Research Article

Patrick Bammer*, Andreas Schröder and Thomas P. Wihler

An *hp*-Adaptive Strategy Based on Locally Predicted Error Reductions

Abstract: We introduce a new *hp*-adaptive strategy for self-adjoint elliptic boundary value problems that does *not* rely on using classical a posteriori error estimators. Instead, our approach is based on a generally applicable prediction strategy for the reduction of the energy error that can be expressed in terms of local modifications of the degrees of freedom in the underlying discrete approximation space. The computations related to the proposed prediction strategy involve low-dimensional linear problems that are computationally inexpensive and highly parallelizable. The mathematical building blocks for this new concept are first developed on an abstract Hilbert space level, before they are employed within the specific context of *hp*-type finite element discretizations. For this particular framework, we discuss an explicit construction of *p*-enrichments and *hp*-refinements by means of an appropriate constraint coefficient technique that can be employed in any dimensions. The applicability and effectiveness of the resulting *hp*-adaptive strategy is illustrated with some 1- and 2-dimensional numerical examples.

1 Introduction

Over the last few decades significant contributions have been made in *a posteriori error* estimation and automatic (adaptive) mesh refinement techniques for finite element methods (FEM), see, e.g., [1, 6, 14, 29–31]. A majority of works in the adaptive FEM literature are focused exclusively on local (usually isotropic) refinements of elements (i.e. on so-called *h*-refinement). In recent years, however, motivated by the pioneering *a priori* results on spectral and variable-order (so-called *hp*-version) FEM of Babuška and co-authors, cf. [2–5, 16, 17], the adaptive selection of locally varying polynomial degrees (i.e. *p*-refinement) has been taken into account as well. In combination with local mesh adaptivity, such approaches lead to *hp-adaptive* schemes; see, e.g., the overview article [24] and the references therein, or the monographs [9, 21, 27, 28] for theoretical and practical aspects of *hp*-FEM. A key advantage of *hp*-approximations is their ability to approximate local singularities in (otherwise smooth) solutions of partial differential equations at high algebraic or even exponential convergence rates, see, e.g., [27].

While it seems reasonable to generalize classical *h*-version *a posteriori* error estimation to the *hp*-framework, it is well-known that the derivation of effective computable *hp*-version error bounds involves considerable technical challenges [11, 22]. Furthermore, in addition to flagging elements for refinement (as in *h*-adaptive

Corresponding author: Patrick Bammer, Fachbereich Mathematik, Paris Lodron Universität Salzburg, Hellbrunner Str. 14, 5020 Salzburg, Austria, e-mail: patrick.bammer@plus.ac.at. https://orcid.org/0009-0009-6041-273X
Andreas Schröder, Fachbereich Mathematik, Paris Lodron Universität Salzburg, Hellbrunner Str. 14, 5020 Salzburg, Austria, e-mail: andreas.schroeder@plus.ac.at. https://orcid.org/0000-0003-3691-0906
Thomas P. Wihler, Mathematisches Institut, Universität Bern, Sidlerstr. 5, CH-3012 Bern, Switzerland, e-mail: thomas.wihler@unibe.ch. https://orcid.org/0000-0003-1232-0637

FEM), the design of *hp*-adaptive strategies requires the development of carefully devised decision procedures that allow to choose between various possible *hp*-refinements of the marked elements; this can be accomplished, for instance, by means of exploiting appropriate smoothness testing strategies [13, 15, 19, 32].

In contrast, the approach proposed in the present paper does neither rely on classical *a posteriori* error estimators nor on smoothness indicators. Instead, we develop a *prediction procedure* for an efficient reduction of the (global) energy error that can be represented explicitly in terms of *hp*-refinements on local discrete spaces. Our methodology is closely related to the energy minimization technique [20], where local problems are employed to compare various elementwise (competitive) *p*- and *hp*-refinements with regards to the potential contribution they may provide to the decay of the global error (see also [9] for a related strategy). Inspired by the low-order approach presented in [18], the main novelty in the current paper consists in a new construction of the approximation spaces for the local problems that *properly conforms to the weak formulation* of the underlying partial differential equation, and thus, provides a mathematically sound structure; this is opposed to [20], where artificial conditions and numerical fluxes along patch boundaries are involved.

More precisely, for the purpose of this work, in order to perform a (local) *hp*-refinement on an element Q in the mesh (with associated polynomial degree p_Q), we first decompose the numerical solution u_{hp} on the current (global) *hp*-space into a local part u_{hp}^{loc} (with local support on or around Q), and a remaining (globally supported) part $\tilde{u}_{hp} = u_{hp} - u_{hp}^{\text{loc}}$. Here, our goal is to suitably replace the local part u_{hp}^{loc} on a modified local *hp*-space that allows for an improved resolution of possible small-scale features of the global solution on Q. To this end, we consider locally supported spaces (spanned by so-called *enrichment functions*), which either provide a polynomial degree $> p_Q$ on Q, or an *hp*-refinement of Q (involving a small number of subelements of Q and an appropriate polynomial degree distribution on this subelements), thereby giving rise to a *p-enrichment* or an *hp-replacement* on Q, respectively. The span of the enrichment functions together with the remaining part of the current solution, $\tilde{u}_{hp}$, defines the local *enrichment* or *replacement space*, on which a potentially refined approximate solution is obtained. Since we explicitly include the remaining part of the current solution in the definition of the local enrichment/replacement space, the new approximate solution represents a *global* solution that allows to rigorously predict the (global) error reduction resulting from replacing the current solution by the new one. Since the local enrichment/replacement spaces are low-dimensional (in comparison to the full *hp*-space), we emphasize that the locally predicted (but globally effective) error reduction can be computed at a negligible cost. Our *hp*-adaptive algorithm passes through all elements of a current mesh (which can be done in parallel), and compares the predicted error reductions for different *p*-enrichments and *hp*-refinements in order to find an optimal enrichment on each element. Then, by using an appropriate marking strategy, all those elements, from which the most substantial (global) error reduction can be expected, will be refined.

The paper is structured as follows: In Section 2 we begin by introducing an abstract framework for linear elliptic problems in Hilbert spaces, and derive some general results on (global) error reductions, which are exploited to devise the error prediction strategy (in terms of low-dimensional modifications) in our new *hp*-adaptive scheme. In Section 3, the abstract setting is applied to *hp*-FEM, whereby we specify two types of enrichment functions on each element of a given *hp*-mesh, namely *p*- and *hp*-enrichment functions; in particular, we present in detail the explicit construction of such enrichments on individual elements based on a recent constraint coefficients technique, see [7, 26]. Furthermore, the proposed *hp*-adaptive procedure is outlined in terms of a general algorithm in Section 4, and some numerical experiments for *hp*-type finite element discretizations in 1D and 2D, which illustrate the effectiveness of our approach, will be presented in Section 5.

Throughout this article, let $\mathbb{N} = \{1, 2, \ldots\}$ denote the positive integers and set $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. For any $n \in \mathbb{N}$ we write $\underline{n}$ and $\underline{n}_0$ for the finite sets $\{1, 2, \ldots, n\}$ and $\{0, 1, \ldots, n\}$, respectively.

2 Abstract Framework

On a (real) Hilbert space $\mathbb{X}$ we consider a bounded, symmetric and coercive bilinear form $a : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ that induces a norm $v \mapsto \|v\|_{\mathbb{X}} := a(v, v)^{\frac{1}{2}}$ on $\mathbb{X}$. We also introduce a bounded linear form $b : \mathbb{X} \to \mathbb{R}$. Then, for any closed subspace $\mathbb{W} \subseteq \mathbb{X}$, by the Riesz representation theorem, there is a unique $u_{\mathbb{W}} \in \mathbb{W}$ such that the weak

formulation

$$a(u_{\mathbb{W}}, w) = b(w) \quad \text{for all } w \in \mathbb{W} \tag{2.1}$$

holds true. Equivalently, defining the residual

$$\rho_{\mathbb{W}}(v) := b(v) - a(u_{\mathbb{W}}, v), \quad v \in \mathbb{X}, \tag{2.2}$$

we have $\rho_{\mathbb{W}}(v) = 0$ for all $v \in \mathbb{W}$. Furthermore, in line with the above notation, we signify by $u_{\mathbb{X}} \in \mathbb{X}$ the (full space) solution of the weak formulation

$$a(u_{\mathbb{X}}, v) = b(v) \quad \text{for all } v \in \mathbb{X}. \tag{2.3}$$

2.1 Low-Dimensional Enrichments

Consider the special case where the subspace $\mathbb{W} \subseteq \mathbb{X}$ introduced above is spanned by finitely many basis functions $\phi_1, \ldots, \phi_N \in \mathbb{W}$, with $N \geq 1$, i.e.

$$\mathbb{W} = \mathrm{span}\{\phi_1, \ldots, \phi_N\}, \quad \dim(\mathbb{W}) = N < \infty.$$

Moreover, for an index subset $\mathcal{J}^{\mathrm{loc}} \subset \underline{N}$, we let

$$\mathbb{W}^{\mathrm{loc}} := \mathrm{span}\left\{\phi_i : i \in \mathcal{J}^{\mathrm{loc}}\right\} \subset \mathbb{W}; \tag{2.4}$$

for instance, in the specific context of *hp*-discretizations to be discussed later on, the spaces $\mathbb{W}$ and $\mathbb{W}^{\mathrm{loc}}$ will take the roles of an *hp*-finite element space and a *locally supported* subspace, respectively. We can then define a linear projection operator

$$\Pi^{\mathrm{loc}} : \mathbb{W} \to \mathbb{W}^{\mathrm{loc}}, \quad v = \sum_{i \in \underline{N}} v_i \, \phi_i \mapsto \Pi^{\mathrm{loc}} v := \sum_{i \in \mathcal{J}^{\mathrm{loc}}} v_i \, \phi_i, \tag{2.5}$$

which enables a decomposition of the solution $u_{\mathbb{W}} \in \mathbb{W}$ of (2.1) into a low dimensional and a remaining part

$$u_{\mathbb{W}}^{\mathrm{loc}} := \Pi^{\mathrm{loc}} u_{\mathbb{W}} \in \mathbb{W}^{\mathrm{loc}}, \quad \tilde{u}_{\mathbb{W}} := u_{\mathbb{W}} - \Pi^{\mathrm{loc}} u_{\mathbb{W}},$$

respectively, with

$$u_{\mathbb{W}} = u_{\mathbb{W}}^{\mathrm{loc}} + \tilde{u}_{\mathbb{W}}. \tag{2.6}$$

We aim to improve the approximation $u_{\mathbb{W}}$ of the (full space) solution $u_{\mathbb{X}} \in \mathbb{X}$ of (2.3) by enriching or replacing the local space $\mathbb{W}^{\mathrm{loc}}$ by a subspace

$$\mathbb{Y} := \mathrm{span}\{\tilde{u}_{\mathbb{W}}, \xi_1, \ldots, \xi_L\} \subseteq \mathbb{X} \tag{2.7}$$

of dimension $L + 1$, where

$$\boldsymbol{\xi} := \{\xi_1, \ldots, \xi_L\} \subseteq \mathbb{X} \tag{2.8}$$

is a (small) set of linearly independent elements in $\mathbb{X}$. If $\mathbb{W}^{\mathrm{loc}} \subset \mathbb{Y}$, then we call $\mathbb{Y}$ a *local enrichment space*, otherwise $\mathbb{Y}$ is a *local replacement space*; the functions $\xi_1, \ldots, \xi_L$ will in both cases be referred to as *(local) enrichment functions*.

We now introduce the low-dimensional problem

$$u_{\mathbb{Y}} \in \mathbb{Y}: \quad a(u_{\mathbb{Y}}, v) = b(v) \quad \text{for all } v \in \mathbb{Y}. \tag{2.9}$$

Since $\dim \mathbb{Y} = L + 1$, there is a unique $\epsilon \in \mathbb{R}$ and some uniquely determined $y_\xi \in \mathrm{span}\{\xi_1, \ldots, \xi_L\}$ such that we can express the solution $u_{\mathbb{Y}}$ of (2.9) in terms of a linear combination

$$u_{\mathbb{Y}} = (1 + \epsilon)\,\tilde{u}_{\mathbb{W}} + y_\xi. \tag{2.10}$$

Then, defining the errors

$$e_{\mathbb{W}} := u_{\mathbb{X}} - u_{\mathbb{W}}, \quad e_{\mathbb{Y}} := u_{\mathbb{X}} - u_{\mathbb{Y}},$$

we prove the following result for the *predicted error reduction* $\Delta e_{\mathbb{W},\mathbb{Y}}$, given by

$$\Delta e_{\mathbb{W},\mathbb{Y}}^2 := \|e_{\mathbb{W}}\|_{\mathbb{X}}^2 - \|e_{\mathbb{Y}}\|_{\mathbb{X}}^2.$$

For a related result in the specific context of a multilevel solver for linear diffusion models, we mention the work [23].

Proposition 1 (Predicted Error Reduction). *For the predicted error reduction we have the identities*

$$\Delta e_{\mathbb{W},\mathbb{Y}}^2 = \|u_{\mathbb{Y}} - u_{\mathbb{W}}\|_{\mathbb{X}}^2 - 2\,\rho_{\mathbb{Y}}(u_{\mathbb{W}}^{\mathrm{loc}}) = \rho_{\mathbb{W}}(y_\xi) - \rho_{\mathbb{Y}}(u_{\mathbb{W}}^{\mathrm{loc}}), \tag{2.11}$$

with $y_\xi \in \mathrm{span}\{\xi_1, \ldots, \xi_L\}$ from (2.10), the residual $\rho_{\mathbb{W}}(\cdot)$ defined in (2.2), and $\rho_{\mathbb{Y}}(\cdot) := b(\cdot) - a(u_{\mathbb{Y}}, \cdot)$.

Proof. We begin by noting the identity $e_{\mathbb{Y}} + (u_{\mathbb{Y}} - u_{\mathbb{W}}) = e_{\mathbb{W}}$, which implies that $\|e_{\mathbb{Y}} + (u_{\mathbb{Y}} - u_{\mathbb{W}})\|_{\mathbb{X}}^2 = \|e_{\mathbb{W}}\|_{\mathbb{X}}^2$. Thus, using that the bilinear form $a(\cdot, \cdot)$ induces the norm $\|\cdot\|_{\mathbb{X}}$, we deduce that

$$\|e_{\mathbb{Y}}\|_{\mathbb{X}}^2 + 2\,a(e_{\mathbb{Y}}, u_{\mathbb{Y}} - u_{\mathbb{W}}) + \|u_{\mathbb{Y}} - u_{\mathbb{W}}\|_{\mathbb{X}}^2 = \|e_{\mathbb{W}}\|_{\mathbb{X}}^2.$$

By Galerkin orthogonality, exploiting that $u_{\mathbb{Y}} - \tilde{u}_{\mathbb{W}} \in \mathbb{Y}$ and $u_{\mathbb{W}}^{\mathrm{loc}} \in \mathbb{W}$, we observe that

$$a(e_{\mathbb{Y}}, u_{\mathbb{Y}} - \tilde{u}_{\mathbb{W}}) = 0 \quad \text{and} \quad a(e_{\mathbb{W}}, u_{\mathbb{W}}^{\mathrm{loc}}) = 0.$$

Hence, recalling (2.6) we obtain

$$a(e_{\mathbb{Y}}, u_{\mathbb{Y}} - u_{\mathbb{W}}) = a(e_{\mathbb{Y}}, u_{\mathbb{Y}} - \tilde{u}_{\mathbb{W}}) - a(e_{\mathbb{Y}}, u_{\mathbb{W}}^{\mathrm{loc}}) = -a(e_{\mathbb{Y}}, u_{\mathbb{W}}^{\mathrm{loc}}).$$

Furthermore, invoking (2.3), it follows that

$$a(e_{\mathbb{Y}}, u_{\mathbb{Y}} - u_{\mathbb{W}}) = -a(u_{\mathbb{X}}, u_{\mathbb{W}}^{\mathrm{loc}}) + a(u_{\mathbb{Y}}, u_{\mathbb{W}}^{\mathrm{loc}}) = -b(u_{\mathbb{W}}^{\mathrm{loc}}) + a(u_{\mathbb{Y}}, u_{\mathbb{W}}^{\mathrm{loc}}) = -\rho_{\mathbb{Y}}(u_{\mathbb{W}}^{\mathrm{loc}}).$$

Thus,

$$\|e_{\mathbb{Y}}\|_{\mathbb{X}}^2 - 2\,\rho_{\mathbb{Y}}(u_{\mathbb{W}}^{\mathrm{loc}}) + \|u_{\mathbb{Y}} - u_{\mathbb{W}}\|_{\mathbb{X}}^2 = \|e_{\mathbb{W}}\|_{\mathbb{X}}^2,$$

which yields the first identity in (2.11). Moreover, recalling (2.6) and (2.10), we have $u_{\mathbb{Y}} - u_{\mathbb{W}} = \epsilon\,\tilde{u}_{\mathbb{W}} + y_\xi - u_{\mathbb{W}}^{\mathrm{loc}}$, and since $\tilde{u}_{\mathbb{W}} \in \mathbb{W} \cap \mathbb{Y}$, we notice the orthogonality property

$$a(u_{\mathbb{Y}} - u_{\mathbb{W}}, \tilde{u}_{\mathbb{W}}) = 0.$$

Thus,

$$\begin{aligned}
\|u_{\mathbb{Y}} - u_{\mathbb{W}}\|_{\mathbb{X}}^2 &= a(u_{\mathbb{Y}} - u_{\mathbb{W}}, u_{\mathbb{Y}} - u_{\mathbb{W}}) \\
&= a(u_{\mathbb{Y}} - u_{\mathbb{W}}, \epsilon\,\tilde{u}_{\mathbb{W}} + y_\xi - u_{\mathbb{W}}^{\mathrm{loc}}) \\
&= a(u_{\mathbb{Y}} - u_{\mathbb{W}}, y_\xi) - a(u_{\mathbb{Y}} - u_{\mathbb{W}}, u_{\mathbb{W}}^{\mathrm{loc}}).
\end{aligned}$$

Then, employing (2.9) with the test function $y_\xi \in \mathbb{Y}$, we have

$$a(u_{\mathbb{Y}} - u_{\mathbb{W}}, y_\xi) = b(y_\xi) - a(u_{\mathbb{W}}, y_\xi) = \rho_{\mathbb{W}}(y_\xi).$$

Similarly, applying (2.1) with the test function $u_{\mathbb{W}}^{\mathrm{loc}} \in \mathbb{W}$, we arrive at

$$-a(u_{\mathbb{Y}} - u_{\mathbb{W}}, u_{\mathbb{W}}^{\mathrm{loc}}) = -a(u_{\mathbb{Y}}, u_{\mathbb{W}}^{\mathrm{loc}}) + b(u_{\mathbb{W}}^{\mathrm{loc}}) = \rho_{\mathbb{Y}}(u_{\mathbb{W}}^{\mathrm{loc}}).$$

Combining the above equalities gives the second identity in (2.11). □

Remark 1 (Energy Reduction Property). If the space $\mathbb{Y}$ is a local enrichment of $\mathbb{W}^{\mathrm{loc}}$, i.e. if $\mathbb{W}^{\mathrm{loc}} \subset \mathbb{Y}$, then $\rho_{\mathbb{Y}}(u_{\mathbb{W}}^{\mathrm{loc}}) = 0$, and the identities (2.11) simplify to

$$\Delta e_{\mathbb{W},\mathbb{Y}}^2 = \|u_{\mathbb{Y}} - u_{\mathbb{W}}\|_{\mathbb{X}}^2 = \rho_{\mathbb{W}}(y_\xi);$$

in particular, in this case, the error resulting from the local enrichment will *not increase*, or will even *decrease* (i.e. $\|e_{\mathbb{W}}\|_{\mathbb{X}}^2 > \|e_{\mathbb{Y}}\|_{\mathbb{X}}^2$) if $u_{\mathbb{Y}} \neq u_{\mathbb{W}}$. In analogous terms, introducing the energy functional

$$\mathsf{E}(v) := \frac{1}{2}\,a(v, v) - b(v), \quad v \in \mathbb{X},$$

and recalling the Dirichlet variational principle, we note that

$$E(u_{\mathbb{W}}) = \min_{w \in \mathbb{W}} E(w) \quad \text{and} \quad E(u_{\mathbb{Y}}) = \min_{v \in \mathbb{Y}} E(v),$$

whence we immediately deduce the *energy reduction* property $E(u_{\mathbb{Y}}) \le E(u_{\mathbb{W}})$ for the solutions $u_{\mathbb{W}} \in \mathbb{W}$ and $u_{\mathbb{Y}} \in \mathbb{Y}$ from (2.1) and (2.9), respectively. In the general case, when $\mathbb{Y}$ is a local replacement space, a minor modification of the proof of Proposition 1, see Appendix A.1, yields the identity

$$\Delta e_{\mathbb{W},\mathbb{Y}}^2 = \|u_{\mathbb{Y}} - \tilde{u}_{\mathbb{W}}\|_{\mathbb{X}}^2 - \|u_{\mathbb{W}}^{\text{loc}}\|_{\mathbb{X}}^2, \tag{2.12}$$

which shows that the error is guaranteed to decrease, i.e. $\Delta e_{\mathbb{W},\mathbb{Y}}^2 > 0$, whenever $\|u_{\mathbb{Y}} - \tilde{u}_{\mathbb{W}}\|_{\mathbb{X}}^2 > \|u_{\mathbb{W}}^{\text{loc}}\|_{\mathbb{X}}^2$.

The *hp*-adaptive approach presented in Section 3 below can be seen as a local reduction strategy. We emphasize, however, that our method, in contrast to more classical approaches, always involves (at least one) global (but computationally inexpensive) mode in order to incorporate a minimal amount of global information on the underlying problem. Similar to the present section, in order to provide maximal flexibility for the choice of local spaces and associated enrichment functions in the *hp*-context, the theoretical building blocks will be derived in abstract form.

2.2 Linear Algebra Representation

We will now illustrate how the difference of the residuals $\rho_{\mathbb{W}}(y_\xi)$ and $\rho_{\mathbb{Y}}(u_{\mathbb{W}}^{\text{loc}})$ from Proposition 1 can be computed by means of linear algebra. To this end, for $\{\xi_1, \ldots, \xi_L\}$ from (2.8), we introduce the matrix $\mathbf{A} = (A_{ij}) \in \mathbb{R}^{L \times L}$ with entries

$$A_{ij} := a(\xi_j, \xi_i), \quad i, j \in \underline{L}, \tag{2.13a}$$

and the (column) vectors $\mathbf{b}, \mathbf{c} \in \mathbb{R}^L$, which are given component-wise by

$$b_i := b(\xi_i), \quad c_i := a(\tilde{u}_{\mathbb{W}}, \xi_i), \quad i \in \underline{L}, \tag{2.13b}$$

respectively. Furthermore, we let

$$a_{00} := \|u_{\mathbb{W}}\|_{\mathbb{X}}^2 - \|u_{\mathbb{W}}^{\text{loc}}\|_{\mathbb{X}}^2 - 2\,\delta \quad \text{with } \delta := b(u_{\mathbb{W}}^{\text{loc}}) - \|u_{\mathbb{W}}^{\text{loc}}\|_{\mathbb{X}}^2. \tag{2.13c}$$

Remark 2. In the finite element context, where the functions $\xi_1, \ldots, \xi_L$ and $u_{\mathbb{W}}^{\text{loc}}$ are supposedly local, we notice that all of the quantities defined in (2.13) are inexpensive to compute (except for the global term $\|u_{\mathbb{W}}\|_{\mathbb{X}}$ which, however, needs to be evaluated once only).

Proposition 2 (Low-Dimensional Computation of Error Reductions). *Suppose that* $\dim \mathbb{Y} = L + 1$ *(in particular, we implicitly assume that $\tilde{u}_{\mathbb{W}} \ne 0$), and consider the (symmetric) linear system*

$$\begin{pmatrix} a_{00} & \mathbf{c}^\top \\ \mathbf{c} & \mathbf{A} \end{pmatrix} \begin{pmatrix} \epsilon \\ \mathbf{y} \end{pmatrix} = \begin{pmatrix} \delta \\ \mathbf{b} - \mathbf{c} \end{pmatrix}, \tag{2.14}$$

with the solution components $\epsilon \in \mathbb{R}$ and $\mathbf{y} \in \mathbb{R}^L$. Then the predicted error change from Proposition 1 can be computed by means of the formula

$$\Delta e_{\mathbb{W},\mathbb{Y}}^2 = \mathbf{y}^\top(\mathbf{b} - \mathbf{c}) - \|u_{\mathbb{W}}^{\text{loc}}\|_{\mathbb{X}}^2 + \epsilon\,\delta. \tag{2.15}$$

Proof. Recalling the decomposition (2.6), we begin by noticing that

$$\|\tilde{u}_{\mathbb{W}}\|_{\mathbb{X}}^2 = \|u_{\mathbb{W}} - u_{\mathbb{W}}^{\text{loc}}\|_{\mathbb{X}}^2 = \|u_{\mathbb{W}}\|_{\mathbb{X}}^2 + \|u_{\mathbb{W}}^{\text{loc}}\|_{\mathbb{X}}^2 - 2\,a(u_{\mathbb{W}}, u_{\mathbb{W}}^{\text{loc}}),$$

which, upon applying (2.1) with the test function $w = u_{\mathbb{W}}^{\text{loc}}$, results in

$$\|\tilde{u}_{\mathbb{W}}\|_{\mathbb{X}}^2 = \|u_{\mathbb{W}}\|_{\mathbb{X}}^2 + \|u_{\mathbb{W}}^{\text{loc}}\|_{\mathbb{X}}^2 - 2\,b(u_{\mathbb{W}}^{\text{loc}}) = a_{00}. \tag{2.16}$$

Similarly, with $w = \tilde{u}_{\mathrm{W}}$ in (2.1), we infer that

$$b(\tilde{u}_{\mathrm{W}}) - a(\tilde{u}_{\mathrm{W}}, \tilde{u}_{\mathrm{W}}) = a(u_{\mathrm{W}}, \tilde{u}_{\mathrm{W}}) - a(\tilde{u}_{\mathrm{W}}, \tilde{u}_{\mathrm{W}}) = a(u_{\mathrm{W}}^{\mathrm{loc}}, \tilde{u}_{\mathrm{W}}) = a(\tilde{u}_{\mathrm{W}}, u_{\mathrm{W}}^{\mathrm{loc}}) = a(u_{\mathrm{W}}, u_{\mathrm{W}}^{\mathrm{loc}}) - a(u_{\mathrm{W}}^{\mathrm{loc}}, u_{\mathrm{W}}^{\mathrm{loc}})$$
$$= b(u_{\mathrm{W}}^{\mathrm{loc}}) - \|u_{\mathrm{W}}^{\mathrm{loc}}\|_{\mathbb{X}}^2,$$

from which we arrive at

$$\delta = b(\tilde{u}_{\mathrm{W}}) - a(\tilde{u}_{\mathrm{W}}, \tilde{u}_{\mathrm{W}}) = b(\tilde{u}_{\mathrm{W}}) - \|\tilde{u}_{\mathrm{W}}\|_{\mathbb{X}}^2. \tag{2.17}$$

Then, using (2.9) with $v = \tilde{u}_{\mathrm{W}} \in \mathbb{Y}$, and applying the representation (2.10), leads to

$$\delta = b(\tilde{u}_{\mathrm{W}}) - a(\tilde{u}_{\mathrm{W}}, \tilde{u}_{\mathrm{W}}) = a(u_{\mathbb{Y}}, \tilde{u}_{\mathrm{W}}) - a(\tilde{u}_{\mathrm{W}}, \tilde{u}_{\mathrm{W}}) = \epsilon\, a(\tilde{u}_{\mathrm{W}}, \tilde{u}_{\mathrm{W}}) + a(y_\xi, \tilde{u}_{\mathrm{W}}) = \epsilon\, \|\tilde{u}_{\mathrm{W}}\|_{\mathbb{X}}^2 + a(y_\xi, \tilde{u}_{\mathrm{W}}),$$

and equivalently, due to (2.16),

$$\epsilon\, a_{00} + a(y_\xi, \tilde{u}_{\mathrm{W}}) = \delta. \tag{2.18}$$

In addition, for $i \in \underline{L}$, testing (2.9) with $v = \xi_i$, and exploiting (2.10), yields

$$(1 + \epsilon)\, a(\tilde{u}_{\mathrm{W}}, \xi_i) + a(y_\xi, \xi_i) = b(\xi_i) \quad \text{for all } i \in \underline{L}. \tag{2.19}$$

Therefore, applying the linear combination

$$y_\xi = \sum_{j \in \underline{L}} y_j\, \xi_j, \tag{2.20}$$

for a uniquely defined coefficient vector $\mathbf{y} := (y_1, \ldots, y_L)^\top \in \mathbb{R}^L$, equations (2.18) and (2.19) transform into

$$\epsilon\, a_{00} + \sum_{j \in \underline{L}} a(\tilde{u}_{\mathrm{W}}, \xi_j)\, y_j = \delta,$$

and

$$\epsilon\, a(\tilde{u}_{\mathrm{W}}, \xi_i) + \sum_{j \in \underline{L}} a(\xi_j, \xi_i)\, y_j = b(\xi_i) - a(\tilde{u}_{\mathrm{W}}, \xi_i) \quad \text{for all } i \in \underline{L},$$

which is the linear system (2.14). Moreover, owing to (2.6), it holds that

$$\rho_{\mathrm{W}}(y_\xi) = b(y_\xi) - a(u_{\mathrm{W}}, y_\xi) = b(y_\xi) - a(\tilde{u}_{\mathrm{W}}, y_\xi) - a(u_{\mathrm{W}}^{\mathrm{loc}}, y_\xi).$$

In addition, invoking (2.10) and (2.1), we have

$$\begin{aligned}
-\rho_{\mathbb{Y}}(u_{\mathrm{W}}^{\mathrm{loc}}) &= a(u_{\mathbb{Y}}, u_{\mathrm{W}}^{\mathrm{loc}}) - b(u_{\mathrm{W}}^{\mathrm{loc}}) \\
&= (1 + \epsilon)\, a(\tilde{u}_{\mathrm{W}}, u_{\mathrm{W}}^{\mathrm{loc}}) + a(y_\xi, u_{\mathrm{W}}^{\mathrm{loc}}) - a(u_{\mathrm{W}}, u_{\mathrm{W}}^{\mathrm{loc}}) \\
&= -a(u_{\mathrm{W}}^{\mathrm{loc}}, u_{\mathrm{W}}^{\mathrm{loc}}) + \epsilon\, a(\tilde{u}_{\mathrm{W}}, u_{\mathrm{W}}^{\mathrm{loc}}) + a(y_\xi, u_{\mathrm{W}}^{\mathrm{loc}}) \\
&= -\|u_{\mathrm{W}}^{\mathrm{loc}}\|_{\mathbb{X}}^2 + \epsilon(a(u_{\mathrm{W}}, \tilde{u}_{\mathrm{W}}) - a(\tilde{u}_{\mathrm{W}}, \tilde{u}_{\mathrm{W}})) + a(y_\xi, u_{\mathrm{W}}^{\mathrm{loc}}) \\
&= -\|u_{\mathrm{W}}^{\mathrm{loc}}\|_{\mathbb{X}}^2 + \epsilon(b(\tilde{u}_{\mathrm{W}}) - \|\tilde{u}_{\mathrm{W}}\|_{\mathbb{X}}^2) + a(y_\xi, u_{\mathrm{W}}^{\mathrm{loc}}).
\end{aligned}$$

Recalling (2.17) implies that

$$-\rho_{\mathbb{Y}}(u_{\mathrm{W}}^{\mathrm{loc}}) = -\|u_{\mathrm{W}}^{\mathrm{loc}}\|_{\mathbb{X}}^2 + \epsilon\, \delta + a(y_\xi, u_{\mathrm{W}}^{\mathrm{loc}}).$$

Thus, upon involving the linear combination (2.20), we finally obtain

$$\rho_{\mathrm{W}}(y_\xi) - \rho_{\mathbb{Y}}(u_{\mathrm{W}}^{\mathrm{loc}}) = b(y_\xi) - a(\tilde{u}_{\mathrm{W}}, y_\xi) - \|u_{\mathrm{W}}^{\mathrm{loc}}\|_{\mathbb{X}}^2 + \epsilon\, \delta = \sum_{j \in \underline{L}} y_j\, (b(\xi_j) - a(\tilde{u}_{\mathrm{W}}, \xi_j)) - \|u_{\mathrm{W}}^{\mathrm{loc}}\|_{\mathbb{X}}^2 + \epsilon\, \delta,$$

which, in view of (2.11), leads to the asserted identity (2.15). ☐

2.3 Assembling Aspects

We aim to apply the abstract framework developed in the previous sections to the finite element context. For this purpose, we let $\mathbb{X} = \mathbb{X}(\Omega)$ be a Hilbert function space defined over a bounded open domain $\Omega \subset \mathbb{R}^d$, $d \geq 1$.

We present the assembling of the matrix $\mathbf{A}$ and the vectors $\mathbf{b}$, $\mathbf{c}$ occurring in the linear system (2.14) in the specific case where $\mathbb{W} \subseteq \mathbb{X}$ is a subspace of finite dimension $N := \dim \mathbb{W}$ with the basis $\{\phi_1, \ldots, \phi_N\}$. Let $\mathcal{D}$ be a decomposition of Ω into closed subsets $K \subseteq \Omega$, i.e.

$$\overline{\Omega} = \bigcup_{K \in \mathcal{D}} K,$$

where $\text{int}(K) \cap \text{int}(K') = \emptyset$ for any $K, K' \in \mathcal{D}$ with $K \neq K'$. For any $K \in \mathcal{D}$, consider the (local) Hilbert space consisting of all restrictions of functions $v \in \mathbb{X}$ to K, i.e. $\mathbb{X}_K := \{v_{|K} : v \in \mathbb{X}\}$. Moreover, let $\{\zeta_1^K, \ldots, \zeta_{M_K}^K\} \subseteq \mathbb{X}_K$ be a set of functions such that there exist representation matrices $\mathbf{C}_K = (c_{ij}^K) \in \mathbb{R}^{N \times M_K}$ and $\mathbf{D}_K = (d_{ij}^K) \in \mathbb{R}^{L \times M_K}$ satisfying

$$\phi_{i|K} = \sum_{j \in \underline{M_K}} c_{ij}^K \zeta_j^K, \quad i \in \underline{N}, \tag{2.21}$$

$$\xi_{i|K} = \sum_{j \in \underline{M_K}} d_{ij}^K \zeta_j^K, \quad i \in \underline{L}, \tag{2.22}$$

with $\{\xi_1, \ldots, \xi_L\}$ from (2.8). Finally, let the bilinear form $a(\cdot, \cdot)$ as well as the linear form $b(\cdot)$ be decomposable in the sense that

$$a(v, w) = \sum_{K \in \mathcal{D}} a_K(v_{|K}, w_{|K}), \quad b(v) = \sum_{K \in \mathcal{D}} b_K(v_{|K}) \quad \text{for all } v, w \in \mathbb{X}, \tag{2.23}$$

for some (local) bilinear forms $a_K : \mathbb{X}_K \times \mathbb{X}_K \to \mathbb{R}$ and (local) linear forms $b_K : \mathbb{X}_K \to \mathbb{R}$. For any $K \in \mathcal{D}$ we introduce the (local) matrix $\mathbf{A}_K = (A_{ij}^K) \in \mathbb{R}^{M_K \times M_K}$ with entries

$$A_{ij}^K := a_K(\zeta_j^K, \zeta_i^K), \quad i, j \in \underline{M_K},$$

and the (local) vector $\mathbf{b}_K = (b_i^K) \in \mathbb{R}^{M_K}$, which is given component-wise by

$$b_i^K := b_K(\zeta_i^K), \quad i \in \underline{M_K}.$$

Let $\tilde{\mathbf{u}} = (u_1, \ldots, u_N)^\top \in \mathbb{R}^N$ denote the uniquely determined coefficient vector of $\tilde{u}_\mathbb{W}$ from (2.6) with respect to the basis $\{\phi_1, \ldots, \phi_N\}$ of $\mathbb{W}$, i.e. $\tilde{u}_\mathbb{W} = \sum_{i \in \underline{N}} u_i \phi_i$. Note that $u_i = 0$ for $i \in \mathcal{J}^{\text{loc}}$. Then the matrix $\mathbf{A} \in \mathbb{R}^{L \times L}$ and the vectors $\mathbf{b}, \mathbf{c} \in \mathbb{R}^L$ from the linear system (2.14) can be assembled by the (local) quantities $\mathbf{A}_K, \mathbf{C}_K, \mathbf{D}_K$, and $\mathbf{b}_K$ as the following result shows.

Proposition 3 (Assembling). *The identities*

$$\mathbf{A} = \sum_{K \in \mathcal{D}} \mathbf{D}_K \mathbf{A}_K \mathbf{D}_K^\top, \quad \mathbf{b} = \sum_{K \in \mathcal{D}} \mathbf{D}_K \mathbf{b}_K, \quad \mathbf{c} = \sum_{K \in \mathcal{D}} \mathbf{D}_K \mathbf{A}_K \mathbf{C}_K^\top \tilde{\mathbf{u}} \tag{2.24}$$

hold true.

Proof. Let $[\cdot]_{ij}$ and $[\cdot]_i$ denote the components of a matrix or a vector, respectively. Then it holds that

$$\left[\sum_{K \in \mathcal{D}} \mathbf{D}_K \mathbf{A}_K \mathbf{D}_K^\top \right]_{ij} = \sum_{K \in \mathcal{D}} [\mathbf{D}_K \mathbf{A}_K \mathbf{D}_K^\top]_{ij} = \sum_{K \in \mathcal{D}} \left(\sum_{\ell \in \underline{M_K}} \sum_{k \in \underline{M_K}} d_{ik}^K a_K(\zeta_\ell^K, \zeta_k^K) d_{j\ell}^K \right)$$

$$= \sum_{K \in \mathcal{D}} a_K \left(\sum_{\ell \in \underline{M_K}} d_{j\ell}^K \zeta_\ell^K, \sum_{k \in \underline{M_K}} d_{ik}^K \zeta_k^K \right)$$

$$= \sum_{K \in \mathcal{D}} a_K(\xi_{j|K}, \xi_{i|K}) = a(\xi_j, \xi_i) = A_{ij},$$

for all $i, j \in \underline{L}$, cf. (2.22) and (2.23) which is the first identity. The second identity follows from

$$\left[\sum_{K \in \mathcal{D}} \mathbf{D}_K \mathbf{b}_K \right]_i = \sum_{K \in \mathcal{D}} [\mathbf{D}_K \mathbf{b}_K]_i = \sum_{K \in \mathcal{D}} \left(\sum_{j \in \underline{M_K}} d_{ij}^K b_K(\zeta_j^K) \right)$$

$$= \sum_{K \in \mathcal{D}} b_K \left(\sum_{j \in \underline{M_K}} d_{ij}^K \zeta_j^K \right)$$

$$= \sum_{K \in \mathcal{D}} b_K(\xi_{i|K}) = b(\xi_i) = b_i,$$

for $i \in \underline{L}$. The last identity follows in an analogous way. $\qquad\square$

3 Application to *hp*-Finite Element Spaces

For $d \geq 1$, let $\Omega \subset \mathbb{R}^d$ be a bounded interval (if $d = 1$), or a bounded and open set with a Lipschitz boundary $\Gamma := \partial\Omega$ that is composed of a finite number of straight faces (if $d \geq 2$). Furthermore, we consider a boundary part $\Gamma_D \subseteq \Gamma$ of positive surface measure, and introduce an associated Hilbert space

$$\mathbb{X} := \{v \in H^1(\Omega) : v = 0 \text{ on } \Gamma_D\}, \tag{3.1}$$

where $H^1(\Omega)$ denotes the usual Sobolev space of all functions in $L^2(\Omega)$, with weak first-order partial derivatives in $L^2(\Omega)$.

The goal of the following subsections is to define appropriate *hp*-finite element approximation subspaces in $\mathbb{X}$. For this purpose, we will begin by introducing a family of hierarchical polynomial spaces on the d-dimensional hypercube signified by $\widehat{Q} := [-1, 1]^d$, which will be referred to as the *reference element*.

3.1 Polynomial Spaces on the Reference Element

For any $j \in \mathbb{N}_0$ we introduce the 1-dimensional functions $\psi_j : [-1, 1] \rightarrow \mathbb{R}$, given by

$$\psi_0(t) := \frac{1}{2}(1 - t), \quad \psi_1(t) := \frac{1}{2}(1 + t), \quad \psi_j(t) := \int_{-1}^{t} L_{j-1}(s)\,ds, \quad j \geq 2, \tag{3.2}$$

where $L_j : [-1, 1] \rightarrow \mathbb{R}$ denotes the j-th Legendre polynomial, normalized such that $L_j(-1) = (-1)^j$, $j \geq 1$; we recall the *Bonnet recursion formula*

$$L_0(t) := 1, \quad L_1(t) := t, \quad j\,L_j(t) = (2j - 1)\,t\,L_{j-1}(t) - (j - 1)\,L_{j-2}(t), \quad j \geq 2,$$

as well as the relation

$$\psi_j(t) = \frac{L_j(t) - L_{j-2}(t)}{2j - 1}, \quad j \geq 2;$$

see, e.g., [30, Section 3.1 and A.4]. The basis functions ψ_0 and ψ_1 are referred to as *nodal* or *external shape functions* whereas the high-order polynomials ψ_j, for $j \geq 2$, are called *internal modes* (based on the fact that $\psi_j(\pm 1) = 0$ for $j \geq 2$). We display the functions $\psi_0, \ldots, \psi_4$ in Figure 1.
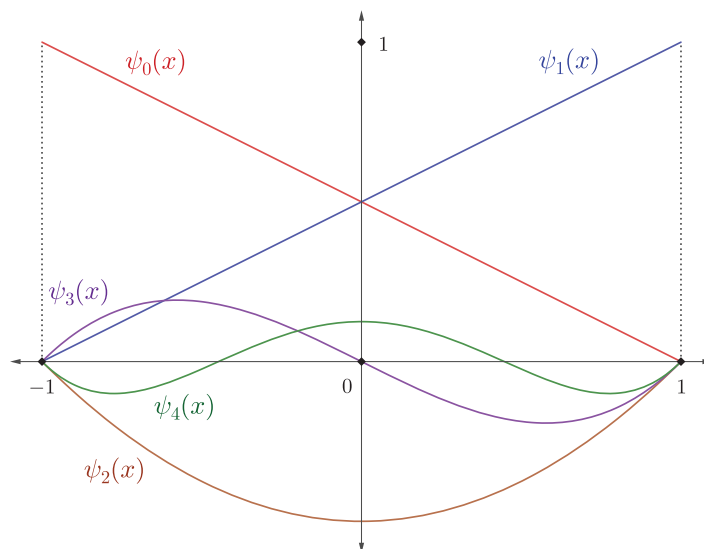

Figure 1: The functions $\psi_0, \ldots, \psi_4$ on $[-1, 1]$.

Furthermore, in the multi-dimensional case, for any multi-index $\boldsymbol{j} = (j_1, \ldots, j_d) \in \mathbb{N}_0^d$, we define the functions $\widehat{\psi}_{\boldsymbol{j}} : \widehat{Q} \to \mathbb{R}$ by

$$\widehat{\psi}_{\boldsymbol{j}}(\mathbf{x}) := \prod_{k \in \underline{d}} \psi_{j_k}(x_k), \quad \mathbf{x} = (x_1, \ldots, x_d)^{\top} \in \widehat{Q}. \tag{3.3}$$

Moreover, for any $k \in \mathbb{N}$, let $\mathbb{Q}_k(\widehat{Q}) := \text{span}\{\widehat{\psi}_{\boldsymbol{j}} : \boldsymbol{j} \in \underline{k}_0^d\}$ signify the usual space of all polynomials up to degree $k \in \mathbb{N}$ in each coordinate direction on the reference element $\widehat{Q}$.

3.2 *hp*-Finite Element Space

The vertices of the reference element $\widehat{Q}$ will be labeled in terms of a multi-index notation. Specifically, to each $\boldsymbol{i} \in \{0, 1\}^d$, we associate a corresponding vertex

$$\widehat{\mathbf{v}}_{\boldsymbol{i}} := (2i_1 - 1, \ldots, 2i_d - 1)^{\top}.$$

Similarly, we consider physical elements $Q \subset \Omega$ whose corner points are indexed by $\mathbf{v}_i \in \mathbb{R}^d$, $i = 1, \ldots, 2^d$; we call Q a *transformed hexahedron* if the map $F_Q : \widehat{Q} \to Q$ defined by

$$F_Q(\mathbf{x}) := \sum_{|\boldsymbol{i}| \leq 1} \widehat{\psi}_{\boldsymbol{i}}(\mathbf{x}) \, \mathbf{v}_{\boldsymbol{i}}, \quad \mathbf{x} = (x_1, \ldots, x_d)^{\top} \in \widehat{Q}, \tag{3.4}$$

is bijective, where we write $|\boldsymbol{i}| := i_1 + \cdots + i_d$ to signify the order of a multi-index $\boldsymbol{i} = (i_1, \ldots, i_d) \in \mathbb{N}_0^d$. Note that it holds $F_Q(\widehat{\mathbf{v}}_{\boldsymbol{i}}) = \mathbf{v}_{\boldsymbol{i}}$ for all $\boldsymbol{i} \in \{0, 1\}^d$. The notation is illustrated in Figure 2 for the 2-dimensional case.

Example 1 (F_Q for $d = 2$). In this case, the transformation F_Q from (3.4) takes the form

$$F_Q(\mathbf{x}) = \widehat{\psi}_{00}(\mathbf{x}) \, \mathbf{v}_{00} + \widehat{\psi}_{10}(\mathbf{x}) \, \mathbf{v}_{10} + \widehat{\psi}_{11}(\mathbf{x}) \, \mathbf{v}_{11} + \widehat{\psi}_{01}(\mathbf{x}) \, \mathbf{v}_{01}, \quad \mathbf{x} = (x_1, x_2)^{\top} \in \widehat{Q},$$

and the tensor product functions $\widehat{\psi}_{\boldsymbol{i}}$ are given by

$$\widehat{\psi}_{00}(\mathbf{x}) = \psi_0(x_1) \, \psi_0(x_2), \quad \widehat{\psi}_{10}(\mathbf{x}) = \psi_1(x_1) \, \psi_0(x_2),$$
$$\widehat{\psi}_{11}(\mathbf{x}) = \psi_1(x_1) \, \psi_1(x_2), \quad \widehat{\psi}_{01}(\mathbf{x}) = \psi_0(x_1) \, \psi_1(x_2).$$

For the purpose of introducing finite element subspaces $\mathbb{W} \subseteq \mathbb{X}$, cf. (3.1), following our abstract framework in Section 2.3, we let $\mathfrak{Q}$ be a decomposition of Ω into transformed hexahedrons with the following additional property: If $Q_1 \cap Q_2 \neq \emptyset$ for $Q_1, Q_2 \in \mathfrak{Q}$ with $Q_1 \neq Q_2$, then $Q_1 \cap Q_2$ represents a $(d - r)$-dimensional face of Q_1 or Q_2 for some $r \in \underline{d}$; we call any faces of dimension zero, one and two *vertices*, *edges* and *faces*, respectively. Furthermore, let us assume that there is no change of the type of boundary conditions within the $(d - r)$-dimensional faces of one element $Q \in \mathfrak{Q}$. We now define the *hp*-finite element space

$$\mathbb{W} := \{v \in \mathbb{X} : v_{|Q} \circ F_Q \in \mathbb{Q}_{p_Q}(\widehat{Q}) \quad \text{for all } Q \in \mathfrak{Q}\}, \tag{3.5}$$

associated with the decomposition $\mathfrak{Q}$ of Ω, where $p_Q \geq 1$ represents an (isotropic) local polynomial degree on each element $Q \in \mathfrak{Q}$.
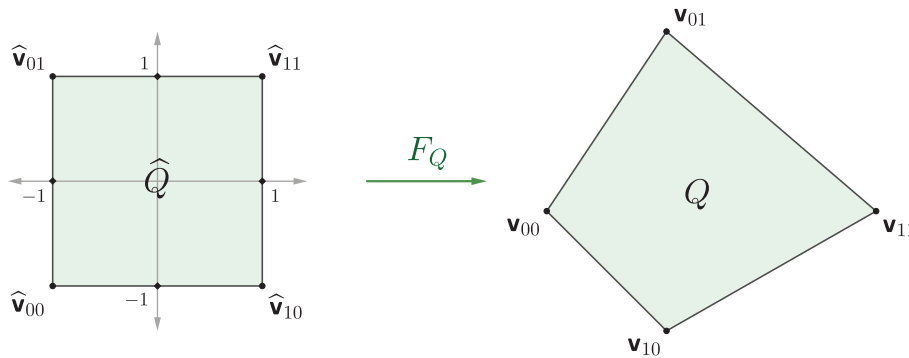

Figure 2: The bijective mapping $F_Q : \widehat{Q} \to Q$ for the 2-dimensional case.

3.3 Constraint Coefficients

In the adaptive finite element procedure we shall represent the functions $\widehat{\psi}_{\boldsymbol{i}}$ from (3.3), defined on the reference element $\widehat{Q}$, in terms of functions defined on sub-hexahedra of $\widehat{Q}$. These representations can be computed efficiently by means of so-called *constraint coefficients*, see [7, 26]. To explain this concept, consider a sub-hexahedron

$$T := \prod_{k \in \underline{d}} I_k \subset \widehat{Q}, \tag{3.6}$$

where $I_k = [a_k, b_k]$, $-1 \le a_k < b_k \le 1$, are 1-dimensional intervals for each $k \in \underline{d}$. Then we define the functions

$$\widetilde{\psi}_{\boldsymbol{j}} : T \to \mathbb{R}, \quad \widetilde{\psi}_{\boldsymbol{j}} := \widehat{\psi}_{\boldsymbol{j}} \circ F_T^{-1}, \quad \boldsymbol{j} \in \mathbb{N}_0, \tag{3.7}$$

with the bijective element map $F_T : \widehat{Q} \to T$, cf. (3.4). Here, due to the tensor structure (3.6) of T, we emphasize that F_T is the composition of a dilation and a translation, wherefore we observe the identity

$$\widetilde{\psi}_{\boldsymbol{j}} = \prod_{k \in \underline{d}} (\psi_{j_k} \circ F_{I_k}^{-1}). \tag{3.8}$$

In particular, we infer that the functions $\widetilde{\psi}_{\boldsymbol{j}}$ in (3.7) constitute a polynomial basis on T. Therefore, there exist uniquely determined *constraint coefficients* $b_{\boldsymbol{ij}}^T \in \mathbb{R}$ such that

$$\widehat{\psi}_{\boldsymbol{i}\,|T} = \sum_{\boldsymbol{j} \le \boldsymbol{i}} b_{\boldsymbol{ij}}^T \widetilde{\psi}_{\boldsymbol{j}} = \sum_{\boldsymbol{j} \le \boldsymbol{i}} b_{\boldsymbol{ij}}^T (\widehat{\psi}_{\boldsymbol{j}} \circ F_T^{-1}), \quad \boldsymbol{i} \in \mathbb{N}_0^d, \tag{3.9}$$

where the sums are taken over all multi-indices $\boldsymbol{j} = (j_1, \ldots, j_d) \in \mathbb{N}_0^d$ with $j_k \le i_k$ for all $k \in \underline{d}$.

Example 2 ($d = 1$). In Figure 3 we illustrate how the restriction of the 1-dimensional function ψ_2 to the interval $I = [-1, 0]$ can be expressed in terms of the functions $\widetilde{\psi}_0$, $\widetilde{\psi}_1$, and $\widetilde{\psi}_2$.

The ensuing result shows that the multi-dimensional constraint coefficients can be expressed in terms of tensor-products of the associated 1-dimensional quantities for which recursion formulas are stated in the Appendix.

Lemma 1. *For $\boldsymbol{i}, \boldsymbol{j} \in \mathbb{N}_0^d$, the constraint coefficients $b_{\boldsymbol{ij}}^T$ from* (3.9) *are given by*

$$b_{\boldsymbol{ij}}^T = \prod_{k \in \underline{d}} b_{i_k, j_k}^{I_k},$$

where T is the sub-hexahedra from (3.6) *and $b_{i_k, j_k}^{I_k}$ are the uniquely determined* 1-*dimensional constraint coefficients from*

$$\psi_{i_k | I_k} = \sum_{j_k=0}^{i_k} b_{i_k, j_k}^{I_k} (\psi_{j_k} \circ F_{I_k}^{-1}).$$

Proof. The argument is based on exploiting the tensor structure of $\widehat{Q}$ and T, and of the functions $\widehat{\psi}_{\boldsymbol{i}}$. Indeed, for $\boldsymbol{i} = (i_1, \ldots, i_d) \in \mathbb{N}_0^d$, using (3.3) and (3.6), and applying the representation (3.9) in the 1-dimensional case, we obtain that

$$\widehat{\psi}_{\boldsymbol{i}\,|T} = \prod_{k \in \underline{d}} \psi_{i_k | I_k} = \prod_{k \in \underline{d}} \left(\sum_{j_k \le i_k} b_{i_k, j_k}^{I_k} (\psi_{j_k} \circ F_{I_k}^{-1}) \right).$$

Then rearranging terms yields

$$\widehat{\psi}_{\boldsymbol{i}\,|T} = \sum_{\boldsymbol{j} \le \boldsymbol{i}} \left(\prod_{k \in \underline{d}} b_{i_k, j_k}^{I_k} (\psi_{j_k} \circ F_{I_k}^{-1}) \right) = \sum_{\boldsymbol{j} \le \boldsymbol{i}} \left(\prod_{k \in \underline{d}} b_{i_k, j_k}^{I_k} \prod_{k \in \underline{d}} (\psi_{j_k} \circ F_{I_k}^{-1}) \right).$$

Owing to (3.8), we conclude that

$$\widehat{\psi}_{\boldsymbol{i}\,|T} = \sum_{\boldsymbol{j} \le \boldsymbol{i}} \left(\prod_{k \in \underline{d}} b_{i_k, j_k}^{I_k} \right) \widetilde{\psi}_{\boldsymbol{j}}.$$

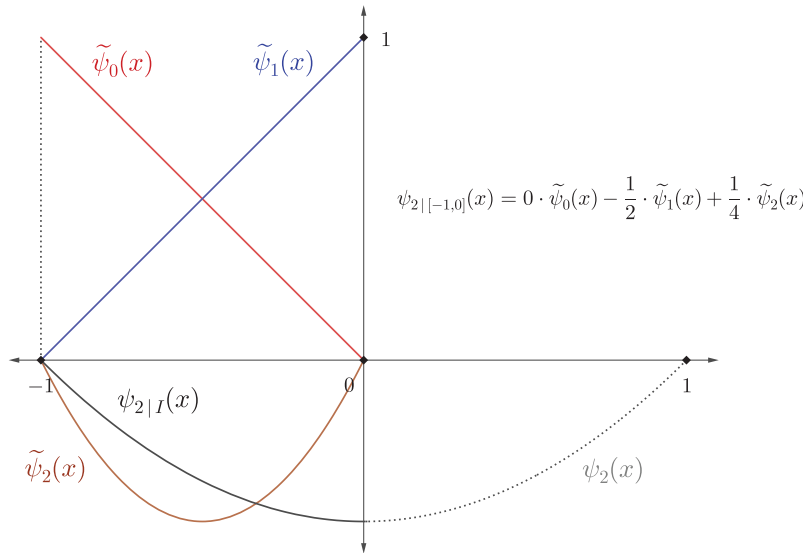The assertion follows from the uniqueness of the coefficients in (3.9). □

Figure 3: The restriction of ψ_2 to I in terms of the functions $\widetilde{\psi}_0, \widetilde{\psi}_1, \widetilde{\psi}_2$ in the 1-dimensional case.

3.4 Enrichment Functions Associated with Individual Elements

We will characterize refinements of elements $Q \in \mathcal{Q}$ in a given mesh $\mathcal{Q}$ via corresponding refinements of the reference element $\widehat{Q}$. To this end, given a (fixed) point $\widehat{\mathbf{z}} \in (-1, 1)^d$ in the interior of $\widehat{Q}$, we call a decomposition $\mathcal{R}(\widehat{Q})$ of $\widehat{Q}$ into the 2^d sub-hexahedra

$$\widehat{T}_{\boldsymbol{i}} := \prod_{k \in \underline{d}} [a_k^{\boldsymbol{i}}, b_k^{\boldsymbol{i}}], \quad \boldsymbol{i} = (i_1, \dots, i_d) \in \{0, 1\}^d,$$

with

$$a_k^{\boldsymbol{i}} := \min\{2\, i_k - 1, z_k\}, \quad b_k^{\boldsymbol{i}} := \max\{2\, i_k - 1, z_k\}, \quad k \in \underline{d},$$

a *refinement of $\widehat{Q}$ with respect to $\widehat{\mathbf{z}}$* $= (z_1, \dots, z_d)^\top$.

Let us now focus on some element $Q \in \mathcal{Q}$. To introduce a refinement of Q, we consider first a refinement $\mathcal{R}(\widehat{Q})$ of $\widehat{Q}$ with respect to $\widehat{\mathbf{z}} \in (-1, 1)^d$, as outlined above, and define

$$\mathcal{R}(Q) := \{T_{\boldsymbol{i}} : \boldsymbol{i} \in \{0, 1\}^d\}, \quad T_{\boldsymbol{i}} := F_Q(\widehat{T}_{\boldsymbol{i}}).$$

We display an illustration of the 2-dimensional situation in Figure 4. Note that the multi-index $\boldsymbol{i}$ encodes the location of the sub-hexahedra $T_{\boldsymbol{i}}$ with respect to the Cartesian reference coordinate system in $\widehat{Q}$; in fact, the k-th entry of $\boldsymbol{i}$ classifies whether $\widehat{T}_{\boldsymbol{i}}$ lies in the left (if $i_k = 0$) or the right (if $i_k = 1$) half of Q with respect to $\widehat{\mathbf{z}}$ along the k-th axial direction. Finally, let us denote by $\mathcal{T}$ the resulting mesh when replacing the element Q by its refinement $\mathcal{R}(Q)$, i.e. $\mathcal{T} := (\mathcal{Q} \setminus \{Q\}) \cup \mathcal{R}(Q)$.

3.4.1 Enrichment Strategies on a Single Element

For an element $Q \in \mathcal{Q}$ with an associated polynomial degree p_Q, two types of local enrichment functions $\xi_1, \dots, \xi_L$, cf. (2.8), will be considered:
(i) For the definition of *p-enrichment functions on Q*, we consider polynomials on the reference element $\widehat{Q}$, with polynomial degrees larger than p_Q, and transform them to the physical element Q.
(ii) For the construction of *hp-enrichment functions on Q*, we consider polynomials on the sub-hexahedra $\widehat{T}_{\boldsymbol{i}} \in \mathcal{R}(\widehat{Q})$, $\boldsymbol{i} \in \{0, 1\}^d$, of a refinement $\mathcal{R}(\widehat{Q})$ of $\widehat{Q}$ with respect to some $\widehat{\mathbf{z}} \in (-1, 1)^d$, which are then transformed to Q.
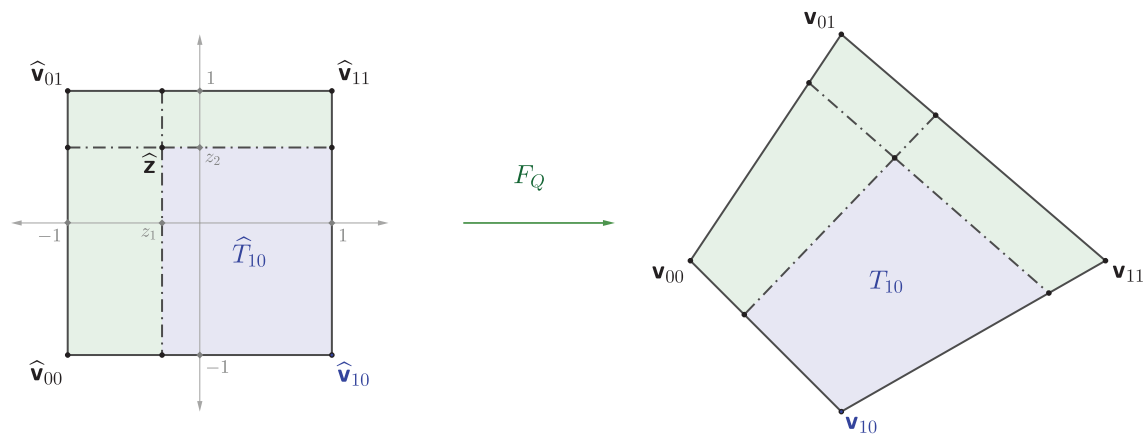
Figure 4: A refinement of the reference element $\widehat{Q}$ with respect to $\widehat{\mathbf{z}} = (z_1, z_2)^\top \in (-1, 1)^2$ and the corresponding refinement of the element $Q = F_Q(\widehat{Q})$.
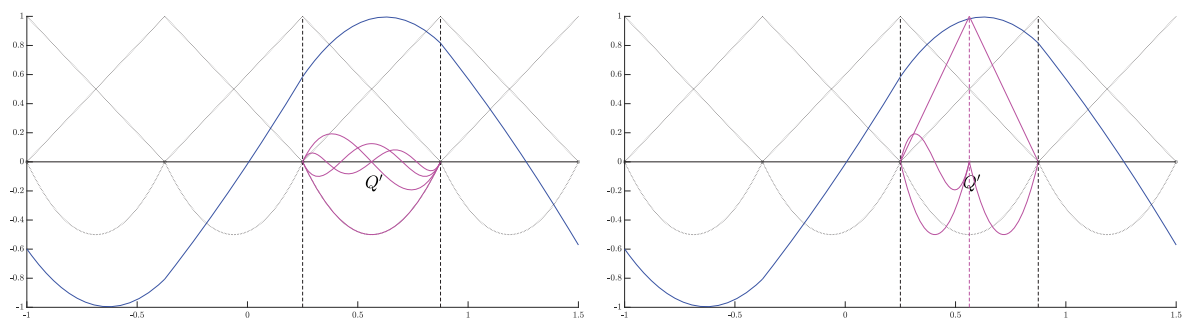

Figure 5: A *p*-enrichment (left) vs. an *hp*-refinement (right). In both figures the discrete solution $u_{\mathbb{W}}$ is highlighted in blue, the in each case four enrichment functions are depicted in magenta and the basis $\{\phi_1, \ldots, \phi_N\}$ of $\mathbb{W}$ is indicated in dotted lines.

We give an illustration of these two scenarios for the *p*-enrichment or *hp*-refinement of a 1-dimensional element in Figure 5. The polynomial functions resulting from the above mappings from $\widehat{Q}$ or from sub-hexahedra of $\widehat{Q}$ to Q will be termed *transformed polynomials*.

In the *hp*-adaptive procedure described in Section 4 we aim to compare different *p*-enrichments and *hp*-refinements on Q; in particular *competitive* refinements, cf. [20], where different enrichments, which generate the same number of degrees of freedom, are compared with each other in view of a maximal potential predicted error reduction. We will express both, the *hp*-enrichment functions (given on the subelements $T_i \in \mathcal{R}(Q)$) as well as the *p*-enrichment functions (given on Q) in terms of the transformed polynomials $\zeta_j^i : T_i \to \mathbb{R}$,

$$\zeta_j^i(\mathbf{x}) := \widehat{\psi}_j \circ F_i^{-1}(\mathbf{x}), \quad \mathbf{x} \in T_i, \tag{3.10}$$

for any $j \in \mathbb{N}_0^d$ and $i \in \{0, 1\}^d$. For ease of notation, we write F_i for the bijective mapping $F_{T_i} : \widehat{Q} \to T_i$, cf. (3.4); similarly, we denote by $\mathbf{A}_i, \mathbf{C}_i, \mathbf{D}_i$, and $\mathbf{b}_i$ the local quantities $\mathbf{A}_{T_i}, \mathbf{C}_{T_i}, \mathbf{D}_{T_i}$ and $\mathbf{b}_{T_i}$ from Section 2.3. We emphasize that the quantities $\mathbf{A}_i, \mathbf{C}_i$ and $\mathbf{b}_i$, for *each* of the different enrichments on Q to be compared, need to be computed *once only*.

3.4.2 *p*-Enrichments on *Q*

For any multi-index $j = (j_1, \ldots, j_d) \in \mathbb{N}_0^d$, let us introduce the functions $\xi_j : \Omega \to \mathbb{R}$ by

$$\xi_j(\mathbf{x}) := \begin{cases} \widehat{\psi}_j \circ F_Q^{-1}(\mathbf{x}) & \text{if } \mathbf{x} \in Q, \\ 0 & \text{if } \mathbf{x} \in \Omega \setminus Q. \end{cases} \tag{3.11}$$

For the *p*-enrichments on Q we choose a finite subset of the functions

$$\mathfrak{E}_p = \{\xi_{\boldsymbol{j}} : \boldsymbol{j} \in \boldsymbol{J}_d\}, \tag{3.12}$$

for some index set

$$\boldsymbol{J}_d \subseteq \{\boldsymbol{j} \in \mathbb{N}_0^d : j_k \geq 2 \text{ for } k \in \underline{d}\} \quad \text{with } |\boldsymbol{J}_d| = L < \infty.$$

We emphasize that we exclusively consider transformed (higher-order bubble-type) polynomials that vanish along the boundary of Q; cf. Remark 7 below for some generalizations to patches. In particular, we have the following result:

Proposition 4. *Any* $\xi \in \mathfrak{E}_p$ *is continuous on* Ω, *and it holds* $\operatorname{supp}(\xi) = Q$.

Obviously, there are various possibilities to choose appropriate *p*-enrichment functions. For instance, if we select all transformed polynomials up to a certain polynomial degree $p_{\max} > p_Q$, we have $\boldsymbol{J}_d = \{2, \ldots, p_{\max}\}^d$, whereas for the so-called *hierarchical surplus* we choose

$$\boldsymbol{J}_d = \{\boldsymbol{j} \in \{2, \ldots, p_{\max}\}^d : |\boldsymbol{j}| \geq p_Q + 2\}.$$

3.4.3 *hp*-Enrichment Functions on Q

In the case of *hp*-refinements, we construct local enrichment functions that can be associated with those r-dimensional faces of the refinement $\mathcal{R}(Q)$ which do not lie on the boundary ∂Q. We will call such faces *internal nodes* of the refinement $\mathcal{R}(Q)$. In this sense, the *vertex* $F_Q(\boldsymbol{\mathfrak{z}})$ is the only 0-dimensional internal node of $\mathcal{R}(Q)$, *edges* in the interior of Q are the 1-dimensional internal nodes of $\mathcal{R}(Q)$, and the *elements* $T_{\boldsymbol{i}}$, for $\boldsymbol{i} \in \{0,1\}^d$, represent the d-dimensional internal nodes of $\mathcal{R}(Q)$.

Indexing of Internal Nodes. Any r-dimensional internal node of $\mathcal{R}(Q)$, $r \in \underline{d}_0$, can be identified uniquely by r axial directions of $\widehat{Q}$, which are represented by an *orientation tuple*

$$\boldsymbol{a} \in D_r := \{(a_1, \ldots, a_r) \in \underline{d}^r : a_1 < \ldots < a_r\},$$

together with a *location tuple* $\boldsymbol{\ell} \in \{0,1\}^r$ that fixes its position with respect to the center point $F_Q(\boldsymbol{\mathfrak{z}})$. For any $\boldsymbol{a} = (a_1, \ldots, a_r) \in D_r$, let us denote by $A(\boldsymbol{a})$ the set of its components, i.e. $A(\boldsymbol{a}) := \{a_k : k \in \underline{r}\}$. Note that a tuple $\boldsymbol{a} \in D_r$ describes the orientation (and, implicitly, contains the dimension r) of an internal node with respect to the Cartesian reference coordinate system in $\widehat{Q} = F_Q^{-1}(Q)$. Moreover, the k-th entry of the location tuple $\boldsymbol{\ell} = (\ell_1, \ldots, \ell_r) \in \{0,1\}^d$ defines whether an internal node with orientation $\boldsymbol{a} = (a_1, \ldots, a_r)$ lies in the left (if $\ell_k = 0$) or in the right (if $\ell_k = 1$) half of the refinement $\mathcal{R}(Q)$ along the k-th axial direction of the reference coordinate system in $\widehat{Q} = F_Q^{-1}(Q)$. Observe further that, for any $d \geq 1$, the only 0-dimensional internal node of $\mathcal{R}(Q)$ is represented by the empty tuples $\boldsymbol{a} = ()$ and $\boldsymbol{\ell} = ()$.

Example 3 (Internal Nodes for the 3-Dimensional Case). The 0-dimension center point is given by the empty tuples $\boldsymbol{a} = ()$ and $\boldsymbol{\ell} = ()$. For the 1-dimensional edges of $\mathcal{R}(\widehat{Q})$ there are three orientations, namely parallel to any of the *x*-, *y*- or *z*-axes; they are described by $\boldsymbol{a} = (1)$, $\boldsymbol{a} = (2)$, and $\boldsymbol{a} = (3)$, respectively. Moreover, there are three orientations for 2-dimensional faces, namely parallel to any of the *xy*-, *xz*- or *yz*-planes; they correspond to the orientation pairs $\boldsymbol{a} = (1, 2)$, $\boldsymbol{a} = (1, 3)$, and $\boldsymbol{a} = (2, 3)$, respectively. Finally, there is a single orientation triple for the eight full-dimensional elements $\widehat{T}_{\boldsymbol{i}}$ given by $\boldsymbol{a} = (1, 2, 3)$. For the interpretation of the location tuple, let us consider, for instance, the four internal nodes of dimension 2, which are parallel to the *xy*-plane, i.e. with orientation $\boldsymbol{a} = (1, 2)$; they are highlighted in Figure 6 (left), and their possible locations are represented by the location pairs listed in Table 1. To give a further example, the left 1-dimensional internal node in *x*-direction of $\mathcal{R}(Q)$, highlighted in Figure 6 (right) is characterized by $\boldsymbol{a} = (1)$ and $\boldsymbol{\ell} = (0)$.

For any $r \in \underline{d}_0$, let us introduce the set

$$\mathcal{N}_r := \{\boldsymbol{n} = (\boldsymbol{a}, \boldsymbol{\ell}) : \boldsymbol{a} = (a_1, \ldots, a_r) \in D_r \text{ and } \boldsymbol{\ell} = (\ell_1, \ldots, \ell_r) \in \{0,1\}^r\}$$
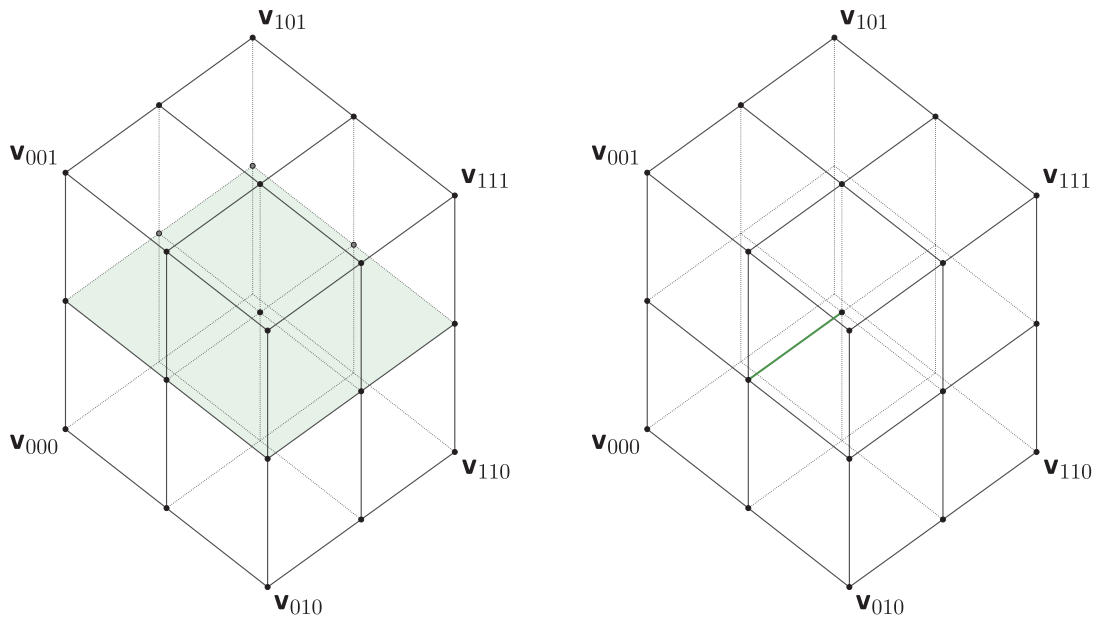
Figure 6: Illustration for Example 3: 3-dimensional example of a refinement $\mathcal{R}(Q)$.

Location Tuple ℓ	Location in Refinement $\mathcal{R}(Q)$
$\ell = (0,0)$	left half in x-, left half in y-direction
$\ell = (1,0)$	right half in x-, left half in y-direction
$\ell = (0,1)$	left half in x-, right half in y-direction
$\ell = (1,1)$	right half in x-, right half in y-direction

Table 1: Possible locations for interior 2-dimensional faces parallel to the xy-plane in the refinement of an element Q, all with orientation tuple $\boldsymbol{a} = (1,2)$.

of all r-dimensional internal nodes in $\mathcal{R}(Q)$; note that the cardinality of $\mathcal{N}_r$ is given by

$$|\mathcal{N}_r| = 2^r \binom{d}{r}.$$

Furthermore, we signify by $\mathcal{N} := \bigcup_{r \in \underline{d}_0} \mathcal{N}_r$ the collection of all internal nodes of any dimension $r = 0, \ldots, d$ in $\mathcal{R}(Q)$. In addition, for any $\boldsymbol{n} = (\boldsymbol{a}, \boldsymbol{\ell}) \in \mathcal{N}$, we define

$$I(\boldsymbol{n}) := \{\boldsymbol{i} = (i_1, \ldots, i_d) \in \{0,1\}^d : i_k = \ell_k \text{ for each } k \in A(\boldsymbol{a})\}$$

to be the set of all multi-indices $\boldsymbol{i} \in \{0,1\}^d$ corresponding to elements $T_{\boldsymbol{i}} \in \mathcal{R}(Q)$ that share the internal node $\boldsymbol{n}$.

Example 4 ($I(\boldsymbol{n})$ for 3-Dimensional Example). In the setting of Example 3, let us consider once more the 1-dimensional edge, which is given by the orientation and location tuples $\boldsymbol{a} = (1)$ and the $\boldsymbol{\ell} = (0)$. Then the indices of all 3-dimensional elements $T_{\boldsymbol{i}} \in \mathcal{R}(Q)$ sharing the internal node $\boldsymbol{n} = \{((1), (0))\} \in \mathcal{N}_1$, which are highlighted in Figure 7, are collected in the set $I(\boldsymbol{n}) = \{(\underline{0},0,0), (\underline{0},1,0), (\underline{0},1,1), (\underline{0},0,1)\}$.

hp-Enrichment Functions on Q. Let $\boldsymbol{n} = (\boldsymbol{a}, \boldsymbol{\ell}) \in \mathcal{N}_r$ be an internal node of $\mathcal{R}(Q)$, for some $r \in \underline{d}_0$, and consider an associated polynomial distribution

$$\boldsymbol{p} = (p_1, \ldots, p_r) \in \{\boldsymbol{p} \in \mathbb{N}_0^r : p_k \geq 2 \text{ for } k \in \underline{r}\};$$

here, the entry p_k corresponds to the coordinate direction $k \in A(\boldsymbol{a})$ of $\boldsymbol{n}$. Then, for each $\boldsymbol{i} \in I(\boldsymbol{n})$, we introduce a d-tuple $\boldsymbol{j}(\boldsymbol{i}, \boldsymbol{p}) = (j_1, \ldots, j_d) \in \mathbb{N}_0^d$ component-wise by

$$j_k := \begin{cases} p_k & \text{if } k \in A(\boldsymbol{a}), \\ 1 - i_k & \text{if } k \notin A(\boldsymbol{a}). \end{cases} \tag{3.13}$$
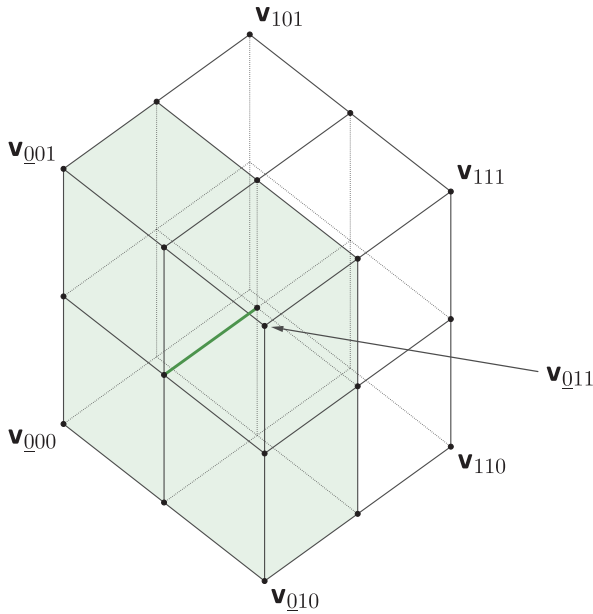
Figure 7: Illustration for Example 4: 1-dimensional edge in a 3D refinement.

Let us now define functions $\xi_{n,p} : \Omega \to \mathbb{R}$, which are associated with the r-dimensional internal node $\boldsymbol{n}$, by

$$\xi_{\boldsymbol{n},\boldsymbol{p}}(\mathbf{x}) := \begin{cases} \zeta_{j(i,\boldsymbol{p})}^{i}(\mathbf{x}) & \text{if } \mathbf{x} \in T_i \text{ for } \boldsymbol{i} \in I(\boldsymbol{n}), \\ 0 & \text{if } \mathbf{x} \in \Omega \setminus T(\boldsymbol{n}), \end{cases} \tag{3.14}$$

where

$$T(\boldsymbol{n}) := \bigcup_{\boldsymbol{i} \in I(\boldsymbol{n})} T_{\boldsymbol{i}} \tag{3.15}$$

represents the support of the functions $\xi_{\boldsymbol{n},\boldsymbol{p}}$. As *hp*-enrichment functions we choose finitely many such functions. More precisely, for any internal node $\boldsymbol{n} = (\boldsymbol{a}, \boldsymbol{\ell}) \in \mathcal{N}_r$, $r \in \underline{d_0}$, we select a finite set

$$\boldsymbol{P}(\boldsymbol{n}) \subseteq \{\boldsymbol{p} \in \mathbb{N}_0^r : p_k \geq 2 \text{ for } k \in \underline{r}\},$$

and let

$$\mathfrak{E}_{hp} := \bigcup_{\boldsymbol{n} \in \mathcal{N}} \mathfrak{E}_{hp,\boldsymbol{n}} \quad \text{with } \mathfrak{E}_{hp,\boldsymbol{n}} := \{\xi_{\boldsymbol{n},\boldsymbol{p}} : \boldsymbol{p} \in \boldsymbol{P}(\boldsymbol{n})\}. \tag{3.16}$$

For the single 0-dimensional internal node of $\mathcal{R}(Q)$, represented by $\boldsymbol{n} = ((), ())$, we note that $\boldsymbol{P}(\boldsymbol{n}) = \{()\}$; hence, the only *hp*-enrichment function $\xi_{\boldsymbol{n},\boldsymbol{p}}$ associated with this node is given by

$$\xi_{((),()),()}(\mathbf{x}) = \begin{cases} \zeta_{1-i}^{i}(\mathbf{x}) & \text{if } \mathbf{x} \in T_{\boldsymbol{i}} \text{ for } \boldsymbol{i} \in \{0,1\}^d, \\ 0 & \text{if } \mathbf{x} \in \Omega \setminus T(\boldsymbol{n}), \end{cases}$$

where $\boldsymbol{1} = (1, \ldots, 1) \in \mathbb{N}^d$.

Proposition 5. *Any $\xi_{\boldsymbol{n},\boldsymbol{p}} \in \mathfrak{E}_{hp}$ is continuous in Ω, and it holds* $\mathrm{supp}(\xi_{\boldsymbol{n},\boldsymbol{p}}) = T(\boldsymbol{n})$.

Proof. Let $\xi_{\boldsymbol{n},\boldsymbol{p}} \in \mathfrak{E}_{hp,\boldsymbol{n}}$ with $\mathfrak{E}_{hp,\boldsymbol{n}}$ from (3.16), be an arbitrary *hp*-enrichment function, for some $\boldsymbol{n} = (\boldsymbol{a}, \boldsymbol{\ell}) \in \mathcal{N}_r$, $r \in \underline{d_0}$, and $\boldsymbol{p} \in \boldsymbol{P}(\boldsymbol{n})$. Let us denote by $f_{\boldsymbol{n}} := \bigcap_{\boldsymbol{i} \in I(\boldsymbol{n})} T_{\boldsymbol{i}}$ the r-dimensional internal node of the refinement $\mathcal{R}(Q)$ characterized by $\boldsymbol{n}$. Noticing that

$$\xi_{\boldsymbol{n},\boldsymbol{p}\,|T_{\boldsymbol{i}}} = \zeta_{j(i,\boldsymbol{p})}^{i} = \widehat{\psi}_{j(i,\boldsymbol{p})} \circ F_{\boldsymbol{i}}^{-1} \quad \text{for all } \boldsymbol{i} \in I(\boldsymbol{n}),$$

and $\xi_{\boldsymbol{n},\boldsymbol{p}\,|T_{\boldsymbol{i}}} \equiv 0$ on elements $T_{\boldsymbol{i}}$ with $\boldsymbol{i} \in \{0,1\}^d \setminus I(\boldsymbol{n})$ as well as on $\Omega \setminus Q$ by (3.14), we immediately obtain the continuity of the function $\xi_{\boldsymbol{n},\boldsymbol{p}}$ on the interior of the subelements $T_{\boldsymbol{i}}$, for $\boldsymbol{i} \in I(\boldsymbol{n})$, and on $\Omega \setminus T(\boldsymbol{n})$. Furthermore,

we have $\text{supp}(\xi_{n,p}) = T(\boldsymbol{n})$, cf. (3.15). Thus, in order to prove the continuity of $\xi_{n,p}$ on Ω, we need to show that $\xi_{n,p} \equiv 0$ on the boundary $\partial T(\boldsymbol{n})$ of the support $T(\boldsymbol{n})$, and that the functions $\zeta^{i}_{j(i,p)}$, for any $\boldsymbol{i} \in I(\boldsymbol{n})$, coincide on the r-dimensional face $f_{\boldsymbol{n}}$.

Since $\boldsymbol{p} = (p_1, \ldots, p_r) \in \boldsymbol{P}(\boldsymbol{n})$, we have $p_k \geq 2$ for any $k \in \underline{r}$. Hence, it holds that $j_{k,i} \geq 2$ for $k \in A(\boldsymbol{a})$, and $j_{k,i} \in \{0, 1\}$ for $k \notin A(\boldsymbol{a})$, where the indices $(j_{1,i}, \ldots, j_{d,i}) = \boldsymbol{j}(i, \boldsymbol{p}) \in \mathbb{N}_0^d$ are defined in (3.13). To ensure that the functions

$$\zeta^{i}_{j(i,p)}(\mathbf{x}) = \prod_{k\in\underline{d}} \psi_{j_{k,i}}(\widehat{x}_k), \quad \widehat{\mathbf{x}} = (\widehat{x}_1, \ldots, \widehat{x}_d)^\top = F_i^{-1}(\mathbf{x}), \quad \boldsymbol{i} \in I(\boldsymbol{n}), \tag{3.17}$$

with the 1-dimensional functions ψ_k from (3.2), coincide for $\mathbf{x} = (x_1, \ldots, x_d)^\top \in f_{\boldsymbol{n}}$, we first show that $\psi_{j_{k,i}}(\widehat{x}_k) = 1$ for any $k \notin A(\boldsymbol{a})$. Indeed, for $\boldsymbol{i} = (i_1, \ldots, i_d) \in I(\boldsymbol{n})$ we have

$$f_{\boldsymbol{n}} = \{F_i(\widehat{\mathbf{x}}) : \widehat{\mathbf{x}} = (\widehat{x}_1, \ldots, \widehat{x}_d)^\top \in \widehat{Q} \text{ and } \widehat{x}_k = 1 - 2\, i_k \text{ for } k \notin A(\boldsymbol{a})\}, \tag{3.18}$$

and because of

$$j_{k,i} = \begin{cases} 1 & \text{if } i_k = 0, \\ 0 & \text{if } i_k = 1, \end{cases} \quad k \notin A(\boldsymbol{a}),$$

cf. (3.13), we obtain

$$\psi_{j_{k,i}}(\widehat{x}_k) = \begin{cases} \psi_1(1) = 1 & \text{if } i_k = 0, \\ \psi_0(-1) = 1 & \text{if } i_k = 1, \end{cases} \quad k \notin A(\boldsymbol{a}),$$

owing to (3.2) and (3.18). Next, for $k \in A(\boldsymbol{a})$ it holds that $\psi_{j_{k,i}}(\widehat{x}_k) = \psi_{p_k}(\widehat{x}_k)$, cf. (3.13). Thus,

$$\zeta^{i}_{j(i,p)}(\mathbf{x}) = \prod_{k\in A(\boldsymbol{a})} \psi_{p_k}(\widehat{x}_k), \quad \boldsymbol{i} \in I(\boldsymbol{n}),$$

which shows the continuity of the enrichment function $\xi_{n,p}$ in the interior of the support $T(\boldsymbol{n})$. Finally, we observe that $\psi_0(1) = \psi_1(-1) = 0$, and $\psi_j(\pm 1) = 0$, for any $j \in \mathbb{N}$ with $j \geq 2$. Then exploiting that

$$\partial T(\boldsymbol{n}) \setminus f_{\boldsymbol{n}} = \bigcup_{i\in I(\boldsymbol{n})} \{F_i(\widehat{\mathbf{x}}) : \widehat{\mathbf{x}} \in \widehat{Q} \text{ and } \exists k \in A(\boldsymbol{a}): \widehat{x}_k = \pm 1 \vee \exists k \notin A(\boldsymbol{a}): \widehat{x}_k = 2\, i_k - 1\},$$

and applying (3.17), we arrive at $\zeta^{i}_{j(i,p)}(\mathbf{x}) = 0$, for any $\mathbf{x} \in \partial T(\boldsymbol{n}) \setminus f_{\boldsymbol{n}}, \boldsymbol{i} \in I(\boldsymbol{n})$, This completes the argument. $\square$

Remark 3 ($\boldsymbol{P}(\boldsymbol{n})$ for Uniform Polynomial Degrees). Again, there are various possibilities to specify *hp*-enrichment functions. Note that in the definition of $\mathfrak{E}_{hp}$, the maximal polynomial degree can differ on each node $\boldsymbol{n} \in \mathbb{N}$, however, we could also consider the special case of a uniform polynomial degree distribution $p_{\text{unif}} \in \{2, \ldots, p_{\max}\}$ on each element $T_i \in \mathcal{R}(Q)$, $\boldsymbol{i} \in \{0, 1\}^d$, obtained by choosing $\boldsymbol{P}(\boldsymbol{n}) = \{2, \ldots, p_{\text{unif}}\}^r$ for any $\boldsymbol{n} \in \mathbb{N}_r$, $r \in \underline{d}$; see Example 5 below for further comments on this particular choice.

3.5 Representation Matrices

Owing to the previous Propositions 4 and 5 the sum over all elements K of a general decomposition $\mathcal{D}$ in (2.24) reduces to a sum over all elements T_i in the refinement $\mathcal{R}(Q)$ in our case. In particular, the representation matrices $\mathbf{C}_i$ and $\mathbf{D}_i$, see the notation in Section 3.4.1, need to be specified only for all $T_i \in \mathcal{R}(Q)$, $\boldsymbol{i} \in \{0, 1\}^d$. Recalling Proposition 3, we remark that these (local) quantities, in turn, are required to assemble the matrix $\mathbf{A}$ and the vectors $\mathbf{b}, \mathbf{c}$ occurring in the linear system (2.14).

In order to compute the matrices $\mathbf{C}_i$ and $\mathbf{D}_i$ on each element $T_i \in \mathcal{R}(Q)$, $\boldsymbol{i} \in \{0, 1\}^d$, we employ the constraint coefficients technique from Section 3.3. To this end, let $\mathfrak{E}_Q := \{\xi_1, \ldots, \xi_L\}$ be a set of *p*- or *hp*-enrichment functions on Q. On each $T_i \in \mathcal{R}(Q)$, $\boldsymbol{i} \in \{0, 1\}^d$, for $\boldsymbol{j} \in \{0, \ldots, p_{\max}\}^d$, we consider the functions ζ^{i}_j from (3.10), which we enumerate by the bijective map $\iota : \{0, \ldots, p_{\max}\}^d \to \underline{M}$, with $M := (p_{\max} + 1)^d$, given by

$$\iota(\boldsymbol{j}) := 1 + \sum_{k\in\underline{d}} (p_{\max} + 1)^{k-1} j_k, \quad \boldsymbol{j} = (j_1, \ldots, j_d), \tag{3.19}$$

i.e. we apply the renumbering $\zeta^i_{\iota(j)} := \zeta^i_j$, for $j \in \{0, \ldots, p_{\max}\}^d$. Note that the functions $\{\zeta^i_1, \ldots, \zeta^i_M\}$, extended by zero to Ω, are discontinuous and represent a basis of a (local) discontinuous Galerkin space on $\mathcal{R}(Q)$. Given a basis $\{\phi_1, \ldots, \phi_N\}$ of the finite element space $\mathbb{W}$, cf. (3.5), recall that the components of the matrix $\mathbf{C}_i = (c^i_{kl}) \in \mathbb{R}^{N \times M}$, for $i \in \{0, 1\}^d$, cf. (2.21), are determined by

$$\phi_{k|T_i} = \sum_{l \in \underline{M}} c^i_{kl} \zeta^i_l, \quad k \in \underline{N},$$

and the components of $\mathbf{D}_i = (d^i_{kl}) \in \mathbb{R}^{L \times M}$, $i \in \{0, 1\}^d$, cf. (2.22), are given by

$$\xi_{k|T_i} = \sum_{l \in \underline{M}} d^i_{kl} \zeta^i_l, \quad k \in \underline{L}. \tag{3.20}$$

3.5.1 Computation of the Matrices $\mathbf{C}_i$

On an element $Q \in \mathfrak{Q}$, for any $j \in \{0, \ldots, p_{\max}\}^d$, we introduce the functions $\psi^Q_1, \ldots, \psi^Q_M$ by

$$\psi^Q_{\iota(j)}(\mathbf{x}) := \widehat{\psi}_j \circ F_Q^{-1}(\mathbf{x}), \quad \mathbf{x} \in Q,$$

with the bijective map ι from (3.19). Furthermore, let $\mathbf{C}_Q = (c^Q_{kl}) \in \mathbb{R}^{N \times M}$ denote the representation matrix, for which it holds

$$\phi_{k|Q} = \sum_{l \in \underline{M}} c^Q_{kl} \psi^Q_l, \quad k \in \underline{N}.$$

Finally, for any $i \in \{0, 1\}^d$, let us define the matrix $\mathbf{B}_i = (b^i_{kl}) \in \mathbb{R}^{M \times M}$ component-wise by $b^i_{kl} := b^{T_i}_{\mathbf{kl}}$, where $\mathbf{k} = \iota(k)$, $\mathbf{l} = \iota(l)$, for any $k, l \in \{0, \ldots, p_{\max}\}^d$, and $b^{T_i}_{\mathbf{kl}}$ are the constrained coefficients from (3.9).

Proposition 6. *For any $i \in \{0, 1\}^d$ it holds that $\mathbf{C}_i = \mathbf{C}_Q \mathbf{B}_i$.*

Proof. For any $k \in \underline{N}$, we have

$$\phi_{k|Q} = \sum_{l \in \underline{M}} c^Q_{kl} \psi^Q_l = \sum_{l \in \underline{M}} c^Q_{kl} (\widehat{\psi}_l \circ F_Q^{-1}).$$

Therefore, we infer that

$$\phi_{k|T_i} \circ F_Q = \sum_{l \in \underline{M}} c^Q_{kl} \widehat{\psi}_{l|\widehat{T}_i}, \quad i \in \{0, 1\}^d.$$

Denoting by $\widehat{F}_i : \widehat{Q} \to \widehat{T}_i$ the bijective map from $\widehat{Q}$ to the sub-hexahedron $\widehat{T}_i$, and noting that $F_i = F_Q \circ \widehat{F}_i$, the restrictions to $\widehat{T}_i$ can be represented by

$$\widehat{\psi}_{l|\widehat{T}_i} = \sum_{r \in \underline{M}} b^{T_i}_{\mathbf{lr}} (\widehat{\psi}_r \circ \widehat{F}_i^{-1})$$

for any $i \in \{0, 1\}^d$, cf. (3.9), where $\mathbf{l} = \iota^{-1}(l)$ and $\mathbf{r} = \iota^{-1}(r)$. Thus,

$$\phi_{k|T_i} = \sum_{l \in \underline{M}} c^Q_{kl} \sum_{r \in \underline{M}} b^{T_i}_{\mathbf{lr}} (\widehat{\psi}_r \circ \widehat{F}_i^{-1} \circ F_Q^{-1}) = \sum_{l \in \underline{M}} c^Q_{kl} \sum_{r \in \underline{M}} b^{T_i}_{\mathbf{lr}} (\widehat{\psi}_r \circ F_i^{-1}).$$

Since $\zeta^i_r = \widehat{\psi}_r \circ F_i^{-1}$, we obtain

$$\phi_{k|T_i} = \sum_{l \in \underline{M}} c^Q_{kl} \sum_{r \in \underline{M}} b^{T_i}_{\mathbf{lr}} \zeta^i_r = \sum_{r \in \underline{M}} \left(\sum_{l \in \underline{M}} c^Q_{kl} b^{T_i}_{\mathbf{lr}} \right) \zeta^i_r,$$

from which we deduce

$$c^i_{kr} = \sum_{l \in \underline{M}} c^Q_{kl} b^{T_i}_{\mathbf{lr}}$$

for any $k \in \underline{N}$ and $r \in \underline{M}$. Thus, $\mathbf{C}_i = \mathbf{C}_Q \mathbf{B}_i$ for any $i \in \{0, 1\}^d$. $\qquad\square$

3.5.2 Computation of D_i for a *p*-Enrichment on *Q*

In this case the enrichment functions $\mathfrak{E}_p = \{\xi_1, \ldots, \xi_L\}$ represent a certain collection of the functions ξ_j in (3.11), with j from a multi-index set

$$J_d \subseteq \{j \in \mathbb{N}_0^d : j_k \geq 2 \text{ for } k \in \underline{d}\}, \quad |J_d| = L.$$

As in the previous section, we apply a renumbering $\xi_{\kappa(j)} := \xi_j$ in terms of a bijective map $\kappa : J_d \to \underline{L}$; for instance, if we choose $J_d := \{2, \ldots, p_{\max}\}^d$, then the resulting *p*-enrichment functions $\xi_1, \ldots, \xi_L$, with $L = (p_{\max} - 1)^d$, can be enumerated by

$$\kappa(j) := 1 + \sum_{k \in \underline{d}} (p_{\max} - 1)^{k-1}(j_k - 2), \quad j = (j_1, \ldots, j_d).$$

Proposition 7. *For any $i \in \{0, 1\}^d$, the matrix $\mathbf{D}_i \in \mathbb{R}^{L \times M}$ is given component-wise by $d_{kl}^i = b_{kl}^{T_i}$ for $k = \kappa(k)$ and $l = \iota(l)$.*

Proof. By the definition (3.11) of *p*-enrichment functions we have $\xi_{k|T_i} = \widehat{\psi}_{k|\widehat{T}_i} \circ F_Q$ for $i \in \{0, 1\}^d$ and any $k \in \underline{L}$. Hence, using the bijective map $\widehat{F}_i : \widehat{Q} \to \widehat{T}_i$, we obtain

$$\widehat{\psi}_{k|\widehat{T}_i} = \sum_{l \in \underline{M}} d_{kl}^i (\zeta_l^i \circ F_Q^{-1}) = \sum_{l \in \underline{M}} d_{kl}^i (\widehat{\psi}_l \circ F_i^{-1} \circ F_Q^{-1}) = \sum_{l \in \underline{M}} d_{kl}^i (\widehat{\psi}_l \circ \widehat{F}_i^{-1})$$

by the representation (3.20). Hence, by virtue of (3.9), we arrive at $d_{kl}^i = b_{kl}^{T_i}$ for $k = \kappa(k)$ and $l = \iota(l)$. $\square$

3.5.3 Computation of D_i for an *hp*-Refinement on *Q*

We consider *hp*-enrichment functions $\mathfrak{E}_{hp} = \{\xi_1, \ldots, \xi_L\}$ from a certain selection of functions $\xi_{n,p}$ in (3.14). Here, each $n = (a, \ell) \in \mathcal{N}$ corresponds to an internal node of the refinement $\mathcal{R}(Q)$, which is characterized by the tuples $a \in D_r$ and $\ell \in \{0, 1\}^r$, for some $r \in \underline{d}_0$, cf. Section 3.4.3. Moreover, $p \in P(n)$ represents the polynomial distribution for the directions $k \in A(a)$, where

$$P(n) \subseteq \{p \in \mathbb{N}_0^r : p_k \geq 2 \text{ for } k \in \underline{r}\} \quad \text{with } |P(n)| < \infty.$$

Given a polynomial degree distribution $(p_i)_{i \in \{0,1\}^d}$ on the subelements $T_i \in \mathcal{R}(Q)$, with $p_i \in \{1, \ldots, p_{\max}\}$, we can associate a *nodal polynomial degree* p_n to any node $n \in \mathcal{N}_r$, with $r \in \underline{d}$, by means of the minimum-rule

$$p_n := \min_{i \in T(n)} p_i$$

In that case, for any node $n \in \mathcal{N}_r$, we set $P(n) = \{2, \ldots, p_n\}^r$ (with the convention $P(((), ())) = \{()\}$ for $r = 0$), and we can enumerate the functions $\xi_{n,p}$ in terms of a bijective map $\nu : \{(n, p) : n \in \mathcal{N}, p \in P(n)\} \to \underline{L}$.

Proposition 8. *For any $i \in \{0, 1\}^d$ the matrix $\mathbf{D}_i \in \mathbb{R}^{L \times M}$ is given component-wise by*

$$d_{kl}^i = \begin{cases} 1 & \text{if } \xi_{n,p} = \zeta_j^i, \\ 0 & \text{otherwise,} \end{cases}$$

for $k = \nu(n, p)$ and $l = \iota(j)$.

Proof. By the definition (3.14) of *hp*-enrichment functions, we have

$$\xi_{n,p\,|T_i} = \zeta_{j(i,p)}^i$$

for any $i \in I(n)$, where $j(i, p) \in \mathbb{N}_0^d$ is defined in (3.13). Since the components of $\mathbf{D}_i$ are determined by

$$\xi_{k|T_i} = \sum_{l \in \underline{M}} d_{kl}^i \zeta_l^i, \quad k = \nu(n, p),$$

we immediately obtain $d_{kl}^i = 1$ if $l = \iota(j(i, p))$, and $d_{kl}^i = 0$ otherwise. In addition, for any $i \in \{0, 1\}^d \setminus I(n)$, we have $\xi_{n,p\,|T_i} \equiv 0$, wherefore the linear independence of the functions $\zeta_1^i, \ldots, \zeta_M^i$ yields $d_{kl}^i = 0$ for any $l \in \underline{M}$. $\square$

Example 5 (Enumeration v for Uniform Polynomial Degrees). Let us assign the same polynomial degree $p_{unif} \in \{2, \ldots, p_{max}\}$ to all elements $T_{\boldsymbol{i}}$ of the refinement $\mathcal{R}(Q)$, $\boldsymbol{i} \in \{0, 1\}^d$, i.e. $\boldsymbol{P}(\boldsymbol{n}) = \{2, \ldots, p_{unif}\}^r$ for any r-dimensional internal node $\boldsymbol{n} \in \mathcal{N}_r$, $r \in \underline{d}_0$. In this case, we observe that there are $(p_{unif} - 1)^r$ hp-enrichment functions on Q that can be associated with $\boldsymbol{n}$. Hence, the total number of hp-refinement functions, which are associated with r-dimensional internal nodes, is given by

$$L_r := \binom{d}{r} 2^r (p_{unif} - 1)^r, \quad r \in \underline{d}_0.$$

Now let us enumerate the r-dimensional internal nodes $\boldsymbol{n} = (\boldsymbol{a}, \boldsymbol{p}) \in \mathcal{N}_r$, $r \in \underline{d}_0$, by the bijective map $v_r : \mathcal{N}_r \to \{1, \ldots, \binom{d}{r} 2^r\}$ given by

$$v_r(\boldsymbol{n}) := 1 + 2^r \sum_{k \in \underline{r}} (a_k - k) + \sum_{k \in \underline{r}} 2^{k-2}(\ell_k + 1), \quad \boldsymbol{a} = (a_1, \ldots, a_r), \quad \boldsymbol{p} = (p_1, \ldots, p_r).$$

Then we can enumerate all hp-enrichment functions $\xi_{\boldsymbol{n},\boldsymbol{p}}$, for $\boldsymbol{n} \in \mathcal{N}$ and $\boldsymbol{p} \in \boldsymbol{P}(\boldsymbol{n})$, by the bijective map $v : \{(\boldsymbol{n}, \boldsymbol{p}) : \boldsymbol{n} \in \mathcal{N}, \boldsymbol{p} \in \boldsymbol{P}(\boldsymbol{n})\} \to \underline{L}$, with $L := L_0 + \cdots + L_d$, defined by

$$v(\boldsymbol{n}, \boldsymbol{p}) := 1 + \sum_{k \in \underline{r}} L_{k-1} + (v_r(\boldsymbol{n}) - 1)(p_{unif} - 1)^r + \sum_{k \in \underline{r}} (p_{unif} - 1)^{k-1}(p_k - 2),$$

where we let $L_{-1} := 0$.

4 *hp*-Adaptivity Based on Locally Predicted Error Reductions

In this section, we will exploit our abstract results in Section 2.1 for the purpose of devising a new adaptive procedure for hp-type finite element discretizations. Our basic idea to hp-refine a given hp-finite element space $\mathbb{W}_{given}$ consists of three essential steps:

Step 1. Firstly, our algorithm aims to *predict* the potential contribution to the (global) energy error reduction from each individual element Q in the given hp-space $\mathbb{W}_{given}$. To this end, for every element Q, with an associated local space $\mathbb{W}^{loc}$, cf. (2.4), we apply various *local p-enrichment* or *hp-replacement spaces* $\mathbb{Y}$, cf. (2.7), as follows:

(p) In the case of p-enrichments, we choose finitely many sets $\boldsymbol{J}_{d,i}$ in order to define suitable collections of p-enrichment functions $\mathfrak{E}_{p,i} := \{\xi_{\boldsymbol{j}}^Q : \boldsymbol{j} \in \boldsymbol{J}_{d,i}\}$, cf. (3.12), and let $\mathbb{Y}_{p,i} := \text{span}\{\bar{u}_{\mathbb{W}}\} + \text{span}\,\mathfrak{E}_{p,i}$ be the associated p-enrichment spaces. We then compute the respective predicted error reductions $\Delta e_{p,i}$ by means of the formula (2.11) (or equivalently (2.15)).

(hp) In the case of hp-refinements, for any node $\boldsymbol{n}$ of a refinement $\mathcal{R}(Q)$ of Q, we choose finitely many sets $\boldsymbol{P}_i(\boldsymbol{n})$ to define collections of (nodal) hp-enrichment functions $\mathfrak{E}_{hp,\boldsymbol{n}}^i := \{\xi_{\boldsymbol{n},\boldsymbol{p}}^Q : \boldsymbol{p} \in \boldsymbol{P}_i(\boldsymbol{n})\}$, cf. (3.16), and let

$$\mathbb{Y}_{hp,i} := \text{span}\{\bar{u}_{\mathbb{W}}\} + \text{span}\,\mathfrak{E}_{hp,i} \quad \text{with } \mathfrak{E}_{hp,i} := \bigcup_{\boldsymbol{n}} \mathfrak{E}_{hp,\boldsymbol{n}}^i$$
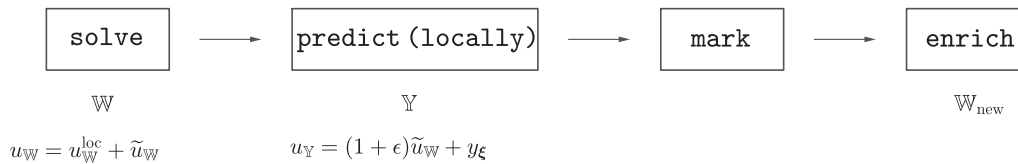
be the associated hp-replacement spaces. We then compute the corresponding predicted error reductions $\Delta e_{hp,i}$ by means of the formula (2.11) (or equivalently (2.15)).

From all these choices of p-enrichments (p) and hp-refinements (hp), we select the one that features the maximal error reduction (signified by Δe_{max}^Q); this will be referred to as the *optimal (local) hp-refinement* on Q. Hereby, the number of different p-enrichments and hp-refinements, respectively, that are performed for each element Q in the prediction step, is technically at the disposal of the user; cf. the ensuing Example 6 for a specific setup in 1d and 2d. Preferably, following the approach proposed in [20], we select the local p-enrichments and hp-refinements in a *competitive* way, i.e. with a comparable number of local degrees of freedom; we will discuss this idea in more detail in the context of the 1-dimensional numerical examples in Section 5 below.

Step 2. Subsequently, we *mark* all elements in the global hp-finite element space $\mathbb{W}_{given}$, from which the most substantial error reductions, as identified in Step 1. for each element, can be expected. This step can be accomplished, for instance, with the aid of a suitable marking strategy, such as Dörfler's criterion [12].

Step 3. Finally, a new *enriched hp*-finite element space $\mathbb{W}_{\text{new}}$ is constructed based on choosing the optimal (local) *hp*-refinement space, cf. Step 1., for each of the marked elements.

Schematically, the proposed *hp*-adaptive procedure has the following structure:

$$\boxed{\text{solve}} \longrightarrow \boxed{\text{predict (locally)}} \longrightarrow \boxed{\text{mark}} \longrightarrow \boxed{\text{enrich}}$$

$$\mathbb{W} \qquad\qquad\qquad \mathbb{Y} \qquad\qquad\qquad\qquad\qquad \mathbb{W}_{\text{new}}$$

$$u_{\mathbb{W}} = u_{\mathbb{W}}^{\text{loc}} + \widetilde{u}_{\mathbb{W}} \qquad\quad u_{\mathbb{Y}} = (1 + \epsilon)\widetilde{u}_{\mathbb{W}} + y_{\xi}$$

Remark 4 (Computational Aspects of the Prediction Step). We emphasize that the prediction step (Step 1) is trivially parallelizable, and hence, inexpensive from a computational point of view compared to the global solve (Step 3). In addition, as all enrichment functions for the different *p*-enrichments and *hp*-refinements on Q can be expressed by the functions $\zeta_1^i, \dots, \zeta_M^i$, cf. (3.10), the matrices $\mathbf{A}_i$, $\mathbf{C}_i$ as well as the vector $\mathbf{b}_i$ from Section 3.5, which are employed in the computation of the corresponding predicted local error reductions, need to be computed *once only* on the subelements T_i of the local refinement $\mathcal{R}(Q)$ for *all* different *p*-enrichments and *hp*-refinements applied on the subelements T_i. Evidently, the computational cost for the prediction step depends strongly on the specific selection of possible *p*-enrichments and *hp*-refinements (as well as on the efficiency of their implementation).

Example 6. We discuss some possible realizations of the choice of *p*-enrichments and *hp*-refinements to be compared on the individual elements (Step 1. above), which will be employed in our numerical examples of Section 5:
(i) In the 1d examples only one *p*-enrichment is considered on each element Q; specifically, we increase the associated local polynomial degree p_Q to $p_Q + 1$, which is represented by the set $J_1^Q := \{2, \dots, p_Q + 1\}$. This *p*-enrichment is then compared to p_Q-many *hp*-refinements which allocate suitable local polynomial degrees p_0, p_1 to subelements T_0, T_1 in a competitive way, so that $p_0 + p_1 = p_Q + 1$.
(ii) In the 2d example in Section 5, on each element Q, we consider an associated *p*-enrichment by which the local polynomial degree p_Q is increased to $p_Q + 1$. The resulting prediction is compared to two different *hp*-refinements: In a first version, the same polynomial degree p_Q is retained on all of the four subelements T_i with $i \in \{0,1\}^2$; in the second version, the polynomial degree is reduced to $p_Q - 1$ on the subelements T_i with $i \in \{0,1\}^2$. In this approach, in contrast to the 1d situation before, we note that the number of *hp*-refinements on an element Q does not increase with the local polynomial degree p_Q.

In Algorithm 1 we outline the technical details for the individual steps in the *hp*-adaptive refinement approach sketched above. We denote by $\mathbb{W}(\mathcal{Q}, \boldsymbol{p}_{\mathcal{Q}})$ the *hp*-finite element space associated with the mesh $\mathcal{Q}$, and by $\boldsymbol{p}_{\mathcal{Q}}$ a corresponding polynomial degree distribution, cf. (3.5). Note that the *solving step*, the *prediction step* (exploiting *p*-enrichments and *hp*-refinements), the *marking step* and the *enrichment step* are implemented in the lines 2, 3–15, 16 and 17 in Algorithm 1, respectively.

Remark 5 (Possible Specifications of *hp*-Adaptive Procedure). (i) Evidently, Algorithm 1 can be turned into a pure *h*-adaptive procedure (by only considering *hp*-refinements that inherit the *fixed low-order* polynomial degree to any subelements), or into a pure *p*-adaptive procedure (whereby no element refinements are taken into account).
(ii) In the context of finite element discretizations for which the global polynomial degree is restricted to a small positive number $p_{\max} \geq 2$ (e.g. for quadratic elements), it is possible to first apply an *h*-adaptive mesh refinement procedure, thereby yielding a locally refined mesh $\mathcal{Q}$, and then to employ *p*-enrichments based on Algorithm 1 in order to determine an effective polynomial distribution $\boldsymbol{p}_{\mathcal{Q}}$ with $\max \boldsymbol{p}_{\mathcal{Q}} \leq p_{\max}$.

Remark 6 (Guaranteed Error Decrease and Control). By Remark 1 the approximation error from one step to another in Algorithm 1 will *not increase* if $\Delta e_{\max}^Q > 0$ for marked elements; this condition can be accommodated if, in addition to employing *hp*-replacements, at least one enrichment is applied on those elements. Furthermore, we note that, at the end of each adaptive step, the use of an appropriate *a posteriori* estimator could

---

Algorithm 1. *hp*-Adaptive Procedure.

1: Choose an initial mesh $\mathcal{Q}_0$ on the computational domain Ω and a starting polynomial degree distribution $\boldsymbol{p}_{\mathcal{Q}_0}$. Set $n := 0$.

2: Solve the weak formulation (2.1) for $u_W \in W(\mathcal{Q}_n, \boldsymbol{p}_{\mathcal{Q}_n})$.   ▷ **solving step**

3: **for** each element $Q \in \mathcal{Q}_n$ **do**   ▷ **prediction step**

4:     Construct the (locally supported) subspace $W_Q^{loc} \subset W(\mathcal{Q}_n, \boldsymbol{p}_{\mathcal{Q}_n})$, and decompose the current solution $u_W = \bar{u}_W + u_W^{loc}$ according to (2.6), where $u_W^{loc} := \Pi_Q^{loc} u_W \in W_Q^{loc}$ is defined via the linear projection operator $\Pi_Q^{loc} : W(\mathcal{Q}_n, \boldsymbol{p}_{\mathcal{Q}_n}) \to W_Q^{loc}$ from (2.5).

5:     **for** finitely many different *p*-enrichments on Q **do**   ▷ *p*-enrichments

6:         Compute the corresponding predicted error reductions $\Delta e_{p,i}$ as outlined in Step 1 (p) above.

7:     **end for**

8:     Set $\Delta e_{p,\max}^Q := \max\{\Delta e_{p,i}\}$ to be the maximal locally predicted error reduction for the *p*-enrichments on the element Q.

9:     Construct the local refinement $\mathcal{R}(Q)$.   ▷ *hp*-refinement

10:     **for** finitely many different *hp*-refinements on Q **do**

11:         Compute the corresponding predicted error reductions $\Delta e_{hp,i}$ as outlined in Step 1 (hp) above, and store the polynomial degree distribution $(\boldsymbol{p}_{T,i})_{T \in \mathcal{R}(Q)}$ for the subelements $T \in \mathcal{R}(Q)$.

12:     **end for**

13:     Set $\Delta e_{hp,\max}^Q := \max\{\Delta e_{hp,i}\}$ to be the maximal locally predicted error reduction for the *hp*-refinements on Q.

14:     Determine $\Delta e_{\max}^Q := \max\{\Delta e_{p,\max}^Q, \Delta e_{hp,\max}^Q\}$ to be the maximal locally predicted error reduction.

15: **end for**

16: Mark a subset $\mathcal{E}_n \subseteq \mathcal{Q}_n$ of elements in the mesh $\mathcal{Q}_n$ to be flagged for enrichment.   ▷ **marking step**

17: For each element $Q \in \mathcal{E}_n$ perform a *p*-enrichment or an *hp*-refinement according to which enrichment leads to the maximal error reduction in 14. This results in a refined mesh $\mathcal{Q}_{n+1}$ with a corresponding polynomial degree distribution $\boldsymbol{p}_{\mathcal{Q}_{n+1}}$.   ▷ **enrichment step**

18: Update $n \leftarrow n + 1$, and start over in line 2.

19: After sufficiently many iterations output the final solution $u_W \in W(\mathcal{Q}_n, \boldsymbol{p}_{\mathcal{Q}_n})$.

---

be taken into account, see, e.g., [8], if some explicit control on the quality of the numerical approximation is desired. Notice that the estimator need not be localizable in terms of individual element contributions as it is not used for the purpose of driving the adaptive process.

5 Numerical Examples

In this section we illustrate the performance of the *hp*-adaptive procedure outlined in Algorithm 1 with some numerical experiments in 1D and 2D.

5.1 Numerical Examples in 1D

In the following 1-dimensional examples on the domain $\Omega := (0, 1)$ we use a basis for the *hp*-finite element spaces W that consists of the usual hat-functions, and, on each element $Q \in \mathcal{Q}$ with $p_Q \geq 2$, of the (elementwise transformed) integrated Legendre polynomials given by $\psi_j^Q := \psi_j \circ F_Q^{-1}$, for $2 \leq j \leq p_Q$, cf. (3.2). In accordance with (3.11) we consider the (extended) functions

$$\xi_j^Q(x) := \begin{cases} \psi_j^Q(x) & \text{if } x \in Q, \\ 0 & \text{if } x \in \Omega \setminus Q. \end{cases}$$

For each element $Q \in \mathcal{Q}$ of the current mesh, the locally supported subspace $\mathbb{W}_Q^{\text{loc}} \subset \mathbb{W}$ is chosen as

$$\mathbb{W}_Q^{\text{loc}} := \{\xi_j^Q : 2 \leq j \leq p_Q\}.$$

In particular, the corresponding linear projection operator $\Pi_Q^{\text{loc}} : \mathbb{W} \to \mathbb{W}_Q^{\text{loc}}$ only retains all higher-order modes on Q. Specifically, we choose

$$\mathfrak{E}_p^Q := \{\xi_j^Q : j \in J_1^Q\} \quad \text{with } J_1^Q := \{2, \ldots, p_Q + 1\}$$

as *p-enrichment* functions, i.e. we increase the polynomial degree p_Q to $p_Q + 1$, thereby using a number of p_Q locally supported degrees of freedom on Q. Moreover, for an *hp-replacement* on Q, we divide the element Q into two equally sized subelements T_0, T_1, for which we allocate some local polynomial degrees p_0, p_1, respectively, that give rise to $p_0 + p_1 + 1$ degrees of freedom within Q:
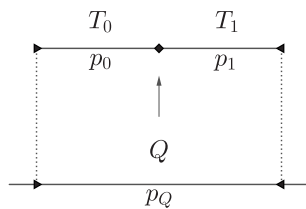
In order to compare *competitive p- and hp-refinements*, we impose the constraint

$$p_0 + p_1 = p_Q + 1. \tag{5.1}$$

The only internal node for the refinement $\mathcal{R}(Q) = \{T_0, T_1\}$ of the element Q is the 0-dimensional midpoint, represented by $\boldsymbol{n}_1 := ((), ())$. The subelements T_1, T_2 are characterized by the tuples $\boldsymbol{n}_2 := ((1), (0))$ and $\boldsymbol{n}_3 := ((1), (1))$. Therefore, in terms of the notation from Section 3.4.3, we have $\mathcal{N} = \{\boldsymbol{n}_1, \boldsymbol{n}_2, \boldsymbol{n}_3\}$, and we choose

$$\boldsymbol{P}^{\mathcal{R}(Q)}(\boldsymbol{n}_1) = \{()\}, \quad \boldsymbol{P}^{\mathcal{R}(Q)}(\boldsymbol{n}_2) = \{2, \ldots, p_1\}, \quad \boldsymbol{P}^{\mathcal{R}(Q)}(\boldsymbol{n}_3) = \{2, \ldots, p_2\}$$

for all possible pairs (p_1, p_2) of local polynomial degree combinations that satisfy (5.1).

5.1.1 Singularly Perturbed Problem with Boundary Layers

As a first example, for (a possibly small) parameter $\varepsilon > 0$, we consider the 1-dimensional singularly perturbed differential equation $-\varepsilon\, u'' + u = 1$ in the domain $\Omega = (0, 1)$ with the homogeneous Dirichlet boundary condition $u(0) = u(1) = 0$. The analytic solution is given by

$$u(x) = \frac{e^{-c} - 1}{e^c - e^{-c}}\, e^{cx} + \frac{1 - e^c}{e^c - e^{-c}}\, e^{-cx} + 1$$

with $c = \varepsilon^{-\frac{1}{2}}$; for very small $0 < \varepsilon \ll 1$, we notice that u exhibits thin boundary layers in the vicinity of the boundary points $x = 0$ and $x = 1$, and takes values of approximately 1 in the interior of the domain Ω. For our numerical experiments, we initiate the *hp*-adaptive procedure of Algorithm 1 with a coarse mesh $\mathcal{Q}_0$ consisting of four elements, and an associated uniform polynomial degree of 1 on each of them. For the marking process we choose the Dörfler marking parameter to be $\frac{1}{2}$.

Figure 8 shows the resulting *hp*-mesh after 28 *hp*-adaptive steps for $\varepsilon = 10^{-5}$; we clearly see that the boundary layers have been resolved on a few small elements (of high polynomial degree), whilst no refinement is employed in the interior of the domain Ω, where the solution is nearly constant.

Moreover, in Figure 9, for the underlying energy norm given by

$$\|v\|_{\mathbb{X}}^2 = \varepsilon \|v'\|_{L^2(\Omega)}^2 + \|v\|_{L^2(\Omega)}^2, \quad v \in \mathbb{X},$$
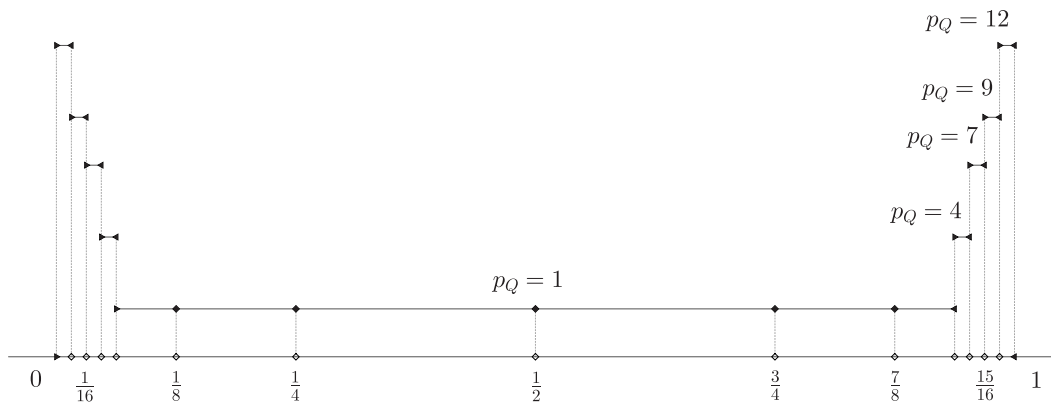
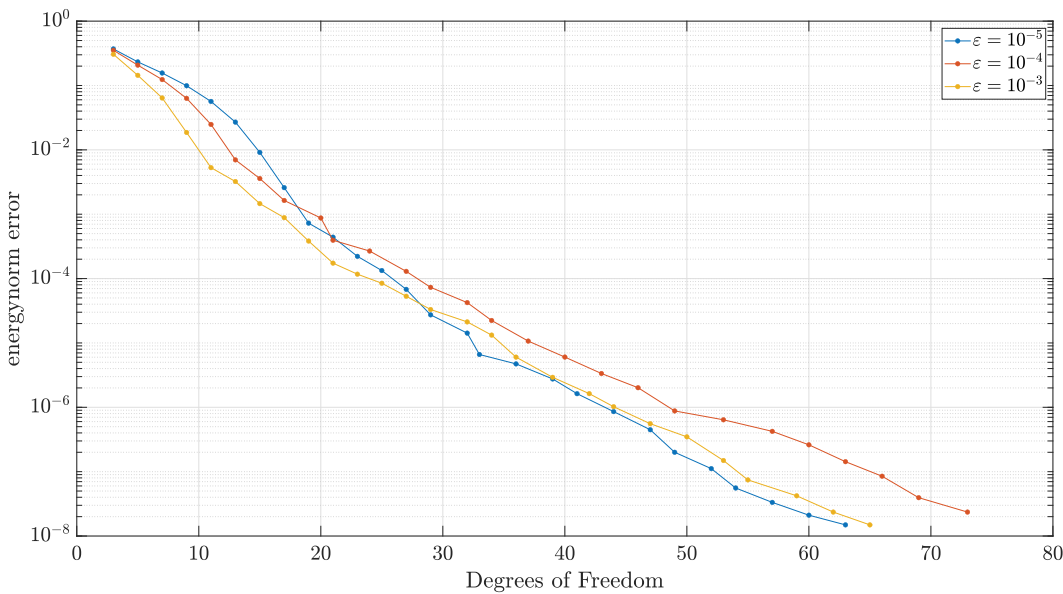**Figure 8:** *hp*-mesh $\mathcal{Q}_{28}$ after 28 adaptive enrichment steps with $\varepsilon = 10^{-5}$.

Figure 9: Singularly perturbed problem in 1D: Performance of the *hp*-adaptive procedure with respect to the energy norm error $\|u - u_W\|_{\mathbb{X}}$ for $\varepsilon = 10^{-j}$ with $j \in \{3, 4, 5\}$.

the error $\|u - u_W\|_{\mathbb{X}}$ is plotted with respect to the number of degrees of freedom in a semilogarithmic scaling for several values of the singular perturbation parameter $\varepsilon \in \{10^{-j} : j = 3, 4, 5\}$. We observe that the *hp*-adaptive procedure is able to achieve an exponential rate of convergence that is fairly robust with respect to ε. In this regard, our results are comparable to alternative *hp*-adaptive strategies proposed in the literature, see, e.g., [32, Expl. 2].

5.1.2 1D-Model Problem with a Boundary Singularity

As a second example, we consider the 1-dimensional boundary value problem $-u'' = f$ in $\Omega = (0, 1)$, with the homogeneous Dirichlet boundary conditions $u(0) = u(1) = 0$; the right-hand side function f is chosen in such a way that the analytic solution is given by $u(x) = x^{\frac{3}{4}} - x$, which features a singularity at $x = 0$ with $u'(x) \to \infty$ as $x \to 0^+$. As in the previous example, we start the *hp*-adaptive procedure of Algorithm 1 with an initial mesh $\mathcal{Q}_0$ consisting of four elements with a uniform polynomial degree distribution of 1, and let the Dörfler marking parameter be $\frac{1}{2}$. The ensuing plot in Figure 10 shows the resulting *hp*-mesh $\mathcal{Q}_{49}$, which contains 51 elements, after 49 *hp*-adaptive enrichment steps.
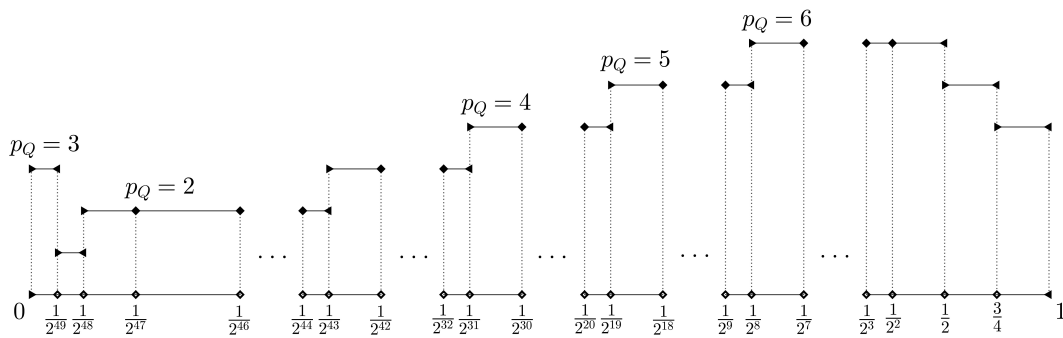
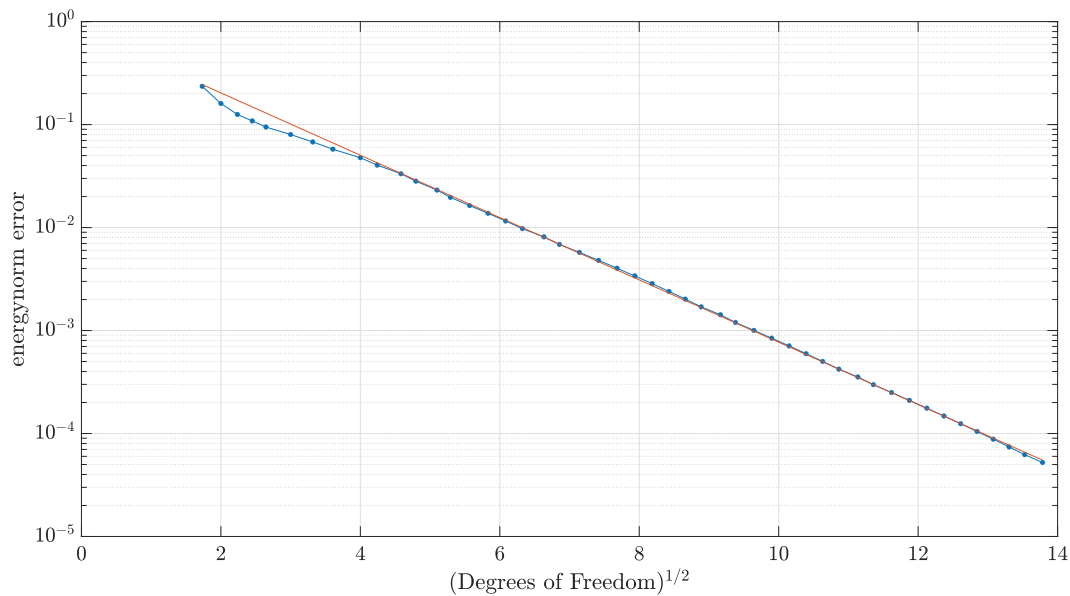**Figure 10:** *hp*-mesh $\mathcal{Q}_{49}$ after 49 *hp*-adaptive enrichment steps.

Figure 11: Algebraic boundary singularity in 1D: Performance of the *hp*-adaptive procedure with respect to the energy norm error $\|u - u_{\mathrm{W}}\|_{\mathbb{X}}$ (measured against the square root of the number of degrees of freedom). In addition to the energy norm error (blue) a regression line (orange) is displayed in order to indicate the exponential decrease.

The mesh is geometrically refined towards the singularity at $x = 0$, and the polynomial degree is increased at an approximately linear rate in dependence of the distance from the origin; this is in line with a priori results on exponentially convergent *hp*-FEM for local algebraic singularities; see, e.g., [27]. Indeed, this is confirmed in Figure 11, where the error in the energy norm, i.e. $\|u - u_{\mathrm{W}}\|_{\mathbb{X}} = \|u' - u'_{\mathrm{W}}\|_{L^2(\mathcal{Q})}$, shows an exponential decay (with respect to the square root of the number of degrees of freedoms) in a semilogarithmic scaling.

5.2 Numerical Example in 2D

In the following 2-dimensional example over the unit square $\Omega := (0, 1)^2$, we use a basis for the *hp*-finite element spaces W that—in addition to the usual hat functions – contains the functions

$$\psi_{\boldsymbol{j}}^{Q} := \widehat{\psi}_{\boldsymbol{j}} \circ F_Q^{-1}, \quad \boldsymbol{j} \in \{2, \ldots, p_Q\}^2,$$

extended by 0 to Ω, cf. (3.3), for elements $Q \in \mathcal{Q}$ with a local polynomial degree $p_Q \geq 2$. In accordance with (3.11) we denote the corresponding extended functions by $\xi_{\boldsymbol{j}}^{Q}$. Recall that we decompose the *hp*-finite element solution $u_{\mathrm{W}} \in \mathrm{W}$ of (2.1) into a local part $u_{\mathrm{W}}^{\mathrm{loc}}$ and a remaining part $\bar{u}_{\mathrm{W}}$, cf. (2.6). If we wish the remaining part $\bar{u}_{\mathrm{W}}$ to be completely independent of the enrichments on Q (in the sense that $\bar{u}_{\mathrm{W}}$ remains unchanged under modifications

in $\mathbb{W}_Q^{\mathrm{loc}}$), we note that the locally supported subspace $\mathbb{W}_Q^{\mathrm{loc}}$ needs to contain all basis functions of $\mathbb{W}$ with support on Q; in particular, all interior bubble functions on Q. For simplicity (see Remark 7), we set

$$\mathbb{W}_Q^{\mathrm{loc}} := \{\xi_j^Q : j \in \{2, \ldots, p_Q\}^2\},$$

and, on each element $Q \in \mathcal{Q}$ (with an associated local polynomial degree p_Q), we choose the p-enrichment functions to be

$$\mathfrak{E}_p^Q := \{\xi_j^Q : j \in J_2^Q\} \quad \text{with } J_2^Q := \{2, \ldots, p_Q + 1\}^2, \tag{5.2}$$

i.e. we increase the local polynomial degree p_Q to $p_Q + 1$ in both coordinate directions, thereby resulting in p_Q^2 many locally supported enrichment functions on Q. Moreover, an hp-refinement on an element Q is based on dividing Q into the four subelements T_i with $i \in \{0, 1\}^2$: In the present experiments we limit ourselves to enrichments that are restricted to single element Q as pointed out in Section 3.4, and do not involve enrichments on a patch around Q; see Remark 7 below for more details on this matter. For simplicity, we compare the p-enrichment (5.2) in two different versions of a single hp-refinement on Q: Firstly, with an hp-refinement that features the same polynomial degree p_Q on all subelements $T_i \in \mathcal{R}(Q)$, and secondly, with an hp-refinement that allocates the reduced polynomial degree $p_Q - 1$ to all subelements $T_i \in \mathcal{R}(Q)$ (if $p_Q > 2$). The internal nodes of the refinement $\mathcal{R}(Q) = \{T_i : i \in \{0, 1\}^2\}$ consist of the midpoint $n_1 := ((), ())$, of the edges

$$n_2 := ((1), (0)), \quad n_3 := ((1), (1)), \quad n_4 := ((2), (0)), \quad n_5 := ((2), (1)),$$

and of the subelements T_i, represented by

$$n_6 := ((1, 2), (0, 0)), \quad n_7 := ((1, 2), (1, 0)), \quad n_8 := ((1, 2), (0, 1)), \quad n_9 := ((1, 2), (1, 1)).$$

When featuring the polynomial degree p_Q on all subelements, we obtain

$$\boldsymbol{P}^{\mathcal{R}(Q)}(\boldsymbol{n}_j) = \begin{cases} \{()\} & \text{for } j = 1, \\ \{(p_1) : p_1 \in \{2, \ldots, p_Q\}\} & \text{for } 2 \leq j \leq 5, \\ \{(p_1, p_2) : p_1, p_2 \in \{2, \ldots, p_Q\}\} & \text{for } 6 \leq j \leq 9, \end{cases}$$

thereby leading to $1 + 4(p_Q - 1)p_Q$ hp-enrichment functions on Q.

Remark 7 (Competitive Refinements for Dimensions $d \geq 2$). In contrast to the 1-dimensional case, if the remaining part $\tilde{u}_{\mathbb{W}}$ of the solution $u_{\mathbb{W}}$ should be independent (in the sense that $\tilde{u}_{\mathbb{W}}$ remains unchanged upon modifications in $\mathbb{W}_Q^{\mathrm{loc}}$) of an enrichment or a refinement associated with an element Q, it is mandatory to include element interface basis functions in the definition of the locally supported subspace $\mathbb{W}_Q^{\mathrm{loc}}$. Indeed, only if degrees of freedom, which are shared by neighboring elements, are taken into account as well, it is possible to compare different p-enrichments and hp-refinements in a truly competitive way. This is due to the fact that, in general, p-enrichments and hp-refinements on a single element Q may influence the new solution on a local patch which also involves neighboring elements of Q; we observe that this effect extends even beyond the direct neighbors of Q if hanging nodes, as can be seen in Figure 12, are present; we refer to, e.g., [10], for a straightforward treatment of hanging nodes. In turn, we note that the enforcement of continuity of enrichment functions along
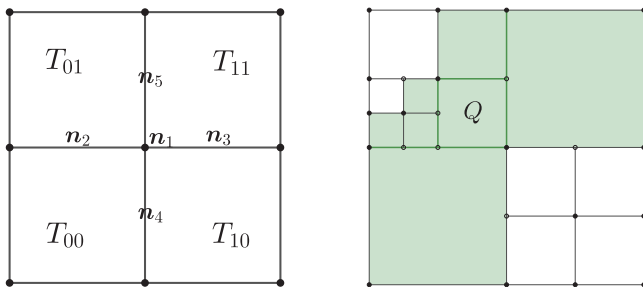

Figure 12: Extended patch around the element Q (right).

element edges and faces of a patch around Q (e.g. by means of the minimum rule for the adjacent polynomial degrees) causes an additional challenge. For simplicity, for the purpose of the present paper, we will *restrict ourselves to the enrichment of internal degrees of freedom only on each element*; competitive enrichments on patches for higher dimension are investigated in forthcoming work. Following the idea of using patches in the context of higher-order FEM, see, e.g., [25], we also include the interface basis functions between neighboring elements in the enrichment step; we remark that this approach, in turn, will generally yield an upper bound on the actual error reduction rather than an exact representation.

Remark 8 (Anisotropic Refinements for Dimensions $d \geq 2$). We note that the abstract framework presented in Section 2 and the constraint coefficient technique from Section 3.3 both apply to anisotropic *hp*-refinements as well; in particular, elements may be refined in a single direction only, with respect to either h or p. Indeed, we have deliberately formulated our *hp*-adaptive approach in a quite general way in order to allow for maximal flexibility and, thereby, a wide variety of practical realizations. The application of anisotropic *hp*-refinements within the adaptive framework will be studied in a forthcoming work.

5.2.1 2D-Poisson Problem with Corner Singularities

Let us consider the 2-dimensional Poisson-problem $-\Delta u = 1$ on $\Omega = (0,1)^2$, subject to the homogeneous Dirichlet boundary condition $u = 0$ on $\Gamma := \partial\Omega$. While an explicit expression for the analytic solution u is unavailable, an eigenfunction expansion yields that

$$\|u\|_{\mathbb{X}}^2 = \|\nabla u\|_{L^2(\Omega)}^2 = \left(\frac{2}{\pi}\right)^6 \sum_{k,l \geq 1 \text{ odd}} \frac{1}{k^2 l^2 (k^2 + l^2)} \approx 0.035144253738788451\ldots.$$

We start the *hp*-adaptive procedure of Algorithm 1 with an initial mesh $\mathbb{Q}_0$ consisting of 16 elements with a uniform polynomial degree distribution of 1 on all elements $Q \in \mathbb{Q}_0$. Moreover, we let the Dörfler parameter to be $\frac{1}{5}$.

Following the a priori error analysis on the exponential convergence of *hp*-FEM for the 2-dimensional Poisson equation with corner singularities in polygons, see, e.g., [27], we depict the energy error $\|u - u_{\mathbb{X}}\|_{\mathbb{X}}$ with respect to the 3rd root of the number of degrees of freedom in Figure 14. The blue line shows the error decay when the polynomial degree p_Q on an *hp*-refined element Q is inherited to the subelements T_i, for $i \in \{0,1\}^2$; we see that this strategy may lead to an unnecessarily high number of local degrees of freedom. The resulting *hp*-mesh of this strategy after 52 adaptive enrichment steps is depicted in Figure 13 (containing 1412 elements). In an alternative scenario, illustrated by the yellow line in Figure 14, the slope of the exponential convergence is considerably improved when allocating a reduced polynomial degree of $p_Q - 1$ to all subelements T_i, with $i \in \{0,1\}^2$, if $p_Q > 2$. The resulting *hp*-mesh after 67 adaptive enrichment steps is shown in Figure 13 (containing 256 elements). In addition, for comparison purposes with a possible benchmark situation, which does not
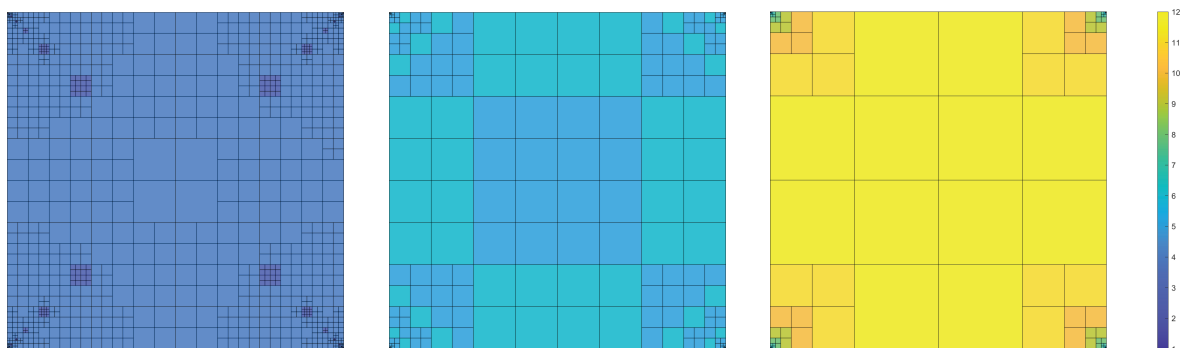

Figure 13: *hp*-mesh $\mathbb{Q}_{52}$ after 52 adaptive enrichment steps applying the first strategy (left), *hp*-mesh $\mathbb{Q}_{67}$ after 67 adaptive enrichment steps applying the second strategy (middle) and geometrically refined mesh towards the corners, with twelve layers and linearly increasing polynomial degrees leading to 192 elements (right).
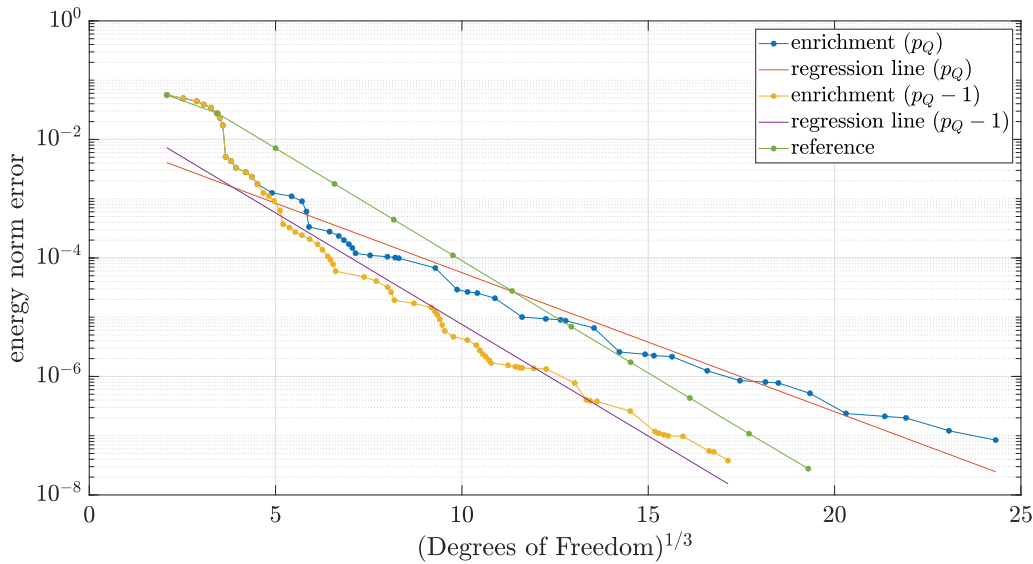
Figure 14: Poisson problem with corner singularities on unit square: Performance of the *hp*-adaptive procedure with respect to the energy norm error $\|u - u_{\mathbb{W}}\|_{\mathbb{X}}$ (measured against the 3rd root of the number of degrees of freedom). In addition to the energy norm errors (blue, yellow) regression lines (orange, purple) are displayed in order to indicate the exponential decrease. Moreover, the energy norm error (green) for a comparable a priori *hp*-refinement strategy has been inserted also.

depend on a specific *hp*-adaptive strategy, we have inserted an additional (green) line in Figure 14 that represents the error decay of a classical *a priori hp*-refinement strategy. More precisely, following the well-known theory of exponential convergence of *hp*-FEM for linear elliptic problems in 2d polygons (see, e.g., [27]), meshes that are geometrically refined towards the corners of Ω, see Figure 13, and that feature a linear polynomial degree distribution away from the corners, are applied. Our results confirm that our proposed *hp*-adaptive algorithm is able to properly resolve the four corner singularities on geometrically refined meshes, and that approximately exponential rates of convergence can be achieved.

A Appendix

A.1 Proof of (2.12)

From Proposition 1 we recall the identity $\Delta e_{\mathbb{W},\mathbb{Y}}^2 = \|u_{\mathbb{Y}} - u_{\mathbb{W}}\|_{\mathbb{X}}^2 - 2\,\rho_{\mathbb{Y}}(u_{\mathbb{W}}^{\mathrm{loc}})$. Furthermore, owing to the decomposition (2.6) of the solution $u_{\mathbb{W}}$, we have

$$\|u_{\mathbb{Y}} - u_{\mathbb{W}}\|_{\mathbb{X}}^2 = \|(u_{\mathbb{Y}} - \tilde{u}_{\mathbb{W}}) - u_{\mathbb{W}}^{\mathrm{loc}})\|_{\mathbb{X}}^2 = \|u_{\mathbb{Y}} - \tilde{u}_{\mathbb{W}}\|_{\mathbb{X}}^2 + \|u_{\mathbb{W}}^{\mathrm{loc}}\|_{\mathbb{X}}^2 - 2\,a(u_{\mathbb{Y}} - \tilde{u}_{\mathbb{W}}, u_{\mathbb{W}}^{\mathrm{loc}}).$$

Moreover, since $u_{\mathbb{W}}^{\mathrm{loc}} \in \mathbb{W}$, the weak formulation (2.1) yields

$$\rho_{\mathbb{Y}}(u_{\mathbb{W}}^{\mathrm{loc}}) = b(u_{\mathbb{W}}^{\mathrm{loc}}) - a(u_{\mathbb{Y}}, u_{\mathbb{W}}^{\mathrm{loc}}) = a(u_{\mathbb{W}}, u_{\mathbb{W}}^{\mathrm{loc}}) - a(u_{\mathbb{Y}}, u_{\mathbb{W}}^{\mathrm{loc}}) = a(u_{\mathbb{W}} - u_{\mathbb{Y}}, u_{\mathbb{W}}^{\mathrm{loc}}).$$

Thus, noticing that $u_{\mathbb{W}} - \tilde{u}_{\mathbb{W}} = u_{\mathbb{W}}^{\mathrm{loc}}$, we obtain

$$\begin{aligned}
\|u_{\mathbb{Y}} - u_{\mathbb{W}}\|_{\mathbb{X}}^2 - 2\rho_{\mathbb{Y}}(u_{\mathbb{W}}^{\mathrm{loc}}) &= \|u_{\mathbb{Y}} - \tilde{u}_{\mathbb{W}}\|_{\mathbb{X}}^2 + \|u_{\mathbb{W}}^{\mathrm{loc}}\|_{\mathbb{X}}^2 - 2\big(a(u_{\mathbb{Y}} - \tilde{u}_{\mathbb{W}}, u_{\mathbb{W}}^{\mathrm{loc}}) - a(u_{\mathbb{W}} - u_{\mathbb{Y}}, u_{\mathbb{W}}^{\mathrm{loc}})\big) \\
&= \|u_{\mathbb{Y}} - \tilde{u}_{\mathbb{W}}\|_{\mathbb{X}}^2 + \|u_{\mathbb{W}}^{\mathrm{loc}}\|_{\mathbb{X}}^2 - 2a(u_{\mathbb{W}} - \tilde{u}_{\mathbb{W}}, u_{\mathbb{W}}^{\mathrm{loc}}) \\
&= \|u_{\mathbb{Y}} - \tilde{u}_{\mathbb{W}}\|_{\mathbb{X}}^2 - \|u_{\mathbb{W}}^{\mathrm{loc}}\|_{\mathbb{X}}^2,
\end{aligned}$$

which gives the assertion.

A.2 Recursion Formulas of the 1D-Constraint Coefficients

We present some recursion formulas for the computation of the 1-dimensional constraint coefficients applied in Section 3.3. For this purpose, let $I = [a, b]$, with $-1 \le a < b \le 1$, be an interval. For $j \in \mathbb{N}_0$, recall that the restrictions of the functions $\psi_j : [-1, 1] \to \mathbb{R}$ from (3.2) to the interval I can be represented as linear combinations of the functions $\widetilde{\psi}_i := \psi_i \circ F_I^{-1}$ on I, for $i \in j_0$:

$$\psi_{i|I} = \sum_{j=0}^{i} b_{i,j}^I \, \widetilde{\psi}_i, \quad i \in \mathbb{N}_0;$$

here, the bijective affine linear transformation $F_I : [-1, 1] \to I$ is given by $F_I(t) = \alpha\, t + \beta$, with $\alpha := \frac{1}{2}(b - a)$ and $\beta := \frac{1}{2}(a + b)$. We note that the uniquely determined *constraint coefficients* $b_{i,j}^I$ in the above representation can be determined recursively (see [7] for a proof):

(i) For $i, j \in \{0, 1\}^2$, the following identities can be derived:

$$b_{0,0}^I = \frac{1}{2}(1 + \alpha - \beta), \quad b_{1,0}^I = \frac{1}{2}(1 - \alpha + \beta), \quad b_{0,1}^I = \frac{1}{2}(1 - \alpha - \beta), \quad b_{1,1}^I = \frac{1}{2}(1 + \alpha + \beta).$$

(ii) Moreover, we find that

$$b_{2,0}^I = \frac{1}{2}((\alpha - \beta)^2 - 1), \quad b_{2,1}^I = \frac{1}{2}((\alpha + \beta)^2 - 1), \quad b_{2,2}^I = \alpha^2.$$

(iii) For $i \ge 3$, it holds $b_{i,i}^I = \alpha\, b_{i-1,i-1}^I$ as well as

$$b_{i,0}^I = \frac{1}{i}\big((2i - 3)(\beta - \alpha)\, b_{i-1,0}^I - (i - 3)\, b_{i-2,0}^I\big),$$

$$b_{i,1}^I = \frac{1}{i}\big((2i - 3)(\alpha + \beta)\, b_{i-1,1}^I - (i - 3)\, b_{i-2,1}^I\big),$$

$$b_{i,2}^I = \frac{1}{i}\Big((2i - 3)\big(\alpha\,\big(\tfrac{1}{5}\, b_{i-1,3}^I - (b_{i-1,0}^I - b_{i-1,1}^I)\big) + \beta\, b_{i-1,2}^I\big) - (i - 3)\, b_{i-2,2}^I\Big).$$

In addition, for $i \ge 4$, we have

$$b_{i,i-1}^I = \frac{2i - 3}{i}\Big(\frac{i - 1}{2i - 5}\, \alpha\, b_{i-1,i-2}^I + \beta\, b_{i-1,i-1}^I\Big).$$

(iv) For $i \ge 5$ and $j \in \{3, \ldots, i - 2\}$, the coefficients are given by

$$b_{i,j}^I = \frac{1}{i}\Big((2i - 3)\big(\alpha\,\big(\tfrac{j}{2j-3}\, b_{i-1,j-1}^I + \tfrac{j-1}{2j+1}\, b_{i-1,j+1}^I\big) + \beta\, b_{i-1,j}^I\big) - (i - 3)\, b_{i-2,j}^I\Big).$$

(v) Finally, for $j \ge 2$ and $i \in \{0, \ldots, j - 1\}$, it holds $b_{i,j} = 0$.

References

[1] M. Ainsworth and J. T. Oden, A posteriori error estimation in finite element analysis, *Comput. Methods Appl. Mech. Engrg.* **142** (1997), no. 1–2, 1–88.

[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** (1988), no. 1, 172–203.

[3] I. Babuška and M. Suri, The *h-p* version of the finite element method with quasi-uniform meshes, *RAIRO Modél. Math. Anal. Numér.* **21** (1987), no. 2, 199–238.

[4]   I. Babuška and M. Suri, The treatment of nonhomogeneous Dirichlet boundary conditions by the *p*-version of the finite element method, *Numer. Math.* **55** (1989), no. 1, 97–121.

[5]   I. Babuška and M. Suri, The *p* and *h-p* versions of the finite element method, basic principles and properties, *SIAM Rev.* **36** (1994), no. 4, 578–632.

[6]   R. Becker and R. Rannacher, An optimal control approach to a posteriori error estimation in finite element methods, *Acta Numer.* **10** (2001), 1–102.

[7]   A. Byfut and A. Schröder, Unsymmetric multi-level hanging nodes and anisotropic polynomial degrees in H^1-conforming higher-order finite element methods, *Comput. Math. Appl.* **73** (2017), no. 9, 2092–2150.

[8]   P. Daniel, A. Ern, I. Smears and M. Vohralík, An adaptive *hp*-refinement strategy with computable guaranteed bound on the error reduction factor, *Comput. Math. Appl.* **76** (2018), no. 5, 967–983.

[9]   L. Demkowicz, *Computing with hp-Adaptive Finite Elements. Vol. 1*, Chapman & Hall/CRC Appl. Math. Nonlinear Sci., Chapman & Hall/CRC, Boca Raton, 2007.

[10] P. Di Stolfo, A. Schröder, N. Zander and S. Kollmannsberger, An easy treatment of hanging nodes in *hp*-finite elements, *Finite Elem. Anal. Des.* **121** (2016), 101–117.

[11] V. Dolejší, A. Ern and M. Vohralík, *hp*-adaptation driven by polynomial-degree-robust a posteriori error estimates for elliptic problems, *SIAM J. Sci. Comput.* **38** (2016), no. 5, A3220–A3246.

[12] W. Dörfler, A convergent adaptive algorithm for Poisson's equation, *SIAM J. Numer. Anal.* **33** (1996), no. 3, 1106–1124.

[13] T. Eibner and J. M. Melenk, An adaptive strategy for *hp*-FEM based on testing for analyticity, *Comput. Mech.* **39** (2007), no. 5, 575–595.

[14] K. Eriksson, D. Estep, P. Hansbo and C. Johnson, Introduction to adaptive methods for differential equations, in: *Acta Numerica 1995*, Cambridge University, Cambridge (1995), 105–158.

[15] T. Fankhauser, T. P. Wihler and M. Wirz, The *hp*-adaptive FEM based on continuous Sobolev embeddings: Isotropic refinements, *Comput. Math. Appl.* **67** (2014), no. 4, 854–868.

[16] B. Guo and I. Babuška, The *hp*-version of the finite element method. Part I: The basic approximation results, *Comput. Mech.* **1** (1986), 21–41.

[17] B. Guo and I. Babuška, The *hp*-version of the finite element method. Part II: General results and applications, *Comput. Mech.* **1** (1986), 203–220.

[18] P. Heid, B. Stamm and T. P. Wihler, Gradient flow finite element discretizations with energy-based adaptivity for the Gross–Pitaevskii equation, *J. Comput. Phys.* **436** (2021), Article ID 110165.

[19] P. Houston and E. Süli, A note on the design of *hp*-adaptive finite element methods for elliptic partial differential equations, *Comput. Methods Appl. Mech. Engrg.* **194** (2005), no. 2–5, 229–243.

[20] P. Houston and T. P. Wihler, Adaptive energy minimisation for *hp*-finite element methods, *Comput. Math. Appl.* **71** (2016), no. 4, 977–990.

[21] G. E. Karniadakis and S. J. Sherwin, *Spectral/hp Element Methods for CFD*, Numer. Math. Sci. Comput., Oxford University, New York, 1999.

[22] J. M. Melenk and B. I. Wohlmuth, On residual-based a posteriori error estimation in *hp*-FEM, *Adv. Comput. Math.* **15** (2001), 311–331.

[23] A. Miraçi, J. Papež and M. Vohralík, A-posteriori-steered *p*-robust multigrid with optimal step-sizes and adaptive number of smoothing steps, *SIAM J. Sci. Comput.* **43** (2021), no. 5, S117–S145.

[24] W. F. Mitchell and M. A. McClain, A comparison of *hp*-adaptive strategies for elliptic partial differential equations, *ACM Trans. Math. Software* **41** (2014), no. 1, Paper No. 2.

[25] J. Schöberl, J. M. Melenk, C. Pechstein and S. Zaglmayr, Additive Schwarz preconditioning for *p*-version triangular and tetrahedral finite elements, *IMA J. Numer. Anal.* **28** (2008), no. 1, 1–24.

[26] A. Schröder, Constraints Coefficients in *hp*-FEM, in: *Numerical Mathematics and Advanced Applications*, Springer, Berlin (2008), 183–190.

[27] C. Schwab, *p- and hp-FEM, Theory and Applications to Solid and Fluid Mechanics*, Oxford University, Oxford, 1998.

[28] P. Šolín, K. Segeth and I. Doležel, *Higher-Order Finite Element Methods*, Stud. Adv. Math., Chapman & Hall/CRC, Boca Raton, 2004.

[29] E. Süli and P. Houston, Adaptive finite element approximation of hyperbolic problems, in: *Error Estimation and Adaptive Discretization Methods in Computational Fluid Dynamics*, Lect. Notes Comput. Sci. Eng. 25, Springer, Berlin (2003), 269–344.

[30] B. Szabó and I. Babuška, *Finite Element Analysis*, J. Wiley & Sons, New York, 1991.

[31] R. Verfürth, *A Review of a Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*, B.G. Teubner, Stuttgart, 1996.

[32] T. P. Wihler, An *hp*-adaptive strategy based on continuous Sobolev embeddings, *J. Comput. Appl. Math.* **235** (2011), no. 8, 2731–2739.