An adaptive hp-DG-FE Method for Elliptic Problems. Convergence and Optimality in the 1D Case

An adaptive hp-DG-FE Method for Elliptic Problems. Convergence and Optimality in the 1D Case

Dedicated to the memory of Professor Ben-yu Guo

Paola Antonietti* Claudio Canuto† Marco Verani‡

Abstract

We propose and analyze an hp-adaptive DG-FEM algorithm, termed **hp-ADFEM**, and a realization of it in one space dimension which is convergent, instance optimal, and h- and p-robust. The procedure consists of iterating two routines: one hinges on Binev's algorithm for the adaptive hp-approximation of a given function, and finds a near-best hp-approximation of the current discrete solution and data to a desired accuracy; the other one improves the discrete solution to a finer but comparable accuracy, by iteratively applying Dörfler marking and h-refinement.

1 Introduction

The design and analysis of adaptive hp-type finite element methods for elliptic problems is significantly more challenging than for h-type methods. Indeed, as demonstrated e.g. by some examples given in [6, Sect.1], one should include in the adaptive procedure the possibility of stepping back from an early choice between h-refinement and p-enrichment: while the true structure of the solution reveals itself along the iterations, one should be able to re-distribute the allocated degrees of freedom between h- and p-resolution. The existence of (rather) pathological situations has not prevented the development of practical hp-adaptive algorithms that work (see e.g. [9] and the references therein), but in most cases these procedures are not supported by a sound mathematical theory, which assesses the optimality, and even the convergence, of the method (unless a-priori assumptions on the structure of the solution are made).

The crucial issue is an approximation problem: how can we build an hp-finite element space of minimal dimension in which a given function can be approximated with a prescribed accuracy? A constructive answer to this question has been given by P. Binev in the past few years (see [5]), who designed a greedy hp-algorithm, which is incremental with respect to the dimension and has instance optimality properties (see Sect. 2.3).

With a good answer to such an approximation problem, one may think of recursively applying the hp-adaptive algorithm to a sequence of Galerkin discrete solutions of the elliptic problem, built in a way to get closer and closer to the exact solution. This idea

*MOX-Dipartimento di Matematica, Politecnico di Milano, P.zza Leonardo Da Vinci 32, I-20133 Milano, Italy (paola.antonietti@polimi.it).

†Dipartimento di Scienze Matematiche, Politecnico di Torino, Corso Duca degli Abruzzi 24, I-10129 Torino, Italy (claudio.canuto@polito.it)

‡MOX-Dipartimento di Matematica, Politecnico di Milano, P.zza Leonardo Da Vinci 32, I-20133 Milano, Italy (marco.verani@polimi.it).

has been implemented in [6], where a general framework for adaptive hp-discretizations has been devised, and an adaptive algorithm termed **hp-AFEM** has been proposed, which guarantees convergence and instance optimality of the sequence of generated Galerkin solutions. The algorithm is both h- and p-optimal in one space dimension, whereas in higher dimensions p-robustness is lost, partly due to the need of going from the non-conforming meshes produced by Binev's algorithm to the conforming ones needed in a continuous Galerkin method, and partly due to the use of a residual-based error estimator (the latter obstruction may be removed by resorting to equilibrated flux estimators, as done in [7]).

Since Binev's algorithm produces non-conforming meshes and discontinuous approximations, it is quite natural to associate to it a Discontinuous, rather than a Continuous, Galerkin discretization of the elliptic problem. The purpose of this paper is to take a step forward in this direction. In particular, hereafter we propose an hp-adaptive DG-FEM algorithm, termed **hp-ADFEM**, and a realization of it in one space dimension which is convergent, instance optimal, and h- and p-robust. No restriction on the relative size of neighboring elements, nor on the polynomial degrees used on them, is required. In building a discrete solution that matches a prescribed accuracy, we extend to the hp-case the approach developed in [4] for h-type DG methods, using in the analysis several results on hp-type a posteriori error estimators (see e.g. [8] and the references therein). The multi-dimensional case is currently under investigation [1]; while our general convergence theorem holds in any dimension, proving p-robustness seems to require a grading property in the distribution of polynomial degrees over the partition, which is not guaranteed by the algorithm proposed in [5].

The paper is organized as follows. In Sect. 2, we introduce our general framework for the hp-approximation of a given function, and we present Binev algorithm. Sect. 3 describes the hp-DG discretizations that we consider, and collects some of their properties. Sect. 4 contains the general convergence and instance optimality result, based on the concatenation of Binev's algorithm and a procedure to compute DG-solutions with polynomial data, matching a prescribed tolerance. Finally, in Sect. 5 we illustrate a possible realization of this procedure, which is based on the classical SOLVE $\rightarrow$ ESTIMATE $\rightarrow$ MARK $\rightarrow$ REFINE paradigm.

The following notation will be used thoughout the paper. By $A \lesssim B$ we will mean that A can be bounded by a multiple of B, independently of parameters which A and B may depend on. Likewise, $A \simeq B$ is defined as both $A \lesssim B$ and $B \lesssim A$.

C.C. wishes to remember the long-lasting friendship and mutual esteem with Professor Ben-yu Guo, a person of great humanity and a devoted scientist.

2 hp-partitions and hp-approximations

Let Ω be a bounded open interval of the real line. In view of the hp-adaptive discretization of a boundary-value problem therein, we introduce some notation concerning partitions in Ω and function spaces built on them.

2.1 Partitions of the domain

We assume that we are given an essentially disjoint initial partition $\mathcal{K}_0$ of $\bar{\Omega}$ into finitely many closed subintervals, which will be the initial geometric elements; the initial subdivision may depend upon the data of the problem at hand. Then, we apply subsequent dyadic subdivisions, by halving each element K that we encounter into two closed subintervals K' and K'' of equal size, the 'children' of K, such that $K = K' \cup K''$ and $|K' \cap K''| = 0$.

The set $\mathfrak{K}$ of all these geometric elements forms an infinite binary 'master tree', having as its roots the elements of the initial partition of $\bar{\Omega}$. A subtree of the master tree is a finite subset of $\mathfrak{K}$ that contains all roots and for each element in the subset both its parent and its sibling are in the subset. The leaves of a subtree form an essentially disjoint partition of $\bar{\Omega}$. The set of all such geometric partitions, or 'h-partitions', will be denoted as $\mathbb{K}$. For $\mathcal{K}, \widetilde{\mathcal{K}} \in \mathbb{K}$, we call $\widetilde{\mathcal{K}}$ a refinement of $\mathcal{K}$, and denoted as $\mathcal{K} \le \widetilde{\mathcal{K}}$, when any $K \in \widetilde{\mathcal{K}}$ is either in $\mathcal{K}$ or has an ancestor in $\mathcal{K}$.

Starting from an h-partition $\mathcal{K} \in \mathbb{K}$, we obtain an *hp-partition* $\mathcal{D}$ by associating an integer $p \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$ to each element $K \in \mathcal{K}$. This integer will represent a polynomial degree, which will identify certain finite dimensional spaces of polynomial functions defined in K. A pair $D = (K_D, p_D) \in \mathfrak{K} \times \mathbb{N}_0$ formed by a geometric element K_D and an integer p_D will be termed an *hp-element*. Thus, a collection $\mathcal{D} = \{D = (K_D, d_D)\}$ of hp-elements is an hp-partition provided $\mathcal{K}(\mathcal{D}) := \{K_D : D \in \mathcal{D}\} \in \mathbb{K}$; the latter will be the associated h-partition. The collection of all hp-partitions is denoted as $\mathbb{D}$. Since $p+1$ is the dimension of the space $\mathbb{P}_p(K)$ of the univariate polynomials of degree $\le p$ in K, we define the *dimension* of the hp-partition $\mathcal{D}$ as the integer

$$\#\mathcal{D} := \sum_{D \in \mathcal{D}} (p_D + 1).$$

For $\mathcal{D}, \widetilde{\mathcal{D}} \in \mathbb{D}$, we call $\widetilde{\mathcal{D}}$ a refinement of $\mathcal{D}$, and write $\mathcal{D} \le \widetilde{\mathcal{D}}$, when both $\mathcal{K}(\mathcal{D}) \le \mathcal{K}(\widetilde{\mathcal{D}})$, and $d_{\widetilde{D}} \ge d_D$, for any $D \in \mathcal{D}$, $\widetilde{D} \in \widetilde{\mathcal{D}}$ with K_D being either equal to $K_{\widetilde{D}}$ or an ancestor of $K_{\widetilde{D}}$.

2.2 Approximation spaces on *hp*-partitions

Let Z be a normed space of vector-valued functions $z : \Omega \to \mathbb{R}^m$ ($m \ge 1$), which is relevant for our application. For any geometric element $K \in \mathfrak{K}$, let Z_K be the space collecting the restrictions $z_{|K}$ to K of all functions $z \in Z$. Then, for any geometric partition $\mathcal{K} \in \mathbb{K}$, we define

$$Z_{\mathcal{K}} := \{z : \Omega \to \mathbb{R}^m \ : \ z_{|K} \in Z_K \ \forall K \in \mathcal{K}\} = \prod_{K \in \mathcal{K}} Z_K \, ; \tag{1}$$

obviously, $Z \subseteq Z_{\mathcal{K}}$. In the sequel, we will work with functions that belong to $Z_{\mathcal{K}}$ for some partition $\mathcal{K} \in \mathbb{K}$; therefore, we set

$$\mathcal{Z} := \bigcup_{\mathcal{K} \in \mathbb{K}} Z_{\mathcal{K}}.$$

We assume that for any $K \in \mathfrak{K}$, the space Z_K contains all polynomial functions of any degree, and this set of functions is dense in Z_K. Then, for $p \in \mathbb{N}_0$ we assume we have chosen finite dimensional spaces $Z_{K,p} \subset Z_K$ of polynomial functions on K of degree related to p, satisfying $Z_{K,p} \subset Z_{K,p+1}$ and $Z_{K,p} \subset Z_{K',p} \times Z_{K'',p}$ (K' and K'' being the children of K). For any $D = (K_D, p_D) \in \mathfrak{K} \times \mathbb{N}_0$, we set $Z_D := Z_{K_D, p_D}$. Then, given an hp-partition $\mathcal{D}$, we define

$$Z_{\mathcal{D}} := \{z : \Omega \to \mathbb{R}^m \ : \ z_{|K_D} \in Z_D \ \forall D \in \mathcal{D}\} = \prod_{D \in \mathcal{D}} Z_D \, , \tag{2}$$

which obviously satisfies $Z_{\mathcal{D}} \subset Z_{\mathcal{K}(\mathcal{D})}$. We will use the notation $z_{\mathcal{D}}$ to indicate a function in $Z_{\mathcal{D}}$. Note that no interelement continuity is imposed in the definition of $Z_{\mathcal{D}}$. Also note that the dimension of $Z_{\mathcal{D}}$ is proportional to the cardinality $\#\mathcal{D}$.

For all $D \in \mathfrak{K} \times \mathbb{N}_0$, we assume a local projector $Q_D : \mathcal{Z} \to Z_D$, and a local error functional $e_D = e_D(z) \ge 0$, that, for any $z \in \mathcal{Z}$ gives a measure for some function of the distance between $z_{|K_D}$ and its local approximation $z_D := Q_D(z)$. We assume that this error functional is non-increasing under both 'h-refinements' and 'p-enrichments', in the sense that

$$\begin{aligned} e_{D'} + e_{D''} \leq e_D \text{ when } K_{D'}, K_{D''} \text{ are the children of } K_D, \text{ and } p_{D'} = p_{D''} = p_D; \\ e_{D'} \leq e_D \text{ when } K_{D'} = K_D \text{ and } p_{D'} \geq p_D. \end{aligned} \tag{3}$$

Given any hp-partition $\mathcal{D} \in \mathbb{D}$, we define the global projector $Q_\mathcal{D} : \mathcal{Z} \to Z_\mathcal{D}$ as $Q_\mathcal{D}(z) := (z_D)_{D \in \mathcal{D}}$, and the global error functional

$$\mathrm{E}_\mathcal{D}(z) := \sum_{D \in \mathcal{D}} e_D(z), \tag{4}$$

which is a measure for the distance between z and its projection $z_\mathcal{D} := Q_\mathcal{D}(z)$. Note that (3) is equivalent to

$$\mathrm{E}_{\widetilde{\mathcal{D}}}(z) \leq \mathrm{E}_\mathcal{D}(z) \quad \forall \widetilde{\mathcal{D}} \geq \mathcal{D}. \tag{5}$$

2.3 The instance optimal *hp*-approximation algorithm

Herafter, we present the greedy algorithm proposed by P. Binev [5] to produce a near-best adaptive hp-approximation of a function $z \in \mathcal{Z}$, based on the associated local error functionals $e_D = e_D(z)$ and global error functional $\mathrm{E}_\mathcal{D} = \mathrm{E}_\mathcal{D}(z)$ introduced above.

Denote by $R \geq 1$ the cardinality of the initial geometric partition K_0. Using property (3), Binev's algorithm builds a sequence of hp-partitions $\mathcal{D}_N$, $N \geq R$, satisfying $\#\mathcal{D}_N = N$; the construction is incremental, in that going from $\mathcal{D}_N$ to $\mathcal{D}_{N+1}$ one exploits the work already done to build $\mathcal{D}_N$. The main feature of the algorithm is its instance optimality, expressed as follows.

Theorem 2.1 ([5]). *For $n \geq R$ let*

$$\sigma_n := \inf_{\#\mathcal{D} \leq n} E_\mathcal{D}$$

be the smallest error achievable with an hp-partition of cardinality $\leq n$. Then, the hp-partitions $\mathcal{D}_N$ produced by Binev's algorithm yield error functionals $E_{\mathcal{D}_N}$ satisfying the bounds

$$\mathrm{E}_{\mathcal{D}_N} \leq \frac{2N}{N-n+1} \sigma_n \quad \forall n \leq N. \qquad \square \tag{6}$$

Binev's construction can be easily used to produce an instance optimal hp-partition for which the error functional is below a given threshold.

Corollary 2.1 ([6]). *Let $B > 1$ arbitrary. Given $\varepsilon > 0$, let $\mathcal{D} \in \mathbb{D}$ be the first partition in Binev's sequence for which $E_\mathcal{D}^{\frac{1}{2}} \leq \varepsilon$. Then, setting $b = \frac{1}{2}(1 - \frac{1}{B}) < 1$, it holds*

$$\#\mathcal{D} \leq B \, \#\hat{\mathcal{D}}$$

for all partitions $\hat{\mathcal{D}} \in \mathbb{D}$ satisfying $E_{\hat{\mathcal{D}}}^{\frac{1}{2}} \leq b\varepsilon$. $\square$

This result motivates the introduction of the following routine, which will constitute one of the two major building blocks of our proposed hp-adaptive algorithm.

- $[\mathcal{D}, z_{\mathcal{D}}] := \textbf{hp-NEARBEST}(\varepsilon, z)$

  The routine **hp-NEARBEST** takes as input $\varepsilon > 0$, and $z \in \mathcal{Z}$, and outputs $\mathcal{D} \in \mathbb{D}$ as well as $z_{\mathcal{D}} \in Z_{\mathcal{D}}$ such that $\mathrm{E}_{\mathcal{D}}(z)^{\frac{1}{2}} \le \varepsilon$ and, for some constants $0 < b < 1 < B$, $\#\mathcal{D} \le B\, \#\widehat{\mathcal{D}}$ for any $\widehat{\mathcal{D}} \in \mathbb{D}$ with $\mathrm{E}_{\widehat{\mathcal{D}}}(z)^{\frac{1}{2}} \le b\varepsilon$.

The approximation $z_{\mathcal{D}}$ of the input z is just the element-wise projection given by the operator $Q_{\mathcal{D}}$ associated with the partition $\mathcal{D}$, i.e., we set

$$z_{\mathcal{D}} := Q_{\mathcal{D}}(z). \tag{7}$$

3 Discontinuous Galerkin *hp*-discretizations

We are interested in solving numerically the model boundary-value problem

$$Au = f \quad \text{in} \ \ \Omega, \qquad u = 0 \quad \text{on} \ \ \partial\Omega \tag{8}$$

with $Au := -(\nu u_x)_x + \xi u$, where $\nu \in L^\infty(\Omega)$ satisfies $\operatorname{essinf}_\Omega \nu > 0$, $\xi \in L^2(\Omega)$ satisfies $\xi \ge 0$ a.e. in Ω, $f \in L^2(\Omega)$. We actually assume that ν, ξ are piecewise-H^1 functions, precisely that $\nu_{|K}, \xi_{|K} \in H^1(K)$ for each element K of the initial partition $\mathcal{K}_0$ introduced in Sect. 2.1; we will write $\nu, \xi \in H^1(\Omega; \mathcal{K}_0)$. It will be convenient to refer to a triple $g := (\nu, \xi, f)$ as to a "data" of our problem; we thus have $g \in G(\Omega) := (H^1(\Omega; \mathcal{K}_0))^2 \times L^2(\Omega)$. The solution $u \in H^1_0(\Omega)$ of Problem (8) for given data g will be indicated by $u(g)$.

The following notation will be useful in the design of a DG discretization of our problem. For any element $K \in \mathfrak{K}$, let $(v, w)_K$ denote the L^2-inner product in K, with corresponding norm $\|v\|_K$. For any geometric partition $\mathcal{K} \in \mathbb{K}$, let us set

$$V_{\mathcal{K}} := \{v \in L^2(\Omega) : v_{|K} \in H^1(K) \ \forall K \in \mathcal{K}\}. \tag{9}$$

For $v \in V_{\mathcal{K}}$, it will be convenient to denote by $\tilde{v}_x$ the function in $L^2(\Omega)$ such that $(\tilde{v}_x)_{|K} = (v_{|K})_x$ for all $K \in \mathcal{K}$; thus, $\|\tilde{v}_x\|^2_\Omega = \sum_{K \in \mathcal{K}} \|(v_{|K})_x\|^2_K$. Let us denote by $\mathcal{E}_{\mathcal{K}}$ the set of all endpoints of elements in $\mathcal{K}$, and let us define the jumps and averages of a piecewise smooth function ϕ on $\mathcal{K}$ as follows: if $e \in \mathcal{E}_{\mathcal{K}}$ is shared by two contiguous elements K^- and K^+, then we set

$$[\![\phi]\!]_e := \phi_{|K^-}(e) - \phi_{|K^+}(e), \qquad \{\!\{\phi\}\!\}_e := \frac{1}{2}(\phi_{|K^-}(e) + \phi_{|K^+}(e)),$$

whereas if e is the left/right boundary point of Ω, we set $[\![\phi]\!]_e = +/-\phi(e)$ and $\{\!\{\phi\}\!\}_e = \phi(e)$.

For any hp-partition $\mathcal{D} \in \mathbb{D}$ let us set

$$V_{\mathcal{D}} := \{v \in L^2(\Omega) : v_{|K_D} \in \mathbb{P}_{p_D}(K_D) \ \ \forall D \in \mathcal{D}\} \subset V_{\mathcal{K}(\mathcal{D})}. \tag{10}$$

If $D \in \mathcal{D}$, let $h_D := |K_D|$ denote the size of the element K_D. If $e \in \mathcal{E}_{\mathcal{D}} := \mathcal{E}_{\mathcal{K}(\mathcal{D})}$, we define the weight

$$\sigma_{\mathcal{D},e} := \max\left(\frac{p^2_{D^-}}{h_{D^-}}, \frac{p^2_{D^+}}{h_{D^+}}\right) \tag{11}$$

if $e \in K_{D^-} \cap K_{D^+}$, and $\sigma_{\mathcal{D},e} := \frac{p^2_D}{h_D}$ if $e \in \partial\Omega \cap K_D$.

It is convenient to introduce the inner product $(\phi, \psi)_{\mathcal{E}_{\mathcal{D}}} := \sum_{e \in \mathcal{E}_{\mathcal{D}}} \phi_e \psi_e$ in $\mathbb{R}^{|\mathcal{E}_{\mathcal{D}}|}$ between two quantities $\phi = (\phi_e)$ and $\psi = (\psi_e)$ indexed in $\mathcal{E}_{\mathcal{D}}$. The corresponding norm will be denoted by $\|\phi\|_{\mathcal{E}_{\mathcal{D}}}$.

At this point, we are ready to introduce the symmetric bilinear form $a_{\mathcal{D}}$ defined on $V_{\mathcal{D}} \times V_{\mathcal{D}}$ as

$$\begin{aligned} a_{\mathcal{D}}(w, v) := {} & (\nu\, \tilde{w}_x, \tilde{v}_x)_\Omega + (\xi\, w, v)_\Omega \\ & - (\{\!\{\nu\, w_x\}\!\}, [\![v]\!])_{\mathcal{E}_{\mathcal{D}}} - (\{\!\{\nu\, v_x\}\!\}, [\![w]\!])_{\mathcal{E}_{\mathcal{D}}} + \gamma\, (\sigma_{\mathcal{D}} [\![w]\!], [\![v]\!])_{\mathcal{E}_{\mathcal{D}}} \ , \end{aligned} \tag{12}$$

where $\gamma > 0$ is a sufficiently large stabilization parameter, as well as the DG-norm defined on $V_{\mathcal{D}}$ as

$$\|v\|_{\mathcal{D}} := \Big((\nu\, \tilde{v}_x, \tilde{v}_x)^2_{\Omega} + \gamma \, \| \, \sigma_{\mathcal{D}}^{1/2} [\![v]\!] \, \|^2_{\mathcal{E}_{\mathcal{D}}} \Big)^{\frac{1}{2}} . \tag{13}$$

It is well-known (see [2, 3]) that $a_{\mathcal{D}}$ is a continuous form with respect to the DG-norm, and it is coercive provided γ is chosen large enough, with coercivity and continuity constants independent of $\mathcal{D}$; in the sequel, we will assume that this condition is satisfied.

Since $a_{\mathcal{D}}$ depends on the choice of coefficients ν and ξ, and since in the adaptive algorithm we will consider a sequence of DG discretizations with changing (piecewise polynomial) data, sometimes we will prefer the more precise notation $a_{\mathcal{D}}(w, v; \nu, \xi)$ to indicate the right-hand side of (12).

Problem 8 with data $g = (\nu, \xi, f) \in G(\Omega)$ is then discretized by the following Simmetric Interior Penalty Discontinuous-Galerkin method ([2]):

$$u_{\mathcal{D}} \in V_{\mathcal{D}} \quad : \quad a_{\mathcal{D}}(u_{\mathcal{D}}, v_{\mathcal{D}}; \nu, \xi) = (f, v_{\mathcal{D}})_{\Omega} \qquad \forall v_{\mathcal{D}} \in V_{\mathcal{D}}. \tag{14}$$

We will write $u_{\mathcal{D}} = u_{\mathcal{D}}(g)$ when we want to stress the dependence of $u_{\mathcal{D}}$ upon the given data g.

3.1 Approximation spaces and error functionals

Hereafter, we specify the choice of approximation spaces and error functionals, introduced in a general setting in Sect. 2.2, that is tailored to the discretization problem of interest.

Since we will deal with approximations of a specific solution of Problem 8, and approximations of the corresponding data, our functions z will be of the form $z = (v, g) = (v, \nu, \xi, f)$. Then, a natural choice for the "base" space Z is $Z = H^1(\Omega) \times G(\Omega) = H^1(\Omega) \times (H^1(\Omega; \mathcal{K}_0))^2 \times L^2(\Omega)$. Note that for $\mathcal{K} \in \mathbb{K}$, the local spaces Z_K that form the global space $Z_{\mathcal{K}}$ according to (1) are given by $Z_K = (H^1(K))^3 \times L^2(K)$.

For any element $K \in \mathfrak{K}$ and integer $p \in \mathbb{N}_0$, we set

$$Z_{K,p} = V_{K,p} \times G_{K,p} \qquad \text{with} \quad V_{K,p} = \mathbb{P}_p(K) \quad \text{and} \quad G_{K,p} = \mathbb{P}_{p+1}(K) \times \mathbb{P}_{p+1}(K) \times \mathbb{P}_{p-1}(K).$$

Then, for any $\mathcal{D} \in \mathbb{D}$, we define $Z_{\mathcal{D}}$ according to (2); it is easily seen that $Z_{\mathcal{D}} =: V_{\mathcal{D}} \times G_{\mathcal{D}}$, where $V_{\mathcal{D}}$ has been already introduced in (10). We will write $z_{\mathcal{D}} = (v_{\mathcal{D}}, g_{\mathcal{D}}) = (v_{\mathcal{D}}, \nu_{\mathcal{D}}, \xi_{\mathcal{D}}, f_{\mathcal{D}})$ for the generic element in $Z_{\mathcal{D}}$.

In order to define the projectors Q_D, let $\Pi^0_{K,p} : L^2(K) \to \mathbb{P}_p(K)$ be the L^2-orthogonal projector, and let $\Pi^1_{K,p} : H^1(K) \to \mathbb{P}_p(K)$ be the projector such that

$$(\Pi^1_{K,p} v)_x = \Pi^0_{K,p} v_x \qquad \text{and} \qquad \int_K \Pi^1_{K,p} v = \int_K v, \qquad \forall v \in H^1(K).$$

The latter definition can be extended to functions v that are just piecewise-H^1 on K, by replacing v_x with $\tilde{v}_x$ in the L^2-projection. Then, for $z = (v, g) = (v, \nu, \xi, f) \in \mathcal{Z}$ and $D = (K_D, p_D)$ we set

$$Q_D(z) = (\Pi^1_{K_D, p_D} v_{|K_D}, \Pi^1_{K_D, p_D+1} \nu_{|K_D}, \Pi^1_{K_D, p_D+1} \xi_{|K_D}, \Pi^0_{K_D, p_D-1} f_{|K_D}).$$

The corresponding local error functional is defined as

$$e_D(z) := e_{1,D}(v) + \frac{1}{\kappa^2} \mathrm{osc}^2_D(g) =: e_{1,D}(v) + \frac{1}{\kappa^2} \big(e_{1,D}(\nu) + e_{1,D}(\xi) + e_{0,D}(f) \big), \tag{15}$$

where for $\varphi = v, \nu, \xi$

$$e_{1,D}(\varphi) := \|(\mathrm{I} - \Pi^0_{K_D, p_D})(\tilde{\varphi}_x)_{|K_D}\|^2_{K_D}, \qquad e_{0,D}(f) := \frac{h_D}{p_D} \|(\mathrm{I} - \Pi^0_{K_D, p_D}) f_{|K_D}\|^2_{K_D},$$

and $\kappa > 0$ is a (sufficiently small) penalization parameter to be chosen later on.

Finally, for a given hp-partition $\mathcal{D} \in \mathbb{D}$, the global projector $Q_{\mathcal{D}} : \mathcal{Z} \to Z_{\mathcal{D}}$ and the global error functional $\mathrm{E}_{\mathcal{D}}(z) = \mathrm{E}_{\mathcal{D}}(v, g)$ are defined as in Sect. 2.2 (see (4)).

We now establish some properties involving the functional $\mathrm{E}_{\mathcal{D}}$, that will be useful in the sequel.

Property 3.1. *There exists a constant $C_0 > 0$ such that for any $z = (v, \nu, \xi, f) \in \mathcal{Z}$ and for any partition $\mathcal{D} \in \mathbb{D}$ one has*

$$\|\nu - \nu_{\mathcal{D}}\|_{L^\infty(\Omega)} + \|\xi - \xi_{\mathcal{D}}\|_{L^\infty(\Omega)} \leq C_0 \kappa\, \mathrm{E}_{\mathcal{D}}(z)^{\frac{1}{2}},$$

where $z_{\mathcal{D}} = (v_{\mathcal{D}}, \nu_{\mathcal{D}}, \xi_{\mathcal{D}}, f_{\mathcal{D}}) = Q_{\mathcal{D}}(z)$.

Proof. For any $\mathcal{D} \in \mathbb{D}$ and any $D \in \mathcal{D}$, set $\psi := (\nu - \nu_{\mathcal{D}})_{|K_D}$. Since by construction ψ vanishes at a point in K_D, we have for any $x \in K_D$

$$|\psi(x)| \leq h_D^{1/2} \|\psi_x\|_{K_D} \leq |\Omega|\, e_{1,D}(\nu)^{1/2},$$

from which the bound for $\|\nu - \nu_{\mathcal{D}}\|_{L^\infty(\Omega)}$ easily follows. The coefficient ξ can be treated similarly. $\square$

At this point, let us fix once and for all the data of interest $g_\star = (\nu_\star, \xi_\star, f_\star) \in G(\Omega)$ for Problem (8), and let $u_\star := u(g_\star)$ be the corresponding solution.
Let us set $\nu_0 := \operatorname{ess\,inf}_\Omega \nu_\star > 0$.

Assumption 3.1. *Let $\mathcal{D}_0$ denote the root partition $\mathcal{K}_0$ endowed with polynomials of degree 1 in each element. Setting $z_0 := (0, g_\star) \in \mathcal{Z}$, we assume that $\mathcal{D}_0$ is chosen to satisfy*

$$C_0 \kappa\, \mathrm{E}_{\mathcal{D}_0}(z_0)^{\frac{1}{2}} \leq \frac{\nu_0}{\lambda}, \qquad \textit{where} \quad \lambda := 2 + \frac{1}{\sqrt{2}}|\Omega|.$$

Recalling (5), this assumption together with Property 3.1 guarantees that for any $\mathcal{D} \in \mathbb{D}$ (which trivially satisfies $\mathcal{D} \geq \mathcal{D}_0$), Problem 8 with approximate data $\nu_{\mathcal{D}}$ and $\xi_{\mathcal{D}}$ is coercive in $H_0^1(\Omega)$, precisely one has

$$(\nu_{\mathcal{D}} v_x, v_x)_\Omega + (\xi_{\mathcal{D}} v, v)_\Omega \geq \frac{\nu_0}{2} \|v_x\|_\Omega^2 \qquad \forall v \in H_0^1(\Omega). \tag{16}$$

This easily follows using the bound $\|v\|_\Omega \leq \frac{1}{2\sqrt{2}} |\Omega|\, \|v_x\|_\Omega$.

The following result is fundamental for establishing the convergence of our adaptive algorithm.

Proposition 3.1. *i) There exists a constant $C_\star > 0$ with the following property: for all $\mathcal{D} \in \mathbb{D}$ and all $z \in \mathcal{Z}$ of the form $z = (v, g_\star)$, let $z_{\mathcal{D}} = (v_{\mathcal{D}}, g_{\mathcal{D}}) := Q_{\mathcal{D}}(z)$, and let $u(g_{\mathcal{D}}) \in H_0^1(\Omega)$ be the solution of Problem 8 with data $g_{\mathcal{D}}$; then, it holds*

$$\|u_\star - u(g_{\mathcal{D}})\|_{H_0^1(\Omega)} \leq C_\star \kappa\, \mathrm{E}_{\mathcal{D}}(z_0)^{\frac{1}{2}} \leq C_\star \kappa\, \mathrm{E}_{\mathcal{D}}(z)^{\frac{1}{2}}, \tag{17}$$

where κ is the penalization parameter introduced in (15).

ii) For all $\mathcal{D} \in \mathbb{D}$, $v \in V_{\mathcal{K}(\mathcal{D})}$, $w \in H_0^1(\Omega)$ and $g \in G(\Omega)$, it holds

$$|\,\mathrm{E}_{\mathcal{D}}(v, g)^{\frac{1}{2}} - \mathrm{E}_{\mathcal{D}}(w, g)^{\frac{1}{2}}| \leq \|v - w\|_{\mathcal{D}}\ . \tag{18}$$

The proof follows step by step the proof of Proposition 3 in [6], to which we refer.

4 The adaptive algorithm *hp*-ADFEM

As anticipated in the Introduction, the algorithm we propose consists in alternating between a stage in which a new *hp*-partition is found, which is near-optimal for the current accuracy, and a stage in which this partition is further refined to guarantee a higher accuracy for the corresponding DG discrete solution; the data used in the latter stage to define the DG problem are approximations of the exact data, provided by the former stage.

The first stage will be accomplished by a call to the routine **hp-NEARBEST** introduced in Sect. 2.3. The second stage will be realized through a routine **DG-SOLVE** that we present now, postponing to Sect. 5 the detailed description of the underlying algorithm and the analysis of its properties. Essentially, starting from a given *hp*-partition and a corresponding data approximation, several DG problems are solved on subsequently refined partitions, whose generation is driven by an a posteriori error estimator, until a contraction property guarantees that the discretization error is brought below a prescribed threshold. In this stage, optimality is not an issue for the output partition, provided its cardinality remains comparable to that of the input partition.

- $[\bar{\mathcal{D}}, \bar{u}]$:=**DG-SOLVE** $(\varepsilon, \mathcal{D}, z_{\mathcal{D}})$

  The routine **DG-SOLVE** takes as input $\varepsilon > 0$, $\mathcal{D} \in \mathbb{D}$, and $z_{\mathcal{D}} = (v_{\mathcal{D}}, g_{\mathcal{D}}) \in Z_{\mathcal{D}}$. It outputs $\bar{\mathcal{D}} \in \mathbb{D}$ with $\mathcal{D} \leq \bar{\mathcal{D}}$ and $\bar{u} := u_{\bar{\mathcal{D}}}(g_{\mathcal{D}}) \in V_{\bar{\mathcal{D}}}$ such that $\|u(g_{\mathcal{D}}) - \bar{u}\|_{\bar{\mathcal{D}}} \leq \varepsilon$.

We recall that $u_{\bar{\mathcal{D}}}(g_{\mathcal{D}})$ denotes the solution of the following DG problem (see (14)): for $g_{\mathcal{D}} = (\nu_{\mathcal{D}}, \xi_{\mathcal{D}}, f_{\mathcal{D}}) \in G_{\mathcal{D}}$,

$$u_{\bar{\mathcal{D}}} \in V_{\bar{\mathcal{D}}} \quad : \quad a_{\bar{\mathcal{D}}}(u_{\bar{\mathcal{D}}}, v_{\bar{\mathcal{D}}}; \nu_{\mathcal{D}}, \xi_{\mathcal{D}}) = (f_{\mathcal{D}}, v_{\bar{\mathcal{D}}})_{\Omega} \qquad \forall v_{\bar{\mathcal{D}}} \in V_{\bar{\mathcal{D}}}. \tag{19}$$

The input function $v_{\mathcal{D}} \in V_{\mathcal{D}}$ may be used in the algorithm to define the starting point of the adaptive iterations.

Assumption 4.1. *Let $b < 1 < B$ be the constants that appear in the statement of the instance optimality property for the routine* **hp-NEARBEST**. *We assume that the penalization parameter κ in* (15) *is chosen small enough, so that it holds*

$$C_\star \kappa < b.$$

We are ready to present our algorithm **hp-ADFEM**. Let us introduce the parameters and the input data.

Parameters: two real numbers $\eta \in (0,1)$, $\omega > 0$ satisfying

$$C_\star \kappa < b(1-\eta) \qquad \text{and} \qquad \omega \in \left(\frac{1}{b}, \frac{1-\eta}{C_\star \kappa}\right).$$

(Note that such a choice of ω is equivalent to $b\omega - 1 > 0$ and $C_\star \kappa \omega + \eta < 1$, which are two quantities that will appear below.)

Input data: $g_\star \in G(\Omega)$, $\varepsilon_0 > 0$, and $\bar{u}_0 \in V_{\bar{\mathcal{D}}_0}$ for some $\bar{\mathcal{D}}_0 \in \mathbb{D}$ such that $\|u_\star - \bar{u}_0\|_{\bar{\mathcal{D}}_0} \leq \varepsilon_0$.

Algorithm hp-ADFEM$(\varepsilon_0, \bar{u}_0, g_\star)$

```
for i = 1,2,... do
   [D_i, (v_{D_i}, g_{D_i})] := hp-NEARBEST(ω ε_{i-1}, (ū_{i-1}, g_⋆))
   [D̄_i, ū_i] := DG-SOLVE (η ε_{i-1}, D_i, (v_{D_i}, g_{D_i}))
   ε_i := (C_⋆ κ ω + η) ε_{i-1}
end do
```

Theorem 4.1. *Under Assumptions 3.1 and 4.1, the sequences* $(\bar{u}_i)$, $(\mathcal{D}_i)$ *produced by* **hp-ADFEM** *satisfy the following properties:*

$$\|u_\star - \bar{u}_i\|_{\bar{\mathcal{D}}_i} \le \varepsilon_i \quad \forall i \ge 0, \qquad \mathrm{E}_{\mathcal{D}_i}(u_\star, g_\star)^{\frac{1}{2}} \le (\omega+1)\varepsilon_{i-1} \quad \forall i \ge 1, \tag{20}$$

and

$$\#\mathcal{D}_i \le B\#\mathcal{D} \quad \textit{for any } \mathcal{D} \in \mathbb{D} \textit{ with } \mathrm{E}_{\mathcal{D}}(u_\star, g_\star)^{\frac{1}{2}} \le (b\omega - 1)\varepsilon_{i-1}. \tag{21}$$

Proof. The bound $\|u_\star - \bar{u}_0\|_{\bar{\mathcal{D}}_0} \le \varepsilon_0$ is valid by assumption. For $i \ge 1$, the tolerances used for **hp-NEARBEST** and **DG-SOLVE**, together with (17) show that

$$\begin{aligned} \|u_\star - \bar{u}_i\|_{\bar{\mathcal{D}}_i} &\le \|u_\star - u(g_{\mathcal{D}_i})\|_{H^1_0(\Omega)} + \|u(g_{\mathcal{D}_i}) - \bar{u}_i\|_{\bar{\mathcal{D}}_i} \\ &\le C_\star \kappa \, \mathrm{E}_{\mathcal{D}_i}(\bar{u}_{i-1}, g_\star)^{\frac{1}{2}} + \mu\varepsilon_{i-1} \le (C_\star\kappa\omega + \mu)\varepsilon_{i-1} = \varepsilon_i. \end{aligned} \tag{22}$$

The first statement follows for all $i \ge 0$. Using this and (18) implies the second assertion

$$\mathrm{E}_{\mathcal{D}_i}(u_\star, g_\star)^{\frac{1}{2}} \le \mathrm{E}_{\mathcal{D}_i}(\bar{u}_{i-1}, g_\star)^{\frac{1}{2}} + \|u_\star - \bar{u}_{i-1}\|_{\bar{\mathcal{D}}_{i-1}} \le (\omega+1)\varepsilon_{i-1} \qquad \forall i \ge 1.$$

Finally, let $\mathcal{D} \in \mathbb{D}$ with $\mathrm{E}_{\mathcal{D}}(u_\star, g_\star)^{\frac{1}{2}} \le (b\omega - 1)\varepsilon_{i-1}$. Then, again by (18), $\mathrm{E}_{\mathcal{D}}(\bar{u}_{i-1}, g_\star)^{\frac{1}{2}} \le b\omega\varepsilon_{i-1}$ and so $\#\mathcal{D}_i \le B\#\mathcal{D}$ because of the optimality property of **hp-NEARBEST**. ∎

The main result of Theorem 4.1 can be summarized by saying that **hp-ADFEM** is *instance optimal* for reducing $\mathrm{E}_{\mathcal{D}}(u_\star, g_\star)$ over $\mathcal{D} \in \mathbb{D}$.

5 The routine DG-SOLVE

The purpose of this section is the description and analysis of a realization of the routine **DG-SOLVE**. It is based on an iterative procedure of the form SOLVE → ESTIMATE → MARK → REFINE, in which ESTIMATE uses a residual-type estimator, whereas REFINE applies a dyadic splitting of each marked element while preserving the polynomial degree. The procedure satisfies a contraction property, which guarantees the reduction of a suitable "error" by a fixed amount at each iteration. Our construction is strongly inspired by [4], whose arguments are hereafter extended to cover the *hp*-case.

In the sequel, the input partition $\mathcal{D}$ will be denoted by $\mathcal{D}_{\text{in}}$, whereas the symbol $\mathcal{D}$ will be used to denote any refinement of $\mathcal{D}_{\text{in}}$ generated by the procedure. Similarly, the input function will be denoted by $z_{\text{in}} = (v_{\text{in}}, g_{\text{in}})$. To avoid cumbersome notation, we will actually write $g_{\text{in}} =: g = (\nu, \xi, f)$, but we will recall that g is a piecewise polynomial approximation on the input partition $\mathcal{D}_{\text{in}}$ of the given data $g_\star = (\nu_\star, \xi_\star, f_\star) \in G(\Omega)$. Coherently, the exact solution of Problem (8) with input data g will be denoted by $u = u(g)$, whereas for any *hp*-partition $\mathcal{D} \le \mathcal{D}_{\text{in}}$, $u_{\mathcal{D}} = u_{\mathcal{D}}(g)$ will be the solution of the corresponding DG Problem (14).

For the analysis of the procedure, following [3], we extend the definition of the DG form $a_{\mathcal{D}}$ given in (12) on $V_{\mathcal{D}} \times V_{\mathcal{D}}$ to the infinite dimensional space $V_{\mathcal{K}(\mathcal{D})} \times V_{\mathcal{K}(\mathcal{D})}$ (recall (9)). To this end, we introduce the lifting operator $L_{\mathcal{D}} : V_{\mathcal{K}(\mathcal{D})} \to V_{\mathcal{D}}$ such that for all $w \in V_{\mathcal{K}(\mathcal{D})}$

$$L_{\mathcal{D}} w \in V_{\mathcal{D}} \ : \ (\nu v, L_{\mathcal{D}} w)_\Omega = (\{\!\{\nu\, v\}\!\}, [\![w]\!])_{\mathcal{E}_{\mathcal{D}}} \quad \forall v \in V_{\mathcal{D}}. \tag{23}$$

Then, on $V_{\mathcal{K}(\mathcal{D})} \times V_{\mathcal{K}(\mathcal{D})}$ we define the bilinear form

$$\begin{aligned} a_{\mathcal{D}}(w, v) := (\nu\, \tilde{w}_x, \tilde{v}_x)_\Omega + (\xi\, w, v)_\Omega \\ - (\nu\, \tilde{w}_x, L_{\mathcal{D}} v)_{\mathcal{E}_{\mathcal{D}}} - (\nu\, \tilde{v}_x, L_{\mathcal{D}} w)_{\mathcal{E}_{\mathcal{D}}} + \gamma\, (\sigma_{\mathcal{D}} [\![w]\!], [\![v]\!])_{\mathcal{E}_{\mathcal{D}}}\ , \end{aligned} \tag{24}$$

which is readily seen to coincide with (12) on $V_{\mathcal{D}} \times V_{\mathcal{D}}$.

The lifting operator satisfies the following stability bound.

Property 5.1. *There exists a constant $C_1 > 0$ independent of $\mathcal{D}$ such that*

$$\|L_{\mathcal{D}} w\|_\Omega \leq C_1 \|\sigma_{\mathcal{D}}^{1/2} [\![w]\!]\|_{\mathcal{E}_{\mathcal{D}}} \qquad \forall w \in V_{\mathcal{K}(\mathcal{D})}. \tag{25}$$

Proof. If K is any interval of length h and e is one of its endpoints, the inverse inequality $|\phi(e)| \lesssim \frac{p}{h^{1/2}} \|\phi\|_K$ holds for any $\phi \in \mathbb{P}_p(K)$. Then, the result easily follows by choosing $v = L_{\mathcal{D}} w$ in (23). ◻

Using (25), one proves the existence of a constant $\gamma_0 > 0$ independent of $\mathcal{D}$ such that for any $\gamma \geq \gamma_0$ the bilinear form $a_{\mathcal{D}}$ is continuous and coercive in $V_{\mathcal{K}(\mathcal{D})}$ with respect to the DG norm $\|v\|_{\mathcal{D}}$, uniformly in $\mathcal{D}$. For future references, let us denote by $0 < \alpha_* \leq \alpha^*$ the coercivity and continuity constants. Since $a_{\mathcal{D}}$ is symmetric, it defines an inner product in $V_{\mathcal{K}(\mathcal{D})}$; the corresponding norm will be denoted by $\|v\|_{a,\mathcal{D}}$ and is uniformly equivalent to the DG norm $\|v\|_{\mathcal{D}}$ introduced in (13).

It is well-known that while the DG-solution $u_{\mathcal{D}} \in V_{\mathcal{D}}$ satisfies the variational equations

$$a_{\mathcal{D}}(u_{\mathcal{D}}, v_{\mathcal{D}}) = (f, v_{\mathcal{D}})_\Omega \qquad \forall v_{\mathcal{D}} \in V_{\mathcal{D}}, \tag{26}$$

the exact solution $u \in H_0^1(\Omega)$ need not satisfy $a_{\mathcal{D}}(u, v) = (f, v)_\Omega$ for *all* $v \in V_{\mathcal{K}(\mathcal{D})}$ (*inconsistency* of the DG formulation). However, we do have the partial consistency property

$$a_{\mathcal{D}}(u, v) = (f, v)_\Omega \qquad \forall v \in H_0^1(\Omega). \tag{27}$$

This motivates the introduction of the conforming subspace $V_{\mathcal{D}}^c := V_{\mathcal{D}} \cap H_0^1(\Omega)$. Then, by subtraction of (26) from (27), we obtain the *partial orthogonality* property

$$a_{\mathcal{D}}(u - u_{\mathcal{D}}, v_{\mathcal{D}}) = 0 \qquad \forall v_{\mathcal{D}} \in V_{\mathcal{D}}^c. \tag{28}$$

It is useful for the sequel to introduce the orthogonal decomposition

$$V_{\mathcal{D}} = V_{\mathcal{D}}^c \oplus V_{\mathcal{D}}^\perp, \tag{29}$$

where $V_{\mathcal{D}}^\perp$ is the orthogonal complement of $V_{\mathcal{D}}^c$ with respect to the inner product $a_{\mathcal{D}}(w, v)$. Any $v_{\mathcal{D}} \in V_{\mathcal{D}}$ will be split according to (29) as $v_{\mathcal{D}} = v_{\mathcal{D}}^c + v_{\mathcal{D}}^\perp$.

Property 5.2. *There exists a constant $C_2 > 0$ independent of $\mathcal{D}$ for which the following bound on the DG discretization error holds:*

$$\|u - u_{\mathcal{D}}\|_{\mathcal{D}} \leq C_2 \left(\inf_{w_{\mathcal{D}} \in V_{\mathcal{D}}^c} \|u - w_{\mathcal{D}}\|_{H_0^1(\Omega)} + \|u_{\mathcal{D}}^\perp\|_{\mathcal{D}} \right).$$

Proof. For any $w_{\mathcal{D}} \in V_{\mathcal{D}}^c$, using (28) we have

$$\begin{aligned} a_{\mathcal{D}}(u_{\mathcal{D}} - w_{\mathcal{D}}, u_{\mathcal{D}} - w_{\mathcal{D}}) &= a_{\mathcal{D}}(u_{\mathcal{D}} - w_{\mathcal{D}}, u_{\mathcal{D}}^c - w_{\mathcal{D}}) + a_{\mathcal{D}}(u_{\mathcal{D}} - w_{\mathcal{D}}, u_{\mathcal{D}}^\perp) \\ &= a_{\mathcal{D}}(u - w_{\mathcal{D}}, u_{\mathcal{D}}^c - w_{\mathcal{D}}) + a_{\mathcal{D}}(u_{\mathcal{D}}^\perp, u_{\mathcal{D}} - w_{\mathcal{D}}) \\ &= a_{\mathcal{D}}(u - w_{\mathcal{D}}, u_{\mathcal{D}} - w_{\mathcal{D}}) - a_{\mathcal{D}}(u_{\mathcal{D}}^\perp, u - u_{\mathcal{D}}), \end{aligned}$$

whence, by the coercivity and continuity of the form $a_{\mathcal{D}}$,

$$\|u_{\mathcal{D}} - w_{\mathcal{D}}\|_{\mathcal{D}}^2 \lesssim \|u - w_{\mathcal{D}}\|_{\mathcal{D}} \|u_{\mathcal{D}} - w_{\mathcal{D}}\|_{\mathcal{D}} + \|u_{\mathcal{D}}^\perp\|_{\mathcal{D}} \|u - u_{\mathcal{D}}\|_{\mathcal{D}}.$$

We conclude by the triangle inequality. ◻

We also introduce an approximation operator $\mathbb{I}_{\mathcal{D}} : V_{\mathcal{K}(\mathcal{D})} \to V_{\mathcal{D}}^c$ that will be useful in the sequel. For any $D \in \mathcal{D}$, set $K_D =: [e_l, e_r]$ and let $\mathcal{P}_D : H^1(K_D) \to \mathbb{P}_{p_D}(K_D)$ be defined as follows:

$$(\mathcal{P}_D v)(x) := v(e_l) + \int_{e_l}^x (\Pi^0_{K_D, p_D - 1} v_x)(s)\, ds$$

(recall that Π^0 means L^2-orthogonal projection). Furthermore, consider the Legendre Gauss-Lobatto grid in K_D containing p_D+1 nodes, and let ψ_{D,e_l} and ψ_{D,e_r} denote the Lagrange basis functions of degree p_D on this grid, associated with the boundary nodes. Then, we define $(\mathbb{I}_{\mathcal{D}}v)_{|K_D} := \mathcal{I}_D v_{|K_D}$, where

$$\mathcal{I}_D v := \mathcal{P}_D v - \tau_{e_l}[\![v]\!]_{e_l}\psi_{D,e_l} + \tau_{e_r}[\![v]\!]_{e_r}\psi_{D,e_r} \tag{30}$$

with $\tau_e = 1$ if $e \in \partial\Omega$, $\tau_e = \frac{1}{2}$ otherwise. Checking that $\mathbb{I}_{\mathcal{D}}v \in V^c_{\mathcal{D}}$ is straightforward.

Property 5.3. *The following error estimates hold for any $v \in H^1_0(\Omega)$:*

$$\|(v - \mathbb{I}_{\mathcal{D}}v)\omega_D^{-1/2}\|_{K_D} \leq \frac{1}{(p_D(p_D+1))^{1/2}}\|v_x\|_{K_D}, \qquad \|(\mathbb{I}_{\mathcal{D}}v)_x\|_{K_D} \leq \|v_x\|_{K_D}, \tag{31}$$

where ω_D is the quadratic bubble function in K_D, defined as $\omega_D(x) = (x-e_l)(e_r-x)$.

The following error estimates hold for any $v \in V_{\mathcal{D}}$:

$$\|v - \mathbb{I}_{\mathcal{D}}v\|_{K_D} \lesssim \frac{h_D^{1/2}}{p_D}([\![v]\!]_{e_l} + [\![v]\!]_{e_r}), \qquad \|(v - \mathbb{I}_{\mathcal{D}}v)_x\|_{K_D} \lesssim \frac{p_D}{h_D^{1/2}}([\![v]\!]_{e_l} + [\![v]\!]_{e_r}). \tag{32}$$

The latter inequality implies the bound

$$\|\tilde{v}_x - (\mathbb{I}_{\mathcal{D}}v)_x\|_\Omega \;\lesssim\; \|\sigma_{\mathcal{D}}^{1/2}[\![v]\!]\|_{\mathcal{E}_{\mathcal{D}}} \qquad \forall v \in V_{\mathcal{D}}. \tag{33}$$

Proof. The first inequality in (31) can be found in [10], whereas the second one is just the stability of the orthogonal projection. The inequalities (32) easily follow from the bounds $\|\psi_{D,e}\|_{K_D} \simeq \frac{h_D^{1/2}}{p_D}$ and $\|(\psi_{D,e})_x\|_{K_D} \lesssim \frac{p_D^2}{h_D}\|\psi_{D,e}\|_{K_D}$. ∎

Corollary 5.1. *There exists a constant $C_3 > 0$ independent of $\mathcal{D}$ such that for any $v = v^c \oplus v^\perp \in V_{\mathcal{D}} = V^c_{\mathcal{D}} \oplus V^\perp_{\mathcal{D}}$ one has*

$$\|v^\perp\|_{\mathcal{D}} \leq C_3\,\gamma^{1/2}\|\sigma_{\mathcal{D}}^{1/2}[\![v]\!]\|_{\mathcal{E}_{\mathcal{D}}}$$

Proof. One has

$$\|v^\perp\|_{\mathcal{D}} \simeq \|v^\perp\|_{a,\mathcal{D}} = \inf_{w \in V^c_{\mathcal{D}}}\|v - w\|_{a,\mathcal{D}} \simeq \inf_{w \in V^c_{\mathcal{D}}}\|v - w\|_{\mathcal{D}} \leq \|v - \mathbb{I}_{\mathcal{D}}v\|_{\mathcal{D}},$$

then one concludes by (33). ∎

5.1 The residual estimator

Given any $v \in V_{\mathcal{D}}$ and any $D \in \mathcal{D}$, let us define the local residual

$$\mathrm{res}_D(v) := (f - Av)_{|K_D};$$

for any $e \in \partial K_D$, let us define the jump of the flux at e

$$J_e(v) = [\![\,\nu v_x\,]\!]_e.$$

Then, the (squared) local error estimator is defined as follows

$$\eta_D^2(v) := \frac{1}{p_D(p_D+1)}\|\mathrm{res}_D(v)\,\omega_D^{1/2}\|^2_{K_D} + \sum_{e \in \partial K_D}\sigma^{-1}_{\mathcal{D},e}J^2_e(v),$$

where ω_D denotes the quadratic bubble function introduced in Property 5.3 above. The (squared) global error estimator is

$$\eta_{\mathcal{D}}^2(v) := \sum_{D\in\mathcal{D}} \eta_D^2(v),$$

whereas its restriction to a subset $\mathcal{D}' \subseteq \mathcal{D}$ of elements will be denoted by

$$\eta_{\mathcal{D}}^2(v;\mathcal{D}') := \sum_{D\in\mathcal{D}'} \eta_D^2(v).$$

We show that $\eta_{\mathcal{D}}(u_{\mathcal{D}})$ is a reliable estimator for our DG problem in two steps.

Proposition 5.1. *There exists a constant $C_4 > 0$ independent of $\mathcal{D}$ such that*

$$a_{\mathcal{D}}(u-u_{\mathcal{D}}, u-u_{\mathcal{D}}) \le C_4 \left(\eta_{\mathcal{D}}^2(u_{\mathcal{D}}) + \gamma \, \|\sigma_{\mathcal{D}}^{1/2} [\![u_{\mathcal{D}}]\!] \|_{\mathcal{E}_{\mathcal{D}}}^2\right).$$

Proof. We adapt the proof of [4], Lemma 3.1, to our *hp* setting. Let us split the DG solution as $u_{\mathcal{D}} = u_{\mathcal{D}}^c + u_{\mathcal{D}}^{\perp}$ and let us set $e := u - u_{\mathcal{D}}$ and $w := u - u_{\mathcal{D}}^c \in H_0^1(\Omega)$, so that $e = w - u_{\mathcal{D}}^{\perp}$. Then, recalling (27) and (28),

$$\begin{aligned} a_{\mathcal{D}}(e,e) &= a_{\mathcal{D}}(e,w) - a_{\mathcal{D}}(e,u_{\mathcal{D}}^{\perp}) = a_{\mathcal{D}}(e, w-\mathbb{I}_{\mathcal{D}}w) - a_{\mathcal{D}}(e,u_{\mathcal{D}}^{\perp}) \\ &= (f, w-\mathbb{I}_{\mathcal{D}}w)_{\Omega} - a_{\mathcal{D}}(u_{\mathcal{D}}, w-\mathbb{I}_{\mathcal{D}}w) - a_{\mathcal{D}}(e,u_{\mathcal{D}}^{\perp}), \end{aligned}$$

Integrating back by parts, we get

$$a_{\mathcal{D}}(u_{\mathcal{D}}, w-\mathbb{I}_{\mathcal{D}}w) = \sum_{D\in\mathcal{D}} (Au_{\mathcal{D}}, w-\mathbb{I}_{\mathcal{D}}w)_{K_D} + (L_{\mathcal{D}}u_{\mathcal{D}}, \nu(w-\mathbb{I}_{\mathcal{D}}w)_x)_{\Omega},$$

whence

$$a_{\mathcal{D}}(e,w) = \sum_{D\in\mathcal{D}} (\mathrm{res}_D(u_{\mathcal{D}}), w-\mathbb{I}_{\mathcal{D}}w)_{K_D} + (L_{\mathcal{D}}u_{\mathcal{D}}, \nu(w-\mathbb{I}_{\mathcal{D}}w)_x)_{\Omega}.$$

Writing $(\mathrm{res}_D(u_{\mathcal{D}}), w-\mathbb{I}_{\mathcal{D}}w)_{K_D} = (\mathrm{res}_D(u_{\mathcal{D}})\omega_D^{1/2}, (w-\mathbb{I}_{\mathcal{D}}w)\omega_D^{-1/2})_{K_D}$ and using (31) as well as (25), we obtain

$$a_{\mathcal{D}}(e,w) \le (\eta_{\mathcal{D}}(u_{\mathcal{D}}) + C_1 \|\sigma_{\mathcal{D}}^{1/2} [\![u_{\mathcal{D}}]\!] \|_{\mathcal{E}_{\mathcal{D}}}) \|w_x\|_{\Omega},$$

where the last norm can be bounded using the coercivity of the form $a_{\mathcal{D}}$:

$$\|w_x\|_{\Omega} = \|w\|_{\mathcal{D}} \le \|e\|_{\mathcal{D}} + \|u_{\mathcal{D}}^{\perp}\|_{\mathcal{D}} \le \alpha_*^{1/2} a_{\mathcal{D}}(e,e)^{1/2} + \|u_{\mathcal{D}}^{\perp}\|_{\mathcal{D}}.$$

By Young's inequality, we obtain for a suitable constant $C > 0$

$$a_{\mathcal{D}}(e,w) \le \frac{1}{4} a_{\mathcal{D}}(e,e) + C\big(\eta_{\mathcal{D}}^2(u_{\mathcal{D}}) + \|u_{\mathcal{D}}^{\perp}\|_{\mathcal{D}}^2 + \|\sigma_{\mathcal{D}}^{1/2} [\![u_{\mathcal{D}}]\!] \|_{\mathcal{E}_{\mathcal{D}}}^2\big).$$

It remains to bound the term $a_{\mathcal{D}}(e,u_{\mathcal{D}}^{\perp})$, which is easily done using the continuity of $a_{\mathcal{D}}$:

$$a_{\mathcal{D}}(e,u_{\mathcal{D}}^{\perp}) \le a_{\mathcal{D}}(e,e)^{1/2} a_{\mathcal{D}}(u_{\mathcal{D}}^{\perp},u_{\mathcal{D}}^{\perp})^{1/2} \le a_{\mathcal{D}}(e,e)^{1/2} (\alpha^*)^{1/2} \|u_{\mathcal{D}}^{\perp}\|_{\mathcal{D}} \le \frac{1}{4} a_{\mathcal{D}}(e,e) + \alpha^* \|u_{\mathcal{D}}^{\perp}\|_{\mathcal{D}}^2.$$

We obtain the desired result by invoking Corollary 5.1. ∎

Proposition 5.2. *There exists a constant $C_5 > 0$ independent of $\mathcal{D}$ such that for any γ large enough, say $\gamma \geq \gamma_1 \geq \gamma_0$, one has*

$$\gamma \, \|\sigma_{\mathcal{D}}^{1/2} [\![u_{\mathcal{D}}]\!]\|_{\mathcal{E}_{\mathcal{D}}} \leq C_5 \eta_{\mathcal{D}}(u_{\mathcal{D}}).$$

Proof. Here, we adapt the proof of [4], Lemma 3.3, to our *hp* setting. By the coercivity of the form $a_{\mathcal{D}}$ applied to $u_{\mathcal{D}} - \mathbb{I}_{\mathcal{D}} u_{\mathcal{D}}$, we have

$$\gamma \, \|\sigma_{\mathcal{D}}^{1/2} [\![u_{\mathcal{D}}]\!]\|^2_{\mathcal{E}_{\mathcal{D}}} \leq \alpha_*^{-1} a_{\mathcal{D}}(u_{\mathcal{D}} - \mathbb{I}_{\mathcal{D}} u_{\mathcal{D}}, u_{\mathcal{D}} - \mathbb{I}_{\mathcal{D}} u_{\mathcal{D}}) \tag{34}$$

since $[\![\mathbb{I}_{\mathcal{D}} u_{\mathcal{D}}]\!] = 0$. For simplicity, let us set $w := u_{\mathcal{D}} - \mathbb{I}_{\mathcal{D}} u_{\mathcal{D}}$ and $v := \mathbb{I}_{\mathcal{D}} u_{\mathcal{D}} \in H_0^1(\Omega)$. Then,

$$a_{\mathcal{D}}(w, w) = (f, w)_\Omega - a_{\mathcal{D}}(v, w)$$

and, using $L_{\mathcal{D}} v = 0$ several times, we have

$$\begin{aligned} a_{\mathcal{D}}(v, w) &= (\nu v_x, \tilde{w}_x)_\Omega + (\xi v, w)_\Omega - (L_{\mathcal{D}} u_{\mathcal{D}}, \nu v_x)_\Omega \\ &= (\nu \tilde{u}_{\mathcal{D},x}, \tilde{w}_x)_\Omega + (\xi u_{\mathcal{D}}, w)_\Omega - \|\nu^{1/2} \tilde{w}_x\|^2_\Omega - \|\xi^{1/2} w\|^2_\Omega - (L_{\mathcal{D}} u_{\mathcal{D}}, \nu v_x)_\Omega. \end{aligned}$$

Using in this identity

$$(\nu v_x, \tilde{w}_x)_\Omega = -\sum_{D \in \mathcal{D}} ((\nu u_{\mathcal{D},x})_x, w)_{K_D} + ([\![\nu u_{\mathcal{D},x}]\!], \{\!\{w\}\!\})_{\mathcal{E}_{\mathcal{D}}} + ([\![w]\!], \{\!\{\nu u_{\mathcal{D},x}\}\!\})_{\mathcal{E}_{\mathcal{D}}}$$

and observing that $([\![w]\!], \{\!\{\nu u_{\mathcal{D},x}\}\!\})_{\mathcal{E}_{\mathcal{D}}} = (L_{\mathcal{D}} w, \nu \tilde{u}_{\mathcal{D},x})_\Omega$, we obtain

$$\begin{aligned} a_{\mathcal{D}}(w, w) = \sum_{D \in \mathcal{D}} (\mathrm{res}_D(u_{\mathcal{D}}), w)_{K_D} + (J_{\mathcal{D}}(u_{\mathcal{D}}), \{\!\{w\}\!\})_{\mathcal{E}_{\mathcal{D}}} \\ + \|\nu^{1/2} \tilde{w}_x\|^2_\Omega + \|\xi^{1/2} w\|^2_\Omega + (L_{\mathcal{D}} u_{\mathcal{D}}, \nu \tilde{w}_x)_\Omega. \end{aligned} \tag{35}$$

By (32) we have

$$\|w\|_{K_D} \leq \sum_{e \in \partial K_D} \frac{h_D^{1/2}}{p_D} |\, [\![u_{\mathcal{D}}]\!]_e \,| = \sum_{e \in \partial K_D} \frac{h_D^{1/2}}{p_D} \sigma_{\mathcal{D},e}^{-1/2} \sigma_{\mathcal{D},e}^{1/2} |\, [\![u_{\mathcal{D}}]\!]_e \,| \leq \frac{h_D}{p_D^2} \sum_{e \in \partial K_D} \sigma_{\mathcal{D},e}^{1/2} |\, [\![u_{\mathcal{D}}]\!]_e \,|,$$

whence

$$\begin{aligned} (\mathrm{res}_D(u_{\mathcal{D}}), w)_{K_D} &\leq \frac{h_D}{p_D^2} \|\mathrm{res}_D(u_{\mathcal{D}})\|_{K_D} \sum_{e \in \partial K_D} \sigma_{\mathcal{D},e}^{1/2} |\, [\![u_{\mathcal{D}}]\!]_e \,| \\ &\lesssim \frac{1}{p_D} \|\mathrm{res}_D(u_{\mathcal{D}}) \omega_D^{1/2}\|_{K_D} \sum_{e \in \partial K_D} \sigma_{\mathcal{D},e}^{1/2} |\, [\![u_{\mathcal{D}}]\!]_e \,| \leq \eta_D(u_{\mathcal{D}}) \sum_{e \in \partial K_D} \sigma_{\mathcal{D},e}^{1/2} |\, [\![u_{\mathcal{D}}]\!]_e \,|, \end{aligned}$$

where we have used the inverse inequality $\|\phi\|_{K_D} \lesssim \frac{p_D}{h_D} \|\phi \, \omega_D^{1/2}\|_{K_D}$ which holds for all polynomials of degree $\simeq p_D$, since $\mathrm{res}_D(u_{\mathcal{D}})$ is such a polynomial. Thus, we obtain

$$\sum_{D \in \mathcal{D}} (\mathrm{res}_D(u_{\mathcal{D}}), w)_{K_D} \lesssim \eta_{\mathcal{D}}(u_{\mathcal{D}}) \, \|\sigma_{\mathcal{D}}^{1/2} [\![u_{\mathcal{D}}]\!]\|_{\mathcal{E}_{\mathcal{D}}}.$$

Concerning the second term on the right-hand side of (35), we observe that by construction of $\mathbb{I}_{\mathcal{D}} u_{\mathcal{D}}$, one has $w(e) = \frac{1}{2} [\![u_{\mathcal{D}}]\!]_e$ at any internal inter-element point e, whereas $w(e) = 0$ at the boundary points of Ω. Thus,

$$\begin{aligned} (J_{\mathcal{D}}(u_{\mathcal{D}}), \{\!\{w\}\!\})_{\mathcal{E}_{\mathcal{D}}} &\lesssim \sum_{e \in \mathcal{E}_{\mathcal{D}}} |J_e(u_{\mathcal{D}})| \, |\, [\![u_{\mathcal{D}}]\!]_e \,| = \sum_{e \in \mathcal{E}_{\mathcal{D}}} \sigma_{\mathcal{D},e}^{-1/2} |J_e(u_{\mathcal{D}})| \, \sigma_{\mathcal{D},e}^{1/2} |\, [\![u_{\mathcal{D}}]\!]_e \,| \\ &\leq \eta_{\mathcal{D}}(u_{\mathcal{D}}) \, \|\sigma_{\mathcal{D}}^{1/2} [\![u_{\mathcal{D}}]\!]\|_{\mathcal{E}_{\mathcal{D}}}. \end{aligned}$$

Finally, using (32) and (25), the three last terms on the right-hand side of (35) can be bounded by $C\|\sigma_{\mathcal{D}}^{1/2}[\![u_{\mathcal{D}}]\!]\|_{\mathcal{E}_{\mathcal{D}}}^2$. Substituting all the previous bounds in (34), we obtain

$$\gamma\,\|\sigma_{\mathcal{D}}^{1/2}[\![u_{\mathcal{D}}]\!]\|_{\mathcal{E}_{\mathcal{D}}}^2 \lesssim \big(\eta_{\mathcal{D}}(u_{\mathcal{D}})\,\|\sigma_{\mathcal{D}}^{1/2}[\![u_{\mathcal{D}}]\!]\|_{\mathcal{E}_{\mathcal{D}}} + \|\sigma_{\mathcal{D}}^{1/2}[\![u_{\mathcal{D}}]\!]\|_{\mathcal{E}_{\mathcal{D}}}^2\big),$$

where the constant implied by the symbol $\lesssim$ is independent of γ. Therefore, choosing γ large enough, we get the desired result. ∎

Corollary 5.2. *There exists a constant $C_6 > 0$ independent of $\mathcal{D}$ such that for any $\gamma \geq \gamma_1$, one has*

$$a_{\mathcal{D}}(u - u_{\mathcal{D}}, u - u_{\mathcal{D}}) \leq C_6\,\eta_{\mathcal{D}}^2(u_{\mathcal{D}}). \qquad ∎$$

5.2 The adaptive iterations

The routine **DG-SOLVE** iterates the mapping

$$(\mathcal{D}, u_{\mathcal{D}}, \eta_{\mathcal{D}}(u_{\mathcal{D}})) \;\to\; (\mathcal{D}_*, u_{\mathcal{D}_*}, \eta_{\mathcal{D}_*}(u_{\mathcal{D}_*}))\,, \tag{36}$$

where $\mathcal{D}_*$ is a refinement of $\mathcal{D}$ obtained by first applying a Dörfler marking to the elements of $\mathcal{D}$ based on the error estimator $\eta_{\mathcal{D}}(u_{\mathcal{D}})$, and then performing a dyadic subdivision to the marked elements and its neighbors.

To be precise, let $\vartheta \in (0,1)$ be the Dörfler parameter. Let us order the local error estimators $\eta_D(u_{\mathcal{D}})$, $D \in \mathcal{D}$, by decreasing value, and let us choose a set $\mathcal{M} \subseteq \mathcal{D}$ of minimal cardinality for which

$$\eta_{\mathcal{D}}(u_{\mathcal{D}}; \mathcal{M}) \geq \vartheta\,\eta_{\mathcal{D}}(u_{\mathcal{D}}). \tag{37}$$

Let $\partial\mathcal{M} \subseteq \mathcal{D}$ denote the set of elements D that share an interface with an element in $\mathcal{M}$. Then, we replace each $D = (K_D, p_D) \in \mathcal{M} \cup \partial\mathcal{M}$ by the two elements $D' = (K'_D, p_D)$ and $D'' = (K''_D, p_D)$, where K'_D and K''_D are the two children of K_D. Thus, the new partition $\mathcal{D}_*$ is defined by

$$\mathcal{D}_* = \{D', D'' : D \in \mathcal{M} \cup \partial\mathcal{M}\} \;\cup\; \{D : D \in \mathcal{D} \setminus (\mathcal{M} \cup \partial\mathcal{M})\}. \tag{38}$$

Our aim is to prove that a suitable combination of (squared) DG error and error estimator, i.e.,

$$\|u - u_{\mathcal{D}}\|_{a,\mathcal{D}}^2 + \beta\,\eta_{\mathcal{D}}^2(u_{\mathcal{D}})$$

for some $\beta > 0$, is reduced by a fixed rate $\varrho \in (0,1)$ in performing the mapping (36). The proof, which extends [4] to our *hp*-setting, will be based on the following results.

Lemma 5.1. *There exists a constant $C_7 > 0$ independent of $\mathcal{D}$ such that for any real $\lambda \in (0,1)$, one has*

$$\eta_{\mathcal{D}_*}^2(u_{\mathcal{D}_*}) \leq (1+\lambda)(1 - \frac{\vartheta^2}{2})\,\eta_{\mathcal{D}}^2(u_{\mathcal{D}}) + \frac{C_7}{\lambda}\|u_{\mathcal{D}_*} - u_{\mathcal{D}}\|_{\mathcal{D}_*}^2.$$

Proof. We first establish a few results about the Lipschitz continuity of the local error estimators. Assume that $v, w \in V_{\mathcal{D}}$ and let $D \in \mathcal{D}$. By Minkowski's inequality,

$$|\eta_D(v) - \eta_D(w)| \leq \left(\frac{1}{p_D^2}\|(\mathrm{res}_D(v) - \mathrm{res}_D(w))\,\omega_D^{1/2}\|_{K_D}^2 + \sum_{e \in \partial K_D} \sigma_{\mathcal{D},e}^{-1}|J_e(v) - J_e(w)|^2\right)^{1/2}.$$

One has

$$\begin{aligned}\|(\mathrm{res}_D(v)-\mathrm{res}_D(w))\,\omega_D^{1/2}\|_{K_D} &\leq \|(\nu(v-w)_x)_x\,\omega_D^{1/2}\|_{K_D}+\|\xi(v-w)\,\omega_D^{1/2}\|_{K_D}\\ &\lesssim p_D\|\nu(v-w)_x\|_{K_D}+h_D\|\xi(v-w)\|_{K_D}\\ &\lesssim p_D\|(v-w)_x\|_{K_D}+h_D\|(v-w)\|_{K_D},\end{aligned}$$

where we have used the inverse inequality $\|\phi_x\,\omega_D^{1/2}\|_{K_D}\lesssim p_D\|\phi\|_{K_D}$, which holds for all polynomial ϕ of degree $\simeq p_D$ in K_D, as well as the bound $\|\omega_D^{1/2}\|_{L^\infty(K_D)}\leq h_D$.

On the other hand, for each $e\in\partial K_D$, let us denote by D' the element in $\mathcal{D}$ sharing the interface e with D. Then,

$$\begin{aligned}|J_e(v)-J_e(w)| &\leq |\nu(v-w)_{x|K_D}(e)|+|\nu(v-w)_{x|K_{D'}}(e)|\\ &\lesssim |(v-w)_{x|K_D}(e)|+|(v-w)_{x|K_{D'}}(e)|\\ &\lesssim \frac{p_D}{h_D^{1/2}}\|(v-w)_x\|_{K_D}+\frac{p_{D'}}{h_{D'}^{1/2}}\|(v-w)_x\|_{K_{D'}}\\ &\leq \sigma_{\mathcal{D},e}^{1/2}\left(\|(v-w)_x\|_{K_D}+\|(v-w)_x\|_{K_{D'}}\right),\end{aligned}$$

where we have used the inverse inequality $|\psi(e)|\lesssim\frac{p_D}{h_D^{1/2}}\|\psi\|_{K_D}$, which holds for all polynomial ψ of degree $\simeq p_D$ in K_D. We conclude that

$$|\eta_D(v)-\eta_D(w)|\lesssim\mathcal{N}_D(v-w),\quad\text{with}\ \ \mathcal{N}_D^2(\phi):=\sum_{D'}\|\phi_x\|_{K_D}^2+\frac{h_D^2}{p_D^2}\|\phi\|_{K_D}^2,$$

where summation is extended to all $D'\in\mathcal{D}$ such that $K_{D'}\cap K_D$ is nonempty; this implies

$$\eta_D^2(v)\leq(1+\lambda)\,\eta_D^2(w)+\frac{C}{\lambda}\mathcal{N}_D^2(v-w)\tag{39}$$

for a suitable constant $C>0$ independent of D.

We now apply these bounds, with $v=u_{\mathcal{D}_*}$ and $w=u_{\mathcal{D}}$, to the partition (38) generated by the refinement procedure. If $D\in\mathcal{M}$, let D_m, $m=1,2$ be the two children in which D is split. We have $\omega_{D_m}(x)\leq\frac{1}{2}\omega_D(x)$ for all $x\in D_m$. By definition of refinement, we have $h_{D_m}=\frac{1}{2}h_D$ as well as $h_{D'_m}=\frac{1}{2}h_{D'}$ for any neighborhood $D'\in\mathcal{D}$ of D, which implies $\sigma^{-1}_{\mathcal{D}_*,e}\leq\frac{1}{2}\sigma^{-1}_{\mathcal{D},e}$ for any $e\in\partial K_D$. Hence, we immediately have $\sum_{m=1}^2\eta^2_{D_m}(u_{\mathcal{D}})\leq\frac{1}{2}\eta_D^2(u_{\mathcal{D}})$ and $\sum_{m=1}^2\mathcal{N}_{D_m}(u_{\mathcal{D}_*}-u_{\mathcal{D}})\leq\mathcal{N}_D(u_{\mathcal{D}_*}-u_{\mathcal{D}})$, whence

$$\sum_{m=1}^2\eta^2_{D_m}(u_{\mathcal{D}_*})\leq\frac{1}{2}(1+\lambda)\,\eta_D^2(u_{\mathcal{D}})+\frac{C}{\lambda}\mathcal{N}_D^2(u_{\mathcal{D}_*}-u_{\mathcal{D}}).$$

If $D\in\partial\mathcal{M}$, we can only say that $\sigma^{-1}_{\mathcal{D}_*,e}\leq\sigma^{-1}_{\mathcal{D},e}$ for any $e\in\partial K_D$, whence

$$\sum_{m=1}^2\eta^2_{D_m}(u_{\mathcal{D}_*})\leq(1+\lambda)\,\eta_D^2(u_{\mathcal{D}})+\frac{C}{\lambda}\mathcal{N}_D^2(u_{\mathcal{D}_*}-u_{\mathcal{D}}).$$

Finally, for any unsplit $D\in\mathcal{D}\setminus(\mathcal{M}\cup\partial\mathcal{M})$, we just have

$$\eta_D^2(u_{\mathcal{D}_*})\leq(1+\lambda)\,\eta_D^2(u_{\mathcal{D}})+\frac{C}{\lambda}\mathcal{N}_D^2(u_{\mathcal{D}_*}-u_{\mathcal{D}}).$$

Summing-up all contributions and using the marking condition, we obtain

$$\begin{aligned}\eta^2_{\mathcal{D}_*}(u_{\mathcal{D}_*}) &\leq (1+\lambda)\left(\eta^2_{\mathcal{D}}(u_{\mathcal{D}})-\frac{1}{2}\eta^2_{\mathcal{D}}(u_{\mathcal{D}};\mathcal{M})\right)+\frac{C}{\lambda}\sum_{D\in\mathcal{D}}\mathcal{N}_D^2(u_{\mathcal{D}_*}-u_{\mathcal{D}})\\ &\leq (1+\lambda)(1-\frac{\vartheta^2}{2})\eta^2_{\mathcal{D}}(u_{\mathcal{D}})+\frac{C}{\lambda}\sum_{D\in\mathcal{D}}\mathcal{N}_D^2(u_{\mathcal{D}_*}-u_{\mathcal{D}}).\end{aligned}$$

It remains to prove that $\sum_{D\in\mathcal{D}} \mathcal{N}_D^2(u_{\mathcal{D}_*}-u_{\mathcal{D}}) \lesssim \|u_{\mathcal{D}_*}-u_{\mathcal{D}}\|^2_{\mathcal{D}_*}$. Setting now $w := u_{\mathcal{D}_*}-u_{\mathcal{D}}$, we have

$$\sum_{D\in\mathcal{D}} \mathcal{N}_D^2(w) = \|\tilde{w}_x\|^2_\Omega + \sum_{D\in\mathcal{D}} \frac{h_D^2}{p_D^2}\|w\|^2_{K_D}.$$

Writing, for a.e. $x\in\Omega$,

$$w(x) = \sum_{e\in\mathcal{E}_{\mathcal{D}_*}, e<x} [\![w]\!]_e + \int_{\min\Omega}^x \tilde{w}_x(s)\,ds = \sum_{e\in\mathcal{E}_{\mathcal{D}_*}, e<x} \sigma_{\mathcal{D}_*,e}^{-1/2}\sigma_{\mathcal{D}_*,e}^{-1/2}[\![w]\!]_e + \int_{\min\Omega}^x \tilde{w}_x(s)\,ds\,,$$

we have

$$w^2(x) \lesssim \Big(\sum_{e\in\mathcal{E}_{\mathcal{D}_*}} \sigma^{-1}_{\mathcal{D}_*,e}\Big) \sum_{e\in\mathcal{E}_{\mathcal{D}_*}} \sigma_{\mathcal{D}_*,e}[\![w]\!]^2_e \;+\; |\Omega|\,\|\tilde{w}_x\|^2_\Omega.$$

Since $\sum_{e\in\mathcal{E}_{\mathcal{D}_*}} \sigma^{-1}_{\mathcal{D}_*,e} \le |\Omega|$, we easily obtain the desired bound. ∎

Lemma 5.2. *There exists a constant $C_8>0$ independent of $\mathcal{D}$ such that for any real $\delta\in(0,1)$ and any $\gamma\ge\gamma_1$, one has*

$$\|u-u_{\mathcal{D}_*}\|^2_{a,\mathcal{D}_*} \le (1+\delta)\|u-u_{\mathcal{D}}\|^2_{a,\mathcal{D}} - \frac{\alpha_*}{2}\|u_{\mathcal{D}_*}-u_{\mathcal{D}}\|^2_{\mathcal{D}_*} + \frac{C_8}{\delta\gamma}\left(\eta^2_{\mathcal{D}_*}(u_{\mathcal{D}_*}) + \eta^2_{\mathcal{D}}(u_{\mathcal{D}})\right).$$

Proof. Let us set $w_* := u-u_{\mathcal{D}_*}$, $w := u-u_{\mathcal{D}}$, $d := u_{\mathcal{D}_*}-u_{\mathcal{D}}$, $d^c := u^c_{\mathcal{D}_*}-u^c_{\mathcal{D}}$ and $d^\perp := u^\perp_{\mathcal{D}_*}-u^\perp_{\mathcal{D}}$. Observing that $a_{\mathcal{D}_*}(w_*,d^c)=0$ by the partial orthogonality property (28), one easily gets

$$\|w_*\|^2_{a,\mathcal{D}_*} = a_{\mathcal{D}_*}(w_*,w_*) = a_{\mathcal{D}_*}(w_*+d^c,w_*+d^c) - a_{\mathcal{D}_*}(d^c,d^c).$$

Using $u_{\mathcal{D}} = u^c_{\mathcal{D}} + u^\perp_{\mathcal{D}}$ and $u_{\mathcal{D}_*} = u^c_{\mathcal{D}_*} + u^\perp_{\mathcal{D}_*}$, one has $w_*+d^c = w-d^\perp$, whence

$$\begin{aligned} a_{\mathcal{D}_*}(w_*+d^c,w_*+d^c) &= a_{\mathcal{D}_*}(w,w) - 2a_{\mathcal{D}_*}(w,d^\perp) + a_{\mathcal{D}_*}(d^\perp,d^\perp) \\ &\le \|w\|^2_{a,\mathcal{D}_*} + 2(\alpha^*)^{1/2}\|w\|_{a,\mathcal{D}_*}\|d^\perp\|_{\mathcal{D}_*} + \alpha^*\|d^\perp\|^2_{\mathcal{D}_*}, \end{aligned}$$

where we have used the uniform continuity of the form $a_{\mathcal{D}_*}$ with respect to the DG-norm. Using the uniform coercivity and the triangle inequality, we get

$$a_{\mathcal{D}_*}(d^c,d^c) \ge \alpha_*\|d^c\|^2_{\mathcal{D}_*} \ge \alpha_*\left(\frac{1}{2}\|d\|^2_{\mathcal{D}_*} - \|d^\perp\|^2_{\mathcal{D}_*}\right).$$

Collecting these inequalities and using Young's inequality, we obtain

$$\|w_*\|^2_{a,\mathcal{D}_*} \le (1+\delta)\|w\|^2_{a,\mathcal{D}_*} - \frac{\alpha_*}{2}\|d\|^2_{\mathcal{D}_*} + \frac{C}{\delta}\|d^\perp\|^2_{\mathcal{D}_*}. \tag{40}$$

At this point, we observe that $\|u^\perp_{\mathcal{D}}\|^2_{\mathcal{D}_*} \le 2\|u^\perp_{\mathcal{D}}\|^2_{\mathcal{D}}$. Indeed, $\|u^\perp_{\mathcal{D}}\|^2_{\mathcal{D}_*} = \|(u^\perp_{\mathcal{D}})^\sim_x\|^2_\Omega + \gamma\sum_{e\in\mathcal{E}_{\mathcal{D}_*}}\sigma_{\mathcal{D}_*,e}[\![u^\perp_{\mathcal{D}}]\!]^2_e$, but the jumps of $u^\perp_{\mathcal{D}}$ occur only at the interfaces $e\in\mathcal{E}_{\mathcal{D}}$, and $\sigma_{\mathcal{D}_*,e}\le 2\sigma_{\mathcal{D},e}$ by definition of the refinement strategy. Thus, using Corollary 5.1, we get

$$\|d^\perp\|^2_{\mathcal{D}_*} \;\lesssim\; \|u^\perp_{\mathcal{D}_*}\|^2_{\mathcal{D}_*} + \|u^\perp_{\mathcal{D}}\|^2_{\mathcal{D}} \;\lesssim\; \gamma\|\sigma^{1/2}_{\mathcal{D}_*}[\![u_{\mathcal{D}_*}]\!]\|^2_{\mathcal{E}_{\mathcal{D}_*}} + \gamma\|\sigma^{1/2}_{\mathcal{D}}[\![u_{\mathcal{D}}]\!]\|^2_{\mathcal{E}_{\mathcal{D}}}. \tag{41}$$

It remains to replace $\|w\|^2_{a,\mathcal{D}_*}$ by $\|w\|^2_{a,\mathcal{D}}$. To this end, let us write

$$a_{\mathcal{D}_*}(w,w) = a_{\mathcal{D}}(w,w) + 2(L_{\mathcal{D}_*}w,\nu\tilde{w}_x)_\Omega - 2(L_{\mathcal{D}}w,\nu\tilde{w}_x)_\Omega - \gamma\|\sigma^{1/2}_{\mathcal{D}}[\![w]\!]\|^2_{\mathcal{E}_{\mathcal{D}}} + \gamma\|\sigma^{1/2}_{\mathcal{D}_*}[\![w]\!]\|^2_{\mathcal{E}_{\mathcal{D}_*}}.$$

Using Property 5.1 and the coercivity of the form $a_{\mathcal{D}}$, one gets

$$(L_{\mathcal{D}_*} w, \nu \tilde{w}_x)_\Omega \;\lesssim\; \|\sigma_{\mathcal{D}_*}^{1/2} [\![w]\!]\|_{\mathcal{E}_{\mathcal{D}_*}} a_{\mathcal{D}}(w,w)^{1/2} \;\lesssim\; \|\sigma_{\mathcal{D}}^{1/2} [\![u_{\mathcal{D}}]\!]\|_{\mathcal{E}_{\mathcal{D}}} a_{\mathcal{D}}(w,w)^{1/2}.$$

A similar bound holds for $(L_{\mathcal{D}} w, \nu \tilde{w}_x)_\Omega$. Therefore, using once more Young's inequality, we arrive at

$$\|w\|^2_{a,\mathcal{D}_*} \le (1+\delta)\|w\|^2_{a,\mathcal{D}} + \frac{C}{\delta}\gamma \|\sigma_{\mathcal{D}}^{1/2} [\![u_{\mathcal{D}}]\!]\|^2_{\mathcal{E}_{\mathcal{D}}}. \tag{42}$$

Replacing (41)-(42) into (40), we obtain

$$\begin{aligned} \|u - u_{\mathcal{D}_*}\|^2_{a,\mathcal{D}_*} \;&\le\; (1+\delta)^2 \|u - u_{\mathcal{D}}\|^2_{a,\mathcal{D}} - \frac{\alpha_*}{2}\|u_{\mathcal{D}_*} - u_{\mathcal{D}}\|^2_{\mathcal{D}_*} \\ &\quad + \frac{C}{\delta}\gamma \left(\|\sigma_{\mathcal{D}_*}^{1/2} [\![u_{\mathcal{D}_*}]\!]\|^2_{\mathcal{E}_{\mathcal{D}_*}} + \|\sigma_{\mathcal{D}}^{1/2} [\![u_{\mathcal{D}}]\!]\|^2_{\mathcal{E}_{\mathcal{D}}} \right). \end{aligned}$$

The desired result follows from Proposition 5.2, after replacing δ by $\delta/3$. ∎

We are ready to establish the main result of this section.

Theorem 5.1. *Consider the mapping* (36) *defined above. There exist constants $\beta > 0$ and $\varrho \in (0,1)$, independent of $\mathcal{D}$, such that, choosing $\gamma > 0$ large enough in the definition* (24)*, one has*

$$\|u - u_{\mathcal{D}_*}\|^2_{a,\mathcal{D}_*} + \beta\, \eta^2_{\mathcal{D}_*}(u_{\mathcal{D}_*}) \le \varrho \left(\|u - u_{\mathcal{D}}\|^2_{a,\mathcal{D}} + \beta\, \eta^2_{\mathcal{D}}(u_{\mathcal{D}}) \right).$$

Proof. Let us simplify our notation by setting $E_*^2 := \|u - u_{\mathcal{D}_*}\|^2_{a,\mathcal{D}_*}$, $E^2 := \|u - u_{\mathcal{D}}\|^2_{a,\mathcal{D}}$, $e_*^2 := \|u_{\mathcal{D}_*} - u_{\mathcal{D}}\|^2_{\mathcal{D}_*}$ and $\eta_*^2 := \eta^2_{\mathcal{D}_*}(u_{\mathcal{D}_*})$, $\eta^2 := \eta^2_{\mathcal{D}}(u_{\mathcal{D}})$. Then, the inequalities of Lemmas 5.2-5.1 read as follows:

$$\begin{aligned} E_*^2 \;&\le\; (1+\delta)E^2 - \frac{\alpha_*}{2} e_*^2 + \frac{C_8}{\delta\gamma}(\eta_*^2 + \eta^2) \\ \eta_*^2 \;&\le\; (1+\lambda)(1 - \frac{\vartheta^2}{2})\,\eta^2 + \frac{C_7}{\lambda} e_*^2. \end{aligned}$$

Thus, for any real $\beta > 0$,

$$\begin{aligned} E_*^2 + \beta\eta_*^2 \;&\le\; (1+\delta)E^2 - \frac{\alpha_*}{2} e_*^2 + \left(\beta + \frac{C_8}{\delta\gamma}\right)\eta_*^2 + \frac{C_8}{\delta\gamma}\eta^2 \\ &\le\; (1+\delta)E^2 - \frac{\alpha_*}{2} e_*^2 + \left(\beta + \frac{C_8}{\delta\gamma}\right)\left((1+\lambda)(1 - \frac{\vartheta^2}{2})\,\eta^2 + \frac{C_7}{\lambda} e_*^2\right) + \frac{C_8}{\delta\gamma}\eta^2. \end{aligned}$$

Writing $1 - \frac{\vartheta^2}{2} = \left(1 - \frac{\vartheta^2}{4}\right) - \frac{\vartheta^2}{4}$ and using $E^2 \le C_6 \eta^2$ from Corollary (5.2), we easily obtain for $\gamma \ge \gamma_1$

$$\begin{aligned} E_*^2 + \beta\eta_*^2 \;&\le\; \left[(1+\delta) - \left(\beta + \frac{C_8}{\delta\gamma}\right)\frac{1+\lambda}{C_6}\frac{\vartheta^2}{4}\right] E^2 \;+\; \left[\left(\beta + \frac{C_8}{\delta\gamma}\right)\frac{C_7}{\lambda} - \frac{\alpha_*}{2}\right] e_*^2 \\ &\quad + \left[(1+\lambda)\left(1 - \frac{\vartheta^2}{4}\right) + \frac{C_8}{\beta\delta\gamma}\left(1 + (1+\lambda)\left(1 - \frac{\vartheta^2}{4}\right)\right)\right]\beta\eta^2 \\ &=:\; \varrho_1 E^2 + \varrho_2\, e_*^2 + \varrho_3\, \beta\eta^2. \end{aligned}$$

At this point, we first choose λ sufficiently small to have $(1+\lambda)\left(1 - \frac{\vartheta^2}{4}\right) < 1$. Next, we choose δ sufficiently small to have $\varrho_1 < 1$ for $\gamma = \gamma_1$, hence for any $\gamma \ge \gamma_1$. Then, the parameter $\beta > 0$ is determined by imposing $\varrho_2 = 0$, which is possible provided γ is large enough, say $\gamma \ge \gamma_2 \ge \gamma_1$. Finally, for γ even larger, say $\gamma \ge \gamma_3 \ge \gamma_2$, the second addend in ϱ_3 can be made so small that $\varrho_3 < 1$. In conclusion, the desired result holds for all $\gamma \ge \gamma_3$ with $\varrho := \max(\varrho_1, \varrho_3)$. ∎

Corollary 5.3. *Denote by* $\{(\mathcal{D}_k, u_{\mathcal{D}_k}, \eta_{\mathcal{D}_k}(u_{\mathcal{D}_k})) : k \geq 0\}$ *the sequence produced by iterating the mapping* (36) *from the input partition* $\mathcal{D}_0 := \mathcal{D}_{in}$. *Then,*

$$\|u - u_{\mathcal{D}_k}\|^2_{\mathcal{D}_k} \leq \alpha_*^{-1} \varrho^k \big(\|u - u_{\mathcal{D}_0}\|^2_{a,\mathcal{D}_0} + \beta\, \eta^2_{\mathcal{D}_0}(u_{\mathcal{D}_0}) \big). \qquad \square$$

The latter result guarantees that the target accuracy $\|u - u_{\mathcal{D}_k}\|^2_{\mathcal{D}_k} \leq \varepsilon^2$ of **DG-SOLVE** can be matched provided the iterations are stopped at a sufficiently large k. In particular, if there exists a constant $C_9 > 0$ such that

$$\|u - u_{\mathcal{D}_0}\|^2_{a,\mathcal{D}_0} + \beta\, \eta^2_{\mathcal{D}_0}(u_{\mathcal{D}_0}) \;\leq\; C_9\, \varepsilon^2, \tag{43}$$

then the number K of iterations in **DG-SOLVE** is bounded independently of ε. In this case, since the mapping (36) at most doubles the cardinality of the partition, i.e., $|\mathcal{D}_*| \leq 2|\mathcal{D}|$, we conclude that the cardinality of the output partition $\mathcal{D}_{\text{out}} := \mathcal{D}_K$ is uniformly bounded by the cardinality of the input partition $\mathcal{D}_{\text{in}}$, precisely

$$|\mathcal{D}_{\text{out}}| \leq 2^K |\mathcal{D}_{\text{in}}|.$$

Remark 5.1. (*Arithmetic complexity*) According to [5], if $N := \#\mathcal{D}$ denotes the cardinality of the current *hp*-partition, the arithmetic complexity of **hp-NEARBEST** is $O(N^2)$ (or $O(N \log N)$ in some specific situations). On the other hand, **DG-SOLVE** performs a bounded numbers of solutions of DG problems, which can be achieved in linear complexity. $\square$

5.3 Initialization

Let us discuss a possible strategy to fulfill (43). Recall that we enter **DG-SOLVE** at iteration i of **hp-ADFEM** with input partition $\mathcal{D}_i$ and data $g_{\mathcal{D}_i}$. This means that, with the notation of **hp-ADFEM**, condition (43) reads

$$\|u(g_{\mathcal{D}_i}) - u_{\mathcal{D}_i}\|^2_{a,\mathcal{D}_i} + \beta\, \eta^2_{\mathcal{D}_i}(u_{\mathcal{D}_i}) \;\leq\; C_9\, \varepsilon_i^2. \tag{44}$$

The first term on the left-hand side can be bounded from above by using the uniform continuity of the form $a_{\mathcal{D}_i}$ and the bounds given in Property 5.2, Corollary 5.1 and Proposition 5.2. This yields

$$\|u(g_{\mathcal{D}_i}) - u_{\mathcal{D}_i}\|^2_{a,\mathcal{D}_i} + \beta\, \eta^2_{\mathcal{D}_i}(u_{\mathcal{D}_i}) \leq C_{10} \inf_{w_{\mathcal{D}_i} \in V^c_{\mathcal{D}_i}} \|u(g_{\mathcal{D}_i}) - w_{\mathcal{D}_i}\|^2_{H^1_0(\Omega)} + C_{11} \eta^2_{\mathcal{D}_i}(u_{\mathcal{D}_i})$$

for constants $C_{10}, C_{11} > 0$ independent of $\mathcal{D}_i$. We now show that the infimum on the right-hand side can be bounded by a multiple of ε_i^2.

Property 5.4. *There exists a constant* $C_{12} > 0$ *independent of* $\mathcal{D}_i$ *such that*

$$\inf_{w_{\mathcal{D}_i} \in V^c_{\mathcal{D}_i}} \|u(g_{\mathcal{D}_i}) - w_{\mathcal{D}_i}\|_{H^1_0(\Omega)} \leq C_{12} \varepsilon_i$$

Proof. For simplicity, set again $u := u(g_{\mathcal{D}_i})$. Then, for any $w_{\mathcal{D}_i} \in V^c_{\mathcal{D}_i}$, let us write $u - w_{\mathcal{D}_i} = (u - u_\star) + (u_\star - \bar{u}_{i-1}) + (\bar{u}_{i-1} - w_{\mathcal{D}_i})$. Using (17), we get

$$\|u - u_\star\|_{H^1_0(\Omega)} = \|u(g_\star) - u(g_{\mathcal{D}_i})\|_{H^1_0(\Omega)} \leq C_\star \kappa\, \mathrm{E}_{\mathcal{D}_i}(\bar{u}_{i-1}, g_\star)^{\frac{1}{2}} \leq C_\star \kappa\, \omega \varepsilon_{i-1}. \tag{45}$$

On the other hand, recalling (20), we have

$$\|u_\star - \bar{u}_{i-1}\|_{\bar{\mathcal{D}}_{i-1}} \leq \varepsilon_{i-1}. \tag{46}$$

Let us define $w_{\mathcal{D}_i}$ as follows. Set $\psi := (\bar{u}_{i-1})^{\sim}_x \in L^2(\Omega)$ and let $q \in L^2(\Omega)$ be the piecewise polynomial function such that $q_{|K_D} = \Pi^0_{K_D,p_D-1}\psi_{|K_D}$ for all $D \in \mathcal{D}_i$. Notice that, recalling the definition (15), we have

$$\|\psi - q\|^2_\Omega = \sum_{D\in\mathcal{D}_i} \|\psi - q\|^2_{K_D} \leq \sum_{D\in\mathcal{D}_i} e_D(\bar{u}_{i-1}, g_\star) = \mathrm{E}_{\mathcal{D}_i}(\bar{u}_{i-1}, g_\star) \leq \omega^2\varepsilon^2_{i-1}.$$

On the other hand, it holds

$$\int_\Omega q = \int_\Omega \psi = \sum_{D\in\bar{\mathcal{D}}_{i-1}} \int_{K_D} \bar{u}_{i-1,x} = -\sum_{e\in\mathcal{E}_{\bar{\mathcal{D}}_{i-1}}} [\![\bar{u}_{i-1}]\!]_e = -\sum_{e\in\mathcal{E}_{\bar{\mathcal{D}}_{i-1}}} \sigma^{-1/2}_{\bar{\mathcal{D}}_{i-1},e}\sigma^{1/2}_{\bar{\mathcal{D}}_{i-1},e}[\![\bar{u}_{i-1}]\!]_e,$$

whence

$$\left(\int_\Omega q\right)^2 \leq \Big(\sum_{e\in\mathcal{E}_{\bar{\mathcal{D}}_{i-1}}} \sigma^{-1}_{\bar{\mathcal{D}}_{i-1},e}\Big) \|\sigma^{1/2}_{\bar{\mathcal{D}}_{i-1}}[\![\bar{u}_{i-1}]\!]\|^2_{\mathcal{E}_{\bar{\mathcal{D}}_{i-1}}} \leq |\Omega|\, \|\sigma^{1/2}_{\bar{\mathcal{D}}_{i-1}}[\![\bar{u}_{i-1}]\!]\|^2_{\mathcal{E}_{\bar{\mathcal{D}}_{i-1}}} \leq \frac{|\Omega|}{\gamma}\,\varepsilon^2_{i-1}$$

by (46). Therefore, if we set

$$w_{\mathcal{D}_i}(x) = \int_{x_0}^x q(s)\,ds - (x - x_0)\int_\Omega q$$

where $x_0 = \min\Omega$, we realize $w_{\mathcal{D}_i} \in V^c_{\mathcal{D}_i}$ and $\|(\bar{u}_{i-1})^{\sim}_x - w_{\mathcal{D}_i,x}\|_\Omega \leq C\varepsilon_{i-1}$. This concludes the proof, since $\varepsilon_{i-1} \simeq \varepsilon_i$. ◻

By Property 5.4, we get the bound

$$\|u(g_{\mathcal{D}_i}) - u_{\mathcal{D}_i}\|^2_{a,\mathcal{D}_i} + \beta\,\eta^2_{\mathcal{D}_i}(u_{\mathcal{D}_i}) \leq C_{13}\varepsilon^2_i + C_{11}\eta^2_{\mathcal{D}_i}(u_{\mathcal{D}_i}).$$

At this point, we may proceed as follows. Assume that we have chosen, once and for all, an absolute constant $\hat{C} > 0$. We check the validity of

$$\eta^2_{\mathcal{D}_i}(u_{\mathcal{D}_i}) \leq \hat{C}\varepsilon^2_i.$$

- In the affirmative case, $u_{\mathcal{D}_i}$ does satisfy condition (44), and we can start the iterations of **DG-SOLVE**.
- In the negative case, we discard $u_{\mathcal{D}_i}$ and compute $\hat{u}^c_{\mathcal{D}_i} \in V^c_{\mathcal{D}_i}$, the (continuous) Galerkin approximation of $u(g_{\mathcal{D}_i})$ on the partition $\mathcal{D}_i$. For such an approximation, it is known that the residual estimator is both reliable and efficient; hence, resorting once more to Property 5.4,

  $$\eta_{\mathcal{D}_i}(\hat{u}^c_{\mathcal{D}_i}) \;\simeq\; \|u(g_{\mathcal{D}_i}) - \hat{u}^c_{\mathcal{D}_i}\|_a \;\simeq\; \|u(g_{\mathcal{D}_i}) - \hat{u}^c_{\mathcal{D}_i}\|_{H^1_0(\Omega)} \leq C_{12}\varepsilon_i.$$

  Therefore, condition (44) is satisfied with $u_{\mathcal{D}_i}$ replaced by $\hat{u}^c_{\mathcal{D}_i}$, and we start the iterations of **DG-SOLVE** from this approximation.

Acknowledgements. Work carried out within the "Progetto di Eccellenza 2018-2022", granted by MIUR (Italian Ministry of University and Research) to the Department of Mathematical Sciences, Politecnico di Torino.

The authors are members of the INdAM research group GNCS, which granted partial support to this research.

References

[1] P. Antonietti, C. Canuto, and M. Verani. Convergence and Optimality of Adaptive *hp*-DG Finite Element Methods for Elliptic Problems, in preparation.

[2] D. Arnold. An interior penalty finite element method with discontinuous elements. *SIAM J. Numer. Anal.* 19 (1982), 742–760

[3] D. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.* 39 (2002), 1749–1779

[4] A. Bonito and R.H. Nochetto. Quasi-optimal convergence rate of an adaptive Discontinuous Galerkin method. *SIAM J. Numer. Anal.* 48 (2010), 734–771

[5] P. Binev. Tree Approximation for *hp*-Adaptivity. *SIAM J. Numer. Anal.* 56 (2018), 3346–3357

[6] C. Canuto, R.H. Nochetto, R. Stevenson, and M. Verani. Convergence and optimality of hp-AFEM. *Numer. Math.* 135 (2017), 1073–1119

[7] C. Canuto, R.H. Nochetto, R. Stevenson, and M. Verani. On *p*-robust saturation for *hp*-AFEM. *Comput. & Math. with Appl.* 73 (2017), 2004–2022

[8] P. Houston, D. Schötzau, and T.P. Wihler Energy norm a posteriori error estimation of *hp*-adaptive discontinuous Galerkin methods for elliptic problems. *Math. Models Methods Appl. Sci.* 17 (2007), 33–62

[9] 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, Art. 2, 39 pp.

[10] Ch. Schwab. *p- and hp- Finite Element Methods.* Oxford University Press, 1998