The Double Adaptivity Method: Numerical Studies in Two Dimensions. Conforming vs. Weakly Conforming Test Spaces.

Jacob Salazar *and Leszek Demkowicz

Oden Institute for Computational Engineering and Sciences
The University of Texas at Austin, 201 E 24th St, Austin, TX 78712, USA

Abstract

We present an extensive numerical investigation on the optimal choice of parameters defining the double adaptivity algorithm presented by Demkowicz et al. [10]. In the original work 1D problems were studied and preliminary 2D results were presented for the lowest order elements only. The present work focuses on two 2D implementations: the first with a conforming test space, and the second one with a weakly conforming test space for arbitrary order elements. The double adaptivity method is set within the Petrov-Galerkin method with optimal test functions and key parameters (e.g. error tolerances, norm weights, refinement factors and enrichment strategies) greatly impact its numerical performance. We present an a-priori error estimate for the weakly conforming method, and compare its performance with the conforming one showing a very similar behavior.

1 Introduction

Consider any well-posed variational problem, possibly in a non-symmetric functional setting:

$$\begin{cases} u \in U \\ b(u,v) \quad = l(v) \quad v \in V \end{cases} \tag{1.1}$$

where U, V are trial and test real Hilbert spaces, $b : U \times V \to \mathbb{R}$ is a bounded bilinear functional, and $l \in V'$.

We are interested in approximating a solution to this problem using finite-dimensional spaces $U_h \subset U$ and $V_h \subset V$ of the same dimension, by solving the discrete problem:

$$\begin{cases} u_h \in U_h \\ b(u_h,v_h) \quad = l(v_h) \quad v_h \in V_h. \end{cases} \tag{1.2}$$

The *ideal* Petrov-Galerkin Method (PGM) with Optimal Test Functions (OTF) [12] guarantees the discrete stability of problem (1.2), the solution u_h is the projection of the exact solution u into U_h in the energy norm [20]:

$$\|u - u_h\|_E = \inf_{w_h \in U_h} \|u - w_h\|_E, \tag{1.3}$$

*Corresponding author. E-mail: jacobess@utexas.edu

where the energy norm is implied by the bilinear form b and test norm V as follows

$$\|u\|_E = \sup_{v \in V} \frac{|b(u,v)|}{\|v\|_V} \,. \tag{1.4}$$

The Optimal Test Function $v_h \in V_h$ corresponding to $u_h \in U_h$ solves the problem:

$$\begin{cases} v_h \in V_h \\ (v_h, \delta v)_V = b(u_h, \delta v) \quad \delta v \in V \end{cases} \tag{1.5}$$

with $(\cdot,\cdot)_V$ being the test inner product. Note that the OTF problem is dependent on the choice of test norm and involves the inversion of the Riesz map of V.

Optimal test norm. Ideally we would like the norm we converge in to be the trial norm $\|\cdot\|_U$. However as noted in (1.4) the energy norm is not necessarily the same as the trial norm. For this to be satisfied we need to make use of the so called optimal test norm :

$$\|v\|_{V,opt} := \sup_{u \in U} \frac{|b(u,v)|}{\|u\|_U} \,. \tag{1.6}$$

In some cases, one can analytically determine the optimal test norm. Amongst those cases we find the *ultraweak* (UW) variational formulation for any system of linear partial diferential equations (PDEs) (in any space dimension); and the primal formulation for the diffusion-convection-reaction equation with constant coefficients (in 1D).

A noticeable disadvantage of using the optimal test norm, in the context of singular perturbation problems, is the development of boundary layers in the OTF that may be difficult to resolve. This limitation is exacerbated in the *practical* discontinuous Petrov-Galerkin (DPG), resulting from replacing the whole space V in problem (1.5) with a discrete subspace usually called the *enriched test space*. The resolution of such functions, becomes problematic for small perturbation parameters [6].

There are two philosophies to avoid this problem:

1. Relinquish the optimality of the test norm and design problem-specific test norms in such a way that the modified OTF do not develop boundary layers; this was the idea behind the robust test norms studied by Chan, Heuer, Bui-Thanh and Demkowicz [6] .
2. Rely on adaptivity to improve the resolution of the OTF (or equivalently, the Riesz representation of the residual when casting the PGM with OTF as a mixed problem). This idea originated from the adaptivity and variational stabilization methods introduced by Cohen, Dahmen and Welper [7] and led to the concept of the Double Adaptivity Method (DAM) presented by Demkowicz, Führer, Heuer and Tian [10].

The DAM makes use of two (an inner and an outer) loops, which adaptively refine the meshes corresponding to the trial and test spaces. Critical to the success of the double adaptivity is the existence of a reliable a-posteriori error estimate for the inner adaptivity loop. Such an estimate, based on the duality theory, has been developed in [10].

The Double Adaptivity Method. Problem (1.1) is cast into a mixed problem by introducing the Riesz representation of the residual ψ (which is zero in the continuous level), and including the corresponding two more terms:

$$\begin{cases} \psi \in V, u \in U & \\ (\psi, v)_V + b(u, v) & = l(v) \quad v \in V \\ b(\delta u, \psi) & = 0 \quad \delta u \in U. \end{cases} \tag{1.7}$$

It can be shown that, upon discretizing the trial space in (1.7), we get a semi-discrete mixed problem:

$$\begin{cases} \psi^h \in V, \widetilde{u}_h \in U_h & \\ (\psi^h, v)_V + b(\widetilde{u}_h, v) & = l(v) \quad v \in V \\ b(\delta u_h, \psi^h) & = 0 \quad \delta u_h \in U_h \,, \end{cases} \tag{1.8}$$

equivalent to the *ideal* PGM with OTF (1.2). Problem (1.8) seeks an approximate solution $\widetilde{u}_h \in U_h$ along with the corresponding exact Riesz representation of the residual $\psi^h \in V$.

Eventually we discretize the test space as well, with $V_h \subset V$, arriving at the *practical* PGM with OTF:

$$\begin{cases} \psi_h \in V_h, u_h \in U_h & \\ (\psi_h, v)_V + b(u_h, v) & = l(v) \quad v \in V_h \\ b(\delta u_h, \psi_h) & = 0 \quad \delta u_h \in U_h \,. \end{cases} \tag{1.9}$$

Function ψ^h provides an a-posteriori error estimate for the ideal PGM with OTF which can drive the adaptive refinements of the trial space. However, what the practical method delivers is its approximation ψ_h [1], which fails to serve as error estimator for the solution $\widetilde{u}_h$ if it is not close to the exact ψ^h.

In order to obtain a good approximation to ψ^h, the DAM proposes to determine an optimal discrete test space V_h starting with an initial test mesh and then adaptively refining it. The adaptive process is guided by an estimate of the difference $\psi^h - \psi_h$, constructed as the difference between the energy functionals corresponding to the primal (original) problem and a dual problem (introduced in the following section).

Related work. Another strategy aiming to improve the resolution of OTF is to use splines along with a Shishkin intra-element submesh for the test pace discretization [19]. This strategy, as well as the use of Robust Test Norms, have been able to resolve convection-dominated diffusion problems with values of ε as small as 10^{-5} but not smaller (in 2 space dimensions) [6] ; while the range of parameters we need to be able to solve, in order to tackle compressible flow problems is of the order of 10^{-7}.

Other methods similar to the PGM with OTF focused on convection-dominated diffusion problem are 1) the non-linear Petro-Galerkin Method developed by Houston, Roggendorf and van der Zee [16] based on minimizing the residual in a dual L^p-norm, which generalizes the Petrov Galerkin framework to Banach spaces, and demonstrate a complete elimination of the Gibbs phenomenon for convection-dominated diffusion problems, and a class of meshes. 2) a similar L^P-DPG method of Li and Demkowicz [18], which generalizes the standard DPG based on Hilbert spaces to Banach which results in solutions less susceptible to Gibbs phenomenon; 3) the Automatic Variationally Stable (AVS) method of Calo, Romkes and Valseth [5] which uses C^0 spaces for all the trial field variables and eliminates the use of independent trial trace variables.

[1] Notice that ψ^h in (1.8) and ψ_h in (1.9) are two different objects, identified by means of a superscript h and an underscript h, respectively.

Goal and scope of this work. This paper, as well as [10], focuses on convection-dominated diffusion as a motivating and non-trivial example of a singular perturbation problem. But the methodology herein presented applies to any well posed system of linear PDEs, including other singular perturbation problems. An efficient implementation of the method results in a general code with broad applications able to solve challenging problems.

In [10], numerical results in 1D and preliminary results in 2D with lowest order polynomials were presented. In this work we tackle *hp* meshes of arbitrary order, and we take on comparing the conforming PGM with OTF against the weakly conforming PGM with OTF. The concept of weakly conforming test functions will be introduced and elaborated in Section 4.1.

Outline. Section 2 presents the main ideas of the DAM and the corresponding practical algorithm. The main focus is on the error estimator err_V defined in (2.12), and the double adaptivity algorithm (1). Having presented the general idea behind the DAM, we proceed with a discussion of the conforming discretization in Section 3, the implementation within pre-existing high-order DPG focused software, and numerical studies on the main parameters defining the algorithm. Then in Section 4 we present the weakly conforming discretization. We present its a-priori error estimate and contrast it with the conforming estimate. Finally we present numerical investigations aimed at comparing the two methodologies numerically.

2 The Double Adaptivity Method

Cohen, Dahmen and Welper [7] originaly proposed the idea of determining an optimal discrete test space V_h using adaptivity, and Demkowicz et al. [10] developed an estimator for $\psi^h - \psi_h$ (the difference in the solution to the Riesz representation of the residual for the semi-discrete problem (1.8) and the solution the fully discrete problem (1.2), to drive the adaptive process. For readers' convenience, we summarize here the derivation of the estimator and the double adaptivity strategy.

Consider the problem (strong formulation):

$$\begin{cases} u \in D(A) \\ Au = f \end{cases} \tag{2.1}$$

where:

$$A : L^2(\Omega) \supset D(A) \to L^2(\Omega), \tag{2.2}$$

is a closed operator representing a system of first order PDEs with BCs reflected in the definition of $D(A)$, Ω is a bounded Lipschitz domain, and A^* is the L^2-adjoint of A. The corresponding *ultraweak* (UW) formulation is obtained by multipliying (2.1) with a test function $v \in D(A^*)$ and integrating by parts:

$$\begin{cases} u \in L^2(\Omega) \\ (u, A^*v) = (f, v) \quad v \in D(A^*). \end{cases} \tag{2.3}$$

The quasi-optimal norm. The optimal test norm in the UW variational formulation according to (1.6) is given by:

$$\|v\|_{V_{opt}} = \|A^*v\|_{L^2}, \tag{2.4}$$

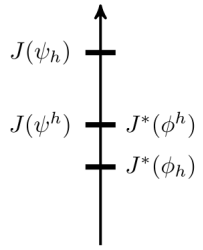


Figure 2.1

frequently called the *adjoint norm*. An equivalent norm is additionally introduced, known as the *adjoint graph norm*:

$$\|v\|^2_{V_{q-opt}} = \|A^*v\|^2_{L^2} + \alpha\|v\|^2_{L^2} \tag{2.5}$$

where $\alpha > 0$; its corresponding energy norm no longer coincides with the L^2 trial norm but remains equivalent. For this reason, the adjoint graph norm has also been named the quasi-optimal test norm.

The addition of the L^2 term makes the test norm *localizable* [20]; therefore, the norm (2.5) is suitable for the PGM with broken test spaces, known as the Discontinuous Petrov Galerkin (DPG) method [13].

The semi-discrete mixed problem (in which (2.3) is embedded, as illustrated in (1.8)) corresponding to the ultraweak formulation with the quasi-optimal test norm (the adjoint graph norm) results in :

$$\begin{cases} \psi^h \in D(A^*), \widetilde{u}_h \in U_h & & & \\ (A^*\psi^h, A^*v) + \alpha(\psi^h, v) + (\widetilde{u}_h, A^*v) & = (f, v) & v \in D(A^*) \\ (A^*\psi^h, \delta u_h) & = 0 & \delta u_h \in U_h. \end{cases} \tag{2.6}$$

The mixed problem (2.6) is equivalent to the constrained minimization of the (primal) functional J:

$$\inf_{\substack{\psi\in D(A^*)\\ A^*\psi\in U_h^\perp}} \underbrace{\frac{1}{2}\|A^*\psi\|^2 + \alpha\frac{1}{2}\|\psi\|^2 - (f,\psi)}_{=:J(\psi)} = \inf_{\substack{\psi\in D(A^*)\\ A^*\psi\in U_h^\perp}} J(\psi) \tag{2.7}$$

where $U_h^\perp$ is the L^2 orthogonal complement of U_h.

Then the following dual problem is derived (see details in [10]) involving the dual functional J^*:

$$-\inf_{\phi\in D(A)} \underbrace{\frac{1}{2}\|\phi^\perp\|^2 + \frac{1}{2\alpha}\|f - A\phi\|^2}_{=:-J^*(\phi)} = \sup_{\phi\in D(A)} J^*(\phi) \tag{2.8}$$

where $\phi^\perp$ is the L^2-projection of ϕ onto the L^2-orthogonal complement of U_h. This dual problem is equivalent to another 'mixed-like' problem:

$$\begin{cases} \phi^h \in D(A), \widetilde{w_h} \in U_h & & & \\ (A\phi^h, A\delta\phi) + \alpha(\phi^h, \delta\phi) & -\alpha(\widetilde{w_h}, \delta\phi) & = (f, A\delta\phi) & \delta\phi \in D(A) \\ -\alpha(\phi^h, \delta w_h) & +\alpha(\widetilde{w}_h, \delta w_h) & = 0 & \delta w_h \in U_h\,. \end{cases} \tag{2.9}$$

Notice that while mixed problem (2.6) corresponds to a saddle problem, (2.9) corresponds to a double minimization problem [10].

Strict convexity of J and strict concavity of J^* imply that its corresponding minimizer and maximizer exist. Minimization of the primal functional yields a discrete minimum $J(\psi_h)$ above the exact one $J(\psi^h)$. Maximization of the dual functional results in a discrete maximizer ϕ_h with corresponding discrete maximum smaller than the continuous one $J^*(\phi_h) \leq J^*(\phi^h)$, see Fig. 2.1. Consequently, the following bounds hold:

$$\left.\begin{array}{l} J(\psi_h) - J(\psi^h) \\ J^*(\phi^h) - J^*(\phi_h) \end{array}\right\} \leq J(\psi_h) - J^*(\phi_h). \tag{2.10}$$

For a quadratic minimization problem

$$\inf_{u \in V} \underbrace{\frac{1}{2} b(u,u) - l(u)}_{J(u)}$$

where form b is symmetric and positive definite, the difference between the minimizer $u \in V$, and an arbitrary element $u_h \in V$, measured in the induced energy norm ($\| \cdot \|_{\mathbf{E}} := b(\cdot,\cdot)^{1/2}$) is related to the energy J by:

$$\|u - u_h\|^2_{\mathbf{E}} = 2(J(u_h) - J(u))\,. \tag{2.11}$$

Both, our primal (2.7) and dual (2.8) energies satisfy the conditions for (2.11), hence it follows:

$$\left.\begin{array}{ll} \|\psi^h - \psi_h\|^2_{\mathbf{E}_1} & = 2J(\psi_h) - 2J(\psi^h) \\ \|\phi_h - \phi^h\|^2_{\mathbf{E}_2} & = 2J^*(\phi^h) - 2J^*(\phi_h) \end{array}\right\} \leq 2J(\psi_h) - 2J^*(\phi_h) =: err_V\,. \tag{2.12}$$

Notice, however, that the energy norms ($\| \cdot \|_{\mathbf{E}_1}$ and $\| \cdot \|_{\mathbf{E}_2}$)for the primal and dual problems are different.

Finally, the right hand side of (2.12) can be expressed as the integral of non-negative terms:

$$2(J(\psi_h) - J^*(\phi_h)) = \frac{1}{\alpha} \int_\Omega \{\alpha(A^*\psi_h - \phi_h^\perp)^2 + (\alpha\psi_h - (f - A\phi_h))^2\}. \tag{2.13}$$

The non-negativeness of the terms makes the local element contributions to the estimate (2.13) good candidates for element error indicators. Finally, note that the coefficient α weighting the L^2 term of the test norm is necessary to derive the dual problem and apply the duality gap estimate, which makes the adjoint graph norm (instead of the optimal adjoint test norm) the default norm for the DAM. Using small values of α gets us closer to the optimal test norm [10], but the duality gap error estimate deteriorates (as observed in the numerical results to follow).

With this, we arrive at the *double adaptivity algorithm*:

Algorithm 1: Double Adaptivity

Input: Initial trial mesh U_h

Output: Solution u_h

(Outer loop) **for** $i = 1$***to*** max_iter_U **do**

  (Re)set the test mesh V_h to coincide with a globally enriched version of the trial mesh U_h

  (Inner loop) **for** $j = 1$***to*** max_iter_V **do**

    Solve the primal and dual problems on the current meshes

    Compute $\|\psi_h\|_V$ and error estimate $err_V := \sqrt{2(J(\psi_h) - J^*(\phi_h))} \geq \|\psi^h - \psi_h\|_V$

    if $err_V/\|\psi_h\|_V < tol_V$ **then**

      Exit inner loop

    Adapt the test mesh V_h using test element contributions of err_V^2

  Compute norm of the solution $\|u_h\|_U$ and error estimate $err_U = \|\psi_h\|_V$

  if $err_U/\|u_h\| < tol_U$ **then**

    Exit outer loop

  Adapt the trial mesh U_h using trial element contributions of err_U^2

The trial error estimate err_U refers to the ψ_h that has satisfied the inner loop convergence tolerance. In the first instruction of the outer loop, "(re)setting the test mesh V_h to coincide with a globally enriched version of the trial mesh U_h" means: first, copying the trial mesh U_h and corresponding data structure into the test mesh V_h while maintaining the appropriate trial energy space (which in this formulation is different from the test one); and then perform a global enrichment to ensure the satisfaction of $\dim V_h > \dim U_h$, see Remark 2 in [10]. Fig. 2.2 illustrates the double adaptivity algorithm refining the trial mesh in the outer loop and the test mesh in the inner loop (for each outer loop step, during the second task, a whole inner loop is run until tol_V is met).

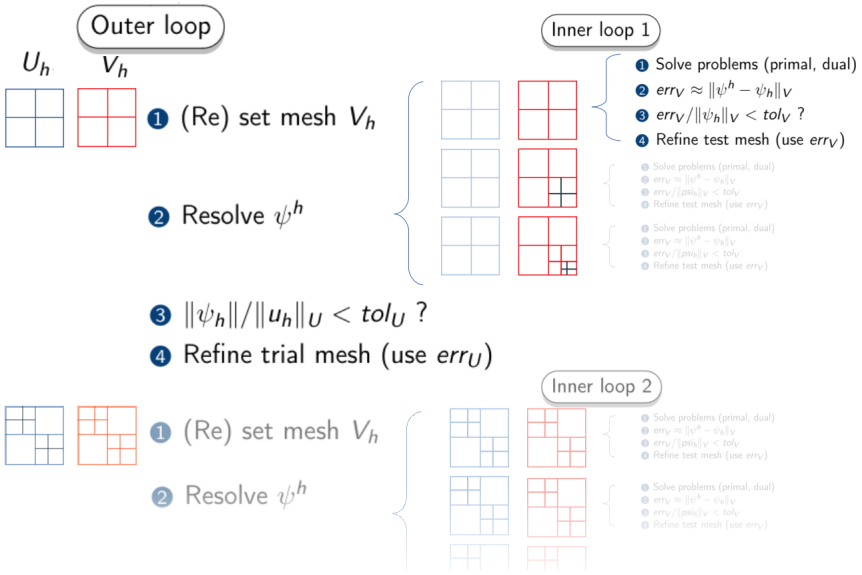


Figure 2.2: This diagram displays the first step of the outer loop and the first two tasks of the second step. It also illustrates the first inner loop, displaying 3 steps (assuming, for illustrative purposes, that 3 were required to satisfy the tolerance). There are 4 main tasks for each step of the outer loop, during the second task (resolve $\|\psi^h\|$) a whole inner loop is required. Each inner loop step contains 4 main tasks as well.

3 Conforming Discretization: Formulation and Numerical Results

As mentioned in the Introduction, the convection-dominated diffusion problem is used to illustrate the methodology.

$$\begin{cases} u &= 0 \quad \text{on } \Gamma \\ -\epsilon\Delta u + \beta \cdot \nabla u &= f \quad \text{in } \Omega\,. \end{cases} \tag{3.1}$$

The original problem is rewritten as a system of first order equations, introducing a new variable σ using the formulation advocated by Broersen and Stevenson [4]:

$$\begin{cases} u &= 0 \quad \text{on } \Gamma \\ \sigma - \epsilon^{\frac{1}{2}}\nabla u &= 0 \quad \text{in } \Omega \\ -\epsilon^{\frac{1}{2}}\mathrm{div}\sigma + \beta \cdot \nabla u &= f \quad \text{in } \Omega\,. \end{cases} \tag{3.2}$$

Introducing the first order operator A and the corresponding energy spaces, Equation (3.2) is cast in the general *strong formulation* (2.1) :

$$\begin{cases} \mathsf{u} \in D(A) \\ A\mathsf{u} = \mathsf{f} \end{cases} \tag{3.3}$$

with $\mathsf{f} = (0, f)$ and:

$$\begin{aligned}
&\mathsf{u} := (\sigma, u) \in D(A) := H(\mathrm{div}, \Omega) \times H_0^1(\Omega) \subset (L^2(\Omega))^N \times L^2(\Omega) \\
&A : D(A) \to (L^2(\Omega))^N \times L^2(\Omega) \\
&A\mathsf{u} = A(\sigma, u) := (\sigma - \epsilon^{\frac{1}{2}} \nabla u, -\epsilon^{\frac{1}{2}} \mathrm{div}\sigma + \beta \cdot \nabla u) \\
&D(A) = \{\mathsf{u} = (\sigma, u) \in (L^2(\Omega))^N \times L^2(\Omega) : A\mathsf{u} \in (L^2(\Omega))^N \times L^2(\Omega),\, u = 0 \,\mathrm{on}\, \Gamma\,.
\end{aligned} \tag{3.4}$$

The UW formulation corresponding to (2.3) results in:

$$\begin{cases}
\mathsf{u} \in (L^2(\Omega))^N \times L^2(\Omega) \\
(\mathsf{u}, A^*\mathsf{v}) = (\mathsf{f}, \mathsf{v}) \quad \mathsf{v} \in D(A^*)\,,
\end{cases} \tag{3.5}$$

with:

$$\begin{aligned}
&D(A^*) := D(A) \\
&\mathsf{v} := (\tau, v) \in D(A^*) \subset (L^2(\Omega))^N \times L^2(\Omega) \\
&A^* : D(A^*) \to (L^2(\Omega))^n \times L^2(\Omega) \\
&A^*\mathsf{v} = A^*(\tau, v) = (\tau + \epsilon^{\frac{1}{2}} \nabla v, \epsilon^{\frac{1}{2}} \mathrm{div}\tau - \mathrm{div}(\beta v))
\end{aligned} \tag{3.6}$$

As mentioned in the previous section, the DAM requires the use of the adjoint graph norm (2.5) for the test space.

Case study. The domain for the case study is the unit square, $\Omega = (0, 1)^2$, and a constant advection vector $\beta = (1, 0)$, homogeneous boundary conditions, and a right-hand side corresponding to the exact solution

$$u(x, y) = \frac{-\frac{1}{\epsilon\pi^2}(1 - e^{-r_2})}{e^{-r_1 x} - e^{r_1(1-x)-r_2}} \cos \pi y \qquad ; \qquad r_{1,2} = \frac{-1 \pm \sqrt{1 + 4\epsilon^2\pi^2}}{-2\epsilon} \tag{3.7}$$

The problem is solved using high order polynomial spaces defined on rectangular elements, corresponding to the exact sequence of tensor product polynomials of degree p.

3.1 Implementation

We use two different meshes defined on the same domain, the 'trial mesh' supporting the variables u_h and w_h, and the test mesh for the variables v_h, ψ_h and ϕ_h, each capable of independent adaptive refinements. Starting from two topologically identical meshes, each of them is independently refined in an adaptive fashion, driven either by err_U (trial mesh) or err_V (test mesh), as described by the pseudocode earlier.

A new implementation different from the one used for preliminary results in [10] was developed, by modifying the *hp2D* Fortran90 code developed by Demkowicz and collaborators [8] to support two independent adaptive meshes. *hp2D* is a flexible, modular FE code that provides a globally conforming discretization of all the usual energy spaces (in the present case, $H^1(\Omega), H(\mathrm{div}, \Omega), L^2(\Omega)$). Among other features, it supports the following:

1. The Nédélec exact sequence of polynomial spaces of the first type for quadrilaterals and triangles.
2. Hierarchical shape functions that allow for the use of high order polynomial discretizations and local p-refinements.
3. 1-irregular meshes (hanging or constrained nodes) and thus local h-adaptive refinements are possible.

4. Coupled multiphysics problems.
5. Projection-based interpolation for geometry and BC data.
6. Static condensation for local element matrices.
7. Adaptive numerical integration, particularly important for problems with boundary layers or singularities.
8. Orientation-embedded shape functions [15].

Two data structure arrays: ELEMS and NODES, contain all the geometrical and functional setting information necessary to assemble and solve a multi-physics *hp*-adaptive finite element problem. ELEMS contains the initial mesh topological data, information on locally supported physics variables, boundary conditions and element-to-nodes connectivities for each element. NODES includes, for each mesh *node*, the node type (vertex, edge or middle node), its corresponding DOF, geometric data, geneological information for *h*-refinements (nodal father and sons) and polynomial order.

The *double-data-structure* technology is used, which consists of two sets of data structures: {ELEMS_U, NODES_U} for the trial mesh and corresponding variables; and {ELEMS_V, NODES_V} for the test mesh and corresponding variables. While the original *hp2d* supports adaptive mesh refinements, it does so for only one mesh (the one determined by the information in ELEMS and NODES) and these data structures are deeply engrained and utilized across the whole code, many times within routines with multiple nesting levels. Modification of all involved routines to handle multiple meshes would entail extensive modification of the whole complex codebase (element stiffness matrix computation, element residual computation, visualization and solver interfaces, to name a few). Instead of such a major endeavour, the pair {ELEMS, NODES} is defined as *pointers*, which allows defining two (or more) independent meshes' variables ({ELEMS_U, NODES_U}, {ELEMS_V, NODES_V}) as targets, while maintaining the same variable name being called by all routines (the pointers themselves). As a results, there is less modification of fundamental routines in exchange for strategical switching of the now pointer-type variables. New subroutines perform the mesh switching and are called accordingly across the whole software.

Among the many processes that are favored by the double-data-structure technology, adaptivity is a key one. Adaptive refinements of either mesh can be performed by using the same pre-existing local refinement subroutines in both cases: using {ELEMS_V, NODES_V} as targets in the inner loop and {ELEMS_U, NODES_U} in the outer loop.

Remark 3.1. The original aim was to use meshes that could be refined fully independently. However during the development of this work, we converged to the strategy where, at the beginning of each inner loop, the test mesh is reset to the current trial mesh (as described in the previous section). This results in an ultimate test mesh being always a submesh of the trial mesh. For this reason the developed double-data-structure technology is perhaps more general than needed. The possible simplifications are taken advantage of in the implementation with weakly conforming spaces (see subsection 4.2).

Remark 3.2. Two solver interfaces were implemented, one for a Frontal solver [17] without pivoting and one for MUMPS [2]. While the MUMPS solver presented no complications, the the Frontal encountered zero value pivots which prevented it from delivering a solution for some polynomial degrees and some meshes. We attribute this to the problem structure where element matrices (particularly for the primal problem) include blocks of zeros which are not filled in before the front tries to reduce the corresponding equations.

3.2 Numerical results

Given the multiple stages it involves (solution of primal and dual problems, estimation of $err_V = ||\psi^h - \psi_h||$, refinement of the test mesh and refinement of the trial mesh), there are a few different strategies and parameters to be selected from that impact the overall performance. The most important ones are explored and the corresponding strategies and values to be used are proposed below:

- Enrichment strategy for the initial test space of the inner loop.
- Inner loop refinement tolerance tol_V (in the resolution of ψ^h).
- Test mesh refinement strategy.
- Refinement factor for the test mesh.
- Refinement factor for the trial mesh.
- L^2 weight α, in the test norm.

These experiments intend to provide a complete set of parameters defining the DOM for 2 dimensional problems.

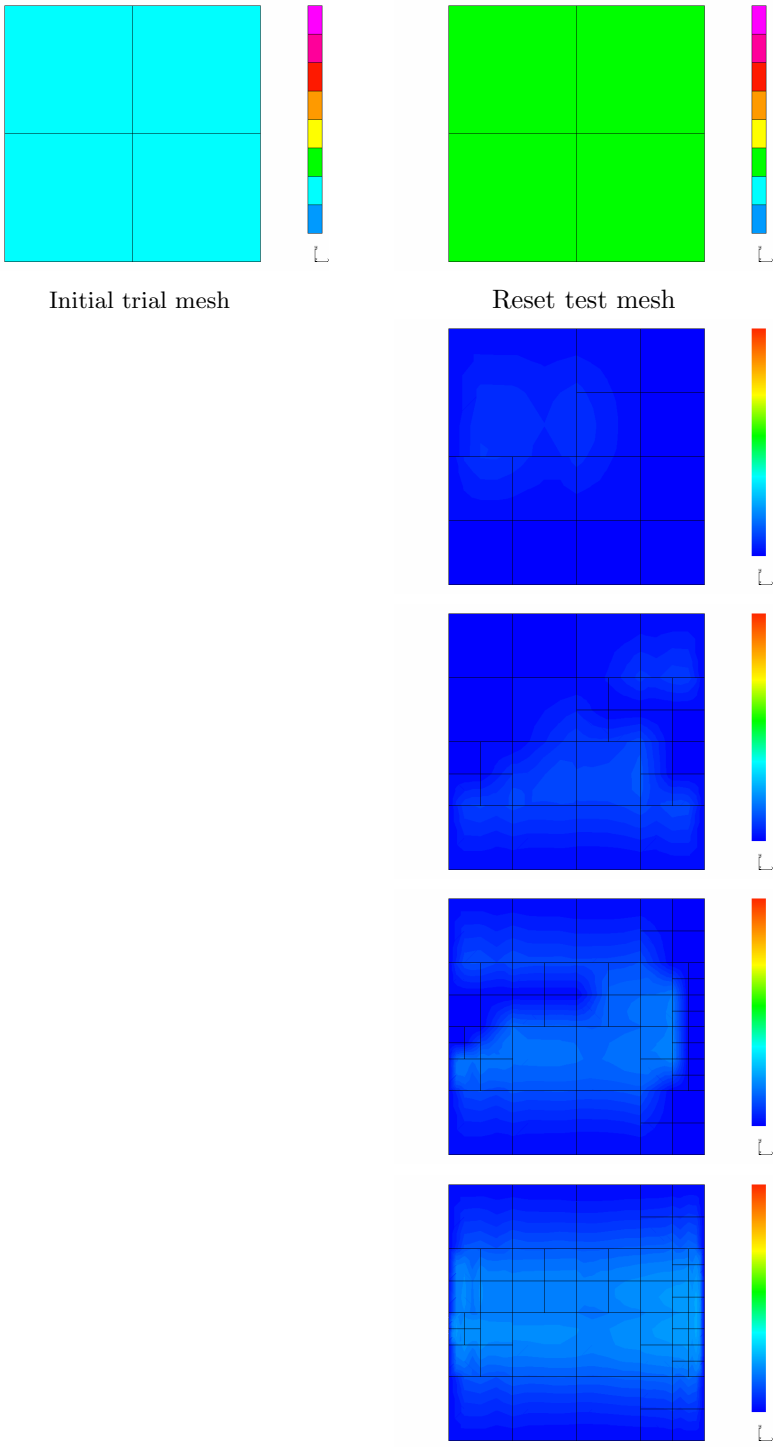
The following figures illustrate the DAM with a problem using $\varepsilon = 10^{-2}$. Fig. 3.1 shows the first outer loop step, and the resolution of its corresponding ψ^h via adaptive refinements of the test mesh. Once the tol_V is reached, the final ψ_h is used to drive the trial mesh adaptivity. The resulting mesh is displayed in the first image (left column) of the Fig. 3.2; in the same figure, the right column displays its corresponding inner loop. Fig. 3.3 shows the first 7 trial mesh adaptive refinements where the resolution of the boundary layer is verified.

3.2.1 Solution convergence

Initially, the convergence of the method is confirmed and the agreement between the numerical residual $\|\psi_h\|_V$ and the numerical error $\|\mathsf{u} - \mathsf{u}_h\|_U$ –which must be equivalent measures of the error according to the theory [11]– is verified. These results make use of the strategies and parameters selected after the numerical experiments detailed in the following sections, which are:

- h-enrichment strategy for the initial test space of the inner loop.
- Inner loop refinement tolerance $tol_V = 0.75$
- Doerfler refinement strategy (test and trial mesh).
- Refinement factor for the test mesh $Fact_D = 0.7$.
- Refinement factor for the trial mesh $Fact_D = 0.7$.
- L^2 weight in the test norm $\alpha = 1$.

In the present notation, p_U and p_V define the *nominal* polynomial degree of the exact sequence of finite elements used to discretize the trial and test spaces correspondingly. Therefore the solution u, which lives in L^2, is approximated with polynomials of degree $p_U - 1$. Fig. 3.4 shows the solution for $\varepsilon = 10^{-2}$ using different polynomial orders. These results numerically verify the agreement between our error estimate $err_U = \|\psi_h\|_V$ and the error measured in the trial norm $\|\mathsf{u} - u_h\|_U$. The plots herein presented use total trial DOF on the x-axis, so a slope of $p_U/2$ is expected for optimal convergence in 2 space dimensions. Clearly, better than optimal convergence rates are observed.

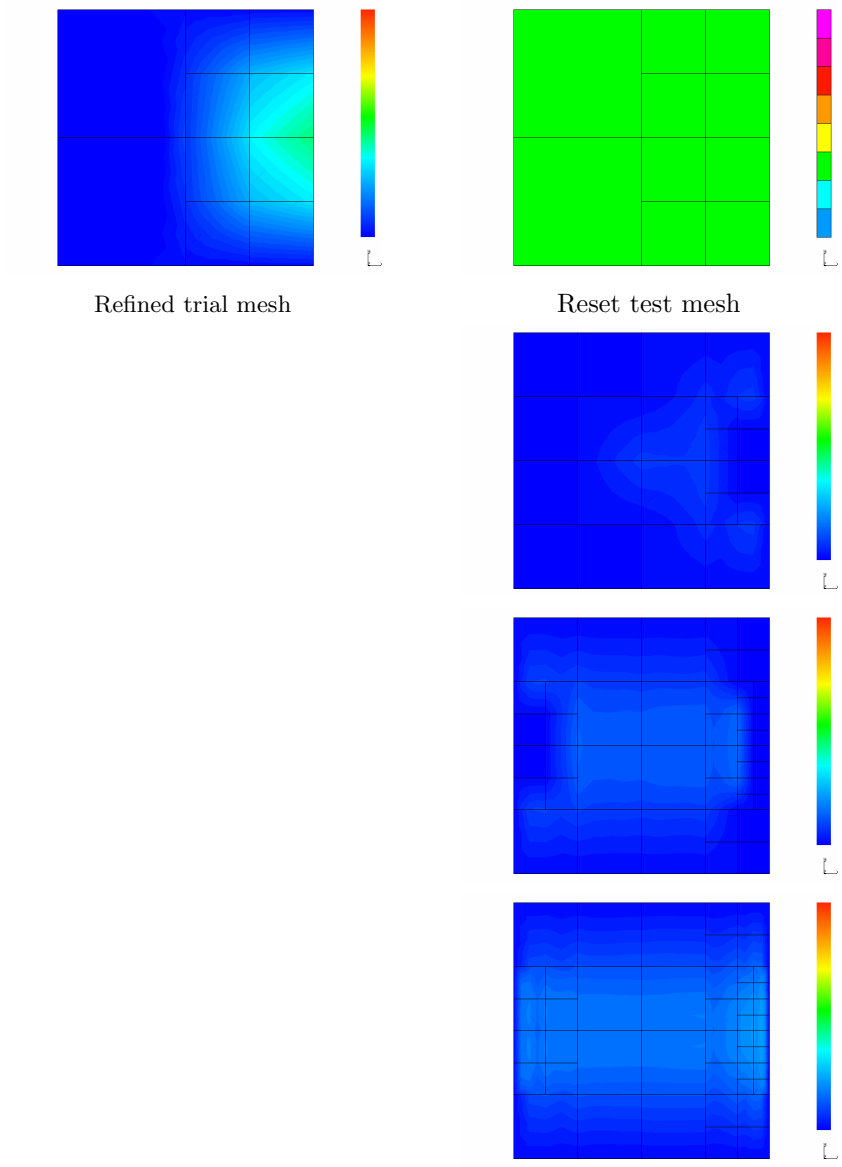


Initial trial mesh

Reset test mesh

Final test mesh (first outer loop step).

Figure 3.1: This figure illustrates the first outer loop step and the corresponding inner loop which takes 4 steps to reach the tolerance tol_V. For the first row, the colors indicate the polynomial degree of the elements, $p = 2$ for the trial elements and $p = 3$ for the test. The following rows display the adaptive refinement of the test mesh, along with the H^1-component of ψ_h using a numerical scale from 0 to 0.1. In this problem, $\varepsilon = 10^{-2}$.



Refined trial mesh

Reset test mesh

Final test mesh (second outer loop step)

Figure 3.2: This figure illustrates the second outer loop step and its corresponding inner loop which takes 3 steps to reach the tolerance tol_V. In the first row, the trial mesh displays the adaptive refinement along with the component u_h of the solution; and the test mesh displays the resetting (notice that the test mesh retains the its polynomial degree, $p = 3$ in this example). The following rows display the adaptive refinement of the test mesh, along with the H^1-component of ψ_h using anumerical scale from 0 to 1. In this problem $\varepsilon = 10^{-2}$.

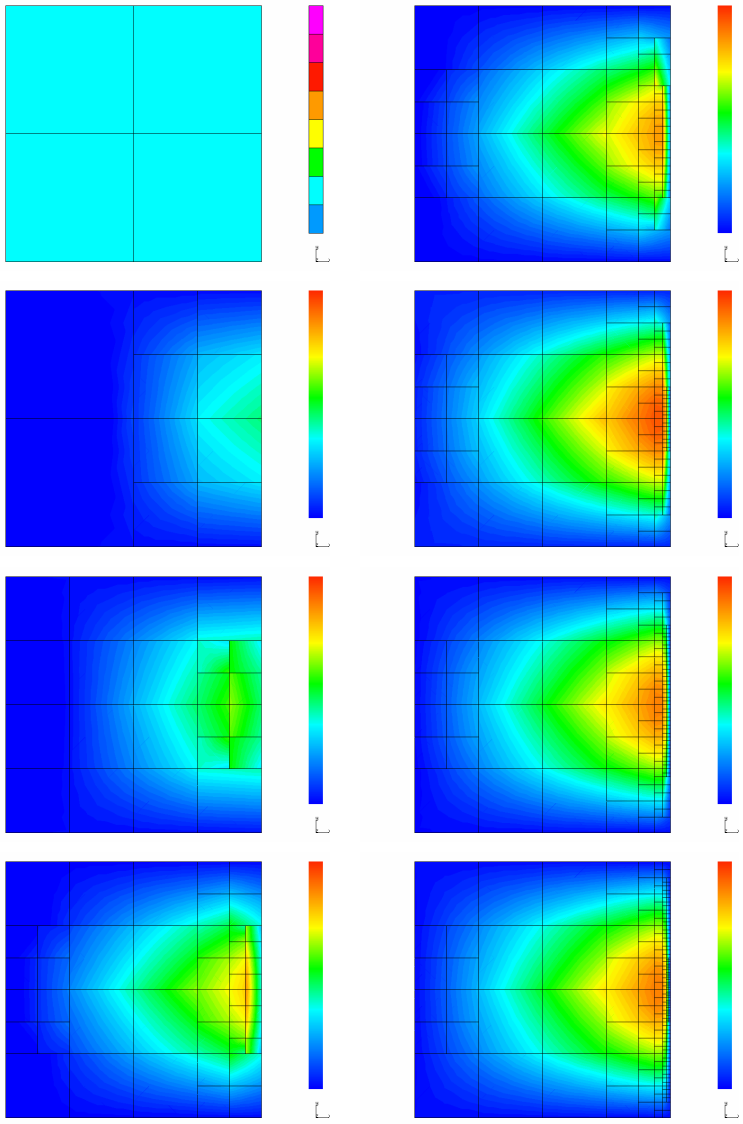
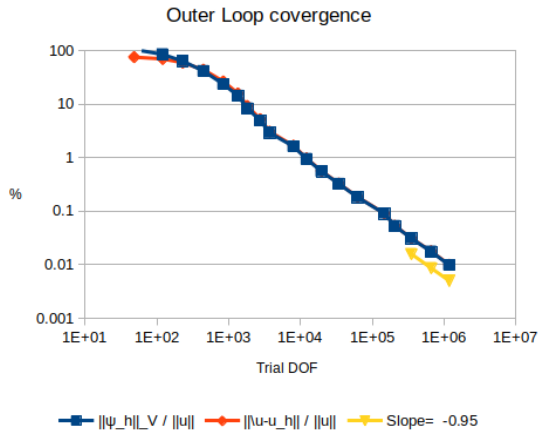
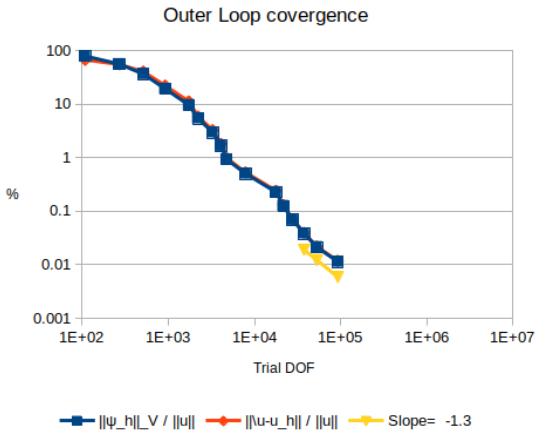


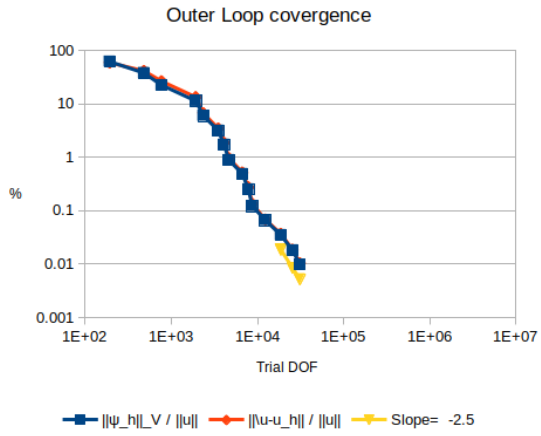
Figure 3.3: The first image (top left) displays the trial mesh and polynomial degree for its elements $p = 2$. The following images (down, and right) display the adaptive refinements along with the u_h component of the solution in a scale from 0 to 1. In this problem $\varepsilon = 10^{-2}$.



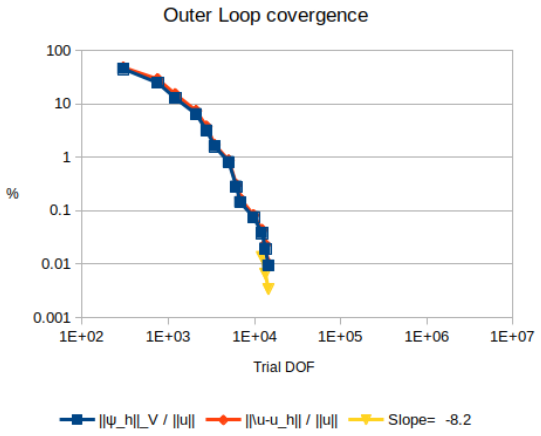
(a) $\epsilon = 10^{-2}$, $p_U = 2$, $p_V = 3$



(b) $\epsilon = 10^{-2}$, $p_U = 3$, $p_V = 4$



(c) $\epsilon = 10^{-2}$, $p_U = 4$, $p_V = 5$



(d) $\epsilon = 10^{-2}$, $p_U = 5$, $p_V = 6$

Figure 3.4: Outer loop convergence of the Riesz representation of the residual ψ in the test norm (normalized) and the L^2 error of the solution u (normalized).

3.2.2 Enrichment strategy

The PGM with OTF methodology requires an enriched test space to facilitate the satisfaction of the discrete Ladyzhenskaya–Babuška–Brezzi (LBB) condition. The DAM requires such an enrichment strategy to be effective at the beginning of the inner loop; later on, the h-refinement of the test mesh will naturally keep increasing its dimension (so preserving the satisfaction of the LBB condition). The initial enrichment can be performed by means of increasing the degree of the polynomial space used to discretize the test space (a common strategy in DPG); or, alternatively, with uniform h-refinements of the test mesh –unlike implementations of DPG with OTF with a single mesh, where enrichment by h-refinements is not possible.

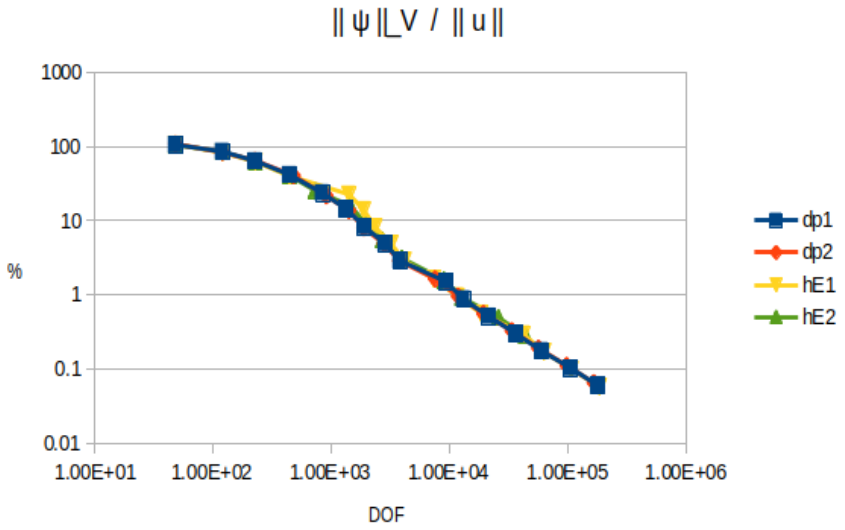


Figure 3.5: Comparison of enrichment strategies, global h-refinement and increasing p_V: $\epsilon = 10^{-2}$, $p_U = 2$, p_V variable. This plot displays the evolution of (normalized) $err_U = ||\psi_h||_V$ as the outer loop adaptively refines the trial mesh, increasing the number of (trial) DOF.

Figures 3.5 and 3.7 present the convergence of $err_U = \psi^h$ (normalized by the L^2 norm of our solution u). The series dp1 and dp2 represent the use of an initial test mesh equal to the trial one and polynomials of order $p_V = p_U + dp1$ and $p_V = p_U + dp2$ with $dp1 = 1$ and $dp2 = 2$; where p_U is the polynomial order used to discretize the trial space. The series hE1 and hE2 represent the use of a test mesh resulting from 1 and 2 global h-refinements of the trial mesh, respectively and $p_V = p_U = 2$.

It is clear that for $\varepsilon = 10^{-2}$ the results are almost exactly the same, and for $\varepsilon = 10^{-3}$ only hE2 shows significantly poorer convergence. However, to better appreciate their differences, the impact in computational cost is illustrated in Figures 3.6 and 3.8. In order to illustrate how computationally expensive inner loops are for each enrichment strategies, we display one data point for each outer loop step. Each data point relates the solution error estimate and the number of test DOF, that the inner loop required to achieve the tolerance tol_V. For a given value of err_U, a smaller number of test DOF implies that the inner loop is less expensive. All series start at about 100% with very similar number of test DOF, and evolve to improve the resolution of err_U by refining the trial and test meshes, thus moving to the left and up). This shows how for a given abscissa, dp1 enrichment presents the smaller number of test DOF, making it a better strategy (less expensive inner loops). This result may falsely be a-priori assumed since global h-refinements 'waste'

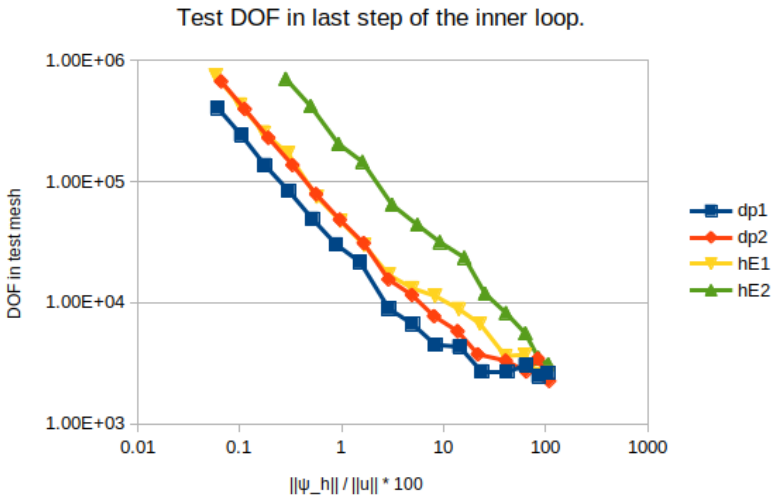


Figure 3.6: Comparison of global enrichment strategies, global h-refinement and increasing p_V: $\epsilon = 10^{-2}$, $p_U = 2$, p_V variable. This plot displays the number of DOF in the test mesh which achieved the desired test tolerance tol_V (last inner loop step) and the value of err_U for each outer loop step.

the extra DOF created where no improved resolution is required. However, we emphasize that the test mesh does not intent to cater to the resolution of u but to the resolution of ψ^h, which may need a quite different mesh to capture its behavior.

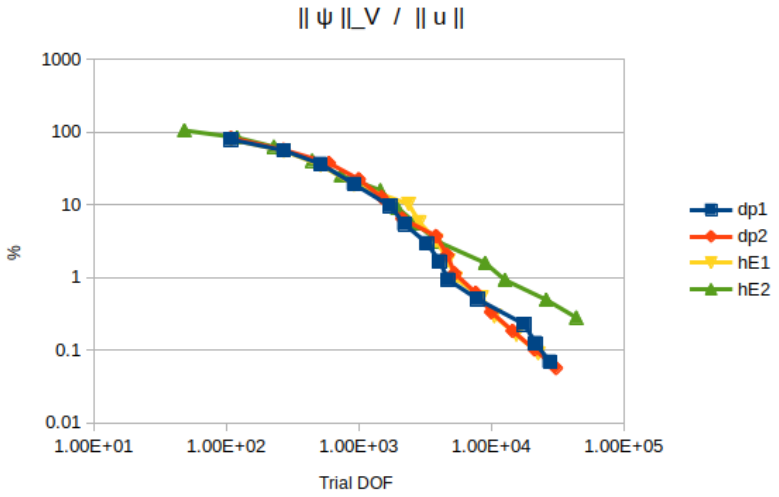


Figure 3.7: Comparison of enrichment strategies, global h-refinement and increasing p_V: $\epsilon = 10^{-3}$, $p_U = 2$, p_V variable. This plot displays the evolution of (normalized) $err_U = ||\psi_h||_V$ as the outer loop adaptively refines the trial mesh, increasing the number of (trial) DOF.

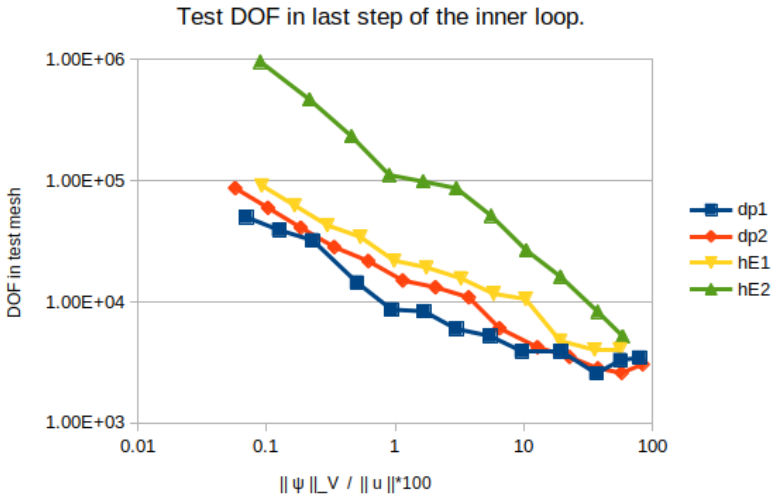


Figure 3.8: Comparison of global enrichment strategies, global h-refinement and increasing p_V: $\epsilon = 10^{-3}$, $p_U = 2$, p_V variable. This plot displays the number of DOF in the test mesh which achieved the desired test tolerance tol_V (last inner loop step) and the value of err_U for each outer loop step.

3.2.3 Inner loop stopping condition.

Next, the impact of tol_V is studied. It is used in the stopping condition of the inner loop:

---

if $err_V/\|\psi_h\|_V < tol_V$ **then**
  Exit inner loop

In principle, it is desirable to get an approximation ψ_h as close to ψ^h as possible; meaning a small value of our error estimate err_V. However, a better approximation implies a greater computational cost during each inner loop. In order to asses the quality in the resolution of ψ^h with respect to the inner loop cost, values of tol_V ranging from 5% to 100% are studied. The case without adaptive h-refinements for the test mesh is included with label *'none'*. In it, a single inner loop step is performed and do not seek for any improvement on the resolution of ψ_h, that is, it coincides with the traditional PGM with OTF. Fig. 3.9 shows the (outer loop) convergence of ψ_h for $\varepsilon = 10^{-2}$, where only very small differences are observed. Notice that the case 'none' displays a small increase at the second refinement step before joining the rest of the series' behavior.

The differences are slightly more visible using $\varepsilon = 10^{-3}$, shown in Figures 3.10 and 3.11 where only a few selected values are displayed. Without improving the resolution of ψ_h the monotone reduction of the error estimate err_U is lost. In practice, this indicates that without the inner loop (i.e., without taking advantage of the err_V estimate) ψ_h is a 'bad' approximation to ψ^h, especially during the first (trial) refinement steps. The largest value which still provides slightly better convergence behavior is tol_V=0.75, so we choose it as the most adequate value to be used. This value is further verified as a good choice in section 3.2.6 where parameter α is studied.

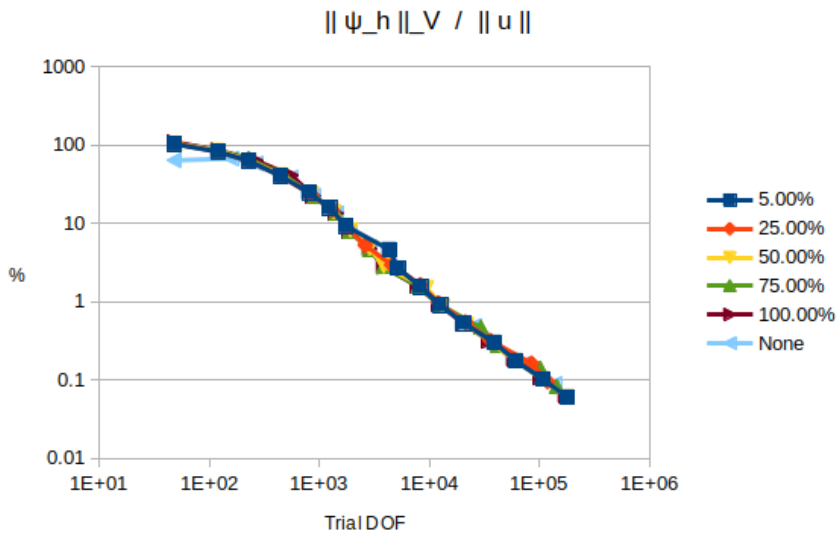


Figure 3.9: Convergence of $\|\psi_h\|_V/\|\mathsf{u}\|$ with respect to total trial DOF, for different values of tol_V, with $\epsilon = 10^{-2}$, $p_U = 2$, $p_V = 3$.

Additionally, Fig. 3.12 compares the convergence of $\|\psi_h\|_V$ and $\|\mathsf{u} - \mathsf{u}_h\|$ for the different tol_V values. For $tol_V \leq 75\%$ both curves become almost indistinguishable right after reaching the ordinate 100%, whereas

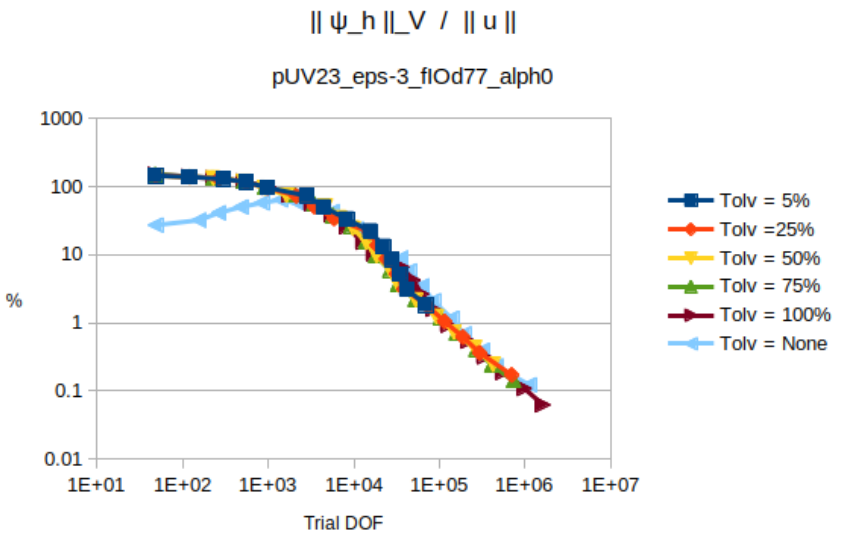


Figure 3.10: Convergence of $\|\psi_h\|_V/\|\mathsf{u}\|$ with respect to total trial DOF, for different values of tol_V, with $\epsilon = 10^{-3}$, $p_U = 2$, $p_V = 3$.

for $tol_V > 75\%$ both curves are congruent after reaching the ordinate 10%, reinforcing the notion of ψ_h not being sufficiently well resolved before exiting the inner loop.

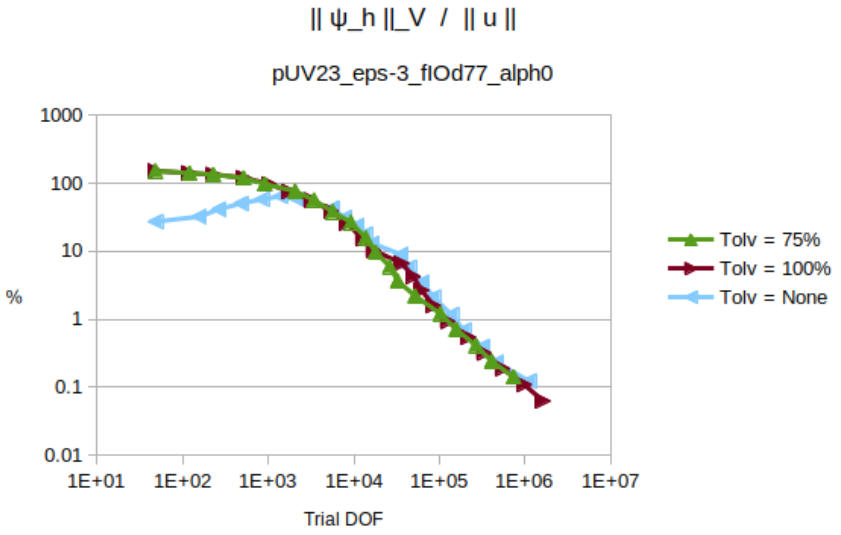
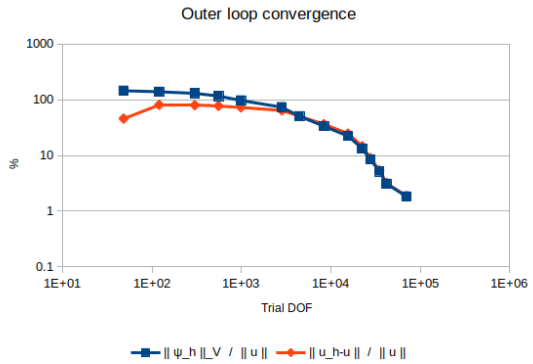
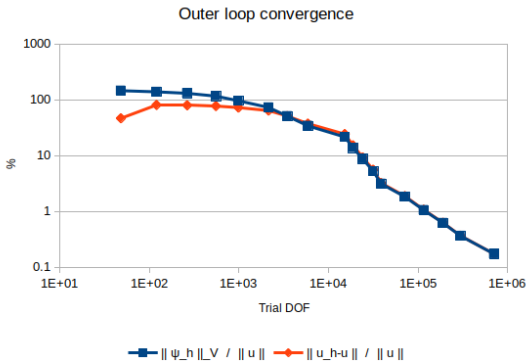


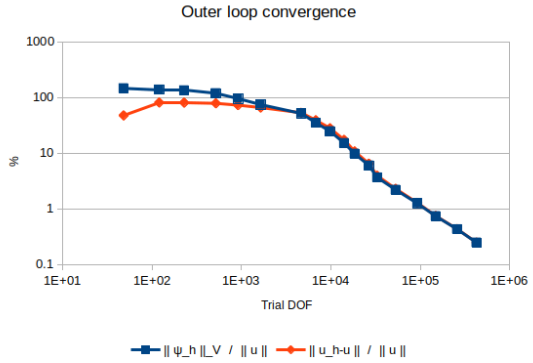
Figure 3.11: Convergence of $\|\psi_h\|_V/\|\mathsf{u}\|$ with respect to total trial DOF, for different values of tol_V, with $\epsilon = 10^{-3}$, $p_U = 2$, $p_V = 3$ (selected values).



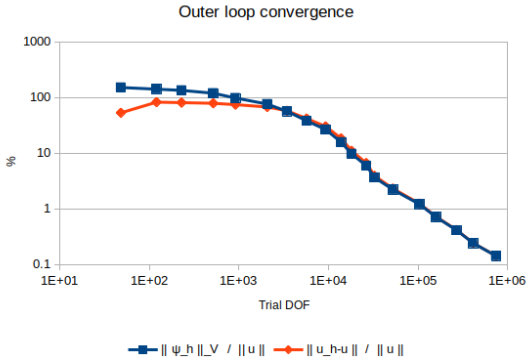
(a) $\epsilon = 10^{-3}$, $p_U = 2$, $p_V = 3$, $tol_V = 5\%$



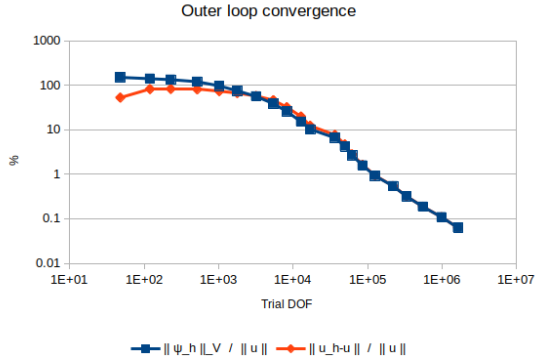
(b) $\epsilon = 10^{-3}$, $p_U = 2$, $p_V = 3$, $tol_V = 25\%$



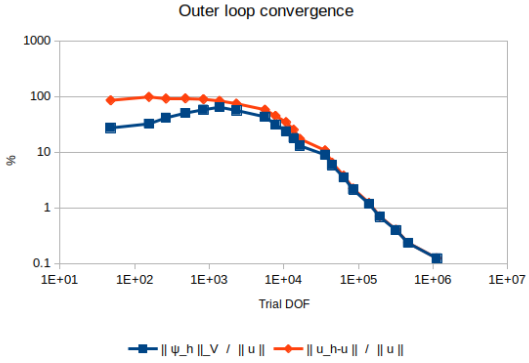
(c) $\epsilon = 10^{-3}$, $p_U = 2$, $p_V = 3$, $tol_V = 50\%$



(d) $\epsilon = 10^{-3}$, $p_U = 2$, $p_V = 3$, $tol_V = 75\%$



(e) $\epsilon = 10^{-3}$, $p_U = 2$, $p_V = 3$, $tol_V = 100\%$



(f) $\epsilon = 10^{-3}$, $p_U = 2$, $p_V = 3$, $tol_V =$ None

Figure 3.12: Outer loop convergence of the Riesz representation of the residual ψ_h in the test norm (normalized) and the L^2 error of the solution u_h (normalized), for different values of tol_V.

3.2.4 Test space: refinement strategies and refinement factor

The general idea of a refinement strategy is to refine elements with large contributions to the error estimate err_V^2. Two strategies are explored: the first one called the "greedy strategy", is based on the "poor man greedy *hp* algorithm" presented in [13] without p-refinement (it refines elements with large error contributions); the second one is the Doerfler strategy presented in [14] (it refines elements that account for a fraction of the total error). Pseudocodes for both strategies are presented next, where coefficients $Fact_G$ and $Fact_D$ are the so called refinement factors.

Algorithm 2: Greedy strategy

Input: Mesh V_h, element's error contributions to err_v^2
Output: Refined test mesh V_h
Find the element with greatest error contribution max_{err}
for $i = 1$ ***to*** *num. elems in* V_h **do**
  if *element* i *contribution to* $err_V^2 > Fact_G * max_{err}$ **then**
    Mark element i
Refine marked elements

Algorithm 3: Doerfler strategy

Input: Mesh V_h, element's error contributions to err_v^2
Output: Refined test mesh V_h
Compute the total error $total_{err}$
while $ref_{error} < Fact_D * total_{err}$ **do**
  From unmarked elements, find and mark element with greatest error contribution max_{err}
  Compute $ref_{error} = \sum_{\text{marked elements}}$ element error contribution
Refine marked elements

Results for the greedy strategy are presented in Fig. 3.13 displaying the convergence of the error estimate err_V for the first inner loop, with respect to the total number of test DOF for multiple values of $Fact_G$. While the total number of DOF required to achieve a particular value of err_V is very similar, they do not represent similar performance. The whole inner loop for a $Fact_G$ value of 0.3 requires only 9 steps, while the 0.9 value requires more than 20; implying a much greater computational cost (recall that both primal and dual problems are resolved at each inner loop step).

Similarly, results for the Doerfler strategy are shown in Fig. 3.14. Here, a smaller value of $Fact_D$ presents a slightly better convergence in err_V than $Fact_V = 0.7$, 0.9, however it also implies a greater number of steps to complete the inner loop. Since $Fact_V = 0.7$ presents only a slightly worse convergence with much fewer steps, this value is favored.

Finally, in Fig. 3.15, both strategies with the best refinement factors are compared. The results are similar with fewer number of steps required by the Doerfler strategy. For this reason, Doerfler with $Fact_D$=0.7 is used in the subsequent experiments.

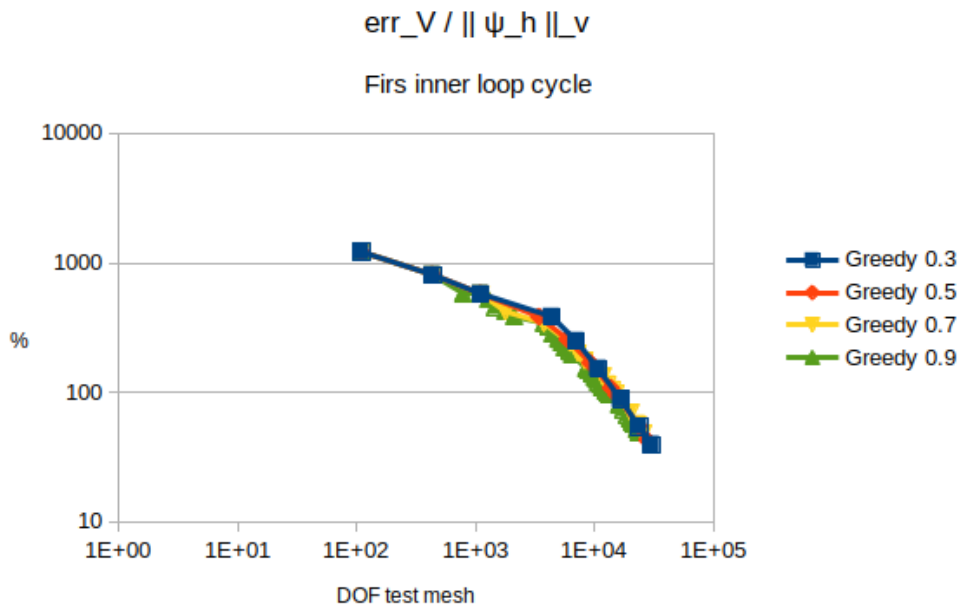


Figure 3.13: Convergence of the gap in energies estimate for the first inner loop cycle using the greedy refinement strategy, $p_U = 2$, $p_V = 3$, $\epsilon = 10^{-3}$.

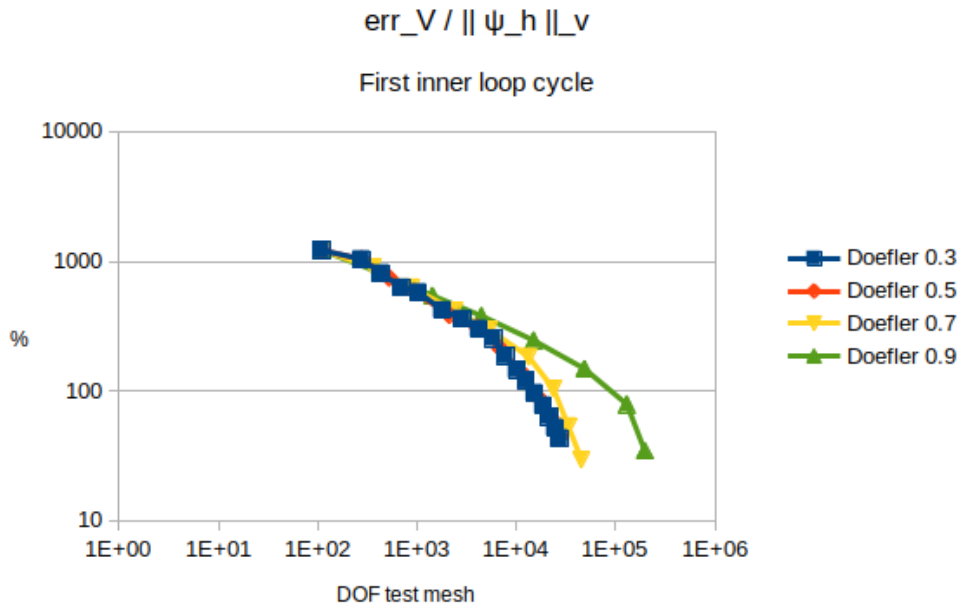


Figure 3.14: Convergence of the gap in energies estimate for the first inner loop cycle using the Doerfler refinement strategy, $p_U = 2$, $p_V = 3$, $\epsilon = 10^{-3}$.

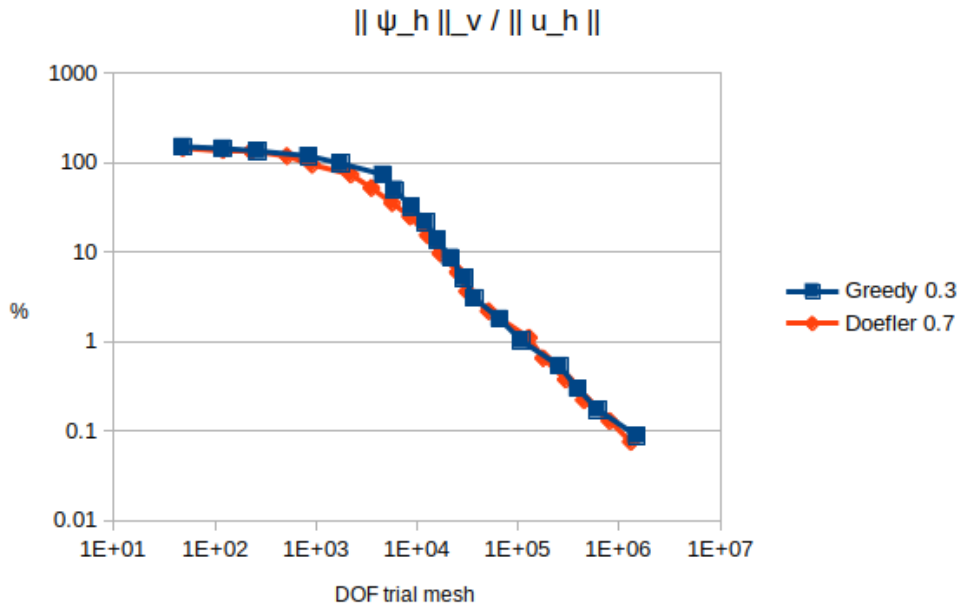


Figure 3.15: Convergence of the Riesz representation of the residual (outer loop) using the greedy and Doerfler refinement strategies, $p_U = 2$, $p_V = 3$, $\epsilon = 10^{-3}$.

3.2.5 Epsilon continuation and trial space refinement factor

Following the results from Section 3.2.4 (focused on the the inner loop), the Doerfler refinement strategy is employed for the outer loop as well, and the effect of the refinement factor $Fact_D$ is studied. Given the fact that we are interested in small values of ε, results of using a continuation strategy on ε are herein shown. Figures 3.16, 3.17 and 3.18 show the convergence results for a refinement factor $Fact_D$=0.7, 0.5, 0.3, respectively, with trial and test spaces of polynomial order $p_U = 2$ and $p_V = 3$. All values start with $\varepsilon = 10^{-2}$ and this value is cut in half once a $err_U/\|\mathsf{u}\|$ value of 1% is reached. To reach this error level when $\varepsilon = 10^{-4}$ all three plots show that a similar number of trial DOF are required. However, a greater value of $Fact_D$ implies much fewer outer loop steps for the same error reduction. For $Fact_D = 0.7$ (Fig. 3.16), three steps are needed for each value of ε to reach 1% (other than $\varepsilon = 10^{-2}$, clearly) whereas for $Fact_D = 0.5$ (Fig. 3.17) four steps are needed and for $Fact_D = 0.3$ (Fig. 3.18) 7 steps are needed, resulting in a much greater computational cost without improved convergence.

Similar results are obtained with higher polynomial orders $p_U = 3$ and $p_V = 4$ as shown in Figures 3.19, 3.20 and 3.20, where using $Fact_D = 0.7$ requires only three or four outer loop steps and $Fact_D = 0.3$ requires seven steps for most ε. For this reason, the value of 0.7 for $Fact_D$ is chosen in the outer loop refinement strategy.

Finally, in Fig. 3.22 the results are shown for high polynomial orders $p_U = 6$, $p_V = 7$, which display a slightly better performance in terms of DOF.

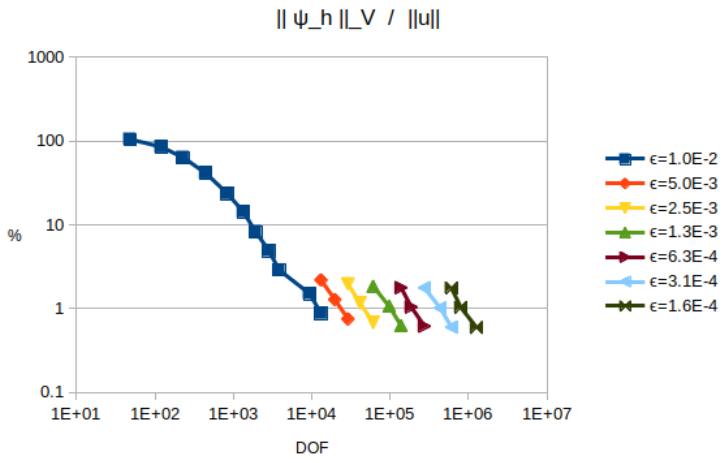


Figure 3.16: Convergence using the continuation technique on ϵ, with a refinement factor of 0.7 for the trial mesh, $p_U = 2$, $p_V = 3$.

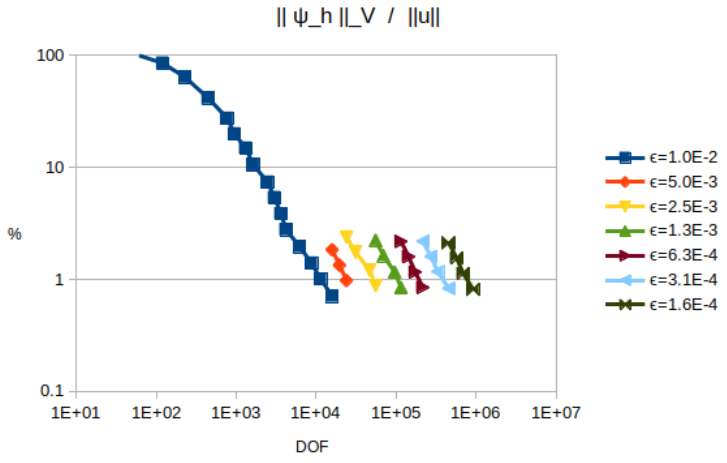


Figure 3.17: Convergence using the continuation technique on ϵ, with a refinement factor of 0.5 for the trial mesh, $p_U = 2$, $p_V = 3$.

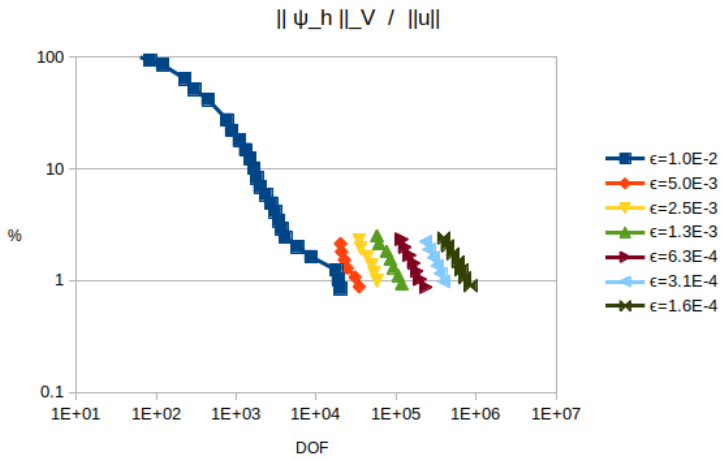


Figure 3.18: Convergence using the continuation technique on ϵ, with a refinement factor of 0.3 for the trial mesh, $p_U = 2$, $p_V = 3$.

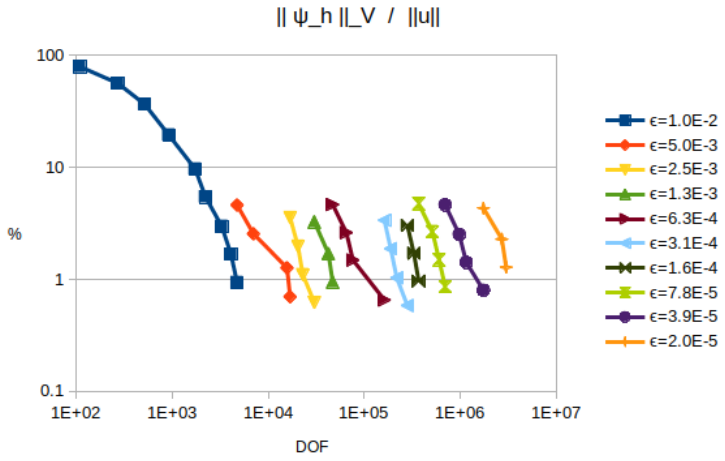


Figure 3.19: Convergence using the continuation technique on ϵ, with a refinement factor of 0.7 for the trial mesh, $p_U = 3$, $p_V = 4$.

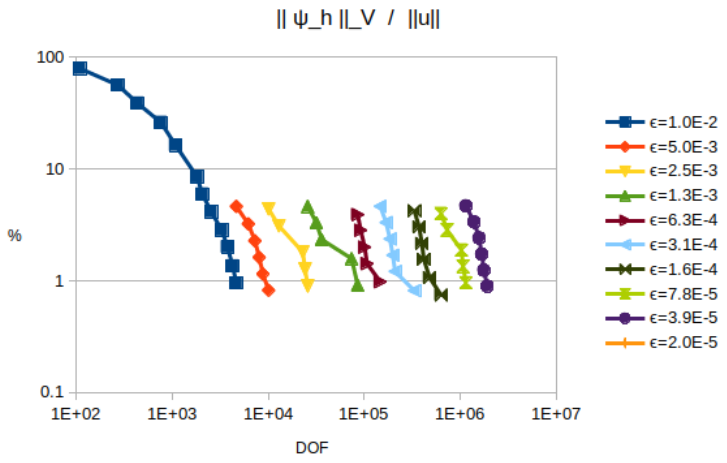


Figure 3.20: Convergence using the continuation technique on ϵ, with a refinement factor of 0.5 for the trial mesh, $p_U = 3$, $p_V = 4$.

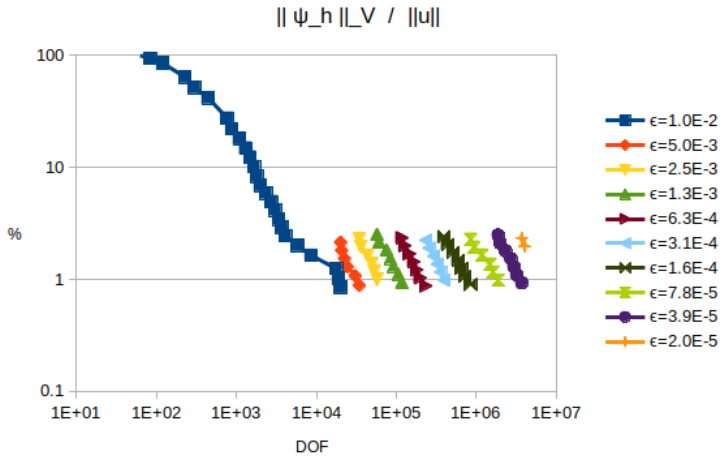


Figure 3.21: Convergence using the continuation technique on ϵ, with a refinement factor of 0.3 for the trial mesh, $p_U = 3$, $p_V = 4$.

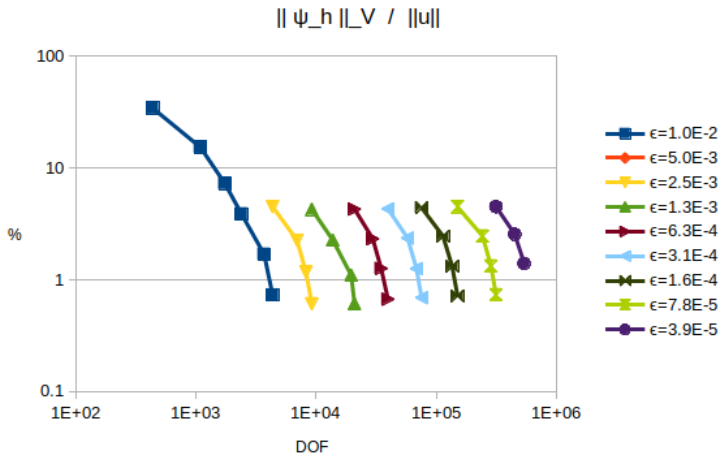


Figure 3.22: Convergence using the continuation technique on ϵ, with a refinement factor of 0.7 for the trial mesh, $p_U = 6$, $p_V = 7$.

3.2.6 Role of parameter α in the energy functionals and the error estimator

In the definition of the test norm required by the DAM, the parameter α is required in order to formulate the dual problem and ultimately obtain the estimator err_V. To determine a good value for α, values of different orders of magnitude are investigated. Clearly, α significantly impacts the value of both energy functionals ((2.7) and (2.8)) and thus our error estimator err_V. Fig. 3.23 presents the primal and dual energies throughout the first inner loop for multiple values of α ranging from 10^{-2} to 10^3. Solid lines correspond to the primal energies J, and dashed lines correspond to dual energies J^*. The plots allow to observe the expected primal and dual energies convergence to a same value (for each pair), which should correspond to $J(\psi^h) = J^*(\phi^h)$, so the series' final value is taken as a good approximation to the exact energy. Fig. 3.24 presents the energy curves normalized with the exact energy.

It is known that two major properties necessary for successful error estimators are efficiency and reliability [1]. They are, respectively :

$$\eta \leq C_1 \|e\|_E \quad , \quad \|e\|_E \leq C_2 \eta \ ,$$

where η is the error estimator, and $\|e\|_E$ is the error in (its corresponding) energy norm, and constants C_1 and C_2 are independent of the mesh. In the context of our estimate $\eta = err_V$ (for both problems), for the primal problem $\|e\|_E^2 = 2J(\psi_h) - 2J(\psi^h)$ and for the dual problem $\|e\|_E^2 = 2J^*(\phi^h) - 2J^*(\phi_h)$. The inequalities in (2.12) show that err_V is a *reliable* error estimate for either problem with constant $C_2 = 1$. Regarding efficiency, there is no theoretical guarantee (see [10, Remark 7]), so we intend to assess that property with numerical experiments. For that purpose, the so called *effectivity index* $\theta = \eta / \|e\|_E$ is herein computed. Having a bounded θ is an indication of efficiency of the estimator η (the bound being a candidate for the above constant C_1). Fig. 3.25 displays the behavior of θ as the test mesh is refined, for multiple values of α. For $\alpha = 10^{-1}, 10^{-2}$ the effectivity index is large, and thus such α values are not favored. Fig. 3.26 more clearly illustrates that for $\alpha = 10^2$ the index θ slightly degrades as the mesh is refined; the curve for $\alpha = 10$ remains somewhat bounded below 1.5, while for $\alpha = 1$ the effectivity index remains bounded below 2 with slight improvements as the mesh is refined. Based on the effectivity index, either values of 1 or 10 could be chosen, then for simplicity $\alpha = 1$ is taken.

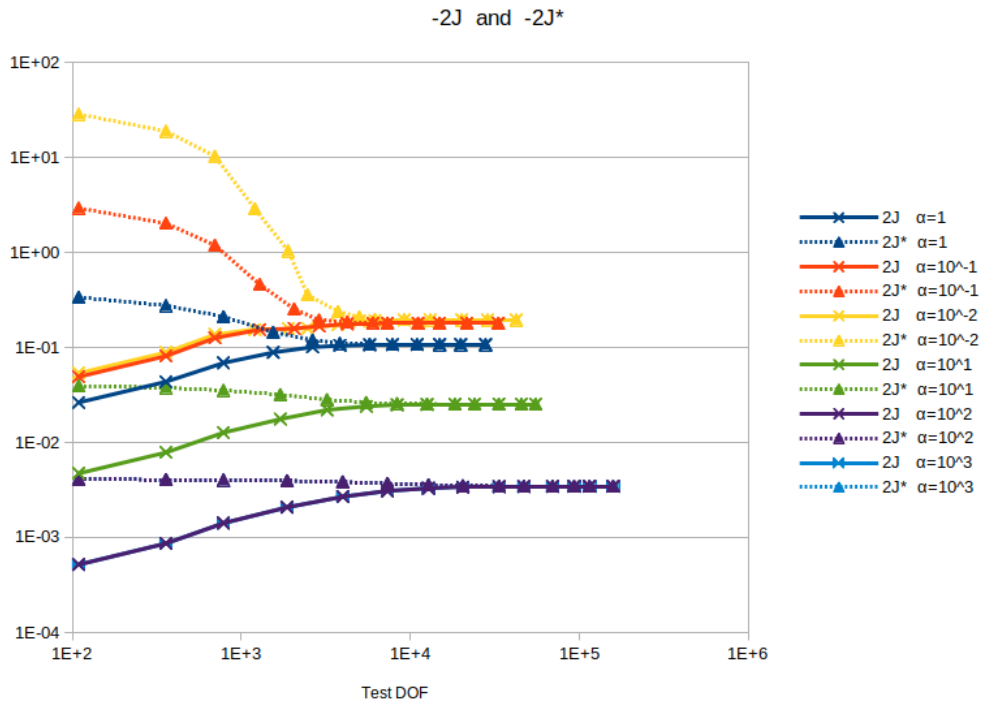


Figure 3.23: Negative of primal and dual energies, $\epsilon = 10^{-2}$, $p_U = 2$, $p_V = 3$

Figures 3.27 (for $\varepsilon = 10^{-2}$) and 3.28 (for $\varepsilon = 10^{-3}$) show the evolution of our error estimator err_V normalized with $\|\psi_h\|$, which is the quantity employed for the inner loop stopping condition. Notice that while the ideal normalization for err_V would be ψ^h, at any point during the inner loop only $\|\psi_h\|$ is known and therefore the latter is used. Here it is observed that high values of α start with smaller err_V but have slower convergence rates, and small values of α have higher convergence rates but starting at larger err_V values. The convergence curves seem to meet close to 100%, and therefore tol_V values close to this are less sensitive to α and thus desirable in terms of our inner loop stopping condition. A tol_V value of 75% satisfies this criterion, and has been favored in the above studies as well.

Finally, in Fig. 3.29, it is verified that smaller values of α are favored as they retain the agreement between err_U (solid line) and the error in the L^2-norm (dashed line); for $\alpha = 10^2$ both curves do not mach as quickly with adaptive refinements. This is coherent with the fact that small α values result in a test norm closer to the 'ideal' one.

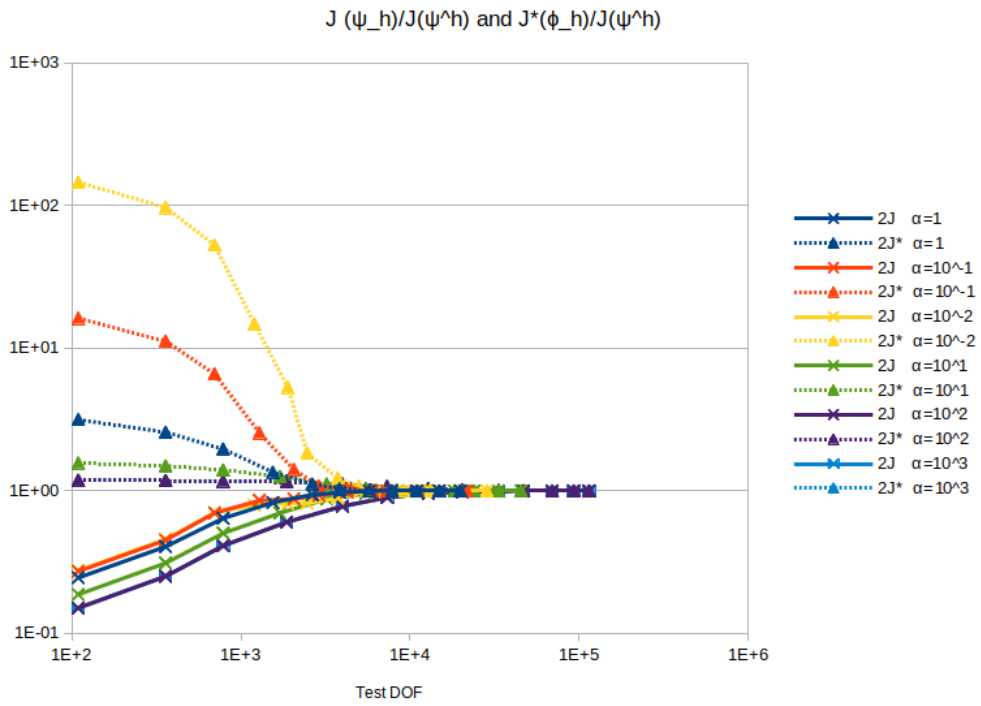


Figure 3.24: Primal and dual energies normalized by the final value of energy, $\epsilon = 10^{-2}$, $p_U = 2$, $p_V = 3$.

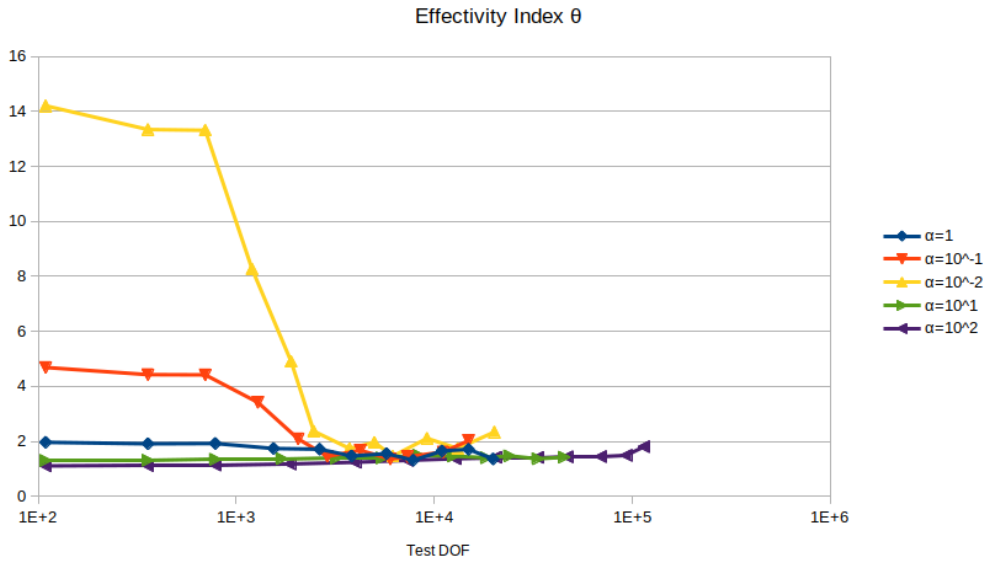


Figure 3.25: Effectivity index for the first inner loop, $\epsilon = 10^{-2}$, $p_U = 2$, $p_V = 3$.

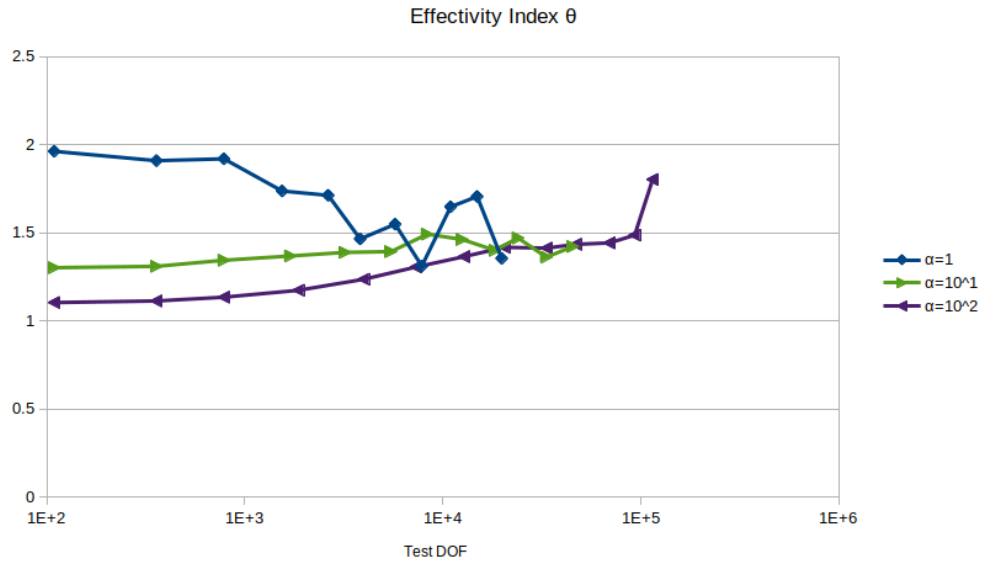


Figure 3.26: Effectivity index for the first inner loop, $\epsilon = 10^{-2}$, $p_U = 2$, $p_V = 3$ (selected values).

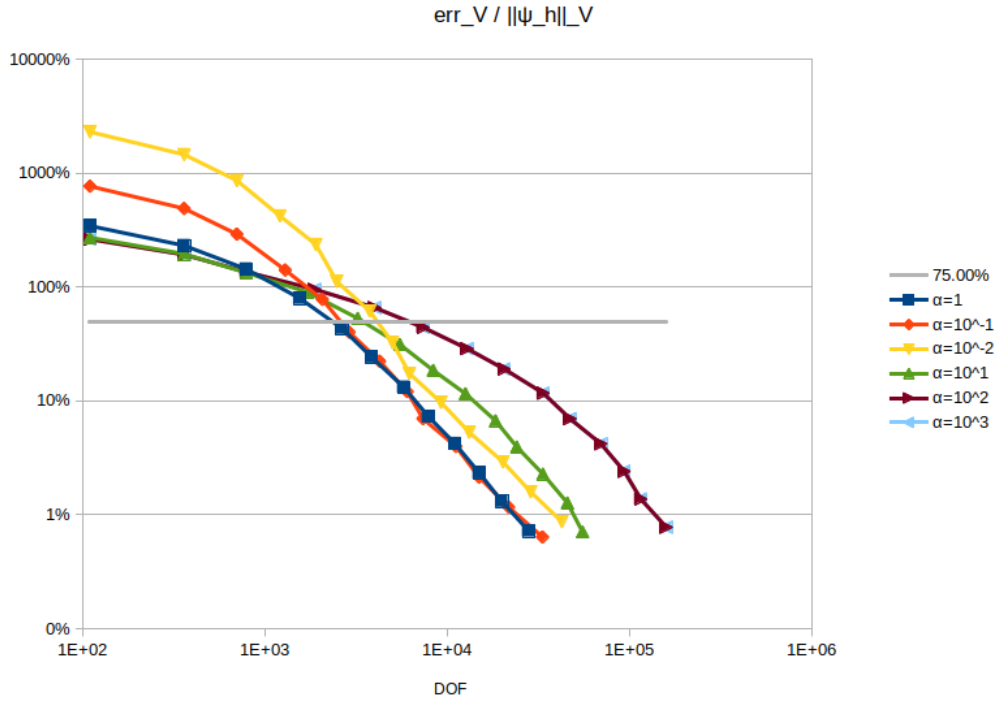


Figure 3.27: Error estimator evolution, normalized by $||\psi_h||_V$ $\epsilon = 10^{-2}$, $p_U = 2$, $p_V = 3$.

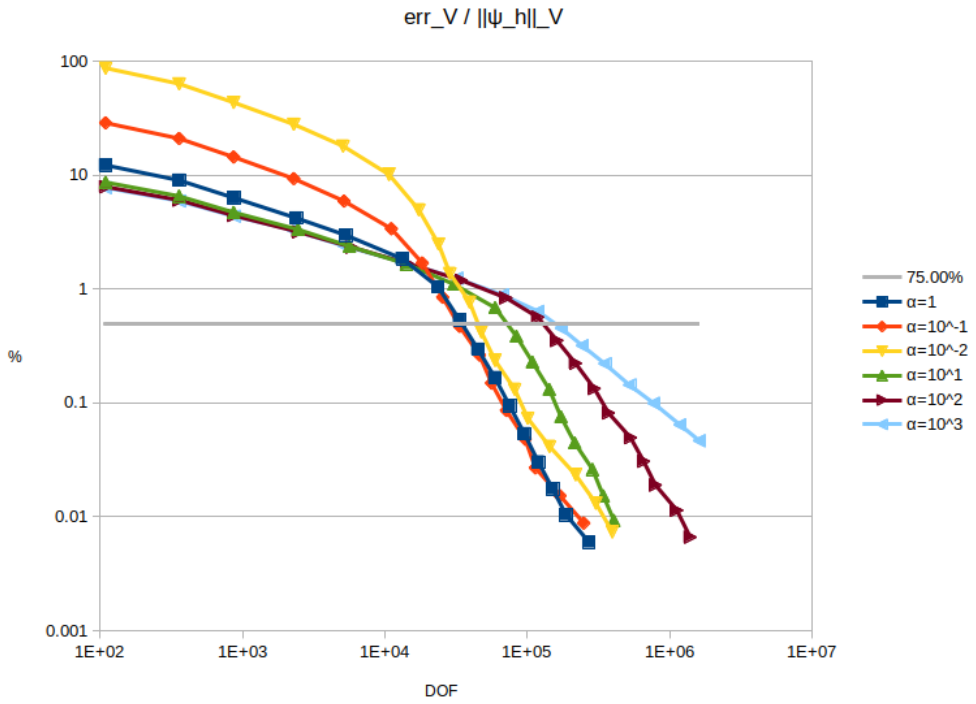


Figure 3.28: Gap between the primal and dual energies, normalized by $\|\psi_h\|_V$ $\epsilon = 10^{-3}$, $p_U = 2$, $p_V = 3$.

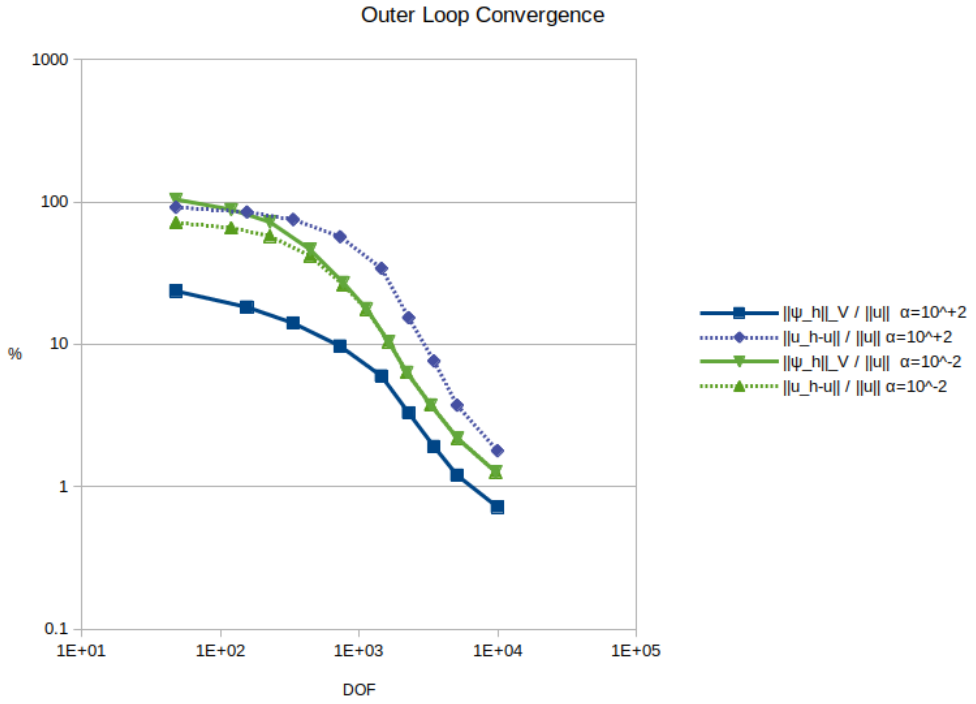


Figure 3.29: Energy norm of the Riesz representation of the residual and L^2 error of the solution u. $\epsilon = 10^{-2}$, $p_U = 2$, $p_V = 3$.

4 Weakly conforming discretization

4.1 Weakly conforming vs. conforming methods

For a ultra-weak variational problem,

$$\begin{cases} u \in \boldsymbol{L}^2(\Omega) \\ (u, A^*v) = (f, v) \quad v \in D(A^*)\,, \end{cases} \tag{4.1}$$

with

$$A : \boldsymbol{L}^2(\Omega) \supset D(A) \to \boldsymbol{L}^2(\Omega)$$

a closed operator representing a system of first order PDEs with BCs reflected in the definition of $D(A)$, and A^* its $\boldsymbol{L}^2$-adjoint, the test space is equiped with the adjoint graph norm,

$$\|v\|_V^2 = (v, v)_V, \qquad (v, \delta v)_V := (A^*v, A^*\delta v) + (v, \delta v)\,.$$

The graph spaces $H_A(\Omega)$ and $H_{A^*}(\Omega)$ are defined as:

$$H_A(\Omega) := \{w \in \boldsymbol{L}^2(\Omega) : Aw \in \boldsymbol{L}^2(\Omega)\}, \qquad H_{A^*}(\Omega) := \{w \in \boldsymbol{L}^2(\Omega) : A^*w \in \boldsymbol{L}^2(\Omega)\}.$$

Recall $D(A)$ and $D(A^*)$ are the corresponding subspaces of $H_A(\Omega)$ and $H_{A^*}(\Omega)$, respectively, which satisfy the appropriate homogeneous boundary conditions.

Conforming discretization. Problem (3.2) can be discretized with the *practical PGM with OTF* as follows. We select conforming subspaces $U_h \subset \boldsymbol{L}^2(\Omega)$, $V_h \subset V := D(A^*)$, $\dim V_h > \dim U_h$, and introduce the subspace $V_h^{\rm opt} \subset V_h$ of discretely optimal test functions v_h corresponding to $w_h \in U_h$, given by:

$$\begin{cases} v_h \in V_h \\ (v_h, \delta v_h)_V = (w_h, A^*\delta v_h) \quad \delta v_h \in V_h \end{cases}$$

The discrete problem looks as follows.

$$\begin{cases} u_h \in U_h \\ (u_h, A^*v_h) = (f, v_h) \quad v_h \in V_h^{\rm opt}\,. \end{cases} \tag{4.2}$$

The Galerkin orthogonality property holds:

$$(u - u_h, A^*v_h) = 0 \quad v_h \in V_h^{\rm opt}\,.$$

We assume the satisfaction of the discrete inf-sup condition,

$$\beta_h \|u_h\| \le \sup_{v_h \in V_h} \frac{|(u_h, A^*v_h)|}{\|v_h\|_V} = \sup_{v_h \in V_h^{\rm opt}} \frac{|(u_h, A^*v_h)|}{\|v_h\|_V} \tag{4.3}$$

Babuška's argument [3] ,

$$\begin{aligned} \|u - u_h\| &\le \|u - w_h\| + \|u_h - w_h\| && \text{(triangle inequality with arbitrary } w_h \in U_h) \\ &\le \|u - w_h\| + \beta_h^{-1} \sup_{v_h \in V_h^{\rm opt}} \frac{|(u_h - w_h, A^*v_h)|}{\|v_h\|_V} && \text{(inf-sup condition ((4.3)))} \\ &= \|u - w_h\| + \beta_h^{-1} \sup_{v_h \in V_h^{\rm opt}} \frac{|(u - w_h, A^*v_h)|}{\|v_h\|_V} && \text{(Galerkin orthogonality)} \\ &\le (1 + \beta_h^{-1})\|u - w_h\|\,, \end{aligned}$$

leads then to the error estimate:

$$\|u - u_h\| \leq (1 + \beta_h^{-1}) \inf_{w_h \in U_h} \|u - w_h\| \,. \tag{4.4}$$

In fact, a more sophisticated argument (for Hilbert spaces) gives a sharper estimate, [9, Lemma 4.1.1] ,

$$\|u - u_h\| \leq \beta_h^{-1} \inf_{w_h \in U_h} \|u - w_h\| \,. \tag{4.5}$$

Weakly conforming discretization. Let the mesh $\mathcal{T}_h$ be a non-overlapping partition of Ω into open elements with Lipschitz boundary, so that

$$\bar{\Omega} = \cup\{\bar{K}; K \in \mathcal{T}_h\}$$

and

$$\Gamma_h = \cup\{\partial K; K \in \mathcal{T}_h\}$$

denotes the mesh skeleton. The resulting broken UW variational formulation is:

$$\begin{cases} u \in \boldsymbol{L}^2(\Omega),\, \hat{u} \in \hat{U}, \\ (u, A_h^* v) + \langle \hat{u}, v \rangle_{\Gamma_h} = (f, v) \quad v \in V_{br}; \end{cases} \tag{4.6}$$

where $\hat{U}$ is the appropriate space of traces defined on Γ_h, A_h^* is the elementwise adjoint operator (i.e., $(A_h^* v)|_K = A^*(v|_K) \in \boldsymbol{L}^2(K)$ for every element $K \in \mathcal{T}_h$), $V_{br} := H_{A^*}(\mathcal{T}_h)$ is the broken test space equipped with the broken adjoint graph norm

$$H_{A^*}(\mathcal{T}_h) := \{w \in \boldsymbol{L}^2(\Omega) : w|_K \in H_{A^*}(K)\ \forall K \in \mathcal{T}_h\},$$

$$\|w\|_{V_{br}}^2 := \|w\|^2 + \underbrace{\sum_{K \in \mathcal{T}_h} \|A^*(w|_K)\|^2}_{\|A_h^* w\|^2}$$

in the understanding that A^* is overloaded to be defined on each element, and $\langle \cdot, \cdot \rangle_{\Gamma_h}$ is the duality pairing between space $\hat{U}$ and the traces of $H_{A^*}(\mathcal{T}_h)$,

$$\langle \hat{u}, v \rangle_{\Gamma_h} := \sum_{K \in \mathcal{T}_h} \langle \hat{u}, \operatorname{tr} v_K \rangle_{\partial K} \ ,$$

where tr is an appropriate local trace operator on $H_{A^*}(K)$. An important detail when introducing broken test spaces is that these contain the conforming test spaces, i.e., $V \subset V_{br}$, and that the norm $\|\cdot\|_{V_{br}}$ restricted to V is equivalent to $\|\cdot\|_V$. From this point on, the broken norm is used for the conforming test spaces as well.

By introducing a finite-dimensional approximation space for traces $\hat{U}_h \subset \hat{U}$ and broken test functions $V_{br,h} \subset V_{br}$, the discretization of the broken problem (4.6) with the PGM scheme with optimal broken test functions results in:

$$\begin{cases} u_h \in U_h,\, \hat{u}_h \in \hat{U}_h \\ (u_h, A_h^* v_h) + \langle \hat{u}_h, v_h \rangle_{\Gamma_h} = (f, v_h) \quad v_h \in V_{br,h}^{\text{opt}} \end{cases} \tag{4.7}$$

where the space of optimal test functions $V_{br,h}^{\text{opt}}$ is the span of test functions v_h defined by the variational problem:

$$\begin{cases} v_h \in V_{br,h} \\ (v_h, \delta v_h)_{V_{br}} = (w_h, A_h^* \delta v_h) + \langle \hat{w}_h, \delta v_h \rangle_{\Gamma_h} \quad \delta v_h \in V_{br,h} \end{cases} \tag{4.8}$$

corresponding to $(w_h, \hat{w}_h) \in U_h \times \hat{U}_h$.

Introducing the subspace of *weakly conforming* test functions,

$$V_{wc,h} = \{ v_h \in V_{br,h} \,:\, \langle \hat{w}_h, v_h \rangle_{\Gamma_h} = 0 \quad \forall\, \hat{w}_h \in \hat{U}_h \} \subset V_{br}\,, \tag{4.9}$$

the formulation (4.7) can be seen as a non-conforming discretization of the original problem (3.2), namely,

$$\begin{cases} u_h \in U_h \\ (u_h, A_h^* v_h) = (f, v_h) \quad v_h \in V_{wc,h}^{\text{opt}} \end{cases} \tag{4.10}$$

where, analogously, the space of optimal weakly conforming test functions $V_{wc,h}^{\text{opt}}$ is the span of the functions v_h associated to each $w_h \in U_h$ that solve the problem

$$\begin{cases} v_h \in V_{wc,h} \\ (v_h, \delta v_h)_{V_{br}} = (w_h, A_h^* \delta v_h) \quad \delta v_h \in V_{wc,h} \end{cases} . \tag{4.11}$$

Remark 4.1. Note that solution to (4.10) coincides with the standard DPG solution of (4.7). To see this, it is convenient to recast both problems as mixed problems. The broken UW problem (4.7) is equivalent to:

$$\begin{cases} \psi_h \in V_h, u_h \in U_h,\, \hat{u_h} \in \hat{U}_h & & & \\ (\psi_h, v_h) + b(u_h, v_h) + \langle \hat{u_h}, v_h \rangle & = (f, v_h) & v_h \in V_h & \text{(a)} \\ b(\delta u_h, \psi_h) & = 0 & \delta u_h \in U_h & \text{(b)} \\ \langle \delta \hat{u_h}, \psi_h \rangle & = 0 & \delta \hat{u_h} \in \hat{U}_h & \text{(c)} \end{cases} \tag{4.12}$$

where $b(u, v) = (u_h, A_h^* v_h)$. This problem has a unique solution assuming that the following inf-sup condition is satisfied:

$$\sup_{v_h \in V_h} \frac{|b(u_h, v_h) + \langle \hat{u_h}, v_h \rangle|}{\|v_h\|_{V_{br}}} \geq \gamma_1 \|(u_h, \hat{u_h})\|\,. \tag{4.13}$$

The weakly conforming problem (4.10) is equivalent to:

$$\begin{cases} \psi \in V_{wc,h}, u_h \in U_h & & & \\ (\psi_h, v_h) + b(u_h, v_h) & = (f, v_h) & v \in V_{wc,h} & \text{(a)} \\ b(\delta u_h, \psi_h) & = 0 & \delta u \in U & \text{(b)} \end{cases} \tag{4.14}$$

and similarly has unique solution given that the inf-sup condition is satisfied:

$$\sup_{v_h \in V_{wc,h}} \frac{|b(u_h, v_h)|}{\|v_h\|_{V_{br}}} \geq \gamma_2 \|u_h\|\,. \tag{4.15}$$

Let $({}^{br}\psi_h, {}^{br}u_h, {}^{br}\hat{u_h})$ be the solution to (4.12); from (4.12)(b) and (c) ${}^{br}\psi \in V_{wc,h}$ and (4.14)(b) is satisfied. From (4.12)(a) $({}^{br}\psi, {}^{br}u, {}^{br}\hat{u})$ satisfy (4.14) (a) since $\langle {}^{br}\hat{u_h}, v_h \rangle = 0$ for $v_h \in V_{wc,h}$ by definition of $V_{wc,h}$.

We have shown that both formulations are equivalent, but the advantage of using the whole broken space $V_{br,h}$ (i.e., the DPG method) is that it enables testing with OTF without ever solving the global weakly conforming OTF problem (4.11).

How to estimate the error for this non-conforming discretization? One obvious answer is simply to invoke the error estimate for the DPG method. Then an error estimate for the group variable $\mathsf{u}_h := (u_h, \hat{u}_h)$ can be obtained. But a more direct argument can also be used, aiming at the error estimation of u_h only. We unpack the argument behind the stability of broken formulation, and follow directly with Babuška's argument for error estimation.

To start, the inf-sup constant corresponding to the weakly-conforming test space is introduced:

$$\widetilde{\beta}_h \|u_h\| \leq \sup_{v_h \in V_{wc,h}} \frac{|(u_h, A_h^* v_h)|}{\|v_h\|_V} = \sup_{v_h \in V_{wc,h}^{\text{opt}}} \frac{|(u_h, A_h^* v_h)|}{\|v_h\|_V} \,. \tag{4.16}$$

Note that, due to the fact that the supremum is taken over a larger space, the new constant $\widetilde{\beta}_h \geq \beta_h$.

By Babuška's argument, the derivation follows:

$$\begin{aligned}
\|u - u_h\| &\leq \|u - w_h\| + \|u_h - w_h\| && \text{(triangle inequality)} \\
&\leq \|u - w_h\| + \widetilde{\beta}_h^{-1} \sup_{v_h \in V_{wc,h}^{\text{opt}}} \frac{|(u_h - w_h, A^* v_h)|}{\|v_h\|_V} && \text{(inf-sup condition ((4.16)))} \\
&\leq \|u - w_h\| + \widetilde{\beta}_h^{-1} \sup_{v_h \in V_{wc,h}^{\text{opt}}} \frac{|(u_h - w_h, A^* v_h) + \langle \hat{u}_h - \hat{w}_h, v_h \rangle_{\Gamma_h}|}{\|v_h\|_V} && \text{(a zero term added)} \\
&\leq \|u - w_h\| + \widetilde{\beta}_h^{-1} \sup_{v_h \in V_{wc,h}^{\text{opt}}} \frac{|(u - w_h, A^* v_h) + \langle \hat{u} - \hat{w}_h, v_h \rangle_{\Gamma_h}|}{\|v_h\|_V} && \text{(Galerkin orthogonality)} \\
&\leq (1 + \widetilde{\beta}_h^{-1}) \|u - w_h\| + \widetilde{\beta}_h^{-1} \sup_{v_h \in V_{wc,h}^{\text{opt}}} \frac{|\langle \hat{u} - \hat{w}_h, v_h \rangle_{\Gamma_h}|}{\|v_h\|_V} \,.
\end{aligned}$$

The second term corresponds precisely to the consistency term in the classical Strang's lemma. It can be bounded with the best approximation error for trace $\hat{u}$ measured in the minimum energy extension norm,

$$\begin{aligned}
\sup_{v_h \in V_{wc,h}^{\text{opt}}} \frac{|\langle \hat{u} - \hat{w}_h, v_h \rangle_{\Gamma_h}|}{\|v_h\|_V} &\leq \sup_{v_h \in V_h(\mathcal{T}_h)} \frac{|\langle \hat{u} - \hat{w}_h, v_h \rangle_{\Gamma_h}|}{\|v_h\|_V} \\
&\leq \|\hat{u} - \hat{w}_h\|_E
\end{aligned}$$

where

$$\begin{aligned}
\|\hat{u}\|_E^2 &= \sum_{K \in \mathcal{T}_h} \|\hat{u}\|_{K,E}^2 \\
\|\hat{u}\|_{K,E}^2 &= \inf_{U \in H_A(K),\, U|_{\partial K} = \hat{u}} \|U\|_{H_A(K)}^2 \\
\|U\|_{H_A(K)}^2 &:= \|AU\|_{L^2(K)}^2 + \|U\|_{L^2(K)}^2 \,.
\end{aligned}$$

Remark 4.2. The a-priori error estimates do not tell us which of the two discretizations versions: conforming or weakly conforming, should deliver a smaller error. It is a 'toss-up'. The weakly conforming version enjoys a better (larger) inf-sup constant but the estimate includes now an additional consistency (trace) term.

A note of caution must be introduced here regarding the error estimates. The duality-based error estimates provides a provable error estimate for $\psi^h - \psi_h$ in the conforming discretization, but we also intend to use it for the weakly conforming one without developing at this time the underlying theory, being aware of the "crime" resulting from the failure of integration-by-parts formula for weakly conforming functions.

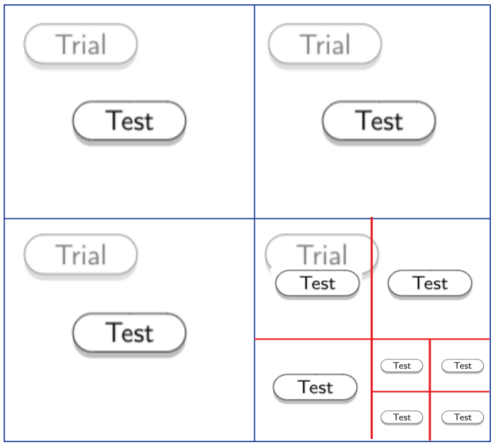


Figure 4.1: Use of a single data structure for both meshes. The elements and nodes corresponding to each are differentiated through flags. This is possible as the test mesh is always a sub-mesh of the trial one.

4.2 Implementation

Given that the weakly conforming discretization is equivalent to that corresponding to broken test spaces, the implementation commonly used for the DPG method is used, with required modifications to incorporate a sub-mesh for the test space. Unlike the implementation for the conforming discretization, the present case uses a single data structure, having an extra flag to identify the trial elements and nodes belonging to the trial mesh as well. Fig. 4.1 displays the conceptually overlapping elements, but using only one data structure. This simplifies the programming implementation since the routines only actually operate on a single data structure (although careful modifications are required for flagging the trial elements). Mesh refinements are performed seamlessly, note that the test mesh is adaptively refined only within the inner loop, for which the trial mesh remains fixed. Then after the trial elements' contributions to err_U^2 are computed the test mesh is no longer needed, the test elements (which are not flagged as trial) can be discarded and the trial mesh is refined. Additionally, while the conforming method is based on monolythic construction of the mixed method matrices, for the weakly conforming method the localization of the test norm is exploited in order to perform static condensation of element test DOF.

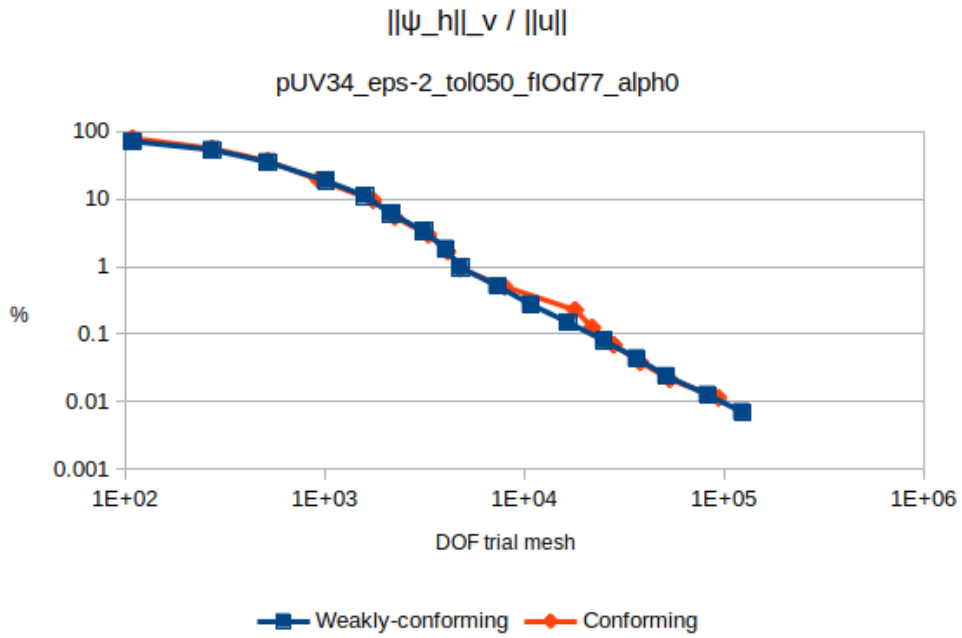


Figure 4.2: Convergence of the Riesz representation of the residual ψ_h for the conforming and weakly-conforming discretizations, $p_U = 3$, $p_V = 4$, $\epsilon = 10^{-2}$.

4.3 Numerical results, comparison with conforming discretization

Since the a-priori error estimation cannot decide which of the two methodologies is superior, we proceed with numerical experiments to compare the conforming and weakly conforming versions of the DAM. For this purpose, values of $\varepsilon = 10^{-2}, 10^{-3}$ are used, and multiple polynomial degrees. Figures 4.2, 4.3 and 4.4 (with $p_U = 3, 4, 5$, respectively and all with $\varepsilon = 10^{-2}$) show the convergence of the Riesz representation of the residual for the outer loop, where both discretizations (conforming and weakly conforming) present the very same convergence since the pre-asymptotic regime. With smaller $\varepsilon = 10^{-3}$ (Figures 4.5 and 4.6) only small differences between conforming and weakly convergence are observed in the pre-asymptotic regime and at very small error values. Overall, there is no significant difference between both methods' convergence.

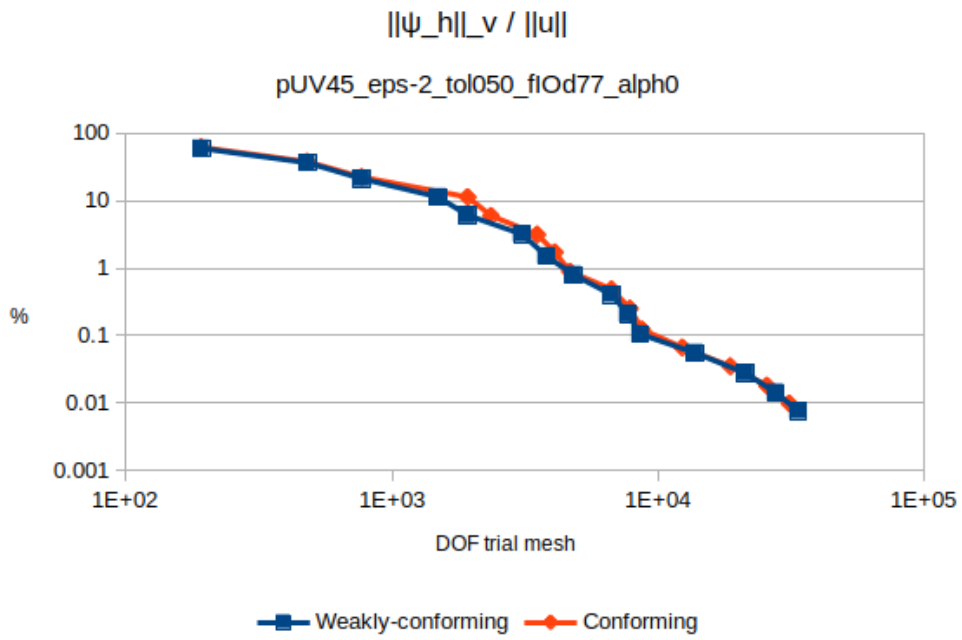


Figure 4.3: Convergence of the Riesz representation of the residual ψ_h for the conforming and weakly-conforming discretizations, $p_U = 4$, $p_V = 5$, $\epsilon = 10^{-2}$.

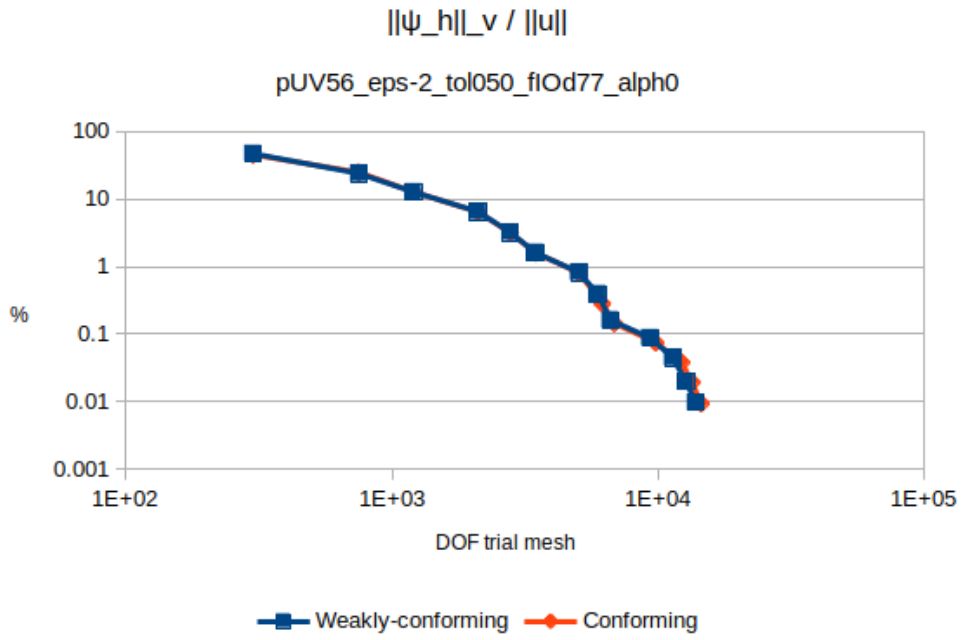


Figure 4.4: Convergence of the Riesz representation of the residual ψ_h for the conforming and weakly-conforming discretizations, $p_U = 5$, $p_V = 6$, $\epsilon = 10^{-2}$.

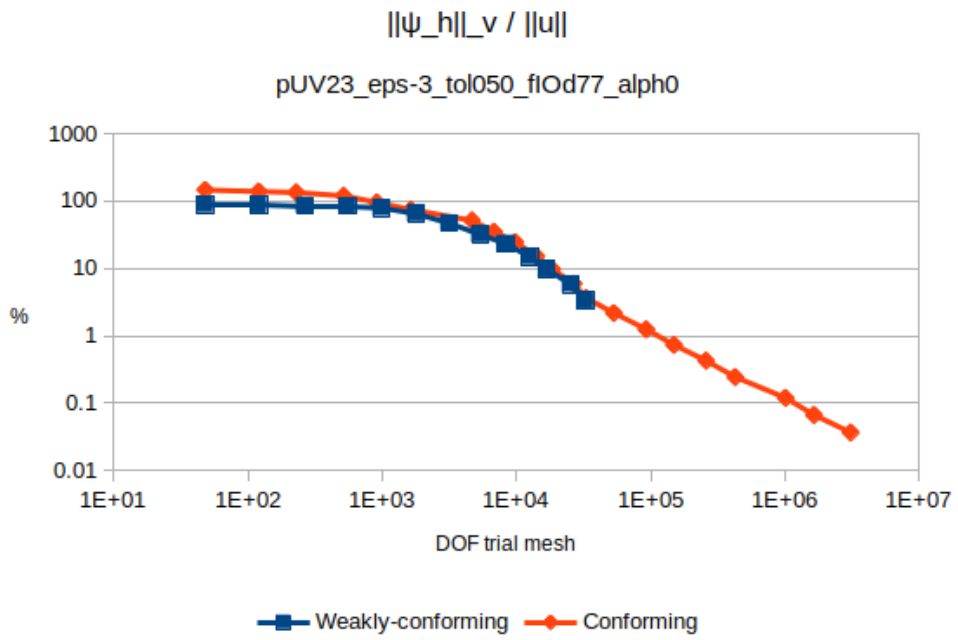


Figure 4.5: Convergence of the Riesz representation of the residual ψ_h for the conforming and weakly-conforming discretizations, $p_U = 2$, $p_V = 3$, $\epsilon = 10^{-3}$.

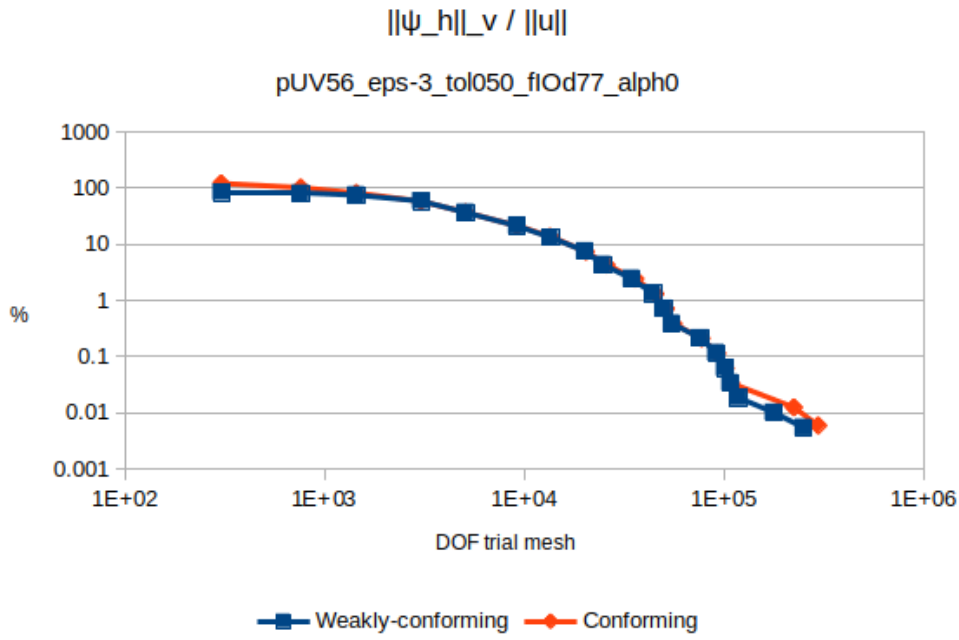


Figure 4.6: Convergence of the Riesz representation of the residual ψ_h for the conforming and weakly-conforming discretizations, $p_U = 5$, $p_V = 6$, $\epsilon = 10^{-3}$.

5 Conclusions

We have presented two implementations of the DAM along with extensive numerical investigations of the parameters that define it. The key idea behind the method is to adaptively refine the test mesh and control the error in the approximation of the Riesz representation of the residual ψ_h, with respect to the ideal one ψ^h, i.e. $\psi^h - \psi_h$. The importance of closely approximating ψ^h relies on its use as error indicator in the PGM with OTF, driving the adaptive refinements of the trial mesh. Before the DAM it was not possible to asses the error incurred by *this* approximation (although Cohen et al. [7] proposed an a-posteriori error estimation of a different Riesz representation of the residual, see Remark 3 in [10]). We propose values for the key parameters of the method, using the convection-dominated diffusion problem as an illustrating example.

We provide an a-priori error analysis for the weakly-conforming discretization (which is equivalent to the broken spaces discretization) without asserting optimality of one methodology over the other. The broken discretizations are frequently used within the PGM with OTF community (partly due to its advantageous element-wise computation of OTF) with the expectation of convergence behavior similar to that of the conforming discretizations. The numerical examples presented verify almost identical convergence curves, which have not been reported before.

Before concluding, we mention notable limitations of our analysis. The DAM was used for the weakly conforming discretization whereas the theory currently only covers conforming ones. Analysis for broken (weakly conforming) discretizations as well as accounting for non-homogeneous boundary conditions remain pending tasks.

References

[1] Ainsworth, M. and Oden, J. T. (1997). A posteriori error estimation in finite element analysis. *Comput. Methods. Appl. Mech. Eng.*, 142(1-2):1–88.

[2] Amestoy, P., Duff, I. S., Koster, J., and L'Excellent, J.-Y. (2001). A fully asynchronous multifrontal solver using distributed dynamic scheduling. *SIAM J. Matrix Anal. Appl.*, 23(1):15–41.

[3] Babuška, I., Kellogg, R. B., and Pitkäranta, J. (1979). Direct and inverse error estimates for finite elements with mesh refinements. *Numer. Math.*, 33(4):447–471.

[4] Broersen, D. and Stevenson, R. P. (2015). A Petrov–Galerkin discretization with optimal test space of a mild-weak formulation of convection–diffusion equations in mixed form. *IMA J. Numer. Anal.*, 35(1):39–73.

[5] Calo, V. M., Romkes, A., and Valseth, E. (2020). Automatic variationally stable analysis for FE computations: an introduction. In *Boundary and Interior Layers, Computational and Asymptotic Methods BAIL 2018*, pages 19–43. Springer.

[6] Chan, J., Heuer, N., Bui-Thanh, T., and Demkowicz, L. (2014). A robust DPG method for convection-dominated diffusion problems II: Adjoint boundary conditions and mesh-dependent test norms. *Comput. Math. Appl.*, 67(4):771–795.

[7] Cohen, A., Dahmen, W., and Welper, G. (2012). Adaptivity and variational stabilization for convection-diffusion equations. *ESAIM Math. Model. Numer. Anal.*, 46(5):1247–1273.

[8] Demkowicz, L. (2006). *Computing with hp Finite Elements. I. One and Two Dimensional Elliptic and Maxwell Problems*. Chapman & Hall/CRC Press, New York.

[9] Demkowicz, L. (2020). Lecture notes on mathematical theory of Finite Elements. Technical Report 11, Oden Institute. `https://users.oden.utexas.edu/~leszek/classes/EM394H/book.pdf`.

[10] Demkowicz, L., Führer, T., Heuer, N., and Tian, X. (2020). The double adaptivity paradigm:(How to circumvent the discrete inf–sup conditions of Babuška and Brezzi). *Comput. Math. with Appl.*

[11] Demkowicz, L. and Gopalakrishnan, J. (2011). A class of discontinuous Petrov–Galerkin methods. II. Optimal test functions. *Numer. Methods Partial. Differ. Equ.*, 27(1):70–105.

[12] Demkowicz, L. and Gopalakrishnan, J. (2015). Discontinuous Petrov-Galerkin (DPG) Method. *ICES Report*, 15-20.

[13] Demkowicz, L., Gopalakrishnan, J., and Niemi, A. H. (2012). A class of discontinuous Petrov–Galerkin methods. Part III: Adaptivity. *Appl. Numer. Math.*, 62(4):396–427.

[14] Dörfler, W. (1995). A robust adaptive strategy for the nonlinear Poisson equation. *Computing*, 55(4):289–304.

[15] Fuentes, F., Keith, B., Demkowicz, L., and Nagaraj, S. (2015). Orientation embedded high order shape functions for the exact sequence elements of all shapes. *Comput. Math. Appl.*, 70(4):353–458.

[16] Houston, P., Roggendorf, S., and van der Zee, K. G. (2020). Eliminating gibbs phenomena: A nonlinear petrov–galerkin method for the convection–diffusion–reaction equation. *Comput. Math. with Appl.*, 80(5):851–873.

[17] Irons, B. M. (1970). A frontal solution program for finite element analysis. *Int. J. Numer. Methods. Eng.*, 2(1):5–32.

[18] Li, J. and Demkowicz, L. (2021). An L^p-DPG method for the convection–diffusion problem. *Comput. Math. with Appl.*, 95:172–185.

[19] Salazar, J., Mora, J., and Demkowicz, L. (2019). Alternative enriched test spaces in the DPG method for singular perturbation problems. *Comput. Methods Appl. Math.*, 19(3):603–630.

[20] Zitelli, J., Muga, I., Demkowicz, L., Gopalakrishnan, J., Pardo, D., and Calo, V. M. (2011). A class of discontinuous Petrov–Galerkin methods. Part IV: The optimal test norm and time-harmonic wave propagation in 1d. *J. Comput. Phys.*, 230(7):2406–2432.