hp-FEM for second moments of elliptic PDEs with stochastic data Part 2: Exponential convergence

Report

Author(s):
Pentenrieder, Bastian; Schwab, Christoph

Publication date:
2010-03

Permanent link:
https://doi.org/10.3929/ethz-a-010403610

Originally published in:
SAM Research Report 2010-09

Funding acknowledgement:
247277 - Automated Urban Parking and Driving (EC)

Eidgenössische
Technische Hochschule
Zürich

Ecole polytechnique fédérale de Zurich
Politecnico federale di Zurigo
Swiss Federal Institute of Technology Zurich

hp-FEM for second moments of elliptic PDEs with stochastic data

Part 2: Exponential convergence

B. Pentenrieder and Ch. Schwab

Research Report No. 2010-09
March 2010

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

hp-FEM FOR SECOND MOMENTS OF ELLIPTIC PDES WITH STOCHASTIC DATA PART 2: EXPONENTIAL CONVERGENCE

BASTIAN PENTENRIEDER AND CHRISTOPH SCHWAB

Seminar for Applied Mathematics, ETH Zentrum, Zürich, Switzerland

ABSTRACT. We prove exponential rates of convergence of a class of hp Galerkin Finite Element approximations of solutions to a model tensor non-hypoelliptic equation in the unit square $\square = (0,1)^2$ which exhibit singularities on $\partial\square$ and on the diagonal $\Delta = \{(x,y) \in \square : x = y\}$, but are otherwise analytic in $\square$. As we explained in the first part [6] of this work, such problems arise as deterministic second moment equations of linear, second order elliptic operator equations $Au = f$ with Gaussian random field data f.

1. INTRODUCTION

The present paper is, together with [6], the second in a series which is devoted to the numerical analysis of a hp-Finite Element Galerkin method for the fast computation of second moments for a model class of linear, elliptic operator equations. Specifically, in the unit interval $D = [0,1]$ and for a constant $b > 0$, we consider the stochastic model equation

$$\left.\begin{aligned} Au(\omega) = -u_{xx}(\cdot,\omega) + b^2 u(\cdot,\omega) = f(\cdot,\omega) \quad \text{in}(H^1(D))' \\ u_x(0,\omega) = u_x(1,\omega) = 0 \end{aligned}\right\} \quad \text{for } \mathbb{P}\text{-a.e. } \omega \in \Omega, \tag{1.1}$$

where the load f and therefore the solution u randomly depend on $\omega \in \Omega$. As in [6], we consider weak solutions $u \in V = H^1(D)$ of (1.1) for data $f \in V' := (H^1(D))'$. As in [6], we assume here that we are given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ over the Hilbert space V' and that $f \in L^2(\Omega, V', d\mathbb{P})$. As was explained in [6], due to the linearity of (1.1), in case that the data f is Gaussian over V', so is the solution over V. The (deterministic) elliptic second order operator A in (1.1) is boundedly invertible from V' to V, and for every Gaussian data f there exists a Gaussian solution u which is characterized by its mean field $\mathcal{M}^1 u$ and its (co)variance $\mathcal{M}^2 u$. Whereas $\mathcal{M}^1 u$ satisfies the deterministic elliptic equation $A\mathcal{M}^1 u = \mathcal{M}^1 f$, in [6] we showed following [9, 10] that the covariance $C_u = \mathcal{M}^2(u - \mathcal{M}^1 u)$ satisfies the deterministic tensorized equation

$$\text{Find} \quad C_u \in V \otimes V: \qquad (A \otimes A)C_u = C_f \quad \text{in} \quad (V \otimes V)' \simeq V' \otimes V'. \tag{1.2}$$

Notice that the deterministic problem (1.2) for the covariance kernel of the Gaussian random solution u is, in fact, a problem with tensor product structure which is posed in the domain $\square = D \times D$, i.e. in a domain which has *twice the dimension of the physical domain* D. Galerkin Finite Element discretizations of this problem are, as we explained in [6], straightforward and converge quasioptimally; however, the convergence rates which can be achieved in terms of the number of degrees of freedom suffer from rapid increase of their number with mesh refinement due to the higher dimension of the domain $\square$. In [9, 10], we approached this problem by the use

Date: March 3, 2010.

of *sparse tensor products* of multilevel Finite Element spaces in the physical domain D. The resulting Galerkin approximations of C_u were shown there to converge at the essentially optimal rate nearly *without* an increase in the number of degrees of freedom, *provided that C_u exhibits a sufficient amount of smoothness in terms of so-called Sobolev spaces of mixed highest derivatives*.
In [6] and in the present paper, we are particularly interested in the case when $C_f(x,y)$ is *stationary*, i.e. when $C_f(x,y) = w(x-y)$ for (x,y) on the unit square

$$\square := \big\{(x,y) \in \mathbb{R}^2 : 0 \le x, y \le 1\big\}$$

depends only on the difference $x - y$ and is featuring a singularity on the diagonal

$$\Delta := \big\{(x,y) \in \square : x = y\big\}$$

while being an analytic function of (x,y) in $\square\backslash\Delta$. Specifically, we considered data f satisfying

Assumption 1.1. We assume the Gaussian random field f to be *stationary*, i.e. its correlation kernel C_f is translation invariant:

$$C_f(x,y) = C_f(x+t, y+t) \quad \text{for all } t \text{ with } (x+t, y+t) \in \square$$

Thus, C_f can be written as a function of the difference $z = x - y$, $z \in [-1,1]$:

$$C_f = C_f(z), \quad z = x - y \tag{1.3}$$

Furthermore, it is assumed

$$C_f\big|_{[0,1]} \in \mathcal{B}^l_{\beta,d}(0,1) \cap C^0([0,1]) \tag{1.4}$$

with some $l \in \mathbb{N}$, $\beta \in [0,1)$ and $d \ge 1$, where $\mathcal{B}^l_{\beta,d}(0,1)$ denotes a certain class of countably normed, weighted Sobolev spaces whose definition will be recalled below for convenience.

In this case, both C_f and C_u in (1.2) could exhibit very low regularity in terms of Sobolev spaces $H^{k,k}(\square)$ of mixed weak k-th derivatives which would entail only very low approximation rates of the sparse tensor product Finite Element Methods proposed in [9, 10].
In [6], we considered in detail the regularity of C_u in (1.2) for covariances C_f corresponding to Gaussian random data f which are analytic in $\square\backslash\Delta$. A key role in the regularity theory of (1.2) for data f with stationary covariances is played by a certain, fourth order ordinary differential equation.

Lemma 1.2. *Let $g \in C^0([-1,1])$ and $b > 0$. Then, the general solution to the ordinary differential equation*

$$v^{(4)}(z) - 2b^2\, v''(z) + b^4\, v(z) = g(z) \qquad \forall z \in (-1,1) \tag{1.5}$$

is given by

$$v(z) = J_g(z) + c_1 \cosh(b\,z) + c_2 \sinh(b\,z) + c_3\, z \cosh(b\,z) + c_4\, z \sinh(b\,z) \tag{1.6}$$

with $c_i \in \mathbb{R}$ and

$$J_g(z) = \int_0^z \left(\frac{z-t}{2\,b^2} \cosh\big(b\,(z-t)\big) - \frac{\sinh\big(b\,(z-t)\big)}{2\,b^3} \right) g(t)\, dt. \tag{1.7}$$

In particular: If g is an even function, J_g is even as well—in this case, v is even, iff $c_2 = c_3 = 0$.

We proved in [6] the following regularity result.

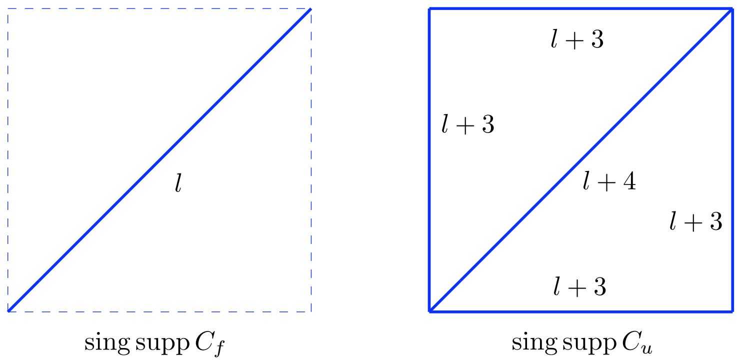

FIGURE 1.1. Illustration of Corollary 1.4. The singularity of C_f on the diagonal (order l) gives rise to singularities in C_u on the diagonal (order $l+4$) and the boundary (order $l+3$).

Theorem 1.3. *Let C_f satisfy Assumption 1.1. With J_{C_f} defined by (1.7), the unique solution C_u to Problem 1.2 admits the representation*

$$C_u(x,y) = C_u^{\Delta}(x-y) + C_u^{\Gamma}(x,y), \tag{1.8}$$

where

$$C_u^{\Delta}(z) = J_{C_f}(z) + \frac{J''_{C_f}(1)}{b^3 \sinh b}\big(2\cosh(b\,z) - b\,z\sinh(b\,z)\big), \tag{1.9a}$$

$$\begin{aligned} C_u^{\Gamma}(x,y) = &-(C_u^{\Delta})'(1-x)\,\frac{\cosh(b\,y)}{b\sinh b} - (C_u^{\Delta})'(x)\,\frac{\cosh\big(b\,(1-y)\big)}{b\sinh b} \\ &- (C_u^{\Delta})'(1-y)\,\frac{\cosh(b\,x)}{b\sinh b} - (C_u^{\Delta})'(y)\,\frac{\cosh\big(b\,(1-x)\big)}{b\sinh b}. \end{aligned} \tag{1.9b}$$

Furthermore, $C_u^{\Delta}(z)$ is an even function of the difference $z = x-y$.

Corollary 1.4 (singularities in C_u)**.** *Let C_f satisfy Assumption 1.1, i.e. in particular $C_f \in \mathcal{B}^l_{\beta,d}(0,1)$. Then, the unique solution C_u to Problem 1.2 admits a splitting*

$$C_u = C_u^{\Delta} + C_u^{\Gamma} \tag{1.10}$$

with

$$C_u^{\Delta}(x,y) = w_1(x-y), \quad w_1(z) \text{ an even function of } z = x-y, \tag{1.11a}$$

$$\begin{aligned} C_u^{\Gamma}(x,y) = &\, w_2(1-x)\,w_3(y) + w_2(x)\,w_3(1-y) \\ &+ w_3(x)\,w_2(1-y) + w_3(1-x)\,w_2(y), \end{aligned} \tag{1.11b}$$

where $w_1 \in \mathcal{B}^{l+4}_{\beta,d}(0,1) \cap C^4([-1,1])$, $w_2 \in \mathcal{B}^{l+3}_{\beta,d}(0,1)$, and w_3 is an analytic function on $[0,1]$ satisfying

$$\max_{x\in[0,1]} \big|w_3^{(k)}(x)\big| \le b^{-1}\coth(b)\,b^k \qquad \forall k \in \mathbb{N}_0. \tag{1.12}$$

Remark 1.5 (enlargement of the singular support). Corollary 1.4 shows

$$\operatorname{sing\,supp} C_u = \Delta \cup \Gamma \supsetneq \Delta = \operatorname{sing\,supp} C_f.$$

This increase of the solution's singular support is a consequence of the non-hypoelliptic nature of the differential operator $A \otimes A$.

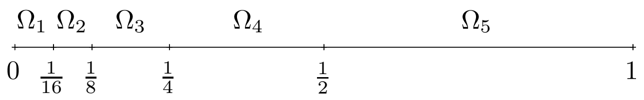

FIGURE 2.1. Partition with grading factor $\sigma = \frac{1}{2}$ and $M = 5$.

This paper is structured as follows: after briefly recapitulating the tool of countably normed, weighted Sobolev spaces in one dimension in Section 2, in Section 3 the space $H^{1,1}(D \times D)$ in which the hp-FE approximations of the covariance are being sought is introduced. In particular, covariance functions $C(x, y)$ corresponding to *stationary* random fields are introduced; these depend only on the difference $x - y$ of the arguments. The proofs of exponential convergence of constrained tensorized hp-approximations for the diagonally singular part C_u^Δ of the solution's covariance are presented in Section 4.

2. Review of hp-approximation in one dimension

Here, we briefly summarize hp-approximation of functions $u : [0, 1] \to \mathbb{R}$ being analytic in $(0, 1)$ and possibly featuring a singularity at $x = 0$ (for the results presented, see also the monograph [8] and the references therein). The behavior of such functions u is conveniently described by countably normed, weighted Sobolev spaces (see e.g. [2]):

Definition 2.1 (spaces $\mathcal{B}^l_{\beta,d}(0, 1)$)**.** Let $0 \le \beta < 1$. Then, for all $l \in \mathbb{N}$ and natural numbers $k \ge l$,

$$|u|_{H^{k,l}_\beta(0,1)} := \left\| x^{\beta+k-l} u^{(k)} \right\|_{L^2(0,1)} \tag{2.1}$$

defines a seminorm. If $u \in H^{l-1}(0, 1)$ and if there exist constants $C > 0$, $d \ge 1$ such that

$$|u|_{H^{k,l}_\beta(0,1)} \le C\, d^{k-l}(k - l)! \quad \forall k \ge l, \tag{2.2}$$

then we write $u \in \mathcal{B}^l_{\beta,d}(0, 1)$, or simply $u \in \mathcal{B}^l_\beta(0, 1)$.

Example 2.2. For $u(x) = x^\gamma$, $\gamma > -\frac{1}{2}$, and $u(x) = \ln(x)$, it is possible to choose l, β and d such that $u \in \mathcal{B}^l_{\beta,d}(0, 1)$.

The following three definitions discretize the domain $I = [0, 1]$ and introduce approximation spaces for functions $u \in \mathcal{B}^l_\beta(0, 1)$:

Definition 2.3 (partition $\mathcal{T}$, elements Ω_j, nodes x_j)**.** Let $M \in \mathbb{N}$. We define a (generic) partition of I by $\mathcal{T} := \{\Omega_j : 1 \le j \le M\}$ with elements $\Omega_j := [x_{j-1}, x_j]$ and nodes

$$0 = x_0 < x_1 < \ldots < x_{M-1} < x_M = 1.$$

To achieve exponential convergence rates for piecewise analytic functions, in hp-approximation one selects the nodes as powers of a *grading factor* $\sigma \in (0, 1)$ (see Figure 2):

$$x_i = \sigma^{M-i}, \quad 1 \le i \le M \tag{2.3}$$

Definition 2.4 (polynomial space $\mathcal{S}^p$)**.** Let $p \in \mathbb{N}_0$. The space of polynomials (in $x \in \mathbb{R}$) of degree at most p is defined by $\mathcal{S}^p := \operatorname{span}\{x^i : i = 0, \ldots, p\}$.

Definition 2.5 (space $S^{\mathbf{p}}(\mathcal{T})$)**.** Let $\mathbf{p} = (p_1, \dots, p_M) \in \mathbb{N}^M$, $M = \#\mathcal{T}$. We define the space of continuous functions on I that are piecewise polynomial with degree vector $\mathbf{p}$ on partition $\mathcal{T}$:

$$S^{\mathbf{p}}(\mathcal{T}) := \left\{ v \in C^0(I) : v\big|_{\Omega_j} \in \mathcal{S}^{p_j} \text{ for all } j = 1, \dots, M \right\}$$

Obviously, $S^{\mathbf{p}}(\mathcal{T}) \subset H^1(I)$.

Next, we provide a functional π_p that assigns to every H^1-function $\hat{u}$ on the reference interval $\widehat{\Omega} := [-1, 1]$ a polynomial of degree p.

Definition 2.6 (Approximation Operator π_p)**.** For $p \in \mathbb{N}$, we define the polynomial approximation of $\hat{u} \in H^1(-1, 1)$ of polynomial degee p by

$$\pi_p : H^1(\widehat{\Omega}) \to \mathcal{S}^p, \qquad (\pi_p \hat{u})(\xi) := \hat{u}(-1) + \int_{-1}^{\xi} \left(\sum_{i=0}^{p-1} b_i L_i(\tilde{\xi}) \right) d\tilde{\xi}, \tag{2.4}$$

where $\sum_{i=0}^{p-1} b_i L_i$ is the truncated Legendre series of $\hat{u}'$ (see e.g. [11]).

We observe that this interpolant is nodally exact at the endpoints (e.g. [8, Theorem 3.14])

$$(\pi_p \hat{u})(\pm 1) = \hat{u}(\pm 1). \tag{2.5}$$

So far, the operator π_p serves as an approximation tool for H^1-functions defined on $\widehat{\Omega} = [-1, 1]$ only. For $u \in H^1(0, 1)$, we obtain local (polynomial) approximations on elements Ω_j in the usual way by linear transformations. These local approximations allow, with (2.5), to assemble a global approximation to u which belongs to the finite element space $S^{\mathbf{p}}(\mathcal{T})$.

Definition 2.7 (element mappings T_{Ω_j})**.** For each element $\Omega_j = [x_{j-1}, x_j]$ in $\mathcal{T}$, we define a bijective linear mapping:

$$T_{\Omega_j} : \widehat{\Omega} \to \Omega_j, \qquad \xi \mapsto x = T_{\Omega_j}(\xi) := \frac{1 - \xi}{2} x_{j-1} + \frac{\xi + 1}{2} x_j$$

Definition 2.8 (local approximation on Ω_j)**.** For $u \in H^1(0, 1)$, we define its local approximation v_j on element Ω_j by:

$$\hat{u}_j := u \circ T_{\Omega_j} \tag{2.6a}$$

$$\hat{v}_j := \pi_{p_j} \hat{u}_j \tag{2.6b}$$

$$v_j := \hat{v}_j \circ T_{\Omega_j}^{-1} \tag{2.6c}$$

Definition 2.9 (global approximation $v_{\mathbf{p}}^{\mathcal{T}}$)**.** Let $u \in H^1(0, 1)$, $\mathcal{T}$ the partition from Definition 2.3 and $\mathbf{p} = (p_1, \dots, p_M) \in \mathbb{N}^M$. Define $v_{\mathbf{p}}^{\mathcal{T}} : [0, 1] \to \mathbb{R}$ element-wise by

$$v_{\mathbf{p}}^{\mathcal{T}}\big|_{\Omega_j} = v_j \quad (1 \le j \le M)$$

with v_j being the local approximation of u from Definition 2.8. In particular, $v_{\mathbf{p}}^{\mathcal{T}}$ is continuous across element transitions due to the construction (2.6) and by (2.5). Thus, $v_{\mathbf{p}}^{\mathcal{T}}$ belongs to the space $S^{\mathbf{p}}(\mathcal{T})$ from Definition 2.5.

The main result of this section is the following:

Theorem 2.10 (hp-approximation)**.** *Let $u \in \mathcal{B}^2_{\beta,d}(0, 1)$. For a fixed grading factor $\sigma \in (0, 1)$, we define $\mathcal{T}$ by (2.3). Then, one can find a $\mu = \mu(\beta, d, \sigma) \ge 1$ such that, with $\mathbf{p} = (p_1, \dots, p_M)$,*

$$p_1 := 1, \quad p_i := \lfloor \mu i \rfloor \;\; (2 \le i \le M), \tag{2.7}$$

the approximation errors of $v_{\mathbf{p}}^{\mathcal{T}} \in S^{\mathbf{p}}(\mathcal{T})$*, as* $M \to \infty$*, are bounded by*

$$\|u - v_{\mathbf{p}}^{\mathcal{T}}\|_{H^1(0,1)} \le c_1 \exp\big(-c_2\sqrt{N}\big), \tag{2.8}$$

where $N := \dim S^{\mathbf{p}}(\mathcal{T})$ *and the constants* $c_1, c_2 > 0$ *are independent of* N*.*

Proof. See e.g. [5, Theorem 4.23 and Corollary 4.25]. $\square$

3. Problem setting

The aim of this paper consists in adapting the 1D result from the previous section so that it allows for exponentially convergent hp approximation of functions on the unit square $\square = D \times D$ depending only on the difference $x - y$ and featuring a singularity on the diagonal $\Delta := \{(x,y) \in \square : x = y\}$. The following notion of stationarity of a function w applies in particular to covariances C_f which are stationary in the sense of Assumption 1.1.

Assumption 3.1. Let $w : \square \to \mathbb{R}$ be "stationary", i.e.

$$w(x,y) = w(x-y) \quad \forall (x,y) \in \square,$$

with an even $w(\cdot)$,

$$w(x-y) = w(y-x) \quad \forall\, x-y \in [-1,1].$$

Furthermore, we assume $w \in H^2(-1,1)$ and $w|_{[0,1]} \in \mathcal{B}^3_\beta(0,1)$. *Notice that we will not distinguish between* w *as a function of* $(x,y) \in \square$ *and* w *as a function of* $x - y \in [-1,1]$*, respectively.*

Such functions w may arise as 2-point correlations of stationary random fields f on the domain $I = [0,1]$. In [5, Chapter 2], some correlation models from the literature on spatial statistics [1, 7] are put into the context of Assumption 3.1. In the important special case where the random field f takes values in the Sobolev space $H^1(I)$, its 2-point correlation lives in the tensor product space $H^1(I) \otimes H^1(I) \cong H^{1,1}(\square)$ (see e.g. [9]). The space $H^{1,1}(\square)$ can be characterized as the set of all functions for which the following norm is finite:

Definition 3.2 ($H^{1,1}$-norm)**.** For a function $w : \square \to \mathbb{R}$, we define the norm

$$\|w\|_{H^{1,1}(\square)} = \Big(\|w\|^2_{L^2(\square)} + \|\partial_x w\|^2_{L^2(\square)} + \|\partial_y w\|^2_{L^2(\square)} + \|\partial_x\partial_y w\|^2_{L^2(\square)}\Big)^{\frac{1}{2}}.$$

In this paper, we will construct approximations to w and measure their errors with respect to the $H^{1,1}$-norm. Notice that, in comparison to the H^1-norm, the $H^{1,1}$-norm features the term $\|\partial_x\partial_y w\|_{L^2(\square)}$, which imposes additional constraints on the interpolants.

Remark 3.3. Many of the intermediary results to be found in the remainder of this work are valid for more general $w : \square \to \mathbb{R}$. Hence, we shall state explicitly whenever Assumption 3.1 is actually required.

Assumption 3.1 suggests to mimick the hp-approximation strategy investigated in Section 2 for functions on $I = [0,1]$ featuring a singularity in $x = 0$. In analogy, the obvious idea is to employ h-refinement on the diagonal of $\square$ (where $x - y = 0$) and p-refinement away from it.

4. hp-APPROXIMABILITY OF w

In the present section, we establish the main result of this paper, namely the exponential rate of convergence of certain hp-Finite Element approximations of "stationary" covariance functions $C(x,y)$ in the sense of Assumption 3.1 which arise, for example, as covariance kernels for stationary processes.
As specified in (1.8) in Theorem 1.3 which we proved in the first part [6] of this work, such functions arise as diagonally singular part $C_u^\Delta(x,y)$ of the covariances $C_u(x,y)$ for solutions u of (1.1) for stationary random inputs f, the second part $C_u^\Gamma(x,y)$ having singular support on $\partial\square$. In tensorized hp approximations, the singularities of $C_u^\Gamma(x,y)$ are easily resolved by tensor product hp FE spaces in $\square$ obtained by tensorization of (suitable adaptations of) the univariate hp FE-spaces $S^{\mathbf{p}}(\mathcal{T})$ defined in Theorem 2.8, providing a convergence rate $\exp(-b'\sqrt[4]{N})$ (the exponent 4 as compared to the univariate error bound (2.8) is a consequence of the so-called curse of dimension). The (straightforward) argument is provided in Theorem 4.30, Corollary 4.32 of [5].
Contrary to this, however, the diagonal singularity of $C_u^\Delta(x,y)$ foils straightforward hp-approximations based on geometric mesh refinement towards the diagonal $\Delta = \operatorname{sing\,supp}(C_u^\Delta) \subset \square$ due to the exponential increase in the number of elements arising in such meshes. Here, we show that this difficulty can be overcome by a form of *constrained hp-approximation* which exploits that C_u^Δ satisfies Assumption 3.1 by enforcing a discrete version of this assumption on all degrees of freedom located in elements located in the geometric position versus Δ. In this section, we prove for this constrained hp approximation of $C_u^\Delta(x,y)$ in the $H^{1,1}(\square)$-norm the error bound $exp(-b'\sqrt[3]{N})$ which was announced in Theorem 6.4 of the first part of this work [6].

4.1. **Domain discretization and related definitions.** The h-refinement towards the diagonal is accomplished by resolving Δ in the fashion of a quad-tree recursion (see e.g. [3]):

Definition 4.1 (partition $\mathcal{Q}_L$, elements Q, side lengths h_Q)**.** Let $\mathcal{Q}_0 := \{\square\}$. The partition $\mathcal{Q}_L$ is obtained from $\mathcal{Q}_{L-1}$ as follows: Start with $\mathcal{Q}_L := \emptyset$. Then, for all elements $Q = [x_1,x_2]\times[y_1,y_2] \in \mathcal{Q}_{L-1}$, check whether $Q \cap \Delta \neq \emptyset$, and if so, set

$$\mathcal{Q}_L := \mathcal{Q}_L \cup \{Q_1,\dots,Q_4\},$$

where

$$Q_1 = \left[x_1, \frac{x_1+x_2}{2}\right]\times\left[y_1,\frac{y_1+y_2}{2}\right], \quad Q_2 = \left[x_1, \frac{x_1+x_2}{2}\right]\times\left[\frac{y_1+y_2}{2}, y_2\right],$$
$$Q_3 = \left[\frac{x_1+x_2}{2}, x_2\right]\times\left[y_1,\frac{y_1+y_2}{2}\right], \quad Q_4 = \left[\frac{x_1+x_2}{2}, x_2\right]\times\left[\frac{y_1+y_2}{2}, y_2\right];$$

else, set $\mathcal{Q}_L := \mathcal{Q}_L \cup \{Q\}$. The side length of an element $Q = [x_1,x_2]\times[y_1,y_2]$ is

$$h_Q := x_2 - x_1 = y_2 - y_1.$$

All elements of partition $\mathcal{Q}_L$ are paraxial squares. However, this convenience is bought at the price of having $\mathcal{O}(2^L)$ many elements in the neighborhood of Δ (see Figure 4.1).
Since we will associate finite element basis functions not only with elements, but also with other geometric entities such as nodes and edges, the following two definitions give a precise notion of these and related terms:

Definition 4.2 (nodes, hanging (inner/boundary) nodes, set $\mathcal{K}_L^{\mathrm{dof}}$)**.** A point $(x,y) \in \square$ is called a *node* (of $\mathcal{Q}_L$), if (x,y) is a vertex of an element $Q \in \mathcal{Q}_L$. A node (x,y) is called *hanging inner node*, if $(x,y) \in \partial Q$ for some $Q \in \mathcal{Q}_L$ without being a vertex

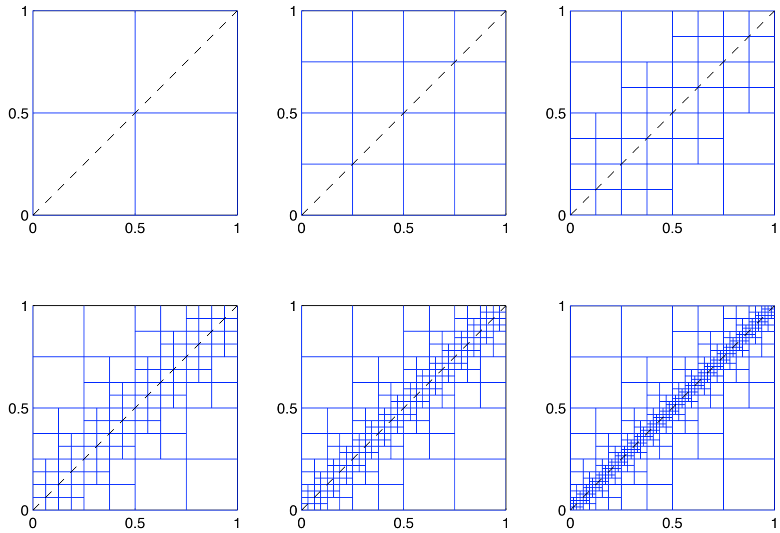

FIGURE 4.1. Partitions $\mathcal{Q}_1$ to $\mathcal{Q}_6$ according to the recursion in Definition 4.1. Observe how a new partition is created from the previous one by refining all those squares that are touched or crossed by Δ (the dashed line).

of Q. A node $(x, y) \in \partial\Box$ is called *hanging boundary node*, if there is a hanging inner node (x_i, y_i) such that $x - y = x_i - y_i$, or if $(x, y) \in \left\{(0, \frac{3}{4}), (\frac{1}{4}, 1), (\frac{3}{4}, 0), (1, \frac{1}{4})\right\}$. A node is called *hanging node*, if it is either a hanging inner node or a hanging boundary node. We refer to the *set of all regular (non-hanging) nodes* by $\mathcal{K}_L^{\mathrm{dof}}$. [1]

Every continuous function on $\Box$ which is piecewise bilinear with respect to $\mathcal{Q}_L$ is uniquely determined by choosing (arbitrary) values for the regular nodes and for the hanging boundary nodes (cf. Figure 4.2). Since the latter ones will be treated in a special way, only the regular nodes have been included in the set $\mathcal{K}_L^{\mathrm{dof}}$.

Definition 4.3 (edges, macro-edges, set $\mathcal{G}_L$)**.** Let K_1 and K_2 be nodes of $\mathcal{Q}_L$. The segment $K_1K_2 = \left\{(1 - \alpha)K_1 + \alpha K_2 : \alpha \in [0, 1]\right\}$ is called *edge* (of $\mathcal{Q}_L$), if it is a side of an element $Q \in \mathcal{Q}_L$ or if it is a macro-edge. A *macro-edge* is a segment K_1K_2 such that (i) its center C is a hanging node and (ii) both CK_1 and CK_2 are sides of elements in $\mathcal{Q}_L$. We refer to the *set of all edges* by $\mathcal{G}_L$.

There is a one-to-one correspondence between hanging nodes and macro-edges (see Figure 4.2). Macro-edges on the boundary are special in the sense that, for them, there are no elements $Q \in \mathcal{Q}_L$ such that they would coincide with one of the sides of Q, whereas for macro-edges in the interior, this is always the case. Correspondingly, hanging boundary nodes are distinct from hanging inner nodes by the fact that they are vertices of two elements $Q \in \mathcal{Q}_L$, but do not lie on the boundary of a third one. Figure 4.2 shows how the classifications of nodes and edges provided by Definitions 4.2 and 4.3, especially the nonstandard definitions of hanging boundary nodes and macro-edges, reflect the stationarity of the function to be approximated.

[1] Throughout, the superscript "dof" denotes "degree of freedom".

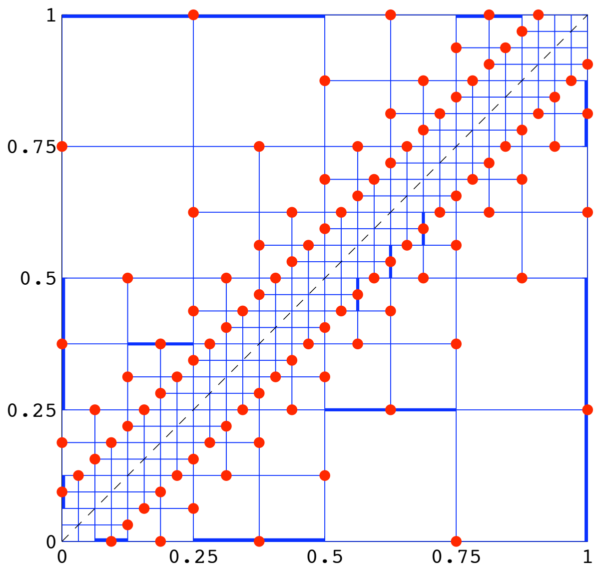

FIGURE 4.2. The partition $\mathcal{Q}_5$ along with Δ (dashed line). All hanging nodes according to Definition 4.2 are marked as dots. The fat lines are examples of the macro-edges from Definition 4.3.

The next two definitions equip the partition $\mathcal{Q}_L$ and the set of all edges $\mathcal{G}_L$ with additional structure. Later, this one will allow to control the growth of the polynomial degrees with increasing distance from the diagonal (p-refinement).

Definition 4.4 (decomposition of $\mathcal{Q}_L$; sets $\mathcal{Q}_L^{(i)}$, $\mathcal{Q}_L^{\mathrm{dof}}$; side lengths h_i)**.** For every partition $\mathcal{Q}_L$, $L \geq 2$, we define a complete decomposition $\mathcal{Q}_L = \bigcup_{i=1}^{L} \mathcal{Q}_L^{(i)}$, where

$$\mathcal{Q}_L^{(1)} := \Big\{ Q \in \mathcal{Q}_L : \mathrm{dist}(Q, \Delta) = 0 \Big\},$$
$$\mathcal{Q}_L^{(i)} := \Big\{ Q \in \mathcal{Q}_L : \mathrm{dist}(Q, \Delta) > 0 \wedge h_Q = 2^{-(L+2-i)} \Big\} \quad (2 \leq i \leq L).$$

For $i \in \{1, 2, \dots, L\}$, we denote by h_i the side length of elements $Q \in \mathcal{Q}_L^{(i)}$:

$$h_1 = 2^{-L}, \qquad h_i = 2^{-(L+2-i)} \ \ (2 \leq i \leq L)$$

Furthermore, all elements with a finite distance to the diagonal are grouped into the set

$$\mathcal{Q}_L^{\mathrm{dof}} := \bigcup_{i=2}^{L} \mathcal{Q}_L^{(i)}.$$

Notice that

$$\frac{h_i}{\sqrt{2}} = \min_{Q \in \mathcal{Q}_L^{(i)}} \mathrm{dist}(Q, \Delta) \quad \text{for all } i \in \{2, 3, \dots, L\}. \tag{4.1}$$

Definition 4.5 (sets $\mathcal{G}_L^{(i)}$, $\dot{\mathcal{G}}_L^{(i)}$, $\mathcal{G}_L^{\mathrm{dof}}$ and $\dot{\mathcal{G}}_L^{\mathrm{dof}}$)**.** For every partition $\mathcal{Q}_L$, $L \geq 2$, define the following subsets of $\mathcal{G}_L$:

$$\mathcal{G}_L^{(i)} := \Big\{ G \in \mathcal{G}_L : \exists Q_1, Q_2 \in \mathcal{Q}_L^{(i)} \text{ such that } Q_1 \cap Q_2 = G \Big\} \quad (2 \leq i \leq L)$$
$$\dot{\mathcal{G}}_L^{(i)} := \Big\{ G \in \mathcal{G}_L : G \text{ is macro-edge of length } 2^{-(L+1-i)} \Big\} \qquad (2 \leq i \leq L)$$

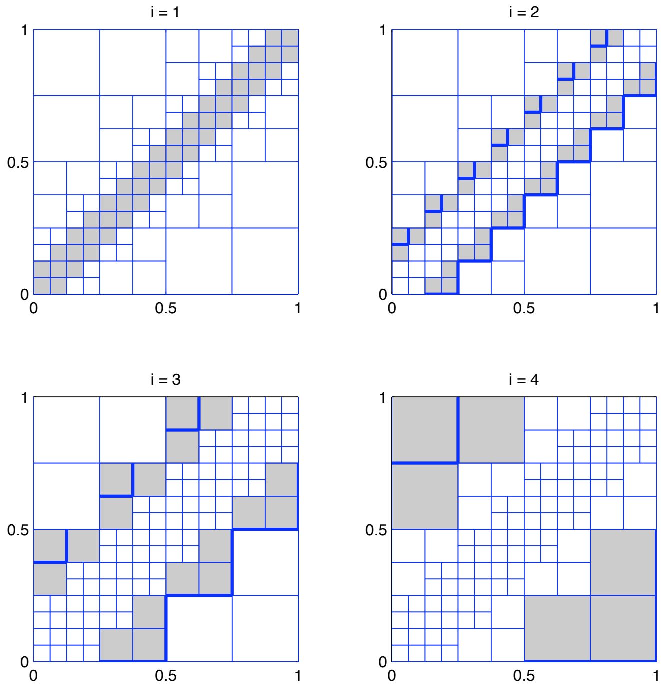

FIGURE 4.3. Decomposition of partition $\mathcal{Q}_4$ into $\mathcal{Q}_4^{(i)}$, $i = 1, \ldots, 4$ (shaded grey). In the graphs, for $i \geq 2$, the bold lines above the diagonal represent edges from $\mathcal{G}_4^{(i)}$, whereas the ones below are macro-edges from $\dot{\mathcal{G}}_4^{(i)}$.

Furthermore, we define:

$$\mathcal{G}_L^{\mathrm{dof}} := \bigcup_{i=2}^{L} \mathcal{G}_L^{(i)}, \qquad \dot{\mathcal{G}}_L^{\mathrm{dof}} := \bigcup_{i=2}^{L} \dot{\mathcal{G}}_L^{(i)}$$

The sets $\mathcal{Q}_L^{(i)}$, $\mathcal{G}_L^{(i)}$ and $\dot{\mathcal{G}}_L^{(i)}$ are visualized in part in Figure 4.3 for the case $L = 4$.

Definition 4.6 (set $\mathcal{F}_L$)**.** For every geometric refinement level $L \geq 2$, we collect all those geometric entities to which basis functions will be assigned into a set

$$\mathcal{F}_L := \mathcal{K}_L^{\mathrm{dof}} \cup \mathcal{G}_L^{\mathrm{dof}} \cup \dot{\mathcal{G}}_L^{\mathrm{dof}} \cup \mathcal{Q}_L^{\mathrm{dof}}.$$

For a reasonable approximation of a *stationary* function $w = w(x - y)$, certain degrees of freedom have to be constrained. The following definition prepares the realization of the constraint by grouping geometric entities that can be mapped into each other by a shift parallel to the diagonal:

Definition 4.7 (equivalence relation $\sim$)**.** On the set $\mathcal{F}_L$, an equivalence relation $\sim$ is defined through

$$F_1 \sim F_2 \quad \Leftrightarrow \quad \exists\, a \in \mathbb{R} : T_a(F_1) = F_2,$$

with $T_a : \mathbb{R}^2 \to \mathbb{R}^2, (x, y) \mapsto (x + a, y + a)$. $[F]$ denotes the equivalence class of F,

$$[F] = \left\{ \bar{F} \in \mathcal{F}_L : \bar{F} \sim F \right\},$$

$\mathcal{F}_L/\!\sim$ denotes the quotient set of $\mathcal{F}_L$ by $\sim$,

$$\mathcal{F}_L/\!\sim\; = \big\{[F] : F \in \mathcal{F}_L\big\}.$$

4.2. **Construction of *hp*-finite element spaces.** In analogy to Definition 2.5 (ansatz space on $I = [0, 1]$) and the setting of Theorem 2.10 (*hp*-approximation in one dimension), we would like to have approximating functions

$$v \in C^0(\Box) : v\big|_Q \in \mathcal{S}^{p_i,p_i} \text{ for all } Q \in \mathcal{Q}_L^{(i)},\ i \in \{1, \dots, L\}, \tag{4.2a}$$

where $\mathcal{S}^{p,p} := \mathcal{S}^p \otimes \mathcal{S}^p$ is the tensor product of the space from Definition 2.4 and the degree vector $\mathbf{p} = (p_1, \dots, p_L)$ is defined by

$$p_1 = 1, \qquad p_i = \lfloor \mu i \rfloor \ \ (2 \le i \le L), \tag{4.2b}$$

again with a slope $\mu \ge 1$. However, using (4.2) as definition of the finite element space would lead to its dimension growing as $\mathcal{O}(2^L)$, which is clearly unsatisfactory with regard to high refinement levels L. For this reason, the present subsection is concerned with the *selection of a subspace* $S^{\Delta}_{\mu,L}$ *of (4.2)* which exploits the fact that the function we want to approximate is stationary. The definition of this subspace is constructive in the sense that we shall obtain it by creating particular basis functions for (4.2) and then defining $S^{\Delta}_{\mu,L}$ as constrained span of these ones. Before starting with this procedure, we state the following:

Proposition 4.8 ($H^{1,1}$-conformity). *Any v satisfying (4.2a) belongs to $H^{1,1}(\Box)$.*

4.2.1. *Shape functions on the reference element.* Here, we provide basis functions for the space $\mathcal{S}^{p,p}$ on the reference element $\widehat{Q} := [-1, 1]^2$. They will be built from one-dimensional polynomials:

Definition 4.9 (basis functions for the polynomial spaces $\mathcal{S}^p$)**.** We denote by

$$\phi_0(\xi) = \frac{1-\xi}{2}, \quad \phi_1(\xi) = \frac{1+\xi}{2},$$

and

$$\phi_i(\xi) = \sqrt{\frac{2i-1}{2}} \int_{-1}^{\xi} L_{i-1}(\tilde{\xi})\, d\tilde{\xi}, \quad i = 2, 3, \dots \tag{4.3}$$

the normalized first antiderivatives of the Legendre polynomials.

For every $p \in \mathbb{N}$, the functions $\phi_0, \phi_1, \dots, \phi_p$ span the space $\mathcal{S}^p$ on the interval $[-1, 1]$. The integrated Legendre polynomials (4.3) are sometimes called *internal* basis functions, since they vanish in both endpoints of $[-1, 1]$ due to the orthogonality of L_{i-1} and $L_0 \equiv 1$ (see e.g. [11]). In contrast, ϕ_0 and ϕ_1 are referred to as *external* basis functions. In particular, they are called *nodal*, because they assume the value 1 in one endpoint of $[-1, 1]$ and 0 in the other:

$$\phi_0(-1) = 1, \qquad \phi_0(1) = 0, \qquad \phi_1(-1) = 0, \qquad \phi_1(1) = 1.$$

With a basis of $\mathcal{S}^p$ at hand, the definition of a basis for the tensor product space $\mathcal{S}^{p,p}$ on $\widehat{Q}$ is straightforward:

Definition 4.10 ((nodal/side/internal) shape functions, basis of $\mathcal{S}^{p,p}$)**.** Let $p \in \mathbb{N}$. We define 4 *nodal shape functions*

$$\begin{aligned} \overset{0}{N}_1(\xi, \eta) &= \phi_0(\xi)\, \phi_0(\eta), & \overset{0}{N}_2(\xi, \eta) &= \phi_0(\xi)\, \phi_1(\eta), \\ \overset{0}{N}_3(\xi, \eta) &= \phi_1(\xi)\, \phi_0(\eta), & \overset{0}{N}_4(\xi, \eta) &= \phi_1(\xi)\, \phi_1(\eta), \end{aligned} \tag{4.4a}$$

$\otimes$	$\phi_0(\eta)$	$\phi_1(\eta)$	$\phi_2(\eta)$	$\phi_3(\eta)$	$\phi_4(\eta)$	$\cdots$
$\phi_0(\xi)$	$\overset{0}{N}_1$	$\overset{0}{N}_2$	$\overset{1}{N}{}_2^{[1]}$	$\overset{1}{N}{}_3^{[1]}$	$\overset{1}{N}{}_4^{[1]}$	$\cdots$
$\phi_1(\xi)$	$\overset{0}{N}_3$	$\overset{0}{N}_4$	$\overset{1}{N}{}_2^{[2]}$	$\overset{1}{N}{}_3^{[2]}$	$\overset{1}{N}{}_4^{[2]}$	$\cdots$
$\phi_2(\xi)$	$\overset{1}{N}{}_2^{[3]}$	$\overset{1}{N}{}_2^{[4]}$	$\overset{2}{N}_{22}$	$\overset{2}{N}_{23}$	$\overset{2}{N}_{24}$	$\cdots$
$\phi_3(\xi)$	$\overset{1}{N}{}_3^{[3]}$	$\overset{1}{N}{}_3^{[4]}$	$\overset{2}{N}_{32}$	$\overset{2}{N}_{33}$	$\overset{2}{N}_{34}$	$\cdots$
$\phi_4(\xi)$	$\overset{1}{N}{}_4^{[3]}$	$\overset{1}{N}{}_4^{[4]}$	$\overset{2}{N}_{42}$	$\overset{2}{N}_{43}$	$\overset{2}{N}_{44}$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

TABLE 4.1. Index scheme for shape functions obtained by tensorization of the ϕ_i in Definition 4.10. The nodal shape functions (4.4a) can be found in the upper left 2-by-2 block, whereas the lower right block holds the internal shape functions (4.4c). Side shape functions (4.4b) are located in the two remaining blocks (lower left and upper right). The first $p+1$ rows and columns of the table constitute a basis of $\mathcal{S}^{p,p}$ on $\widehat{Q} = [-1,1]^2$.

$4\,(p-1)$ *side shape functions*

$$\begin{aligned} \overset{1}{N}{}_i^{[1]}(\xi,\eta) &= \phi_0(\xi)\,\phi_i(\eta), \\ \overset{1}{N}{}_i^{[2]}(\xi,\eta) &= \phi_1(\xi)\,\phi_i(\eta), \\ \overset{1}{N}{}_i^{[3]}(\xi,\eta) &= \phi_i(\xi)\,\phi_0(\eta), \\ \overset{1}{N}{}_i^{[4]}(\xi,\eta) &= \phi_i(\xi)\,\phi_1(\eta), \end{aligned} \qquad (2 \le i \le p) \tag{4.4b}$$

and $(p-1)^2$ *internal shape functions*

$$\overset{2}{N}_{ij}(\xi,\eta) = \phi_i(\xi)\,\phi_j(\eta), \quad 2 \le i,j \le p. \tag{4.4c}$$

Altogether, the functions (4.4) yield $(p+1)^2$ *shape functions* forming a basis of $\mathcal{S}^{p,p}$ on $\widehat{Q}$.

Table 4.1 illustrates how the shape functions are built from the ϕ_i.

4.2.2. *Global basis functions.* Now, we will assemble basis functions for the space (4.2) from the shape functions given above. In order to do that, we need to transport the shape functions, which were defined on a reference element $\widehat{Q}$, to the actual elements of the partition. This will be accomplished by the following maps:

Definition 4.11 (element mappings T_Q)**.** For every element $Q = [x_1,x_2]\times[y_1,y_2] \in \mathcal{Q}_L$, we define a linear mapping

$$T_Q : \widehat{Q} \to Q, \qquad \begin{pmatrix}\xi\\ \eta\end{pmatrix} \mapsto \frac{1}{2}\begin{pmatrix}x_1+x_2\\ y_1+y_2\end{pmatrix} + \frac{h_Q}{2}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\begin{pmatrix}\xi\\ \eta\end{pmatrix},$$

where $h_Q = x_2 - x_1 = y_2 - y_1$ is the side length of Q.

Obviously, T_Q maps the lower/upper left/right vertex of $\widehat{Q}$ into the lower/upper left/right vertex of Q, respectively:

$$\begin{aligned} T_Q(-1,-1) &= (x_1,y_1), \qquad & T_Q(-1,+1) &= (x_1,y_2) \\ T_Q(+1,-1) &= (x_2,y_1), \qquad & T_Q(+1,+1) &= (x_2,y_2) \end{aligned}$$

The below Definitions 4.12, 4.13, 4.14 and 4.15 will present particular basis functions for (4.2), grouped into four categories: *node-associated*, *macro-edge-associated*, *edge-associated* and *element-associated*. Due to the continuity requirement in (4.2a) and the presence of hanging nodes in $\mathcal{Q}_L$, the actual assembly of the node- and macro-edge-associated basis functions is a rather tricky task. In order to reduce technicalities to a minimum, these ones are defined implicitly:

Definition 4.12 (node-associated basis functions ψ^K)**.** For every $K \in \mathcal{K}_L^{\mathrm{dof}}$, we define a (unique) function $\psi^K : \square \to \mathbb{R}$ by:

$$\begin{aligned} &(i) && \psi^K \in C^0(\square) \\ &(ii) && \psi^K\big|_Q \in \mathcal{S}^{1,1} \quad \forall Q \in \mathcal{Q}_L \\ &(iii) && \psi^K(\hat{K}) = \delta_{K,\hat{K}} \quad \forall \hat{K} \in \mathcal{K}_L^{\mathrm{dof}} \\ &(iv) && \psi^K\big|_G \text{ is linear on all macro-edges } G \subset \partial\square \end{aligned}$$

Properties (iii) and (iv) prescribe values for non-hanging nodes and hanging boundary nodes, respectively. Since, by (i) and (ii), the function ψ^K is continuous and piecewise bilinear, these prescribed values determine ψ^K uniquely. In practice, the ψ^K are assembled from nodal shape functions and proper element mappings.

Definition 4.13 (macro-edge-associated basis functions ψ_k^G)**.** For every $G \in \dot{\mathcal{G}}_L^{\mathrm{dof}}$, let K_G denote the hanging node in its center. With ϕ_k given by (4.3), we define families $(\psi_k^G)_{k\geq 2}$ of (unique) functions $\psi_k^G : \square \to \mathbb{R}$ by:

$$(i) \qquad \psi_k^G \in C^0(\square)$$

$$(ii) \qquad \psi_k^G = 0 \begin{cases} \text{in all elements } Q \in \mathcal{Q}_L \text{ with } \operatorname{dist}(K_G, Q) > 0 \\ \text{on all edges } \hat{G} \in \mathcal{G}_L \text{ with } \operatorname{dist}(K_G, \hat{G}) > 0 \end{cases}$$

$$(iii) \qquad \psi_k^G\big|_G(x,y) = \begin{cases} \phi_k\left(\frac{x-x_1}{x_2-x_1} - \frac{x_2-x}{x_2-x_1}\right) & \text{if } G = [x_1,x_2] \times \{y\} \\ \phi_k\left(\frac{y-y_1}{y_2-y_1} - \frac{y_2-y}{y_2-y_1}\right) & \text{if } G = \{x\} \times [y_1,y_2] \end{cases}$$

$$(iv) \qquad \text{for } Q \in \mathcal{Q}_L \text{ with } \operatorname{dist}(K_G, Q) = 0 \,:\, \psi_k^G\big|_Q \in \begin{cases} \mathcal{S}^{k,1} & \text{if } G = [x_1,x_2] \times \{y\} \\ \mathcal{S}^{1,k} & \text{if } G = \{x\} \times [y_1,y_2] \end{cases}$$

In practice, the ψ_k^G are assembled from nodal and side shape functions.

Definition 4.14 (edge-associated basis functions ψ_k^G)**.** For every $G \in \mathcal{G}_L^{\mathrm{dof}}$, there are exactly two elements in $\mathcal{Q}_L$ which G is a side of. Case (a): If G is parallel to the x-axis, denote the element above G by Q_G^N and the one below by Q_G^S. For all $k \geq 2$, we define $\psi_k^G : \square \to \mathbb{R}$ by:

$$\psi_k^G(x,y) = \begin{cases} \overset{1}{N}{}_k^{[3]}\left(\left(T_{Q_G^N}\right)^{-1}(x,y)\right) & \text{if } (x,y) \in Q_G^N \\ \overset{1}{N}{}_k^{[4]}\left(\left(T_{Q_G^S}\right)^{-1}(x,y)\right) & \text{if } (x,y) \in Q_G^S \\ 0 & \text{else} \end{cases}$$

Case (b): If G is parallel to the y-axis, denote the element to the left of G by Q_G^W and the one to the right by Q_G^E. For all $k \geq 2$, we define $\psi_k^G : \square \to \mathbb{R}$ by:

$$\psi_k^G(x,y) = \begin{cases} \overset{1}{N}{}_k^{[2]}\left(\left(T_{Q_G^W}\right)^{-1}(x,y)\right) & \text{if } (x,y) \in Q_G^W \\ \overset{1}{N}{}_k^{[1]}\left(\left(T_{Q_G^E}\right)^{-1}(x,y)\right) & \text{if } (x,y) \in Q_G^E \\ 0 & \text{else} \end{cases}$$

Definition 4.15 (element-associated basis functions ψ_{kl}^{Q})**.** For every element $Q \in \mathcal{Q}_L^{\text{dof}}$, we define a family $(\psi_{kl}^{Q})_{k,l\geq 2}$ of functions $\psi_{kl}^{Q} : \square \to \mathbb{R}$ by:

$$\psi_{kl}^{Q}(x,y) = \begin{cases} \overset{2}{N}_{kl}\left(T_Q^{-1}(x,y)\right) & \text{if } (x,y) \in Q \\ 0 & \text{else} \end{cases}$$

4.2.3. *Finite element spaces.* With the set $\mathcal{F}_L = \mathcal{K}_L^{\text{dof}} \cup \dot{\mathcal{G}}_L^{\text{dof}} \cup \mathcal{G}_L^{\text{dof}} \cup \mathcal{Q}_L^{\text{dof}}$, the space (4.2) can now be written as

$$\left\{ v = \sum_{F\in\mathcal{F}_L} \sum_{j\in I_\mu(F)} c_j^F \psi_j^F \,:\, c_j^F \in \mathbb{R} \right\}. \tag{4.5}$$

Herein, the ψ_j^F represent the basis functions that were associated with the geometric entities $F \in \mathcal{F}_L$ in Definitions 4.12–4.15. The index set $I_\mu(F)$ is chosen such that the polynomial degrees (4.2b) are obtained:

- $F = K \in \mathcal{K}_L^{\text{dof}}$: There is only one basis function per non-hanging node K. Thus, ψ_j^F equals ψ^K, and we define $I_\mu(K) \equiv \{1\}$ (or any other set containing exactly one element).
- $F = G \in \dot{\mathcal{G}}_L^{\text{dof}}$: If $G \in \dot{\mathcal{G}}_L^{(i)}$, $2 \leq i \leq L$, define $I_\mu(G) = \{2, 3, \ldots, \lfloor \mu i \rfloor\}$.
- $F = G \in \mathcal{G}_L^{\text{dof}}$: If $G \in \mathcal{G}_L^{(i)}$, $2 \leq i \leq L$, define $I_\mu(G) = \{2, 3, \ldots, \lfloor \mu i \rfloor\}$.
- $F = Q \in \mathcal{Q}_L^{\text{dof}}$: In this case, j represents a double index (k,l). The corresponding index set for an element $Q \in \mathcal{Q}_L^{(i)}$, $2 \leq i \leq L$, is given by $I_\mu(Q) = \{(k,l) \in \mathbb{N}^2 : 2 \leq k, l \leq \lfloor \mu i \rfloor\}$.

Notice that two sets $I_\mu(F_1)$ and $I_\mu(F_2)$ are guaranteed to be identical, if $F_1 \sim F_2$, where $\sim$ denotes the equivalence relation from Definition 4.7. For this reason, we may use the sloppy notation $I_\mu([F])$ instead of $I_\mu(F)$ in the following. Another way of writing (4.5) thus is:

$$\left\{ v = \sum_{[F]\in\mathcal{F}_L/\sim} \sum_{j\in I_\mu([F])} \sum_{\bar{F}\in[F]} c_j^{\bar{F}} \psi_j^{\bar{F}} \,:\, c_j^{\bar{F}} \in \mathbb{R} \right\} \tag{4.6}$$

Definition 4.16 (space $S_{\mu,L}^{\Delta}$)**.** We constrain the coefficients $c_j^{\bar{F}}$ in (4.6) by the rule

$$c_{j_1}^{\bar{F}_1} = c_{j_2}^{\bar{F}_2} \quad \Leftrightarrow \quad \bar{F}_1 \sim \bar{F}_2 \,\wedge\, j_1 = j_2. \tag{4.7}$$

The resulting space is called $S_{\mu,L}^{\Delta}$ and can be written as

$$S_{\mu,L}^{\Delta} = \left\{ v = \sum_{[F]\in\mathcal{F}_L/\sim} \sum_{j\in I_\mu([F])} c_j^{[F]} \sum_{\bar{F}\in[F]} \psi_j^{\bar{F}} \,:\, c_j^{[F]} \in \mathbb{R} \right\}. \tag{4.8}$$

In particular, $S_{\mu,L}^{\Delta} \subset H^{1,1}(\square)$ due to Proposition 4.8.

In (4.6), every geometric entity had its own set of coefficients, whereas in (4.8), it is one set for the whole equivalence class. This makes a significant difference concerning the dimension of the space:

Lemma 4.17 (number of degrees of freedom in $S_{\mu,L}^{\Delta}$)**.** *For a fixed $\mu \in [1,\infty)$, the dimension of the space $S_{\mu,L}^{\Delta}$ scales as follows:*

$$N := \dim S_{\mu,L}^{\Delta} = \mathcal{O}(L^3)$$

Proof. See [5, Lemma 5.20]. $\square$

4.3. **Interpolation operator and basic approximation results.** For the results presented here, see also [8, Section 4.5.4].

Definition 4.18 (operator Π_p)**.** Let $p \in \mathbb{N}$, $\mathcal{S}^{p,p} := \mathcal{S}^p \otimes \mathcal{S}^p$ the tensor product of the space from Definition 2.4 and $\widehat{Q} = [-1,1]^2$ the reference element. Then, we define the mapping $\Pi_p : H^{1,1}(\widehat{Q}) \to \mathcal{S}^{p,p}$ as the tensor product analog of the one-dimensional projector π_p from Definition 2.6:

$$\Pi_p := \pi_p^{(x)} \otimes \pi_p^{(y)}$$

Lemma 4.19 (interpolation properties of Π_p)**.** *For every $\hat{w} \in H^{1,1}(\widehat{Q})$, $\Pi_p \hat{w}$ interpolates $\hat{w}$ in the vertices of $\widehat{Q}$:*

$$(\Pi_p \hat{w})(\pm 1, \pm 1) = \hat{w}(\pm 1, \pm 1) \tag{4.9a}$$

Furthermore, it holds

$$(\Pi_p \hat{w})(\cdot, \pm 1) = (\pi_p^{(x)} \hat{w})(\cdot, \pm 1), \tag{4.9b}$$

$$(\Pi_p \hat{w})(\pm 1, \cdot) = (\pi_p^{(y)} \hat{w})(\pm 1, \cdot), \tag{4.9c}$$

i.e. $\Pi_p \hat{w}$ evaluated on a side γ of $\widehat{Q}$ equals the corresponding 1D projection of $\hat{w}|_\gamma$.

Proof. See [8, Lemma 4.67]. □

The following three lemmas analyze the different components of the $H^{1,1}$-approximation error of Π_p. Throughout, we assume $\hat{w} \in H^{k+1}(\widehat{Q})$ with a $k \in \mathbb{N}$, so that in particular $\hat{w} \in H^{1,1}(\widehat{Q})$ is guaranteed (cf. Definition 3.2).

Lemma 4.20 (L^2-error of Π_p)**.** *Let $k \in \mathbb{N}$, $\hat{w} \in H^{k+1}(\widehat{Q})$. Then, the projector Π_p satisfies*

$$\begin{aligned}\|\hat{w} - \Pi_p \hat{w}\|^2_{L^2(\widehat{Q})} \leq{}& \frac{2}{p(p+1)} \frac{(p-s)!}{(p+s)!} \|\partial_\xi^{s+1} \hat{w}\|^2_{L^2(\widehat{Q})} \\ &+ \frac{4}{p(p+1)} \frac{(p-s)!}{(p+s)!} \|\partial_\eta^{s+1} \hat{w}\|^2_{L^2(\widehat{Q})} \\ &+ \frac{4}{p^2(p+1)^2} \frac{(p-s+1)!}{(p+s-1)!} \|\partial_\xi \partial_\eta^s \hat{w}\|^2_{L^2(\widehat{Q})}\end{aligned}$$

for any integer s with $1 \leq s \leq \min\{p, k\}$.

Proof. See [5, Lemma 5.24]. □

Lemma 4.21 (H^1-error of Π_p)**.** *Let $k \in \mathbb{N}$ and $\hat{w} \in H^{k+1}(\widehat{Q})$. Then, the projector Π_p satisfies*

$$\begin{aligned}\|\partial_\xi(\hat{w} - \Pi_p \hat{w})\|^2_{L^2(\widehat{Q})} \leq{}& 2 \frac{(p-s)!}{(p+s)!} \|\partial_\xi^{s+1} \hat{w}\|^2_{L^2(\widehat{Q})} \\ &+ \frac{2}{p(p+1)} \frac{(p-s+1)!}{(p+s-1)!} \|\partial_\xi \partial_\eta^s \hat{w}\|^2_{L^2(\widehat{Q})}\end{aligned}$$

for any integer s with $1 \leq s \leq \min\{p, k\}$. An analogous estimate holds for $\|\partial_\eta(\hat{w} - \Pi_p \hat{w})\|_{L^2(\widehat{Q})}$.

Proof. See [8, Lemma 4.67]. □

Lemma 4.22 ($H^{1,1}$-error of Π_p)**.** *Let $k \in \mathbb{N}$ and $\hat{w} \in H^{k+1}(\widehat{Q})$. Then, the projector Π_p satisfies*

$$\|\partial_\xi \partial_\eta(\hat{w} - \Pi_p \hat{w})\|^2_{L^2(\widehat{Q})} \leq 2 \frac{(p-s)!}{(p+s)!} \|\partial_\eta \partial_\xi^{s+1} \hat{w}\|^2_{L^2(\widehat{Q})} + 2 \frac{(p-s)!}{(p+s)!} \|\partial_\xi \partial_\eta^{s+1} \hat{w}\|^2_{L^2(\widehat{Q})}$$

for any integer s with $0 \leq s \leq \min\{p, k-1\}$.

Proof. See [5, Lemma 5.26]. □

4.4. **Construction of the approximating function.** This subsection defines a function v_L^Δ suited for the approximation of a $w : \Box \to \mathbb{R}$ satisfying Assumption 3.1. The construction of v_L^Δ consists of two steps: In the first one, we build an approximation $\tilde{v}_L^\Delta$ from element-wise projections in analogy to the procedure in the one-dimensional case, which was described in Section 2. Due to the presence of hanging nodes in the partition, $\tilde{v}_L^\Delta$ will in general be discontinuous. For this reason, in a second step, we add particular functions in order to *lift* the discontinuities and thus get a continuous approximation v_L^Δ. The finally obtained v_L^Δ is guaranteed to belong to the space $S_{\mu,L}^\Delta$ from Definition 4.16, provided that the function to be approximated is stationary.

4.4.1. *Discontinuous approximation.*

Definition 4.23 (local approximation on an element Q)**.** Let $T_Q : \widehat{Q} \to Q$ be the element mapping introduced in Definition 4.11. With $\mathbf{p} = (p_1, \ldots, p_L)$ as in (4.2b), a staggered degree vector $\tilde{\mathbf{p}} = (\tilde{p}_1, \ldots, \tilde{p}_L)$ shall be given by

$$\tilde{p}_1 = \tilde{p}_2 = p_1 = 1, \qquad \tilde{p}_i = p_{i-1} = \lfloor \mu(i-1) \rfloor \;\; (3 \le i \le L). \tag{4.10}$$

Then, for $w \in H^{1,1}(\Box)$, we define its local approximation v_Q on $Q \in \mathcal{Q}_L^{(i)}$, $i \in \{1, \ldots, L\}$, by:

$$\hat{w}_Q := w \circ T_Q \tag{4.11a}$$

$$\hat{v}_Q := \Pi_{\tilde{p}_i} \hat{w}_Q \tag{4.11b}$$

$$v_Q := \hat{v}_Q \circ T_Q^{-1} \tag{4.11c}$$

Definition 4.24 (discontinuous interpolant $\tilde{v}_L^\Delta$)**.** For $w \in H^{1,1}(\Box)$, we define an approximation $\tilde{v}_L^\Delta : \Box \to \mathbb{R}$ element-wise by

$$\tilde{v}_L^\Delta\big|_Q = v_Q \quad (Q \in \mathcal{Q}_L),$$

where v_Q is the local approximation of w on element Q from Definition 4.23.

Remark 4.25 (discontinuities in $\tilde{v}_L^\Delta$). From the construction (4.11) and the interpolation properties (4.9b), (4.9c) of Π_p, it becomes clear that $\tilde{v}_L^\Delta$ is ambiguous on a macro-edge, whenever the situation in Figure 4.4 appears: Two small elements Q_k ($k = 2, 3$) with $\tilde{v}_L^\Delta|_{Q_k} = v_{Q_k} \in \mathcal{S}^{p',p'}$ border on a larger element Q_1, on which we have $\tilde{v}_L^\Delta|_{Q_1} = v_{Q_1} \in \mathcal{S}^{p,p}$. Since the larger element must belong to a $\mathcal{Q}_L^{(i+1)}$ with $i \ge 2$ ($\mathcal{Q}_L^{(1)}$ and $\mathcal{Q}_L^{(2)}$ contain only elements of minimum size), its polynomial degree $p = \tilde{p}_{i+1} = p_i$ is higher than $p' = \tilde{p}_i = p_{i-1}$. A jump in $\tilde{v}_L^\Delta$ occurs along the macro-edge $\gamma := \overline{\gamma_{12} \cup \gamma_{13}}$ (see Figure 4.4). This jump lies in the space $H_0^1(\gamma) \cap \big(\mathcal{S}^{p_i}(\gamma_{12}) \cap \mathcal{S}^{p_i}(\gamma_{13})\big)$, i.e. it is piecewise polynomial of degree p_i and vanishes in the two endpoints of γ.

Remark 4.26 ("stationarity" of $\tilde{v}_L^\Delta$). For a stationary $w = w(x-y)$, the function $\tilde{v}_L^\Delta$ will in general *not* be stationary. However, because of (4.11a) and (4.11b), stationarity of w at least implies

$$Q_1 \sim Q_2 \quad \Rightarrow \quad \hat{v}_{Q_1} = \hat{v}_{Q_2},$$

where $\sim$ denotes the equivalence relation from Definition 4.7. Thus, $\tilde{v}_L^\Delta$ is identical on all elements Q of the same equivalence class—in the sense that the local approximations $v_Q = \hat{v}_Q \circ T_Q^{-1}$ are just shifted copies of each other.

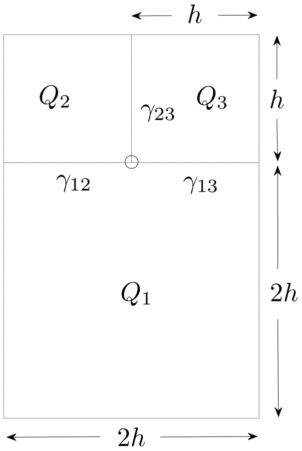

FIGURE 4.4. Hanging node ($\circ$) and adjacent elements. Notice that, for macro-edges lying on the boundary of $\square$, there is no element $Q_1 \in \mathcal{Q}_L$, whereas the elements Q_2 and Q_3 always exist.

4.4.2. *Lifting of the discontinuities.* As explained in Remark 4.25, the approximation $\tilde{v}_L^\Delta$ exhibits discontinuities along macro-edges and thus cannot belong to the finite element space $S_{\mu,L}^\Delta \subset C^0(\square)$. The apparent remedy to eliminate the discontinuities is to add particular functions, so-called *polynomial trace liftings*, to $\tilde{v}_L^\Delta$. Before we make this precise, we need the following:

Definition 4.27 (jump)**.** Let $\tilde{v}_L^\Delta$ be the discontinuous interpolant from Definition 4.24, and $\dot{\mathcal{G}}_L^{\mathrm{dof}} = \bigcup_{i=2}^L \dot{\mathcal{G}}_L^{(i)}$ shall denote the set of all macro-edges introduced in Definition 4.5. Furthermore, we recall the degree vector $\mathbf{p} = (p_1, \ldots, p_L)$ from (4.2b):

$$p_1 = 1, \qquad p_i = \lfloor \mu i \rfloor \;\; (2 \le i \le L)$$

Case (i): With every $\gamma \in \dot{\mathcal{G}}_L^{(i)}$ $(2 \le i \le L)$, $\gamma \not\subseteq \partial\square$, we associate the jump

$$\big[\tilde{v}_L^\Delta\big]_\gamma := \big(\tilde{v}_L^\Delta\big)_+^\gamma - \big(\tilde{v}_L^\Delta\big)_-^\gamma\,, \tag{4.12}$$

where $(\tilde{v}_L^\Delta)_+^\gamma$ is the limit of $\tilde{v}_L^\Delta$ when approaching γ from the large element in Figure 4.4, and $(\tilde{v}_L^\Delta)_-^\gamma$ the limit of $\tilde{v}_L^\Delta$ when approaching γ from the two small elements. In particular, it holds:

$$\big[\tilde{v}_L^\Delta\big]_\gamma \in H_0^1(\gamma) \cap \big(\mathcal{S}^{p_i}(\gamma_{12}) \cap \mathcal{S}^{p_i}(\gamma_{13})\big)$$

Case (ii): For a $\gamma \in \dot{\mathcal{G}}_L^{(i)}$ $(2 \le i \le L)$ with $\gamma \subset \partial\square$, the partition $\mathcal{Q}_L$ does not feature a large element as the one depicted in Figure 4.4. Thus, the expression $(\tilde{v}_L^\Delta)_+^\gamma$ in (4.12) is not meaningful. As a substitute for $(\tilde{v}_L^\Delta)_+^\gamma$, we define v_γ by

$$\hat{w}_\gamma := w\big|_\gamma \circ T_\gamma, \quad \hat{v}_\gamma := \pi_{p_i}\hat{w}_\gamma, \quad v_\gamma := \hat{v}_\gamma \circ T_\gamma^{-1},$$

where T_γ maps $\widehat{\Omega} = [-1, 1]$ to γ linearly in analogy to Definition 4.11, and π_{p_i} is the one-dimensional projection (2.4) of $H^1(\widehat{\Omega})$ onto $\mathcal{S}^{p_i}$. In this way, we associate a "virtual" jump

$$\left[\tilde{v}_L^{\Delta}\right]_\gamma := v_\gamma - \left(\tilde{v}_L^{\Delta}\right)_-^\gamma \left[\tilde{v}_L^{\Delta}\right]_\gamma \in H_0^1(\gamma) \cap \left(\mathcal{S}^{p_i}(\gamma_{12}) \cap \mathcal{S}^{p_i}(\gamma_{13})\right)$$

with every $\gamma \in \dot{\mathcal{G}}_L^{\mathrm{dof}}$, $\gamma \subset \partial\square$.

Remark 4.28 ("stationarity" of jumps). Thanks to the construction in case (ii) of Definition 4.27, we have the notion of a jump $[\tilde{v}_L^{\Delta}]_\gamma$ for *all* macro-edges $\gamma \in \dot{\mathcal{G}}_L^{\mathrm{dof}}$, regardless whether they lie on the boundary or not. For a stationary $w = w(x-y)$, it is furthermore guaranteed that

$$\gamma_1 \sim \gamma_2 \quad \Rightarrow \quad \left[\tilde{v}_L^{\Delta}\right]_{\gamma_1} = \left[\tilde{v}_L^{\Delta}\right]_{\gamma_2},$$

i.e. the jump is the same for all those macro-edges that belong to the same equivalence class.

Lemma 4.29 (polynomial trace lifting V_γ)**.** *Let $[\tilde{v}_L^{\Delta}]_\gamma$ be the jump at a macro-edge $\gamma \in \dot{\mathcal{G}}_L^{(i)}$, $i \in \{2, 3, \ldots, L\}$, from Definition 4.27. $Q_2, Q_3 \in \mathcal{Q}_L^{(i)}$ shall be the two small elements bordering on γ as in Figure 4.4. Then, there exists a function $V_\gamma : Q_2 \cup Q_3 \to \mathbb{R}$ with properties*

- *$V_\gamma \in \mathcal{S}^{p_i,1}(Q_2) \cap \mathcal{S}^{p_i,1}(Q_3)$, if γ is parallel to the x-axis,* *$V_\gamma \in \mathcal{S}^{1,p_i}(Q_2) \cap \mathcal{S}^{1,p_i}(Q_3)$, if γ is parallel to the y-axis,*
- *$V_\gamma \in C^0(Q_2 \cup Q_3)$,*
- $V_\gamma = \begin{cases} \left[\tilde{v}_L^{\Delta}\right]_\gamma & \text{on } \gamma, \\ 0 & \text{on } \partial(Q_2 \cup Q_3) \backslash \gamma, \end{cases}$

and

$$\|V_\gamma\|^2_{H^{1,1}(Q_2 \cup Q_3)} \le C\, h^{-1} \left|[\tilde{v}_L^{\Delta}]_\gamma\right|^2_{H^1(\gamma)}, \tag{4.13}$$

where $h = h_{Q_2} = h_{Q_3}$.

Proof. Let $g(z)$ be an affine linear function in the coordinate direction z perpendicular to γ, i.e. $z = x$, if γ is oriented vertically, and $z = y$, if horizontally. $g(z)$ shall be such that it is equal to 1 on γ and equal to 0 on that side of $Q_2 \cup Q_3$ which lies opposite of γ. The interval of the real line where g assumes values $g(z) \in [0, 1]$ shall be $[z_1, z_2]$. Notice that $h = z_2 - z_1$. Now, we define:

$$V_\gamma := \begin{cases} \left[\tilde{v}_L^{\Delta}\right]_\gamma \otimes g & (\gamma \text{ oriented horizontally}) \\ g \otimes \left[\tilde{v}_L^{\Delta}\right]_\gamma & (\gamma \text{ oriented vertically}) \end{cases}$$

Since $[\tilde{v}_L^{\Delta}]_\gamma \in H_0^1(\gamma) \cap \left(\mathcal{S}^{p_i}(\gamma_{12}) \cap \mathcal{S}^{p_i}(\gamma_{13})\right)$, V_γ obviously satisfies the first three properties in the lemma. Concerning the bound (4.13), its left hand side can be rewritten as follows:

$$\begin{aligned} \|V_\gamma\|^2_{H^{1,1}(Q_2 \cup Q_3)} &= \|g\|^2_{H^1([z_1,z_2])} \|[\tilde{v}_L^{\Delta}]_\gamma\|^2_{H^1(\gamma)} \\ &= \left(\|g\|^2_{L^2([z_1,z_2])} + |g|^2_{H^1([z_1,z_2])}\right)\left(\|[\tilde{v}_L^{\Delta}]_\gamma\|^2_{L^2(\gamma)} + \left|[\tilde{v}_L^{\Delta}]_\gamma\right|^2_{H^1(\gamma)}\right) \\ &= \left(\frac{h}{3} + \frac{1}{h}\right)\left(\|[\tilde{v}_L^{\Delta}]_\gamma\|^2_{L^2(\gamma)} + \left|[\tilde{v}_L^{\Delta}]_\gamma\right|^2_{H^1(\gamma)}\right) \end{aligned}$$

The length of γ being $2h$, application of the Poincaré–Friedrichs inequality provides

$$\|[\tilde{v}_L^{\Delta}]_\gamma\|^2_{L^2(\gamma)} \le (2h)^2 \left|[\tilde{v}_L^{\Delta}]_\gamma\right|^2_{H^1(\gamma)}.$$

Substituting this estimate into the above expression gives the claim (4.13). $\square$

Remark 4.30. The reason for having h^{-1} instead of h in (4.13) is that the trace lifting function V_γ is measured in the $H^{1,1}$-norm instead of the H^1-norm. For the convergence analysis to follow later, this constitutes a major difference to the standard case which is described e.g. in [8, Section 4.5.4] or [4, Section 3.3].

Definition 4.31 (continuous interpolant v_L^Δ)**.** For $w \in H^{1,1}(\Box)$, let $\tilde{v}_L^\Delta$ be the discontinuous interpolant from Definition 4.24. With every macro-edge $\gamma \in \dot{\mathcal{G}}_L^{\mathrm{dof}}$, we associate the trace lifting V_γ constructed in the proof of Lemma 4.29. Then, a continuous interpolant v_L^Δ is defined through:

$$v_L^\Delta := \tilde{v}_L^\Delta + \sum_{\gamma \in \dot{\mathcal{G}}_L^{\mathrm{dof}}} V_\gamma$$

Proposition 4.32. *For a stationary $w = w(x-y)$, the interpolant v_L^Δ belongs to the space $S_{\mu,L}^\Delta$ from Definition 4.16.*

Proof. See [5, Proposition 5.36]. $\Box$

4.5. **Local error estimates.** In this subsection, we give estimates for the $H^{1,1}$-error of the local approximations v_Q from Definition 4.23. Because of

$$\begin{aligned} \|w - v_Q\|_{H^{1,1}(Q)}^2 = \|w - v_Q\|_{L^2(Q)}^2 + \|\partial_x(w - v_Q)\|_{L^2(Q)}^2 \\ + \|\partial_y(w - v_Q)\|_{L^2(Q)}^2 + \|\partial_x\partial_y(w - v_Q)\|_{L^2(Q)}^2, \end{aligned} \tag{4.14}$$

we obtain these estimates from finding upper bounds for the individual summands in (4.14). For this reason, the first three lemmas of this section are dedicated to exactly that purpose. They are nothing but scaled versions of the approximation results in Lemmas 4.20, 4.21 and 4.22:

Lemma 4.33. *Let $Q \in \mathcal{Q}_L^{(i)}$, $i \in \{1, \dots, L\}$, and $w \in H^{k+1}(Q)$ with $k \in \mathbb{N}$. Then, for the local approximation $v_Q \in \mathcal{S}^{\tilde{p}_i,\tilde{p}_i}$, it holds*

$$\begin{aligned} \|w - v_Q\|_{L^2(Q)}^2 \le & \frac{2}{\tilde{p}_i(\tilde{p}_i+1)} \frac{(\tilde{p}_i - s)!}{(\tilde{p}_i + s)!} \left(\frac{h_i}{2}\right)^{2s+2} \|\partial_x^{s+1} w\|_{L^2(Q)}^2 \\ & + \frac{4}{\tilde{p}_i(\tilde{p}_i+1)} \frac{(\tilde{p}_i - s)!}{(\tilde{p}_i + s)!} \left(\frac{h_i}{2}\right)^{2s+2} \|\partial_y^{s+1} w\|_{L^2(Q)}^2 \\ & + \frac{4}{\tilde{p}_i^2(\tilde{p}_i+1)^2} \frac{(\tilde{p}_i - s + 1)!}{(\tilde{p}_i + s - 1)!} \left(\frac{h_i}{2}\right)^{2s+2} \|\partial_x \partial_y^s w\|_{L^2(Q)}^2 \end{aligned}$$

for any integer s with $1 \le s \le \min\{\tilde{p}_i, k\}$.

Proof. The element mapping $T_Q : \widehat{Q} \to Q$ from Definition 4.11 has Jacobian $\left(\frac{h_Q}{2}\right)^2 = \left(\frac{h_i}{2}\right)^2$. Thus, the claim follows from scaling Q to the reference element $\widehat{Q}$, employing the error estimate of Lemma 4.20 on $\widehat{Q}$ and retransforming from $\widehat{Q}$ to Q. $\Box$

Lemma 4.34. *Let $Q \in \mathcal{Q}_L^{(i)}$, $i \in \{1, \dots, L\}$, and $w \in H^{k+1}(Q)$ with $k \in \mathbb{N}$. Then, for the local approximation $v_Q \in \mathcal{S}^{\tilde{p}_i,\tilde{p}_i}$, we have*

$$\begin{aligned} \|\partial_x(w - v_Q)\|_{L^2(Q)}^2 \le & \, 2 \frac{(\tilde{p}_i - s)!}{(\tilde{p}_i + s)!} \left(\frac{h_i}{2}\right)^{2s} \|\partial_x^{s+1} w\|_{L^2(Q)}^2 \\ & + \frac{2}{\tilde{p}_i(\tilde{p}_i+1)} \frac{(\tilde{p}_i - s + 1)!}{(\tilde{p}_i + s - 1)!} \left(\frac{h_i}{2}\right)^{2s} \|\partial_x \partial_y^s w\|_{L^2(Q)}^2 \end{aligned}$$

for any integer s with $1 \le s \le \min\{\tilde{p}_i, k\}$. A corresponding estimate holds for $\|\partial_y(w - v_Q)\|_{L^2(Q)}$.

Proof. Analogous to the proof of Lemma 4.33, employing Lemma 4.21 in the middle step. $\Box$

Lemma 4.35. *Let $Q \in \mathcal{Q}_L^{(i)}$, $i \in \{1, \ldots, L\}$, and $w \in H^{k+1}(Q)$ with $k \in \mathbb{N}$. Then, for the local approximation $v_Q \in \mathcal{S}^{\tilde{p}_i, \tilde{p}_i}$, it holds*

$$\begin{aligned} \|\partial_x \partial_y (w - v_Q)\|_{L^2(Q)}^2 \leq 2\, \frac{(\tilde{p}_i - s)!}{(\tilde{p}_i + s)!} \left(\frac{h_i}{2}\right)^{2s} \|\partial_y \partial_x^{s+1} w\|_{L^2(Q)}^2 \\ + 2\, \frac{(\tilde{p}_i - s)!}{(\tilde{p}_i + s)!} \left(\frac{h_i}{2}\right)^{2s} \|\partial_x \partial_y^{s+1} w\|_{L^2(Q)}^2 \end{aligned}$$

for any integer s with $0 \leq s \leq \min\{\tilde{p}_i, k-1\}$.

Proof. Analogous to the proof of Lemma 4.33, employing Lemma 4.22 in the middle step. □

Lemma 4.36 ($H^{1,1}$-approximation error of v_Q)**.** *Let $Q \in \mathcal{Q}_L^{(i)}$, $i \in \{1, \ldots, L\}$, and w shall be in $H^{k+1}(Q)$ for every $k \in \mathbb{N}$. Then, the $H^{1,1}$-approximation error of $v_Q \in \mathcal{S}^{\tilde{p}_i, \tilde{p}_i}$ satisfies*

$$\begin{aligned} \|w - v_Q\|_{H^{1,1}(Q)}^2 \leq C\, \frac{(\tilde{p}_i - s)!}{(\tilde{p}_i + s)!} \left(\frac{h_i}{2}\right)^{2s} \Big(\|\partial_x^{s+2} w\|_{L^2(Q)}^2 + \|\partial_y^{s+2} w\|_{L^2(Q)}^2 \\ + \|\partial_y \partial_x^{s+1} w\|_{L^2(Q)}^2 + \|\partial_x \partial_y^{s+1} w\|_{L^2(Q)}^2 \Big) \end{aligned}$$

for all $s \in \{0, 1, \ldots, \tilde{p}_i - 1\}$, where the constant C does not depend on $\tilde{p}_i$, s and h_i.

Proof. See [5, Lemma 5.40]. □

Remark 4.37. The advantage of the preceding lemma is that it provides a clear and compact upper bound for the local $H^{1,1}$-approximation error featuring only partial derivatives of same order. The price we have to pay for this convenience lies in the fact that, due to $\tilde{p}_1 = \tilde{p}_2 = 1$, it is impossible to obtain from it a positive power of h_i for elements $Q \in \mathcal{Q}_L^{(1)} \cup \mathcal{Q}_L^{(2)}$. However, as long as $w \in H^{k+1}(Q)$ with $k \geq 2$, this problem can be overcome by uniformly setting $s = 1$ in Lemmas 4.33, 4.34 and 4.35. This choice leads to

$$\|w - v_Q\|_{H^{1,1}(Q)}^2 \leq C\, h_i^2 \Big(|w|_{H^2(Q)}^2 + \|\partial_y \partial_x^2 w\|_{L^2(Q)}^2 + \|\partial_x \partial_y^2 w\|_{L^2(Q)}^2 \Big) \tag{4.15}$$

for all $Q \in \mathcal{Q}_L^{(i)}$, $i = 1, 2$, if $w \in H^3(Q)$.

We are particularly interested in the approximation of stationary functions $w = w(x-y)$, where $w(x-y)$ is an even function in $H^2(-1,1)$ whose restriction to the interval $[0,1]$ lies in the space $\mathcal{B}_\beta^3(0,1)$ (cf. Assumption 3.1). In this case, the derivatives of w on elements $Q \in \mathcal{Q}_L^{(1)}$ are in general square-integrable only up to order 2. Already the third derivatives of w need not be in $L^2(Q)$ anymore. Thus, in Lemma 4.35, we have $k = 1$ and one can only select $s = 0$, for which reason no positive power of h_1 will occur on the right hand side. The following lemma fixes this problem:

Lemma 4.38. *Let $Q \in \mathcal{Q}_L^{(1)}$. The function $w : \square \to \mathbb{R}$ shall satisfy Assumption 3.1. Then, for the local approximation $v_Q \in \mathcal{S}^{1,1}$, it holds*

$$\|\partial_x \partial_y (w - v_Q)\|_{L^2(Q)}^2 \leq C\, h_1^{3-2\beta} |w|_{H_\beta^{3,3}(0,1)}^2$$

with a constant C depending only on β. In particular, plugging this result along with those from Lemmas 4.33, 4.34 into (4.14) yields:

$$\|w - v_Q\|_{H^{1,1}(Q)}^2 \leq \bar{C} \left(h_1^2\, |w|_{H^2(Q)}^2 + h_1^{3-2\beta} |w|_{H_\beta^{3,3}(0,1)}^2 \right)$$

Proof. v_Q is the bilinear function which interpolates w in the four vertices of $Q = [a_1, b_1] \times [a_2, b_2]$. It can be written as

$$v_Q(x,y) = w(a_1,a_2) + \frac{x-a_1}{h_1}\int_{a_1}^{b_1} \partial_\xi w(\xi,a_2)\,d\xi + \frac{y-a_2}{h_1}\int_{a_2}^{b_2} \partial_\eta w(a_1,\eta)\,d\eta \\ + \frac{x-a_1}{h_1}\,\frac{y-a_2}{h_1}\int_{a_1}^{b_1}\int_{a_2}^{b_2} \partial_\xi\partial_\eta w(\xi,\eta)\,d\eta\,d\xi.$$

From this representation, it follows that (the constant) $\partial_x\partial_y v_Q$ coincides with the mean value of $\partial_x\partial_y w$ on element Q:

$$\partial_x\partial_y v_Q = \frac{1}{h_1^2}\int_{a_1}^{b_1}\int_{a_2}^{b_2} \partial_\xi\partial_\eta w(\xi,\eta)\,d\eta\,d\xi$$

Thus, we obtain:

$$\|\partial_x\partial_y(w - v_Q)\|^2_{L^2(Q)} = \iint_Q \left(\partial_x\partial_y w(x,y) - \frac{1}{h_1^2}\iint_Q \partial_\xi\partial_\eta w(\xi,\eta)\,d\eta\,d\xi\right)^2 dy\,dx \\ = \frac{1}{h_1^4}\iint_Q \left(\iint_Q \big(\partial_x\partial_y w(x,y) - \partial_\xi\partial_\eta w(\xi,\eta)\big) d\eta\,d\xi\right)^2 dy\,dx$$

The Cauchy–Schwarz inequality for the inner integral gives:

$$\|\partial_x\partial_y(w - v_Q)\|^2_{L^2(Q)} \le \frac{1}{h_1^4}\iint_Q\iint_Q 1^2 d\eta\,d\xi \\ \cdot \iint_Q \big(\partial_x\partial_y w(x,y) - \partial_\xi\partial_\eta w(\xi,\eta)\big)^2 d\eta\,d\xi\,dy\,dx \tag{4.16}$$

$$= \frac{1}{h_1^2}\iint_Q\iint_Q \big(\partial_x\partial_y w(x,y) - \partial_\xi\partial_\eta w(\xi,\eta)\big)^2 d\eta\,d\xi\,dy\,dx \tag{4.17}$$

Since $w = w(x-y)$, the mixed derivatives become

$$\partial_x\partial_y w(x,y) = -w''(x-y), \qquad \partial_\xi\partial_\eta w(\xi,\eta) = -w''(\xi-\eta).$$

Because of $Q \in \mathcal{Q}_L^{(1)}$, this knowledge allows us to replace Q as the domain of integration in (4.17), without loss of generality, by $Q_o := [0,h_1]\times[0,h_1]$, $Q_a := [0,h_1]\times[h_1,2h_1]$ or $Q_b := [h_1,2h_1]\times[0,h_1]$, respectively—depending on whether the element Q lies on, above or below the diagonal Δ. To avoid a case differentiation, we simply use $[0,2h_1]\times[0,2h_1]$ which is a superset of Q_o, Q_a and Q_b at the same time:

$$\|\partial_x\partial_y(w - v_Q)\|^2_{L^2(Q)} \le \frac{1}{h_1^2}\int_0^{2h_1}\int_0^{2h_1}\int_0^{2h_1}\int_0^{2h_1} \big(w''(x-y) - w''(\xi-\eta)\big)^2 d\eta\,d\xi\,dy\,dx$$

Since $w(\cdot)$ is even (cf. Assumption 3.1), $w''(\cdot)$ is as well. With this and with the substitutions $z = x - y$ and $\zeta = \xi - \eta$, we estimate further:

$$\begin{aligned}
\|\partial_x \partial_y (w - v_Q)\|^2_{L^2(Q)} &\leq \frac{4}{h_1^2} \int_0^{2h_1} \int_0^{x} \int_0^{2h_1} \int_0^{\xi} \big(w''(x-y) - w''(\xi - \eta)\big)^2 d\eta \, d\xi \, dy \, dx \\
&= \frac{4}{h_1^2} \int_0^{2h_1} \int_0^{2h_1} \underbrace{(2h_1 - z)}_{\leq 2h_1} \underbrace{(2h_1 - \zeta)}_{\leq 2h_1} \big(w''(z) - w''(\zeta)\big)^2 d\zeta \, dz \\
&\leq 16 \int_0^{2h_1} \int_0^{2h_1} \big(w''(z) - w''(\zeta)\big)^2 d\zeta \, dz \\
&= 32 \int_0^{2h_1} \int_0^{z} \big(w''(z) - w''(\zeta)\big)^2 d\zeta \, dz
\end{aligned}$$

Now, w'' is continuously differentiable inside $(0,1)$. Thus, the fundamental theorem of calculus yields:

$$\|\partial_x \partial_y (w - v_Q)\|^2_{L^2(Q)} \leq 32 \int_0^{2h_1} \int_0^{z} \left(\int_\zeta^z w^{(3)}(t) \, dt \right)^2 d\zeta \, dz$$

We insert the factor $1 = t^{-\beta} \, t^{\beta}$ and apply the Cauchy–Schwarz inequality once more in order to obtain:

$$\begin{aligned}
\|\partial_x \partial_y (w - v_Q)\|^2_{L^2(Q)} &\leq 32 \int_0^{2h_1} \int_0^{z} \left(\int_\zeta^z t^{-\beta} \, t^{\beta} w^{(3)}(t) \, dt \right)^2 d\zeta \, dz \\
&\leq 32 \int_0^{2h_1} \int_0^{z} \int_\zeta^z t^{-2\beta} dt \int_\zeta^z t^{2\beta} w^{(3)}(t)^2 dt \, d\zeta \, dz
\end{aligned}$$

With the trivial bound

$$\int_\zeta^z t^{2\beta} w^{(3)}(t)^2 dt \leq \int_0^{2h_1} t^{2\beta} w^{(3)}(t)^2 dt \leq |w|^2_{H^{3,3}_\beta(0,1)},$$

we finally arrive at

$$\|\partial_x \partial_y (w - v_Q)\|^2_{L^2(Q)} \leq 32 \, |w|^2_{H^{3,3}_\beta(0,1)} \int_0^{2h_1} \int_0^{z} \int_\zeta^{z} t^{-2\beta} dt \, d\zeta \, dz.$$

The remaining integral is finite for any $\beta < 1$, since

$$\int_0^{2h_1} \int_0^{z} \int_\zeta^{z} t^{-2\beta} dt \, d\zeta \, dz = \frac{2^{2-2\beta}}{(1-\beta)(3-2\beta)} \, h_1^{3-2\beta}. \qquad \square$$

4.6. **Auxiliary results for the convergence analysis.**

Notation 4.39. From now on, we will write Q_L^i in order to denote the subset of $\square$ which is obtained as union of all elements $Q \in \mathcal{Q}_L^{(i)}$:

$$Q_L^i := \bigcup_{Q \in \mathcal{Q}_L^{(i)}} Q$$

Lemma 4.40. *Let $w : \square \to \mathbb{R}$ satisfy Assumption 3.1. Then, with $(\alpha_1, \alpha_2) \in \mathbb{N}_0^2$, $k = \alpha_1 + \alpha_2$, it holds:*

$$\text{if } k \leq 2: \quad \|\partial_x^{\alpha_1}\partial_y^{\alpha_2} w\|^2_{L^2(Q_L^i)} \leq |w|^2_{H^k(-1,1)} \qquad \forall\, i \in \{1, 2, \dots, L\} \tag{4.18a}$$

$$\text{if } k \geq 3: \quad \|\partial_x^{\alpha_1}\partial_y^{\alpha_2} w\|^2_{L^2(Q_L^i)} \leq 2\, h_i^{2(3-\beta-k)} |w|^2_{H_\beta^{k,3}(0,1)} \quad \forall\, i \in \{2, 3, \dots, L\} \tag{4.18b}$$

Proof. See [5, Lemma 5.44]. □

Lemma 4.41 (boundedness of particular trace operators)**.** *(i) Let $[0, h]$ be an interval on the real line. Then, for sufficiently smooth functions $g : [0, h] \to \mathbb{R}$, the trace operator $g \mapsto g(0)$ satisfies:*

$$|g(0)|^2 \leq 2\, h^{-1} \|g\|^2_{L^2(0,h)} + 2\, h\, \|g'\|^2_{L^2(0,h)} \tag{4.19a}$$

(ii) Let $R = [0, h_x] \times [0, h_y]$ be a rectangle in the Cartesian plane with left side $\gamma = \{0\} \times [0, h_y]$. Then, for sufficiently smooth functions $g : R \to \mathbb{R}$, the trace operator $g \mapsto g|_\gamma$ satisfies:

$$\|g\|^2_{L^2(\gamma)} \leq 2\, h_x^{-1} \|g\|^2_{L^2(R)} + 2\, h_x \|\partial_x g\|^2_{L^2(R)} \tag{4.19b}$$

Proof. See [5, Lemma 5.45] □

Remark 4.42. In comparison to the classical (and more general) trace theorems, the advantage of the estimates in Lemma 4.41 lies in the right hand side: Instead of a bound $C\, \|g\|^2_{H^1}$ with domain-dependent constant C, we have a weighted sum of (semi)norms. In particular, the scaling of the summands with respect to the domain size h and h_x, respectively, is given explicitly. It differs depending on the order of the derivative. This fact will become important in the proof of the next lemma.

Lemma 4.43. *Let $w : \square \to \mathbb{R}$ satisfy Assumption 3.1. V_γ shall denote the trace lifting associated with a macro-edge $\gamma \in \dot{\mathcal{G}}_L^{(i)}$, $i \in \{2, 3, \dots, L\}$. By $Q_2, Q_3 \in \mathcal{Q}_L^{(i)}$, we refer to the two small elements bordering on γ (cf. Figure 4.4). Then, for $i \geq 3$, it holds:*

(4.20a)

$$\|V_\gamma\|^2_{H^{1,1}(Q_2 \cup Q_3)} \leq C\, \frac{(\tilde{p}_i - t)!}{(\tilde{p}_i + t)!} \left(\frac{1}{2}\right)^{2t} h_i^{3-2\beta} |w|^2_{H_\beta^{t+1,3}(0,1)} \quad \forall t \in \{2, 3, \dots, \tilde{p}_i\}$$

In case $i = 2$, we have:

$$\|V_\gamma\|^2_{H^{1,1}(Q_2 \cup Q_3)} \leq C\, h_2^{2-\beta} \left(|w|^2_{H^2(-1,1)} + |w|^2_{H_\beta^{3,3}(0,1)} \right) \tag{4.20b}$$

Proof. From Lemma 4.29, we recall the bound:

$$\|V_\gamma\|^2_{H^{1,1}(Q_2 \cup Q_3)} \leq C_1\, h_i^{-1} \left| [\tilde{v}_L^\Delta]_\gamma \right|^2_{H^1(\gamma)} \tag{4.21}$$

Using the notation introduced in Definition 4.27, we can represent the jump $[\tilde{v}_L^\Delta]_\gamma$, which occurs in $\tilde{v}_L^\Delta$ across the macro-edge γ, as follows:

$$\left[\tilde{v}_L^\Delta\right]_\gamma = \begin{cases} v_\gamma - v_{\gamma_{12}} & \text{on } \gamma_{12} \\ v_\gamma - v_{\gamma_{13}} & \text{on } \gamma_{13} \end{cases}$$

Hence, the H^1-seminorm of the jump can be estimated by

$$\begin{aligned} \left|[\tilde{v}_L^\Delta]_\gamma\right|^2_{H^1(\gamma)} &= |v_\gamma - w + w - v_{\gamma_{12}}|^2_{H^1(\gamma_{12})} + |v_\gamma - w + w - v_{\gamma_{13}}|^2_{H^1(\gamma_{13})} \\ &\leq 2\, |w - v_\gamma|^2_{H^1(\gamma_{12})} + 2\, |w - v_{\gamma_{12}}|^2_{H^1(\gamma_{12})} \\ &\quad + 2\, |w - v_\gamma|^2_{H^1(\gamma_{13})} + 2\, |w - v_{\gamma_{13}}|^2_{H^1(\gamma_{13})} \\ &= 2\, |w - v_\gamma|^2_{H^1(\gamma)} + 2\, |w - v_{\gamma_{12}}|^2_{H^1(\gamma_{12})} + 2\, |w - v_{\gamma_{13}}|^2_{H^1(\gamma_{13})}. \end{aligned} \tag{4.22}$$

For $v_\gamma \in \mathcal{S}^{p_i}(\gamma)$, $v_{\gamma_{12}} \in \mathcal{S}^{\tilde{p}_i}(\gamma_{12})$ and $v_{\gamma_{13}} \in \mathcal{S}^{\tilde{p}_i}(\gamma_{13})$ with lengths $2h_i$ of γ and h_i of γ_{12} and γ_{13}, respectively, we apply the one-dimensional error estimate [8, (3.3.29)] and obtain:

$$\begin{aligned}
|w - v_\gamma|^2_{H^1(\gamma)} &\leq \frac{(p_i - s)!}{(p_i + s)!}\, h_i^{2s} |w|^2_{H^{s+1}(\gamma)} && (0 \leq s \leq p_i)\\
|w - v_{\gamma_{12}}|^2_{H^1(\gamma_{12})} &\leq \frac{(\tilde{p}_i - \tilde{s})!}{(\tilde{p}_i + \tilde{s})!} \left(\frac{h_i}{2}\right)^{2\tilde{s}} |w|^2_{H^{\tilde{s}+1}(\gamma_{12})} && (0 \leq \tilde{s} \leq \tilde{p}_i)\\
|w - v_{\gamma_{13}}|^2_{H^1(\gamma_{13})} &\leq \frac{(\tilde{p}_i - \tilde{s})!}{(\tilde{p}_i + \tilde{s})!} \left(\frac{h_i}{2}\right)^{2\tilde{s}} |w|^2_{H^{\tilde{s}+1}(\gamma_{13})} && (0 \leq \tilde{s} \leq \tilde{p}_i)
\end{aligned}$$

Due to $\tilde{p}_i \leq p_i$, one can select $s = \tilde{s} = t \in \{0, 1, \ldots, \tilde{p}_i\}$. Substituting the above bounds into (4.22) along with

$$\frac{(p_i - t)!}{(p_i + t)!} \leq \frac{(\tilde{p}_i - t)!}{(\tilde{p}_i + t)!} \quad \text{and} \quad \left(\frac{h_i}{2}\right)^{2t} \leq h_i^{2t}$$

yields:

$$\left|[\tilde{v}_L^\Delta]_\gamma\right|^2_{H^1(\gamma)} \leq 4\, \frac{(\tilde{p}_i - t)!}{(\tilde{p}_i + t)!}\, h_i^{2t} |w|^2_{H^{t+1}(\gamma)}$$

With (4.21), we arrive at the intermediate result:

$$\|V_\gamma\|^2_{H^{1,1}(Q_2 \cup Q_3)} \leq C_2\, \frac{(\tilde{p}_i - t)!}{(\tilde{p}_i + t)!}\, h_i^{2t-1} |w|^2_{H^{t+1}(\gamma)} \quad \forall t \in \{0, 1, \ldots, \tilde{p}_i\} \tag{4.23}$$

Proof of (4.20a). For $i \geq 3$, we have $\tilde{p}_i \geq 2$. Thus, we can restrict the range for t to the non-empty set $\{2, 3, \ldots, \tilde{p}_i\}$. Without loss of generality, it is assumed that γ is located below the diagonal Δ, i.e. $x > y$ for all $(x, y) \in \gamma$. Furthermore, all $(x, y) \in \gamma$ satisfy $x - y \geq 2h_i$ (cf. Figure 4.3). Thus, with the substitution $z = x - y$, we can estimate:

$$\begin{aligned}
|w|^2_{H^{t+1}(\gamma)} &\leq \int_{2h_i}^1 w^{(t+1)}(z)^2 dz = \int_{2h_i}^1 z^{-2(\beta+t-2)} z^{2(\beta+t-2)} w^{(t+1)}(z)^2 dz\\
&\leq (2h_i)^{-2(\beta+t-2)} \int_0^1 z^{2(\beta+t-2)} w^{(t+1)}(z)^2 dz\\
&= \left(\frac{1}{2}\right)^{2(\beta+t-2)} h_i^{-2(\beta+t-2)} |w|^2_{H^{t+1,3}_\beta(0,1)}
\end{aligned}$$

Inserting this into (4.23) provides:

$$\|V_\gamma\|^2_{H^{1,1}(Q_2 \cup Q_3)} \leq C_2 \left(\frac{1}{2}\right)^{2\beta-4} \frac{(\tilde{p}_i - t)!}{(\tilde{p}_i + t)!} \left(\frac{1}{2}\right)^{2t} h_i^{3-2\beta} |w|^2_{H^{t+1,3}_\beta(0,1)}$$

Setting $C = C_2 \left(\frac{1}{2}\right)^{2\beta-4}$, we have shown (4.20a).

Proof of (4.20b). Because of $\tilde{p}_2 = 1$, the only admissible values for t in (4.23) are 0 and 1. We choose the latter and obtain:

$$\|V_\gamma\|^2_{H^{1,1}(Q_2 \cup Q_3)} \leq C_3\, h_2 |w|^2_{H^2(\gamma)} \tag{4.24}$$

Without loss of generality, we assume:

$$\begin{aligned}
\gamma &= \{\bar{x}\} \times [\bar{y}_1, \bar{y}_2] && \text{(i.e. } \gamma \text{ is vertical)}\\
\forall (x, y) \in \gamma &: x - y \geq 2h_2 && \text{(i.e. } \gamma \text{ lies below the diagonal } \Delta)
\end{aligned}$$

A rectangle whose left side coincides with γ shall be defined by:

$$R := \left[\bar{x}, \bar{x} + \tfrac{1}{2} h_2^\beta\right] \times \left[\bar{y}_1, \bar{y}_2\right]$$

We observe that

$$x - y \in [2h_2, 1] \quad \forall (x, y) \in R. \tag{4.25}$$

Hence, the function w is well-defined on the whole rectangle R (even if $R \not\subseteq \square$), and we can apply (4.19b) with $h_x = \frac{1}{2} h_2^{\beta}$:

$$|w|_{H^2(\gamma)}^2 = \|\partial_y^2 w\|_{L^2(\gamma)}^2 \leq 4\, h_2^{-\beta} \|\partial_y^2 w\|_{L^2(R)}^2 + h_2^{\beta} \|\partial_x \partial_y^2 w\|_{L^2(R)}^2$$

Inserting this upper bound into (4.24) produces:

$$\|V_\gamma\|_{H^{1,1}(Q_2 \cup Q_3)}^2 \leq C_4 \left(h_2^{1-\beta} \|\partial_y^2 w\|_{L^2(R)}^2 + h_2^{1+\beta} \|\partial_x \partial_y^2 w\|_{L^2(R)}^2 \right) \tag{4.26}$$

It remains to investigate the integrals $\|\partial_y^2 w\|_{L^2(R)}^2$ and $\|\partial_x \partial_y^2 w\|_{L^2(R)}^2$. With $\alpha \in \{0, 1\}$, we use the substitution $z = x - y$ and employ (4.25) in order to obtain:

$$\begin{aligned} \|\partial_x^\alpha \partial_y^2 w\|_{L^2(R)}^2 &= \iint_R w^{(2+\alpha)}(x-y)^2 dy\, dx \\ &\leq 2h_2 \int_{2h_2}^{1} w^{(2+\alpha)}(z)^2 dz \end{aligned}$$

The factor $2h_2$ in front of the integral is due to the height of R being equal to the length of γ. Further:

$$\|\partial_y^2 w\|_{L^2(R)}^2 \leq 2h_2 |w|_{H^2(-1,1)}^2 \tag{4.27}$$

$$\begin{aligned} \|\partial_x \partial_y^2 w\|_{L^2(R)}^2 &\leq 2h_2 \int_{2h_2}^{1} z^{-2\beta} z^{2\beta} w^{(3)}(z)^2 dz \\ &\leq (2h_2)^{1-2\beta} |w|_{H_\beta^{3,3}(0,1)}^2 \end{aligned} \tag{4.28}$$

Substituting (4.27) and (4.28) into (4.26) finally yields

$$\|V_\gamma\|_{H^{1,1}(Q_2 \cup Q_3)}^2 \leq C_5\, h_2^{2-\beta} \left(|w|_{H^2(-1,1)}^2 + |w|_{H_\beta^{3,3}(0,1)}^2 \right). \qquad \square$$

4.7. **Convergence analysis.**

Definition 4.44 (broken norm)**.** Let $\mathcal{Q}_L$ be the partition from Definition 4.1. Then, the broken $H^{1,1}$-norm with respect to $\mathcal{Q}_L$ is defined by:

$$\|v\|_{H^{1,1}(\square, \mathcal{Q}_L)} := \sqrt{\sum_{Q \in \mathcal{Q}_L} \|v\|_{H^{1,1}(Q)}^2}$$

On the space $H^{1,1}(\square)$, the broken norm $\|\cdot\|_{H^{1,1}(\square, \mathcal{Q}_L)}$ coincides with the standard norm:

$$\|v\|_{H^{1,1}(\square, \mathcal{Q}_L)} = \|v\|_{H^{1,1}(\square)} \quad \forall v \in H^{1,1}(\square)$$

Furthermore, whenever two functions v_1, v_2 have supports

$$\operatorname{supp} v_i \subset \bigcup_{Q \in I_i} Q \quad \text{with} \quad I_1, I_2 \subset \mathcal{Q}_L,\ \ I_1 \cap I_2 = \emptyset, \tag{4.29}$$

it holds:

$$\|v_1 + v_2\|_{H^{1,1}(\square, \mathcal{Q}_L)}^2 = \|v_1\|_{H^{1,1}(\square, \mathcal{Q}_L)}^2 + \|v_2\|_{H^{1,1}(\square, \mathcal{Q}_L)}^2 \tag{4.30}$$

Theorem 4.45. *Let $w : \square \to \mathbb{R}$ satisfy Assumption 3.1. $S_{\mu,L}^{\Delta}$ shall denote the finite element spaces constructed in Section 4.2. Then, one can select a constant $\mu \geq 1$ such that the corresponding approximations $v_L^{\Delta} \in S_{\mu,L}^{\Delta}$ from Section 4.4 exhibit exponential convergence towards w in $H^{1,1}(\square)$:*

$$\|w - v_L^{\Delta}\|_{H^{1,1}(\square)} \leq c_1 \exp(-c_2\, L), \tag{4.31}$$

where the constants $c_1 > 0$, $c_2 = \frac{1-\beta}{2} \ln(2)$ are independent of L.

Proof.

$$\begin{aligned}
\|w - v_L^{\Delta}\|^2_{H^{1,1}(\square)} &= \|w - v_L^{\Delta}\|^2_{H^{1,1}(\square,\mathcal{Q}_L)} \\
&\le 2\,\|w - \tilde{v}_L^{\Delta}\|^2_{H^{1,1}(\square,\mathcal{Q}_L)} + 2\,\|\tilde{v}_L^{\Delta} - v_L^{\Delta}\|^2_{H^{1,1}(\square,\mathcal{Q}_L)} \\
&= 2\,\|w - \tilde{v}_L^{\Delta}\|^2_{H^{1,1}(\square,\mathcal{Q}_L)} + 2\,\Big\| \sum_{\gamma\in\dot{\mathcal{G}}_L^{\mathrm{dof}}} V_\gamma \Big\|^2_{H^{1,1}(\square,\mathcal{Q}_L)} \qquad (4.32)
\end{aligned}$$

Splitting the set $\dot{\mathcal{G}}_L^{\mathrm{dof}}$ into horizontal and vertical macro-edges, we rewrite the summation over all $\gamma \in \dot{\mathcal{G}}_L^{\mathrm{dof}}$:

$$\sum_{\gamma\in\dot{\mathcal{G}}_L^{\mathrm{dof}}} V_\gamma = \sum_{\substack{\gamma\in\dot{\mathcal{G}}_L^{\mathrm{dof}}:\\ \gamma \text{ horiz.}}} V_\gamma + \sum_{\substack{\gamma\in\dot{\mathcal{G}}_L^{\mathrm{dof}}:\\ \gamma \text{ vert.}}} V_\gamma$$

The support of a trace lifting V_γ is $Q_2^\gamma \cup Q_3^\gamma$, where Q_2^γ, Q_3^γ are the two small elements bordering on γ as in Figure 4.4. For every pair of horizontal macro-edges γ_1, γ_2, the supports of V_{γ_1} and V_{γ_2} (i.e. $Q_2^{\gamma_1} \cup Q_3^{\gamma_1}$ and $Q_2^{\gamma_2} \cup Q_3^{\gamma_2}$) are guaranteed to be disjoint in the sense of (4.29) (cf. Figure 4.3). The same statement is true for arbitrary pairs of vertical macro-edges. Thus, iterated application of (4.30) yields:

$$\begin{aligned}
\Big\| \sum_{\gamma\in\dot{\mathcal{G}}_L^{\mathrm{dof}}} V_\gamma \Big\|^2_{H^{1,1}(\square,\mathcal{Q}_L)} &\le 2\,\Big\| \sum_{\substack{\gamma\in\dot{\mathcal{G}}_L^{\mathrm{dof}}:\\ \gamma \text{ horiz.}}} V_\gamma \Big\|^2_{H^{1,1}(\square,\mathcal{Q}_L)} + 2\,\Big\| \sum_{\substack{\gamma\in\dot{\mathcal{G}}_L^{\mathrm{dof}}:\\ \gamma \text{ vert.}}} V_\gamma \Big\|^2_{H^{1,1}(\square,\mathcal{Q}_L)} \\
&= 2 \sum_{\substack{\gamma\in\dot{\mathcal{G}}_L^{\mathrm{dof}}:\\ \gamma \text{ horiz.}}} \|V_\gamma\|^2_{H^{1,1}(\square,\mathcal{Q}_L)} + 2 \sum_{\substack{\gamma\in\dot{\mathcal{G}}_L^{\mathrm{dof}}:\\ \gamma \text{ vert.}}} \|V_\gamma\|^2_{H^{1,1}(\square,\mathcal{Q}_L)} \\
&= 2 \sum_{\gamma\in\dot{\mathcal{G}}_L^{\mathrm{dof}}} \|V_\gamma\|^2_{H^{1,1}(Q_2^\gamma \cup Q_3^\gamma)}
\end{aligned}$$

Substituting this into (4.32) provides a first intermediate result:

$$\begin{aligned}
\|w - v_L^{\Delta}\|^2_{H^{1,1}(\square)} &\le 2\,\|w - \tilde{v}_L^{\Delta}\|^2_{H^{1,1}(\square,\mathcal{Q}_L)} + 4 \sum_{\gamma\in\dot{\mathcal{G}}_L^{\mathrm{dof}}} \|V_\gamma\|^2_{H^{1,1}(Q_2^\gamma \cup Q_3^\gamma)} \\
&= 2 \sum_{\substack{Q\in\mathcal{Q}_L^{(i)},\\ i\in\{1,\dots,L\}}} \|w - v_Q\|^2_{H^{1,1}(Q)} + 4 \sum_{\substack{\gamma\in\dot{\mathcal{G}}_L^{(i)},\\ i\in\{2,\dots,L\}}} \|V_\gamma\|^2_{H^{1,1}(Q_2^\gamma \cup Q_3^\gamma)} \qquad (4.33)
\end{aligned}$$

For the individual summands in (4.33), we can employ Lemma 4.36, estimate (4.15), Lemma 4.38, Lemma 4.40 and Lemma 4.43 in order to obtain

(4.34)
$$\begin{aligned}
\|w - v_L^{\Delta}\|^2_{H^{1,1}(\square)} &\le C_1 \left(|w|^2_{H^2(-1,1)} + |w|^2_{H^{3,3}_\beta(0,1)} \right) \left(\frac{1}{2}\right)^{(1-\beta)L} \\
&\quad + C_2 \underbrace{\sum_{i=3}^{L} \frac{(\tilde{p}_i - s_i)!}{(\tilde{p}_i + s_i)!} \left(\frac{1}{2}\right)^{2s_i} \left(\frac{1}{2}\right)^{2(1-\beta)(L+2-i)} |w|^2_{H^{s_i+2,3}_\beta(0,1)}}_{(\star)}
\end{aligned}$$

with $s_i \in \{1, 2, \dots, \tilde{p}_i - 1\}$. (For details, see the proof of [5, Theorem 5.49].) The first summand on the right hand side has its origin in the local approximation errors on elements $Q \in \mathcal{Q}_L^{(1)} \cup \mathcal{Q}_L^{(2)}$ and in the trace liftings associated with macro-edges $\gamma \in \dot{\mathcal{G}}_L^{(2)}$ (the latter ones building the dominant part). It is of the form $C_0(\frac{1}{2})^{(1-\beta)L} = C_0 \exp\big(-(1-\beta)\ln(2)\,L\big)$ with a constant C_0 depending on the function w, but not on L. In the proof of [5, Theorem 5.49], it is shown that the sum ($\star$) can be bounded by an analogous expression $C_3(\frac{1}{2})^{(1-\beta)L}$, where C_3 does not depend on L. $\square$

Corollary 4.46. *In the setting of Theorem 4.45, the convergence rate of the approximation error with respect to the number of degrees of freedom in the ansatz space $S^{\Delta}_{\mu,L}$ is bounded by*

$$\|w - v^{\Delta}_L\|_{H^{1,1}(\square)} \leq c_1 \exp\left(-c_2 \sqrt[3]{N}\right),$$

where $N := \dim S^{\Delta}_{\mu,L}$ and c_1, c_2 are positive constants independent of N.

Proof. For a fixed μ, Lemma 4.17 yields

$$N = \dim S^{\Delta}_{\mu,L} = \mathcal{O}(L^3),$$

i.e. $L \geq \sqrt[3]{\frac{N}{C}}$ for $L \to \infty$ with a constant $C > 0$. Inserting this into (4.31) gives the claim. $\square$

References

[1] P. Abrahamsen. A review of Gaussian random fields and correlation functions. Technical Report 917, Norwegian Computing Center, Oslo, 1997.

[2] I. Babuška and B.Q. Guo. Regularity of the solution of elliptic problems with piecewise analytic data. I. Boundary value problems for linear elliptic equation of second order. *SIAM J. Math. Anal.*, 19(1):172–203, 1988.

[3] H.-J. Bungartz, M. Griebel, and Ch. Zenger. *Einführung in die Computergraphik. Grundlagen, Geometrische Modellierung, Algorithmen*. Mathematische Grundlagen der Informatik. Friedr. Vieweg & Sohn, Braunschweig, second edition, 2002.

[4] P. Houston, Ch. Schwab, and E. Süli. Stabilized *hp*-finite element methods for first-order hyperbolic problems. *SIAM J. Numer. Anal.*, 37(5):1618–1643 (electronic), 2000.

[5] B. Pentenrieder. *Exponential convergence of hp-FEM for second moments of elliptic PDEs with stochastic data*. PhD thesis, ETH Zürich Diss. No. 18729, 2009.

[6] B. Pentenrieder and Ch. Schwab. *hp*-FEM for second moments of elliptic PDEs with stochastic data. Part 1: Analytic regularity. Research Report 2010-??, Seminar for Applied Mathematics, ETH Zürich, 2010. In review.

[7] M. Schlather. Introduction to positive definite functions and to unconditional simulation of random fields. Technical Report ST-99-10, Lancaster University, UK, 1999.

[8] Ch. Schwab. *p- and hp-finite element methods. Theory and applications in solid and fluid mechanics*. Numerical Mathematics and Scientific Computation. The Clarendon Press Oxford University Press, New York, 1998.

[9] Ch. Schwab and R.A. Todor. Sparse finite elements for elliptic problems with stochastic loading. *Numer. Math.*, 95(4):707–734, 2003.

[10] Ch. Schwab and R.A. Todor. Sparse finite elements for stochastic elliptic problems - higher order moments. *Computing*, 71:43–63, 2003.

[11] G. Szegö. *Orthogonal Polynomials*. American Mathematical Society Colloquium Publications, v. 23. American Mathematical Society, New York, 1939.

Research Reports

No.	Authors/Title
10-09	*B. Pentenrieder and C. Schwab* *hp*-FEM for second moments of elliptic PDEs with stochastic data. Part 2: Exponential convergence
10-08	*B. Pentenrieder and C. Schwab* *hp*-FEM for second moments of elliptic PDEs with stochastic data. Part 1: Analytic regularity
10-07	*C. Jerez-Hanckes and J.-C. Nédélec* Asymptotics for Helmholtz and Maxwell solutions in 3-D open waveguides
10-06	*C. Schwab and O. Reichmann* Numerical analysis of additive, Lévy and Feller processes with applications to option pricing
10-05	*C. Schwab and R. Stevenson* Fast evaluation of nonlinear functionals of tensor product wavelet expansions
10-04	*B.N. Khoromskij and C. Schwab* Tensor-structured Galerkin approximation of parametric and stochastic elliptic PDEs
10-03	*A. Cohen, R. DeVore and C. Schwab* Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDEs
10-02	*V. Gradinaru, G.A. Hagedorn, A. Joye* Tunneling dynamics and spawning with adaptive semi-classical wavepackets
10-01	*N. Hilber, S. Kehtari, C. Schwab and C. Winter* Wavelet finite element method for option pricing in highdimensional diffusion market models
09-41	*E. Kokiopoulou, D. Kressner, N. Paragios, P. Frossard* Optimal image alignment with random projections of manifolds: algorithm and geometric analysis
09-40	*P. Benner, P. Ezzatti, D. Kressner, E.S. Quintana-Ortí, A. Remón* A mixed-precision algorithm for the solution of Lyapunov equations on hybrid CPU-GPU platforms
09-39	*V. Wheatley, P. Huguenot, H. Kumar* On the role of Riemann solvers in discontinuous Galerkin methods for magnetohydrodynamics