{-1/2}|\widetilde{w}|_{H^{3/2}(B_{(1+\rho)r})}.$$

Using, for example, the Besicovitch covering theorem [11], we can cover $\mathcal{C}$ by overlapping balls $B_{r_i}(x_i)$ with centers x_i and radii $r_i \sim (h+\delta_\Gamma(x_i))$ such that the stretched balls $B_{r_i(1+\rho)}(x_i)$ have a finite overlap property (see Lemma A.1 for details). A covering argument then shows

$$\|\sqrt{\delta_{\gamma_0}+h}\,\nabla^2\widetilde{w}\|_{L^2(\mathcal{C})} \leq C|\widetilde{w}|_{H^{3/2}(\mathcal{C}')} \leq C\sqrt{|\ln h|}\|w\|_{B^{3/2}_{2,\infty}(\Omega)}. \tag{5.9}$$

Since $\widetilde{w}$ is obtained by a translation of w, we arrive at

$$\|\sqrt{\delta_{\gamma_0}}\,\nabla^2 w\|_{L^2(\mathcal{C}_h)} \leq C\sqrt{|\ln h|}\|w\|_{B^{3/2}_{2,\infty}(\Omega)},$$

where $\mathcal{C}_h := \{(x,\phi(x)+t)\colon 2h/C_1 \leq t \leq D-2h/C_1\}$.

Finally, covering Ω_i by such cylinders allows us to conclude the proof of (5.4). The estimate (5.5) merely expresses interior regularity for harmonic functions. ∎

5.1.2. FEM a priori estimates. An important ingredient of the proof of Theorem 2.1, which provides estimates for $\|u-u_h\|_{L^2(S_h)}$, is the analysis of $\|\nabla(w-w_h)\|_{L^2(\Gamma;L^1)}$, where $w = T^D v$ solves (5.1) with v supported by the strip S_h, and w_h is the mortar approximation of w. In view of the support properties of v, the L^1 integral appearing in $\|\nabla(w-w_h)\|_{L^2(\Gamma;L^1)}$ is split into an integral over $(0,\tilde{c}h)$ and

$(\tilde{c}h, D)$ for suitable $\tilde{c} > 0$, see (3.1). These two integrals are handled differently and are covered in Lemma 5.5. In contrast to the case of the integral over $(0, \tilde{h})$, the case of the integral over $(\tilde{c}h, D)$ does not exploit the support properties of v in the dual problem (5.1) and we have to consider the two cases $k = 1$ and $k > 1$ separately. Instead, it uses local approximation properties of the FEM as discussed in [22, 23]. Indeed, a key ingredient of the proof of Lemma 5.5 rests on the following result that can be found, for example, in [22, Sec. 5.4]: For two balls $B_r \subset B_{r'} \subset \mathcal{C}'$ with the same center and radii r, r' we have the local estimate

$$\|\nabla(w - w_h)\|_{L^2(B_r)} \le C\left(\|\nabla(w - I_h^k w)\|_{L^2(B_{r'})} + \frac{1}{r' - r}\|w - w_h\|_{L^2(B_{r'})}\right). \quad (5.10)$$

Based on similar covering arguments as those employed to reach (5.9), we obtain from (5.10) for arbitrary $\varepsilon > 0$ the estimate

$$\|\delta_\Gamma^{1/2+\varepsilon}\nabla(w - w_h)\|_{L^2(\mathcal{C}\setminus S_{\tilde{c}h})} \le C\left(\|\delta_\Gamma^{1/2+\varepsilon}\nabla(w - I_h^k w)\|_{L^2(\mathcal{C}'\setminus S_{\tilde{c}_1 h})} + \|\delta_\Gamma^{-1/2+\varepsilon}(w - w_h)\|_{L^2(\mathcal{C}'\setminus S_{\tilde{c}_1 h})}\right); \quad (5.11)$$

here, $\tilde{c}_1 \ge 1$ is arbitrary and $\tilde{c} > 1$ is assumed to be sufficiently large (it depends on $\tilde{c}_1$ but is independent of h).
This weighted estimate (5.11) leads to estimates in anisotropic norms. To see this, we write for $\varepsilon > 0$ (see also (3.1a))

$$\begin{aligned}\Big(\int_{\tau=\tilde{c}h}^{D}\|\nabla(w - w_h)\|_{L^2(\gamma_\tau)}\,d\tau\Big)^2 &\le \int_{\tau=\tilde{c}h}^{D}\tau^{-(1+2\varepsilon)}\,d\tau\int_{\tau=\tilde{c}h}^{D}\tau^{1+2\varepsilon}\|\nabla(w - w_h)\|^2_{L^2(\gamma_\tau)}\,d\tau\\ &\le Ch^{-2\varepsilon}\|\delta_\Gamma^{1/2+\varepsilon}\nabla(w - w_h)\|^2_{L^2(\mathcal{C}\setminus S_{\tilde{c}h})}\\ &\le Ch^{-2\varepsilon}\Big(\|\delta_\Gamma^{1/2+\varepsilon}\nabla(w - I_h^k w)\|^2_{L^2(\mathcal{C}'\setminus S_{\tilde{c}_1 h})} + \|\delta_\Gamma^{-1/2+\varepsilon}(w - w_h)\|^2_{L^2(\mathcal{C}'\setminus S_{\tilde{c}_1 h})}\Big) \quad (5.12)\end{aligned}$$

According to (3.1b), the limiting case $\varepsilon = 0$ takes the form

$$\begin{aligned}&\Big(\int_{\tau=\tilde{c}h}^{D}\|\nabla(w - w_h)\|_{L^2(\gamma_\tau)}\,d\tau\Big)^2\\ &\qquad\le C|\ln h|\left(\|\delta_\Gamma^{1/2}\nabla(w - I_h^k w)\|^2_{L^2(\mathcal{C}'\setminus S_h)} + \|\delta_\Gamma^{-1/2}(w - w_h)\|^2_{L^2(\mathcal{C}'\setminus S_h)}\right). \quad (5.13)\end{aligned}$$

With these preparations in hand, we formulate
LEMMA 5.5. *Let Ω be convex. Then, for $v \in L^2(S_h) \subset L^2(\Omega)$ and $w := T^D(v)$ (see (5.1)) and the mortar approximation w_h of w, there holds*

$$\|\nabla(w - w_h)\|_{L^2(\Gamma;L^1)} \le Ch^{3/2}\,(1 + \delta_{k,1}|\ln h|)\,\|v\|_{L^2(\Omega)}.$$

Proof. Let $\mathcal{C} \subset \mathcal{C}' \subset \Omega_i$ be cylinders as in Section 3. The Cauchy–Schwarz inequality and Fubini–Tonelli imply for arbitrary but fixed $\tilde{c} > 0$

$$\begin{aligned}\int_{\tau=0}^{\tilde{c}h}\|\nabla(w - w_h)\|_{L^2(\gamma_\tau)}\,d\tau &\le \sqrt{\tilde{c}h}\sqrt{\int_{\tau=0}^{\tilde{c}h}\|\nabla(w - w_h)\|^2_{L^2(\gamma_\tau)}\,d\tau}\\ &\le \sqrt{\tilde{c}h}\|\nabla(w - w_h)\|_{L^2(S(0,\tilde{c}h))} \le C\sqrt{h}\|\nabla(w - w_h)\|_{L^2(\Omega_i)}.\end{aligned}$$

Since Ω is convex, we have $\|w\|_{H^2(\Omega)} \leq C\|v\|_{L^2(\Omega)}$, and standard mortar estimates yield $\sum_{i=1}^{M} \|w - w_h\|^2_{H^1(\Omega_i)} \leq Ch^2\|w\|^2_{H^2(\Omega)} \leq Ch^2\|v\|^2_{L^2(\Omega)}$. Thus,

$$\int_{\tau=0}^{\tilde{c}h} \|\nabla(w - w_h)\|_{L^2(\gamma_\tau)} \leq Ch^{3/2}\|v\|_{L^2(\Omega)}. \tag{5.14}$$

For the integral $\int_{\tau=\tilde{c}h}^{D}$, we consider the cases $k = 1$ and $k > 1$ separately. We start with the case $k = 1$. We recall the regularity assertion $\|w\|_{B^{3/2}_{2,\infty}(\Omega)} \leq C\sqrt{h}\|v\|_{L^2(\Omega)}$ of Lemma 5.2 and note that therefore we have estimates for $\nabla^2 w$ in a weighted Sobolev space given by (5.4). Combining this observation with (5.13) yields together with the approximation properties of the interpolation operator $I_h^k = I_h^1$

$$\int_{\tau=\tilde{c}h}^{D} \|\nabla(w - w_h)\|_{L^2(\gamma_\tau)}\, d\tau \leq Ch^{3/2}|\ln h|\|v\|_{L^2(\Omega)} + Ch^{-1/2}\|w - w_h\|_{L^2(\Omega)}. \tag{5.15}$$

The convexity of Ω implies (see, e.g., [3, Rem. 2.8])

$$\|w - w_h\|_{L^2(\Omega)} \leq Ch^2\|w\|_{H^2(\Omega)} \leq Ch^2\|v\|_{L^2(\Omega)}. \tag{5.16}$$

Inserting this in (5.15) and combining the result with (5.14), we get for the cylinder $\mathcal{C}$ that $\|w - w_h\|_{L^2(\gamma_0;L^1)} \leq Ch^{3/2}|\ln h|\|v\|_{L^2(\Omega)}$. By summing over all cylinders, we obtain the desired estimate.
For the case $k > 1$, we use (5.12) with a fixed $\varepsilon \in (0, 1/2)$. From the approximation properties of I_h^k, we have for suitable $\tilde{c}_2$, $\tilde{c}_3$:

$$\|\delta_\Gamma^{1/2+\varepsilon}\nabla(w - I_h^k w)\|_{L^2(\mathcal{C}\setminus S_{\tilde{c}_1 h})} \leq Ch^2\|\delta_\Gamma^{1/2+\varepsilon}\nabla^3 w\|_{L^2(\mathcal{C}\setminus S_{\tilde{c}_2 h})} \leq Ch^2\|\delta_\Gamma^{-1/2+\varepsilon}\nabla^2 w\|_{L^2(\mathcal{C}\setminus S_{\tilde{c}_3 h})},$$

where, in the last step we used (5.5). Finally, the assumption $\varepsilon \in (0, 1/2)$ implies $\delta_\Gamma^{-1/2+\varepsilon} \leq h^{-1/2+\varepsilon}$. Inserting these consideration in (5.12), we obtain

$$\begin{aligned}\int_{\tau=\tilde{c}h}^{D} \|\nabla(w - w_h)\|_{L^2(\gamma_\tau)} &\leq Ch^{-\varepsilon}\left(h^2 h^{-1/2+\varepsilon}\|\nabla^2 w\|_{L^2(\mathcal{C}\setminus S_{\tilde{c}_3 h})} + h^{-1/2+\varepsilon}\|w - w_h\|_{L^2(\Omega)}\right) \\ &\leq Ch^{3/2}\|\nabla^2 w\|_{L^2(\Omega)},\end{aligned}$$

where we used convexity of Ω. ∎

5.2. Local dual problem with Neumann data. The regularity theory and convergence estimates of Section 5.1 are useful for the proof of Theorem 2.1 and, in turn, the estimate for $\|\lambda-\lambda_h\|_{L^2(\Gamma)}$ in Theorem 2.4. For the estimate $\|\lambda-\lambda_h\|_{H^{-1/2}(\gamma_l)}$ of Theorem 2.4, we need to consider a local Neumann problem instead of the global Dirichlet problem (5.1). Since most of the arguments run parallel to those of Section 5.1, we will be brief.
We consider the problem: Given v, find $\widetilde{w}_v$ such that

$$-\Delta\widetilde{w}_v = v - \frac{1}{|\Omega_i|}\int_{\Omega_i} v\, dx \quad \text{in } \Omega_i, \qquad \partial_n\widetilde{w}_v = 0 \quad \text{on } \partial\Omega_i, \qquad \int_{\Omega_i} \widetilde{w}_v = 0, \tag{5.17}$$

where $|\Omega_i|$ denotes the measure of Ω_i. Since the right-hand side $v - 1/|\Omega_i|\int_{\Omega_i} v$ has vanishing mean, (5.17) has a unique solution. We denote by $T^N : v \mapsto u$ the corresponding solution operator. As is customary in elliptic regularity theory, for functions v that are merely in $(H^1(\Omega_i))'$, the integral $\int_{\Omega_i} 1v\,dx$ is understood as a duality pairing so that T^N is in fact an operator $(H^1(\Omega_i))' \to H^1(\Omega_i)$. Concerning its regularity properties, we have analogously to Lemma 5.2:

LEMMA 5.6. *Assume that Ω_i is convex. Let $S_{h,i}$ be as in (2.6b). Let T^N be the solution operator for (5.17). Then T^N is a bounded linear operator $(B^{1/2}_{2,1}(\Omega_i))' \to B^{3/2}_{3,\infty}(\Omega_i)$. Additionally, if $v \in L^2(\Omega_i)$ satisfies $\operatorname{supp} v \subset \overline{S_{h,i}}$, then $\|T^N v\|_{B^{3/2}_{2,\infty}(\Omega_i)} \le C\sqrt{h}\|v\|_{L^2(\Omega_i)}$, where $C > 0$ is independent of v and h.*

Proof. Lax–Milgram provides in the standard way that $T^N : (H^1(\Omega_i))' \to H^1(\Omega_i)$ is bounded and linear; by convexity we have furthermore that $T^N : L^2(\Omega_i) \to H^2(\Omega_i)$ is bounded. Reasoning in exactly the same way as in the proof of Lemma 5.2 then yields the result. □

The analog of Lemma 5.5 is

LEMMA 5.7. *Let Ω_i be convex, $S_{h,i}$ be given by (2.6b). Assume $u \in B^{k+3/2}_{2,1}(\Omega_i)$ and that $u_{h,i} \in V_i$ satisfies the orthogonality condition*

$$\int_{\Omega_i} \nabla(u - u_{h,i}) \cdot \nabla v\,dx = 0 \qquad \forall v \in V_i.$$

Then,

$$\inf_{m\in\mathbb{R}} \|u - u_{h,i} - m\|_{L^2(S_{h,i})} \le Ch^{k+3/2}\,(1 + \delta_{k,1}|\ln h|)\,\|u\|_{B^{k+3/2}_{2,1}(\Omega_i)}, \tag{5.18}$$

$$|u - u_{h,i}|_{H^{1/2}(\partial\Omega_i)} \le Ch^{k+1/2}\,(1 + \delta_{k,1}|\ln h|)\,\|u\|_{B^{k+3/2}_{2,1}(\Omega_i)}. \tag{5.19}$$

where $C > 0$ is independent of h and u.

Proof. The proof follows from the developments in Section 5.1.2. Let m be the average of $u - u_{h,i}$ over $S_{h,i}$. For any $v \in L^2(\Omega_i)$ with $\operatorname{supp} v \subset \overline{S_{h,i}}$, let m_v be its average over Ω_i. Let $v_h \in \{z \in V_{h,i} \colon \int_{\Omega_i} z\,dx = 0\}$ be the Ritz projection of $T^N v$, where T^N is the solution operator for (5.17). Then, by the standard Aubin–Nitsche argument

$$(u - u_{h,i} - m, v)_{L^2(\Omega_i)} = (u - u_{h,i} - m, v - m_v)_{L^2(\Omega_i)} = \int_{\Omega_i} \nabla(u - u_{h,i}) \cdot \nabla T^N v\,dx$$
$$= \int_{\Omega_i} \nabla(u - u_{h,i}) \cdot \nabla(T^N v - v_h)\,dx = \int_{\Omega_i} \nabla(u - I^k_h u) \cdot \nabla(T^N v - v_h)\,dx.$$

We note that the regularity assertion of Lemma 5.6 for the Neumann problem is of the same type as that of Lemma 5.2 for the Dirichlet problem. Therefore, the same arguments as those used in Lemma 5.5 can be employed, leading to

$$\|T^N v - v_h\|_{L^2(\partial\Omega_i;L^1)} \le Ch^{3/2}(1 + \delta_{k,1}|\ln h|)\|v\|_{L^2(\Omega_i)}.$$

Finally, the arguments of the proof of Lemma 4.1 yield $\|\nabla(u - I^k_h u)\|_{L^2(\partial\Omega_i;L^\infty)} \le Ch^k\|u\|_{B^{k+3/2}_{2,1}(\Omega_i)}$, which allows us to conclude the validity of (5.18). The estimate (5.19) follows from (5.18) by the triangle inequality and inverse estimates: for arbitrary $m \in \mathbb{R}$ we can estimate $|u - u_{h,i}|_{H^{1/2}(\partial\Omega_i)} = |u - u_{h,i} - m|_{H^{1/2}(\partial\Omega_i)} \le |u - I^k_h u|_{H^{1/2}(\partial\Omega_i)} + |I^k_h u - u_{h,i} - m|_{H^{1/2}(\partial\Omega_i)}$. The approximation properties of I^k_h allow us to estimate the first term in the desired way. Inverse estimates yield $|I^k_h u - u_{h,i} - m|_{H^{1/2}(\partial\Omega_i)} \le Ch^{-1}\|I^k_h u - u_{h,i} - m\|_{L^2(S_{ch,i})}$ for suitable $c > 0$. Inserting again u by means of the triangle inequality, using the approximation properties of I^k_h and (5.18) allows us to conclude the proof. □

6. Proof of Theorems 2.1, 2.4. With the technical tools of Sections 3–5 in hand, we can show the main results. As alluded to in Section 2.5, the starting point is the representation of the L^2-error in a strip by a dual solution w. Then, Galerkin orthogonalities are used, bearing in mind the mortar setting. Next, the anisotropic norms of Section 3 come into play in the form of a Cauchy-Schwarz like estimate. Finally, Section 4 is used to bound the interpolation error for the solution u, and Section 5 allows us to control the finite element error of the dual solution w.

6.1. Proof of Theorem 2.1. We start with some notation. For $v \in L^2(S_h) \subset L^2(\Omega)$, let $w = T^D v$ be the solution of the dual problem (5.1). Correspondingly, we let $\lambda_w \in L^2(\Gamma)$ be defined by $\lambda_w|_{\gamma_l} := -\partial_{n_l} w|_{\Omega_{s(l)}}$. The function $w_h \in V_h^{-1}$ stands for the nonconforming mortar finite element approximation of w (i.e., the solution of (2.9) with $l(z) = (v, z)_{L^2(\Omega)}$). Due to (2.9b), we have $w_h \in V_h$.
For our nonconforming mortar method, the classical Galerkin orthogonalities for $u-u_h$ and $w - w_h$ do not hold anymore and have to be replaced with

$$a(u - u_h, \chi_h) + b(\chi_h, \lambda) = 0, \quad \forall \chi_h \in V_h, \tag{6.1a}$$

$$a(w - w_h, \chi_h) + b(\chi_h, \lambda_w) = 0, \quad \forall \chi_h \in V_h, \tag{6.1b}$$

where the second term in (6.1a) and (6.1b) measures the nonconformity of the finite element approximation. We are now in position of prove Theorem 2.1:
Proof Theorem 2.1. For $v \in L^2(S_h)$, let w and w_h be as defined above. Then the L^2-norm of the error $e_h := u - u_h$ restricted to S_h can be expressed as

$$\|e_h\|_{L^2(S_h)} = \sup_{v\in L^2(S_h), \|v\|_{L^2(S_h)}=1} (e_h, v)_{L^2(S_h)} = \sup_{v\in L^2(S_h), \|v\|_{L^2(S_h)}=1} (e_h, -\Delta w)_{L^2(\Omega)}.$$

Using Green's formula, we find with the aid of (6.1) for all $\mu_h, \tilde{\mu}_h \in M_h$ and the operator P_h of (4.2)

$$\begin{aligned}(e_h, -\Delta w)_{L^2(\Omega)} &= a(w, u - u_h) + b(u - u_h, \lambda_w) - a(u - u_h, w_h) - b(w_h, \lambda) \\ &\quad -a(w - w_h, P_h u - u_h) - b(P_h u - u_h, \lambda_w) \\ &= a(w - w_h, u - P_h u) \\ &\qquad + b(w - w_h, \lambda - \mu_h) + b(u - P_h u, \lambda_w - \tilde{\mu}_h). \end{aligned} \tag{6.2}$$

Now we consider the three terms on the right hand side separately and start with the two contributions resulting from the consistency error. First, the assumption $u \in B_{2,1}^{k+\frac{3}{2}}$ implies by the trace theorem and the fact that the subdomains Ω_i are polygonal/polyhedral that $\lambda \in H^k(\gamma_l)$ edgewise/facewise, [21, Thm. 2.9.3]. The convexity of Ω implies $w \in H^2(\Omega)$ with $\|w\|_{H^2(\Omega)} \le C\|v\|_{L^2(\Omega)}$ and additionally $\sum_{i=1}^M \|w - w_h\|^2_{H^1(\Omega_i)} \le Ch^2\|w\|^2_{H^2(\Omega)}$ as well as the L^2-estimate (5.16). Together with the approximation property Assumption (A2) and the multiplicative trace inequality, we get

$$\begin{aligned}\inf_{\mu_h\in M_h} b(w - w_h, \lambda - \mu_h) &\le \inf_{\mu_h\in M_h} \|[w - w_h]\|_{L^2(\Gamma)}\|\lambda - \mu_h\|_{L^2(\Gamma)} \\ &\le Ch^k\|\lambda\|_{H^k(\Gamma)} \left(\sum_{i=1}^M \|w - w_h\|_{L^2(\Omega_i)}\|w - w_h\|_{H^1(\Omega_i)}\right)^{1/2} \\ &\le Ch^k\|u\|_{B_{2,1}^{k+\frac{3}{2}}} h^{\frac{3}{2}}\|w\|_{H^2(\Omega)} \le Ch^{k+\frac{3}{2}}\|u\|_{B_{2,1}^{k+\frac{3}{2}}}. \end{aligned} \tag{6.3}$$

For the second consistency term in (6.2), we use the approximation properties of P_h given in (4.3b) and our convexity assumption (which implies $\lambda_w \in H^{1/2}$ edgewise/facewise with corresponding bounds that can be controlled in terms of $\|v\|_{L^2(\Omega)} = 1$):

$$\inf_{\tilde{\mu}_h \in M_h} b(u - P_h u, \lambda_w - \tilde{\mu}_h) \le \inf_{\tilde{\mu}_h \in M_h} \|[u - P_h u]\|_{L^2(\Gamma)} \|\lambda_w - \tilde{\mu}_h\|_{L^2(\Gamma)}$$
$$\le C h^{k+1+\frac{1}{2}} \|\lambda_w\|_{H^{\frac{1}{2}}(\Gamma)} \|u\|_{B^{k+\frac{3}{2}}_{2,1}} \le C h^{k+\frac{3}{2}} \|u\|_{B^{k+\frac{3}{2}}_{2,1}}. \quad (6.4)$$

The first term on the right of (6.2) can be bounded by

$$a(w - w_h, u - P_h u) \le C \|\nabla(w - w_h)\|_{L^2(\Gamma; L^1)} \|\nabla(u - P_h u)\|_{L^2(\Gamma; L^\infty)}$$

in view of the Hölder type inequality (3.2). Then, the upper bounds (6.3) and (6.4) in combination with Lemma 5.5 and Lemma 4.1 yield the result. □

Corollary 6.1. *Assume Ω to be a convex polygon/polyhedron. Then*

$$\|\partial_n u - \partial_n u_h\|_{L^2(\Gamma)} \le C h^k \left(1 + \delta_{k,1} |\ln h|\right) \|u\|_{B^{k+3/2}_{2,1}}.$$

Proof. Fix one subdomain Ω_i. By the triangle inequality, we get $\|\partial_n(u - u_h)\|_{L^2(\partial\Omega_i)} \le \|\partial_n(u - I_h^k u)\|_{L^2(\partial\Omega_i)} + \|\partial_n(I_h^k u - u_h)\|_{L^2(\partial\Omega_i)}$. The approximation properties (4.1) then imply

$$\|\partial_n(u - I_h^k u)\|^2_{L^2(\partial\Omega_i)} \le C h^{k/2} \|u\|_{H^{k+1}(S_h)} h^{(k-1)/2} \|u\|_{H^{k+1}(S_h)} \le C h^{k-1/2} \|u\|_{H^{k+1}(S_h)}.$$

Moreover [16, Lemma 2.1] yields $\|u\|_{H^{k+1}(S_h)} \le C h^{1/2} \|u\|_{B^{k+3/2}_{2,1}}$ and thus $\|\partial_n(u - I_h^k u)\|_{L^2(\partial\Omega_i)} \le C h^k \|u\|_{B^{k+3/2}_{2,1}}$. Using standard inverse estimates, we obtain $\|\partial_n(u_h - I_h^k u)\|_{L^2(\partial\Omega_i)} \le C h^{-3/2} \|u_h - I_h^k u\|_{L^2(S_h)}$. The triangle inequality, Lemma 5.5, and (4.1) allow us to conclude the argument. □

6.2. Proof of Theorem 2.4. Since the exact solution (u, λ) satisfies $a(u, v) + b(v, \lambda) = l(v)$ for all $v \in \prod_{i=1}^M H^1(\Omega_i)$ and the discrete saddle point solution (u_h, λ_h) satisfies (2.9a), we have the relation

$$a(u - u_h, v_h) + b(v_h, \lambda - \lambda_h) = 0 \qquad \forall v_h \in V_h^{-1}, \quad (6.5)$$

which will be important to transfer estimates for $\|u - u_h\|_{L^2(S_h)}$ to estimates for $\lambda - \lambda_h$.

Proof of Theorem 2.4. We start with the *a priori* bound in the L^2-norm, whose proof is based on Theorem 2.1 and the triangle inequality. Using (2.3a), (2.3b) and (4.3a), (4.3b), we get with the aid of (6.5) (recall the definition of E_h in (2.8))

$$\|\lambda - \lambda_h\|_{L^2(\Gamma)} \le C \Big(\inf_{\mu_h \in M_h} \|\lambda - \mu_h\|_{L^2(\Gamma)} + \sup_{w_h \in W_h} \frac{\int_\Gamma (\lambda - \lambda_h) w_h \, ds}{\|w_h\|_{L^2(\Gamma)}} \Big)$$
$$\le C \Big(h^k \|\lambda\|_{H^k(\Gamma)} + \sup_{w_h \in W_h} \frac{b(E_h w_h, \lambda - \lambda_h)}{\|w_h\|_{L^2(\Gamma)}} \Big)$$
$$\le C \Big(h^k \|u\|_{B^{k+\frac{3}{2}}_{2,1}} + \sup_{w_h \in W_h} \frac{a(E_h w_h, u_h - u)}{\|w_h\|_{L^2(\Gamma)}} \Big)$$
$$\le C h^k \|u\|_{B^{k+\frac{3}{2}}_{2,1}} + C \frac{1}{\sqrt{h}} \|u_h - u\|_{H^1(S_h)}$$

$$\leq Ck^k \|u\|_{B_{2,1}^{k+\frac{3}{2}}} + C\frac{1}{\sqrt{h}}(\|P_h u - u\|_{H^1(S_h)} + \frac{1}{h}\|P_h u - u_h\|_{L^2(S_h)})$$
$$\leq C(h^k \|u\|_{B_{2,1}^{k+\frac{3}{2}}} + h^{-1/2+k}|u|_{H^{k+1}(S_{2h})} + \frac{1}{h^{\frac{3}{2}}}\|u - u_h\|_{L^2(S_h)})$$
$$\leq Ch^k |\ln h| \|u\|_{B_{2,1}^{k+\frac{3}{2}}},$$

where in the last step we employed [16, Lemma 2.1] to bound $\|\nabla^{k+1}u\|_{L^2(S_{2h})} \leq C\sqrt{h}\|\nabla^{k+1}u\|_{B_{2,1}^{\frac{1}{2}}} \leq C\sqrt{h}\|u\|_{B_{2,1}^{k+\frac{3}{2}}}$ and used Theorem 2.1 to estimate $\|u - u_h\|_{L^2(S_h)}$. This shows the desired $L^2(\Gamma)$ estimate.

For the $H^{-1/2}$-estimate, we focus on one interface γ_l where $\Omega_{s(l)}$ is convex. In contrast to the weighted L^2-norm, the trivial extension E_h is not stable with respect to the $H_{00}^{\frac{1}{2}}(\gamma_l)$-norm. Thus we have to work with a different extension operator. Here we first extend $w_h \in W_{s(l)}$ trivially to an element on $\partial\Omega_{s(l)}$ and then apply the discrete harmonic extension operator onto $V_{s(l)}$. The resulting element is denoted by $H_{s(l)}w_h$ and is trivially extended to the other subdomains. We note that $H_{s(l)}w_h \in V_h^{-1}$. We denote by $u_{h,s(l)} \in V_{s(l)}$ the solution of

$$\int_{\Omega_{s(l)}} \nabla(u - u_{h,s(l)}) \cdot \nabla v \, dx = 0 \qquad \forall v \in V_{s(l)},$$

which is unique if we impose the additional condition that $u - u_{h,s(l)}$ has vanishing mean. $u_{h,s(l)}$ is viewed as an element of $\prod_{i=1}^M V_i$ by the trivial extension. Following the lines of the L^2-estimate and using the uniform inf-sup stability in the $H^{-1/2}$-norm, we find using the facts that $H_{s(l)}w_h$ and $u_h - u_{h,s(l)}$ are discrete harmonic on $\Omega_{s(l)}$:

$$\begin{aligned}
\|\lambda - \lambda_h\|_{H^{-\frac{1}{2}}(\gamma_l)} &\leq C\Big(\inf_{\mu_h \in M_{s(l)}} \|\lambda - \mu_h\|_{H^{-\frac{1}{2}}(\gamma_l)} + \sup_{w_h \in W_{s(l)}} \frac{\int_\Gamma (\lambda - \lambda_h) w_h ds}{\|w_h\|_{H_{00}^{\frac{1}{2}}(\gamma_l)}}\Big) \\
&\leq C\Big(h^{k+\frac{1}{2}}\|u\|_{B_{2,1}^{k+\frac{3}{2}}} + \sup_{w_h \in W_{s(l)}} \frac{a(H_{s(l)}w_h, u_h - u)}{\|w_h\|_{H_{00}^{\frac{1}{2}}(\gamma_l)}}\Big) \\
&= C\Big(h^{k+\frac{1}{2}}\|u\|_{B_{2,1}^{k+\frac{3}{2}}} + \sup_{w_h \in W_{s(l)}} \frac{a(H_{s(l)}w_h, u_h - u_{h,s(l)})}{\|w_h\|_{H_{00}^{\frac{1}{2}}(\gamma_l)}}\Big) \\
&\leq Ch^{k+\frac{1}{2}}\|u\|_{B_{2,1}^{k+\frac{3}{2}}} + C|u_h - u_{h,s(l)}|_{H^{\frac{1}{2}}(\partial\Omega_{s(l)})} \\
&\leq Ch^{k+\frac{1}{2}}\|u\|_{B_{2,1}^{k+\frac{3}{2}}} + C\inf_{m\in\mathbb{R}} \frac{1}{h}\|u_h - u_{h,s(l)} - m\|_{L^2(S_{h,s(l)})} \\
&\leq Ch^{k+\frac{1}{2}}\|u\|_{B_{2,1}^{k+\frac{3}{2}}} + C\frac{1}{h}\|u - u_h\|_{L^2(S_{h,s(l)})} \\
&\quad + C\inf_{m\in\mathbb{R}} \frac{1}{h}\|u - u_{h,s(l)} - m\|_{L^2(S_{h,s(l)})}.
\end{aligned}$$

Here, we used the strips $S_{h,i}$ defined in (2.6b). Finally the result follows from Lemma 5.7 and Theorem 2.1. □

Appendix A. An application of Besicovitch's covering theorem.

LEMMA A.1. *Let $\Omega \subset \mathbb{R}^d$ be bounded open and $M \subset \overline{\Omega}$ be a closed set. Fix $c \in (0,1)$ and $\varepsilon \in (0,1)$ such that*

$$1 - c(1+\varepsilon) =: c_0 > 0.$$

For each $x \in \Omega$, let $B_x := \overline{B}_{c\,\mathrm{dist}(x,M)}(x)$ be the closed ball of radius $c\,\mathrm{dist}(x,M)$ centered at x, and let $\widehat{B}_x := \overline{B}_{(1+\varepsilon)c\,\mathrm{dist}(x,M)}(x)$ denote the stretched (closed) ball of radius $(1+\varepsilon)c\,\mathrm{dist}(x,M)$ also centered at x.
Then there exists a countable set $x_i \in \Omega$, $i \in \mathbb{N}$, and a constant $N \in \mathbb{N}$ depending solely on the spatial dimension d with the following properties:

1. *(covering property) $\cup_{i\in\mathbb{N}} B_{x_i} \supset \Omega$*
2. *(finite overlap) for each $x \in \Omega$, there holds $\mathrm{card}\{i \,|\, x \in \widehat{B}_{x_i}\} \le N$.*

Proof. By the Besicovitch covering theorem (see, e.g., [26, Thm. 1.3.5] or [11, Sec. 1.5.2]) there exists N', which depends solely on d, and there exist N' families $\mathcal{G}_i$, $i = 1,\ldots,N'$, of balls with the following properties:

1. Each $\mathcal{G}_i$ consists of a countable set of closed balls $B_{ij} := \overline{B}_{c\,\mathrm{dist}(x_{ij},M)}(x_{ij})$ (the countably family is, for convenience, indexed by $j \in \mathbb{N}$).
2. For each i, the elements of $\mathcal{G}_i$ are pairwise disjoint.
3. The balls cover Ω, i.e., $\Omega \subset \cup_{i=1}^{N'} \cup_{j\in\mathbb{N}} B_{ij}$.

Hence, we have obtained the sets B_{x_i} that cover Ω. In order to see the finite overlap property of the stretched balls $\widehat{B}_{x_i}$, we proceed in several steps. We recall the definition of $c_0 > 0$ and introduce c_1 by

$$c_0 := 1 - c(1+\varepsilon), \qquad c_1 := 1 + c(1+\varepsilon).$$

1. step: For $x \in \Omega$ and i such that $x \in \widehat{B}_{x_i}$ there holds

$$c_0\,\mathrm{dist}(x_i,M) \le \mathrm{dist}(x,M) \le c_1\,\mathrm{dist}(x_i,M).$$

This follows from the triangle inequality in the following way. Since $x \in \widehat{B}_{x_i}$, we have $|x - x_i| \le c(1+\varepsilon)\,\mathrm{dist}(x_i,M)$ and thus

$$\mathrm{dist}(x_i,M) \le \mathrm{dist}(x,M) + |x - x_i| \le \mathrm{dist}(x,M) + c(1+\varepsilon)\,\mathrm{dist}(x_i,M),$$

which implies

$$(1 - c(1+\varepsilon))\,\mathrm{dist}(x_i,M) \le \mathrm{dist}(x,M).$$

Conversely, we have

$$\begin{aligned}\mathrm{dist}(x,M) &\le \mathrm{dist}(x_i,M) + |x - x_i| \le \mathrm{dist}(x_i,M) + c(1+\varepsilon)\,\mathrm{dist}(x_i,M)\\ &= (1 + c(1+\varepsilon))\,\mathrm{dist}(x_i,M)\end{aligned}$$

2. step Fix $x \in \Omega$. Consider one of the families $\mathcal{G}_i$, i.e., fix $i \in \{1,\ldots,N'\}$. Then the balls B_{ij} of $\mathcal{G}_i$ are pairwise disjoint. Define the set of indices

$$\mathcal{I}_x := \{j \,|\, x \in \widehat{B}_{ij}\}$$

of stretched balls containing x. By the first step, for any $j \in \mathcal{I}_x$, we have that the radius $r_j = c\,\mathrm{dist}(x_{ij},M)$ of the ball B_{ij} satisfies $r_j \sim \mathrm{dist}(x,M) =: r_x$ with the implies constants depending solely on c_0 and c_1. In order to estimate the cardinality of $\mathcal{I}_x$, we write

$$\begin{aligned}\mathrm{card}\,\mathcal{I}_x &= \sum_{j\in\mathcal{I}_x} 1 = \sum_{j\in\mathcal{I}_x} \frac{|\widehat{B}_{ij}|}{|\widehat{B}_{ij}|} \sim \sum_{j\in\mathcal{I}_x} \frac{|B_{ij}|}{r_x^d} = \frac{1}{r_x^d}|\cup_{j\in\mathcal{I}_x} B_{ij}|\\ &\le \frac{1}{r_x^d}|\cup_{j\in\mathcal{I}_x} \widehat{B}_{ij}| \le \frac{1}{r_x^d}|B_{Cr_x}(x)| \le C'.\end{aligned}$$

Here, we exploited first the fact that for fixed i the balls B_{ij}, $j \in \mathbb{N}$ are pairwise disjoint, then the fact that all the stretched balls $\widehat{B}_{ij}$, $j \in \mathcal{I}_x$ contain the point x and have radius comparable to $r_x = \mathrm{dist}(x, M)$. The constant $C > 0$ thus depends solely on c_0 and c_1 and therefore also the constant C'.

3. step: The second step shows that for each family $\mathcal{G}_i$, the cardinality $\mathcal{I}_x$ is bounded by C'. Hence, any $x \in \Omega$ can be in at most $N'C'$ of the balls $\{\widehat{B}_{ij} \,|\, i \in \{1, \ldots, N'\}, j \in \mathbb{N}\}$, which concludes the argument. □

REFERENCES

[1] T. Apel, J. Pfefferer, and A. Rösch. Finite element error estimates on the boundary with application to optimal control. in preparation.

[2] C. Bacuta, J.H. Bramble, and J. Xu. Regularity estimates for elliptic boundary value problems in Besov spaces. *Math. Comp.*, 72(244):1577–1595 (electronic), 2003.

[3] F. Ben Belgacem. The mortar finite element method with Lagrange multipliers. *Numer. Math.*, 84:173–197, 1999.

[4] F. Ben Belgacem, P. Seshaiyer, and M. Suri. Optimal convergence rates of *hp* mortar finite element methods for second-order elliptic problems. *M2AN, Math. Model. Numer. Anal.*, 34:591–608, 2000.

[5] C. Bernardi, Y. Maday, and A. Patera. Domain decomposition by the mortar element method. In H. Kaper et al., editor, *Asymptotic and numerical methods for partial differential equations with critical parameters*, pages 269–286. Reidel, Dordrecht, 1993.

[6] C. Bernardi, Y. Maday, and A. Patera. A new nonconforming approach to domain decomposition: The mortar element method. In H. Brezis and J.-L. Lions, editors, *Nonlinear partial differential equations and their applications*, volume XI of *Collège de France Seminar*, pages 13–51. Pitman, 1994.

[7] D. Braess and W. Dahmen. Stability estimates of the mortar finite element method for 3-dimensional problems. *East-West J. Numer. Math.*, 6:249–263, 1998.

[8] D. Braess and W. Dahmen. The mortar element method revisited. What are the right norms? In Debit et al., editor, *Domain decomposition methods in science and engineering*, Theory and Engineering Applications of Computational Methods., pages 27–40. Internat. Center Numer. Methods Eng. (CIMNE), 2002.

[9] F. Brezzi and M. Fortin. *Mixed and hybrid finite element methods.* Springer, New York, 1991.

[10] R.A. DeVore and G.G. Lorentz. *Constructive approximation.* Grundlehren der Mathematischen Wissenschaften. 303. Berlin: Springer-Verlag, 1993.

[11] L.C. Evans and R.F. Gariepy. *Measure Theory and Fine Properties of Functions.* CRC Press, 1992.

[12] D. Gilbarg and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order.* Grundlagen der mathematischen Wissenschaften 224. Springer, 1977.

[13] J. Kadlec. On the regularity of the solution of the Poisson problem on a domain with boundary locally similar to the boundary of a convex open set. *Czech. Math. J.*, 14(89):386–393, 1964.

[14] C. Kim, R. Lazarov, J. Pasciak, and P. Vassilevski. Multiplier spaces for the mortar finite element method in three dimensions. *SIAM J. Numer. Anal.*, 39:519–538, 2001.

[15] B. Lamichhane and B. Wohlmuth. Biorthogonal bases with local support and approximation properties. *Math. Comp.*, 76:233–249, 2007.

[16] J. Li, J.M. Melenk, B. Wohlmuth, and J. Zou. Optimal a priori estimates for higher order finite elements for elliptic interface problems. *Appl. Numer. Math.*, 60:19–37, 2010.

[17] J.M. Melenk, H. Rezaijafari, and B. Wohlmuth. Quasi-optimal a priori estimates for fluxes in mixed finite element methods and applications to the stokes–darcy coupling. Technical Report 05/2012, Institute for Analysis and Scientific Computing, TU Wien, 2012.

[18] R.A. Nicolaides. Existence, uniqueness and approximation for generalized saddle point problems. *SIAM J. Numer. Anal.*, 19:349–357, 1982.

[19] J. Nitsche. Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind. *Abhandlungen aus den Mathematischen Seminaren der Universität Hamburg*, 36:9–15, 1971.

[20] L. Tartar. *An introduction to Sobolev spaces and interpolation spaces*, volume 3 of *Lecture Notes of the Unione Matematica Italiana*. Springer, Berlin, 2007.

[21] H. Triebel. *Interpolation theory, function spaces, differential operators.* Johann Ambrosius

Barth, Heidelberg, second edition, 1995.

[22] L. Wahlbin. Local behavior in finite element methods. In P.G. Ciarlet and J.L. Lions, editors, *Handbook of numerical Analysis. Volume II: Finite element methods (Part 1)*, pages 353–522. North Holland, 1991.

[23] L. Wahlbin. *Superconvergence in Galerkin finite element methods*, volume 1605 of *Lecture Notes in Mathematics*. Springer Verlag, 1995.

[24] B. Wohlmuth. A mortar finite element method using dual spaces for the Lagrange multiplier. *SIAM J. Numer. Anal.*, 38:989–1012, 2000.

[25] B. Wohlmuth. *Discretization Methods and Iterative Solvers Based on Domain Decomposition*. Springer, Berlin Heidelberg, 2001.

[26] W.P. Ziemer. *Weakly Differentiable Functions*. Springer Verlag, 1989.