



HAL
open science

A posteriori-steered p -robust multigrid with optimal step-sizes and adaptive number of smoothing steps

Ani Miraçi, Jan Papež, Martin Vohralík

► To cite this version:

Ani Miraçi, Jan Papež, Martin Vohralík. A posteriori-steered p -robust multigrid with optimal step-sizes and adaptive number of smoothing steps. 2020. hal-02494538v1

HAL Id: hal-02494538

<https://hal.science/hal-02494538v1>

Preprint submitted on 28 Feb 2020 (v1), last revised 22 Nov 2021 (v5)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A posteriori-steered p -robust multigrid with optimal step-sizes and adaptive number of smoothing steps*

Ani Miraçi^{†‡} Jan Papež^{†‡} Martin Vohralík^{†‡}

February 28, 2020

Abstract

In this work, we develop a multigrid solver that is steered by a posteriori estimates of the algebraic error. We adopt the context of a second order elliptic diffusion problem discretized by the conforming finite element method of arbitrary polynomial degree $p \geq 1$. Our solver in particular features an optimal (adaptive) choice of the step-size in the smoothing correction on each level. Developing our previous work [HAL Preprint 02070981, 2019], we show the two following results and their equivalence: 1) the solver contracts the algebraic error independently of the polynomial degree p ; 2) the estimator represents a two-sided p -robust bound on the algebraic error. The p -robustness results are obtained by careful application of the work done in Schöberl *et al.* [IMA J. Numer. Anal., 28 (2008), pp. 1–24] for one given mesh, combined with a multilevel stable decomposition for piecewise affine polynomials on quasi-uniform/bisection grids given in Xu *et al.* [Springer, Berlin, 2009, pp. 599–659]. We consider either quasi-uniform or graded bisection meshes and show robustness with respect to the number of mesh levels J for H^2 -regular solutions. Our solver relies on zero pre- and one post-smoothing by an overlapping Schwarz (block-Jacobi) method. We also present a simple and effective way for the solver to adaptively choose the number of post-smoothing steps, which yields yet improved error reduction. We present numerical tests confirming the p -robust behavior of the solver and illustrating the adaptive number of smoothing steps. Moreover, the tests indicate numerical robustness with respect to the number of levels J even in low regularity settings, as well as robustness with respect to the jumps in diffusion coefficient.

Key words: finite element method, multigrid method, Schwarz method, block-Jacobi smoother, a posteriori estimate, stable decomposition, p -robustness

1 Introduction

Multilevel methods, including multigrid, have shown their versatility as solvers and/or preconditioners of large sparse linear systems arising from numerical discretizations of partial differential equations; we refer to pioneering works such as e.g. Brandt *et al.* [6], Bramble *et al.* [4], Hackbusch [15], Bank *et al.* [2], Ruge and Stüben [24], or Oswald [20]. We also refer to survey works that thoroughly treat subspace correction methods in Xu in [33], robust multigrid methods with respect to non-smooth coefficients in Chan and Wan [8], multigrid solvers for high-order discretizations in Sundar *et al.* [30], see also references therein.

In this work, we develop a multilevel solver for algebraic linear systems arising from the discretization using conforming finite elements of arbitrary polynomial degree $p \geq 1$. One iteration of our solver can be viewed as a V-cycle employing zero pre- and one post-smoothing step, with the levelwise smoother being overlapping additive Schwarz (block Jacobi) associated to patches of elements that share a common vertex. An important difference to the classical V-cycle is the use of optimal step-size at the error correction stage

*This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (grant agreement No 647134 GATIPOR). The authors are grateful to Inria Sophia Antipolis - Méditerranée “NEF” computation cluster for providing resources and support.

[†]Inria, 2 rue Simone Iff, 75589 Paris, France

[‡]Université Paris-Est, CERMICS (ENPC), 77455 Marne-la-Vallée, France

on each level, which yields a minimal residual in the subsequent level. The idea of an optimal step-size in the error correction is not new; in fact, a weighting of multigrid error corrections concept appears as early as in Brandt [5]. Then, the minimal residual approach is used e.g. in Canuto and Quarteroni [7], though not in a multigrid setting. The interest in this minimal residual approach used for the multigrid error correction has been pointed out in Heinrichs [16], where this choice resulted in better numerical performance of the solver. Another version of multigrid solvers with a changing step-size error correction be found in the form of a scaled residual used in Růde [22].

We prove that the multilevel solver we introduce contracts the error p -robustly on each iteration. Notable works in treating p -robustness include e.g. Kanschäat [17] and Lucero Lorca and Kanschäat [18] using multilevel preconditioners for rectangular/hexahedral, as well as Antonietti and Pennesi [1] for more general meshes. A p -robust stable decomposition on triangular/tetrahedral meshes was presented in Schöberl *et al.* [26] and it consequently leads to a (one-mesh) p -robust preconditioner. This result plays an important part in the analysis of our previous work [19], which presented a p -robust multilevel solver, that we also use in this work. Similarly to [19], the construction of the solver is tightly related to an a posteriori estimator on the algebraic error. We show that this estimator is p -robustly efficient and that this claim is equivalent with the p -robust contraction of the solver. The construction of the estimator is done easily and without extra costs, stemming from the multilevel solver (see also Růde [23] for the link between efficient estimators and subspace decomposition type solvers). These results hold for quasi-uniform meshes as well as possibly highly graded ones. In this case, we show that the dependence with respect to the number of mesh levels J is linear. Moreover, under additional full-elliptic regularity, we show that the above results are also J -robust. Numerical experiments support the theoretical findings. Moreover, the tests suggest independence of the jump in the diffusion coefficient, and that the numerical J -robustness holds even in low-regularity test cases.

A crucial immediate consequence of using the optimal step-size in the error correction stage is the fact that the error contraction of the previous levels is explicitly known. Indeed, thanks to it, we obtain the formula representing the error decrease:

$$\|\mathcal{K}^{\frac{1}{2}}\nabla(u_J - u_J^{i+1})\|^2 = \|\mathcal{K}^{\frac{1}{2}}\nabla(u_J - u_J^i)\|^2 - \sum_{j=0}^J (\lambda_j^i \|\mathcal{K}^{\frac{1}{2}}\nabla\rho_j^i\|)^2, \quad (1.1)$$

where \mathcal{K} is the diffusion coefficient, $j \in \{1, \dots, J\}$ is the level counter, u_J is the true algebraic solution, u_J^{i+1} denotes the iterate given by the solver from u_J^i with levelwise corrections ρ_j^i and levelwise optimal step-sizes λ_j^i . This is the foundation of a simple and efficient strategy to choose adaptively the number of post-smoothing steps per level. The essence and particularity of our strategy relies on a posteriori-steered decision-making of the number of smoothing steps. Thus, as formula (1.1) explains, at each level we are decreasing the error by a positive computable quantity. After one mandatory smoothing step at each level, if the given decrease is bigger than a user-prescribed portion of the decrease made by the previous levels (in the spirit of Dörfler [10]), we can decide to do another smoothing step before going to the next level. The idea of employing a variable number of smoothing steps per level has also been explored e.g. in Bramble and Pasciak [3] where a generalized V-cycle is proposed. It uses a number of smoothing steps which is bigger on coarser grids and smaller on finer ones. Closely related to the subject, in Thekale *et al.* [31] a new approach is presented: it suggests an adaptive number of multigrid cycles per level in order to optimize the costs of the full multigrid method.

Compared to previous work [19], we point out the novelties and improvements. While in [19] a global optimal step-size was used in the proposed solver, here the levelwise step-sizes offer a better numerical performance, as well as nicer development of the analysis. In fact, the current analysis gives independence with respect the number of levels in full-regularity setting, which was not possible within the previous construction of the solver. Moreover, the levelwise optimal step-sizes allow for the error representation (1.1), which other than making the analysis simpler, is also very useful for the developed adaptive number of smoothings strategy. Finally and importantly, the solver proposed in this work does not need any additional damping parameters, whose tuning can be cumbersome.

The practicality of a solver is largely determined by its behavior with respect to discretization parameters such as possible discontinuities in the diffusion coefficient of the problem. As for studies on the behavior of solvers with respect to discontinuities of the diffusion coefficient, see for e.g.: Vassilevski [32], where an multilevel preconditioner is presented, both robust with respect to the number of levels and the discontinuity of the diffusion coefficient when the discontinuities take place across edges/faces; Dryja *et al.* [11], where

the proposed multilevel solver is robust with respect to both the number of mesh levels and the jump in the diffusion coefficient under a quasi-monotonicity assumption; Graham and Hagger [13], where number of iterations grows with discontinuity only logarithmically for unstructured meshes; Xu and Zhu [35], where the convergence rate of the proposed preconditioner is uniform with respect to the large jump and meshsize on quasi-uniform meshes, then extended on bisection grids in Chen *et al.* [9]; Scheichl *et al.* [25], where a robust multilevel preconditioner is presented by using stable splittings based on weighted quasi-interpolants and weighted Poincaré-type inequalities, in particular the coarse grid needn't be aligned with the discontinuities of the diffusion coefficient; Spillane *et al.* [28],[29], where a two level robust method with respect to the diffusion jump is presented thanks to the construction of a coarse space generated from the solution of local eigenvalue problems.

In Section 2 of this manuscript, we present the setting and notation we will be working with, and in Section 3 we present the motivation, in the spirit of Papež *et al.* [21], leading us to consider this particular multilevel solver. The multilevel solver is then presented in Section 4, and the a posteriori error estimator is given in Section 5. We collect in Section 6 the assumptions we work with and the main results of the manuscript. In Section 7, we present the algorithm of the solver with the adaptive choice of number of post-smoothing steps. Section 8 gives the results obtained by the numerical experiments. The proofs of our main results are given in Section 9, and we give our concluding remarks in Section 10.

2 Setting

We consider in this work a second order elliptic diffusion problem defined over $\Omega \subset \mathbb{R}^d$, $d \in \{1, 2, 3\}$, an open bounded polytope with a Lipschitz-continuous boundary.

2.1 Model problem, finite element discretization, and algebraic system

Let $f \in L^2(\Omega)$ be the source term, $\mathcal{K} \in [L^\infty(\Omega)]^{d \times d}$ a symmetric positive definite diffusion coefficient. We search for $u \in H_0^1(\Omega)$ such that

$$(\mathcal{K}\nabla u, \nabla v) = (f, v), \quad \forall v \in H_0^1(\Omega), \quad (2.1)$$

where (\cdot, \cdot) is the $L^2(\Omega)$ or $[L^2(\Omega)]^d$ scalar product. The existence and uniqueness of the solution of (2.1) follows from the Riesz representation theorem.

We discretize the above continuous problem by fixing \mathcal{T}_J a matching simplicial mesh of Ω , and an integer $p \geq 1$, in order to introduce the finite element space of continuous piecewise p -degree polynomials

$$V_J^p := \mathbb{P}_p(\mathcal{T}_J) \cap H_0^1(\Omega), \quad (2.2)$$

where $\mathbb{P}_p(\mathcal{T}_J) := \{v_J \in L^2(\Omega), v_J|_K \in \mathbb{P}_p(K) \forall K \in \mathcal{T}_J\}$. The discrete problem now consists in finding $u_J \in V_J^p$ such that

$$(\mathcal{K}\nabla u_J, \nabla v_J) = (f, v_J) \quad \forall v_J \in V_J^p. \quad (2.3)$$

If one introduces a basis of V_J^p , then the discrete problem is equivalent to solving a system of linear algebraic equations, whose matrix is symmetric and positive definite. Note, however, that such a linear system depends on the choice of the basis functions. To avoid this dependence, we work instead with a functional description of the problem; we in particular define the algebraic residual functional on V_J^p by

$$v_J \mapsto (f, v_J) - (\mathcal{K}\nabla u_J^i, \nabla v_J) \in \mathbb{R}, \quad v_J \in V_J^p. \quad (2.4)$$

2.2 A hierarchy of meshes and spaces

We work with a hierarchy of matching simplicial meshes $\{\mathcal{T}_j\}_{0 \leq j \leq J}$, where \mathcal{T}_J has been introduced in Section 2.1, and where \mathcal{T}_j is a refinement of \mathcal{T}_{j-1} , $1 \leq j \leq J$. We also introduce a hierarchy of finite element spaces associated to the mesh hierarchy. For this purpose, for $j \in \{1, \dots, J\}$, fix p_j the polynomial degree associated to mesh level j such that $1 \leq p_1 \leq \dots \leq p_{J-1} \leq p_J = p$. In particular, let:

$$\text{for } j = 0 : \quad V_0^1 := \mathbb{P}_1(\mathcal{T}_0) \cap H_0^1(\Omega) \quad (\text{lowest-order space}), \quad (2.5a)$$

$$\text{for } 1 \leq j \leq J - 1 : \quad V_j^{p_j} := \mathbb{P}_{p_j}(\mathcal{T}_j) \cap H_0^1(\Omega) \quad (p_j\text{-th order spaces}), \quad (2.5b)$$

where $\mathbb{P}_{p_j}(\mathcal{T}_j) := \{v_j \in L^2(\Omega), v_j \in \mathbb{P}_{p_j}(K) \ \forall K \in \mathcal{T}_j\}$. Note that $V_0^1 \subset V_1^{p_1} \subset \dots \subset V_{J-1}^{p_{J-1}} \subset V_J^p$. Let \mathcal{V}_j be the set of vertices of the mesh \mathcal{T}_j . We denote by $\psi_{j,\mathbf{a}}$ the standard hat function associated to the vertex $\mathbf{a} \in \mathcal{V}_j$, $0 \leq j \leq J$. This is the piecewise affine function with respect to the mesh \mathcal{T}_j , that takes value 1 in the vertex \mathbf{a} and vanishes in all other vertices of \mathcal{V}_j .

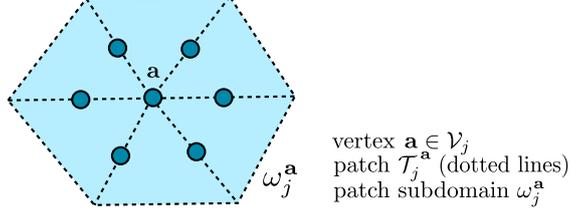


Figure 1: Illustration of degrees of freedom ($p_j = 2$) for the space $V_j^{\mathbf{a}}$ associated to the patch $\mathcal{T}_j^{\mathbf{a}}$.

For the following, we need to define the notion of patches of elements, illustrated in Figure 1. Given a vertex $\mathbf{a} \in \mathcal{V}_j$, $j \in \{1, \dots, J\}$, we denote the patch related to \mathbf{a} by $\mathcal{T}_j^{\mathbf{a}}$, the corresponding open patch subdomain by $\omega_j^{\mathbf{a}}$, and the associated local space $V_j^{\mathbf{a}}$. Let \mathcal{V}_K be the set of vertices of element K . Then

$$\mathcal{T}_j^{\mathbf{a}} := \{K \in \mathcal{T}_j, \mathbf{a} \in \mathcal{V}_K\}, \quad (2.6)$$

$$V_j^{\mathbf{a}} := \mathbb{P}_{p_j}(\mathcal{T}_j) \cap H_0^1(\omega_j^{\mathbf{a}}), \quad j \in \{1, \dots, J\}. \quad (2.7)$$

As done in previous work of the authors [19], we could present a larger version of the patches and the main results of this manuscript would still be valid. However, to make the presentation easier, we only consider here the version with small patches.

3 Motivation: level-wise orthogonal decomposition of the error

Let us first motivate our multilevel construction. In the spirit of Papež *et al.* [21], consider for a given $u_J^i \in V_J^p$, the following (infeasible in practice but illustrative) hierarchical construction $\tilde{\rho}_{J,\text{alg}}^i \in V_J^p$

$$\tilde{\rho}_{J,\text{alg}}^i := \rho_0^i + \sum_{j=1}^J \tilde{\rho}_j^i, \quad (3.1)$$

where ρ_0^i is given as a solution to a global lowest-order problem on the coarsest mesh

$$(\mathcal{K} \nabla \rho_0^i, \nabla v_0) = (f, v_0) - (\mathcal{K} \nabla u_J^i, \nabla v_0) \quad \forall v_0 \in V_0, \quad (3.2)$$

and for all $j \in \{1, \dots, J\}$, $\tilde{\rho}_j^i \in V_j^p$ are the solutions of

$$(\mathcal{K} \nabla \tilde{\rho}_j^i, \nabla v_j) = (f, v_j) - (\mathcal{K} \nabla u_J^i, \nabla v_j) - \sum_{k=0}^{j-1} (\mathcal{K} \nabla \tilde{\rho}_k^i, \nabla v_j) \quad \forall v_j \in V_j^p. \quad (3.3)$$

This construction (see also [21]) returns the algebraic error, i.e. we actually have $\tilde{\rho}_{J,\text{alg}}^i = u_J - u_J^i$. This, in turn, means that $\tilde{\rho}_{J,\text{alg}}^i$ satisfies

$$(\mathcal{K} \nabla \tilde{\rho}_{J,\text{alg}}^i, \nabla v_J) = (f, v_J) - (\mathcal{K} \nabla u_J^i, \nabla v_J) \quad \forall v_J \in V_J^p. \quad (3.4)$$

Moreover, there holds $(\mathcal{K} \nabla \tilde{\rho}_j^i, \nabla \tilde{\rho}_k^i) = 0$, $0 \leq k, j \leq J$; $j \neq k$. These observations lead to the orthogonal decomposition

$$\|\mathcal{K}^{\frac{1}{2}} \nabla (u_J - u_J^i)\|^2 = \|\mathcal{K}^{\frac{1}{2}} \nabla \tilde{\rho}_{J,\text{alg}}^i\|^2 = \sum_{j=0}^J \|\mathcal{K}^{\frac{1}{2}} \nabla \tilde{\rho}_j^i\|^2. \quad (3.5)$$

4 Multilevel solver

Let us now introduce our local constructions inspired by (3.1)–(3.3) and produce level-wise approximations of the error. The construction relies on the use of coarse solution of (3.2), which is cheap enough to be done in practice, and on local contributions arising from all the finer mesh levels. These local contributions are defined on patches of elements on each level. We go through the levels adding gradually level-wise updates to the current approximation as described below. The final update gives us the new approximation.

Definition 4.1 (Multilevel solver). *1. Initialize $u_J^0 \in V_0^1$ as the zero function.*

2. Let $i \geq 0$ be the iteration counter and $u_J^i \in V_J^p$ the current approximation.

(a) Let ρ_0^i be defined by (3.2), $\lambda_0^i = 1$, and $u_{J,0}^i := u_J^i + \lambda_0^i \rho_0^i$.

(b) For $j \in \{1, \dots, J\}$:

define the local contributions $\rho_{j,\mathbf{a}}^i \in V_j^{\mathbf{a}}$ as solutions of patch problems, for all vertices $\mathbf{a} \in \mathcal{V}_j$

$$(\mathcal{K} \nabla \rho_{j,\mathbf{a}}^i, \nabla v_{j,\mathbf{a}})_{\omega_j^{\mathbf{a}}} = (f, v_{j,\mathbf{a}})_{\omega_j^{\mathbf{a}}} - (\mathcal{K} \nabla u_{J,j-1}^i, \nabla v_{j,\mathbf{a}})_{\omega_j^{\mathbf{a}}} \quad \forall v_{j,\mathbf{a}} \in V_j^{\mathbf{a}}, \quad (4.1)$$

and the level j descent direction $\rho_j^i \in V_j^p$ by

$$\rho_j^i := \sum_{\mathbf{a} \in \mathcal{V}_j} \rho_{j,\mathbf{a}}^i. \quad (4.2)$$

If $\rho_j^i \neq 0$, define the optimal step-size on level j

$$\lambda_j^i := \frac{(f, \rho_j^i) - (\mathcal{K} \nabla u_{J,j-1}^i, \nabla \rho_j^i)}{\|\mathcal{K}^{\frac{1}{2}} \nabla \rho_j^i\|^2}, \quad (4.3)$$

otherwise set $\lambda_j^i := 0$. The level update is given by

$$u_{J,j}^i := u_{J,j-1}^i + \lambda_j^i \rho_j^i, \quad (4.4)$$

and the final update is $u_J^{i+1} := u_{J,J}^i \in V_J^p$. If $u_J^{i+1} = u_J^i$, then we stop the solver.

Note that by definition $\lambda_0^i = 1$ and we thus have for $\rho_0^i \neq 0$,

$$\frac{(f, \rho_0^i) - (\mathcal{K} \nabla u_J^i, \nabla \rho_0^i)}{\|\mathcal{K}^{\frac{1}{2}} \nabla \rho_0^i\|^2} \stackrel{(3.2)}{=} 1 = \lambda_0^i.$$

Remark 4.2 (Compact writing of the iteration update). *Let $u_J^i \in V_J^p$, using the convention $u_{J,-1}^i := u_J^i$, and $\frac{0}{0} = 0$, the new iterate after a step of the solver described in Definition 4.1 is*

$$u_J^{i+1} = u_J^i + \sum_{j=0}^J \lambda_j^i \rho_j^i = u_J^i + \sum_{j=0}^J \frac{(f, \rho_j^i) - (\mathcal{K} \nabla u_{J,j-1}^i, \nabla \rho_j^i)}{\|\mathcal{K}^{\frac{1}{2}} \nabla \rho_j^i\|^2} \rho_j^i. \quad (4.5)$$

In the lemma below, we justify rigorously the use and choice of each level's step size (4.3).

Lemma 4.3 (Level-wise optimal step-size). *Let $u_J^i \in V_J^p$ be arbitrary, let $j \in \{1, \dots, J\}$ and let $u_{J,j-1}^i, \rho_j^i$ be given by the solver described in Definition 4.1. Consider the level's update $u_{J,j}^i$ given by*

$$u_{J,j}^i = u_{J,j-1}^i + \lambda \rho_j^i, \quad \text{for } \lambda \in \mathbb{R}.$$

The choice of $\lambda = [(f, \rho_j^i) - (\mathcal{K} \nabla u_{J,j-1}^i, \nabla \rho_j^i)] / \|\mathcal{K}^{\frac{1}{2}} \nabla \rho_j^i\|^2$ leads to the best decrease of the algebraic error with respect to the energy norm.

Proof. We write the algebraic error associated to the new update as a function of λ

$$\|\mathcal{K}^{\frac{1}{2}}\nabla(u_J - u_{J,j}^i)\|^2 = \|\mathcal{K}^{\frac{1}{2}}\nabla(u_J - u_{J,j-1}^i)\|^2 - 2\lambda(\mathcal{K}\nabla(u_J - u_{J,j-1}^i), \nabla\rho_j^i) + \lambda^2\|\mathcal{K}^{\frac{1}{2}}\nabla\rho_j^i\|^2.$$

We realize that this function has a minimum at

$$\lambda_{\min} = \frac{(\mathcal{K}\nabla(u_J - u_{J,j-1}^i), \nabla\rho_j^i)}{\|\mathcal{K}^{\frac{1}{2}}\nabla\rho_j^i\|^2} \stackrel{(2.3)}{=} \frac{(f, \rho_j^i) - (\mathcal{K}\nabla u_{J,j-1}^i, \nabla\rho_j^i)}{\|\mathcal{K}^{\frac{1}{2}}\nabla\rho_j^i\|^2},$$

which gives us the expression (4.3). \square

Remark 4.4 (Construction of the new iterate). *The construction of u_J^{i+1} from u_J^i by the solver of Definition 4.1 can be seen as one iteration of a V-cycle multigrid, with no pre-smoothing and a single post-smoothing step, with a variable step-size in the error correction. The smoother on each level corresponds to additive Schwarz with patch subdomains where the local problems (4.1) are defined. In particular, for $p = 1$ (i.e. for all $j \in \{1, \dots, J\}$, $p_j = 1$), this corresponds to one-step Jacobi (diagonal) smoother, whereas when $p_j > 1$, $j \in \{1, \dots, J\}$, the smoother is block-Jacobi.*

The main motivation for choosing appropriate level-wise step-sizes is explained in the theorem below.

Theorem 4.5 (Error representation of one solver step). *For $u_J^i \in V_J^p$, let $u_J^{i+1} \in V_J^p$ be given by Definition 4.1. Then*

$$\|\mathcal{K}^{\frac{1}{2}}\nabla(u_J - u_J^{i+1})\|^2 = \|\mathcal{K}^{\frac{1}{2}}\nabla(u_J - u_J^i)\|^2 - \sum_{j=0}^J (\lambda_j^i \|\mathcal{K}^{\frac{1}{2}}\nabla\rho_j^i\|)^2. \quad (4.6)$$

Proof. We obtain the result by going through the levels from finest to coarsest, and using the relation of each level's update with its associated optimal step-size

$$\begin{aligned} \|\mathcal{K}^{\frac{1}{2}}\nabla(u_J - u_J^{i+1})\|^2 &\stackrel{(4.4)}{=} \|\mathcal{K}^{\frac{1}{2}}\nabla(u_J - (u_{J,J-1}^i + \lambda_J^i \rho_J^i))\|^2 \\ &\stackrel{(2.3)}{=} \|\mathcal{K}^{\frac{1}{2}}\nabla(u_J - u_{J,J-1}^i)\|^2 - (\lambda_J^i \|\mathcal{K}^{\frac{1}{2}}\nabla\rho_J^i\|)^2 = \dots \\ &= \|\mathcal{K}^{\frac{1}{2}}\nabla(u_J - (u_J^i + \lambda_0^i \rho_0^i))\|^2 - \sum_{j=1}^J (\lambda_j^i \|\mathcal{K}^{\frac{1}{2}}\nabla\rho_j^i\|)^2 \\ &\stackrel{(3.2)}{=} \|\mathcal{K}^{\frac{1}{2}}\nabla(u_J - u_J^i)\|^2 - \sum_{j=0}^J (\lambda_j^i \|\mathcal{K}^{\frac{1}{2}}\nabla\rho_j^i\|)^2. \end{aligned}$$

\square

5 A posteriori estimator on the the algebraic error

We present below how the solver introduced in Section 4 induces an a posteriori estimator η_{alg}^i .

Definition 5.1 (Algebraic error estimator). *Let $u_J^i \in V_J^p$ be arbitrary, and let $u_J^{i+1} \in V_J^p$ be the update at the end of one step of the solver introduced in Definition 4.1. We define the algebraic error estimator*

$$\eta_{\text{alg}}^i := \left(\sum_{j=0}^J (\lambda_j^i \|\mathcal{K}^{\frac{1}{2}}\nabla\rho_j^i\|)^2 \right)^{\frac{1}{2}}. \quad (5.1)$$

Following Theorem 4.5, the estimator η_{alg}^i is immediately a guaranteed lower bound on the algebraic error.

Lemma 5.2 (Guaranteed lower bound on the algebraic error). *There holds:*

$$\|\mathcal{K}^{\frac{1}{2}}\nabla(u_J - u_J^i)\| \geq \eta_{\text{alg}}^i. \quad (5.2)$$

6 Main results

In this section, we present the main results concerning our multilevel solver of Definition 4.1 and our a posteriori estimator η_{alg}^i of Definition 5.1. As in [19], these two results are equivalent. We first list below the assumptions used in the main results.

6.1 Mesh assumptions

For $j \in \{1, \dots, J\}$, we denote in the following $h_K := \text{diam}(K)$ for $K \in \mathcal{T}_j$ and $h_j = \max_{K \in \mathcal{T}_j} h_K$. We shall always assume that our meshes are shape-regular:

Assumption 6.1 (Shape regularity). *There exists $\kappa_{\mathcal{T}} > 0$ such that*

$$\max_{K \in \mathcal{T}_j} \frac{h_K}{\rho_K} \leq \kappa_{\mathcal{T}} \text{ for all } 0 \leq j \leq J, \quad (6.1)$$

where ρ_K denotes the diameter of the largest ball inscribed in K .

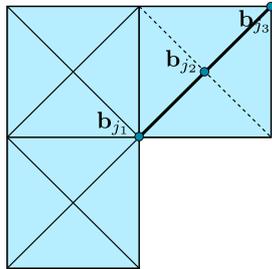
Then we work in one of the three settings below. In the first setting: we work with a hierarchy of quasi-uniform meshes with a bounded refinement factor between consecutive levels. This setting is described by the following assumption:

Assumption 6.2 (Refinement strength and mesh quasi-uniformity). *There exists $0 < C_{\text{ref}} \leq 1$, a fixed positive real number such that for any $j \in \{1, \dots, J\}$, for all $K \in \mathcal{T}_{j-1}$, and for any $K^* \in \mathcal{T}_j$ such that $K^* \subset K$, there holds*

$$C_{\text{ref}} h_K \leq h_{K^*} \leq h_K. \quad (6.2)$$

There further exists C_{qu} , a fixed positive real number such that for any $j \in \{0, \dots, J\}$ and for all $K \in \mathcal{T}_j$, there holds

$$C_{\text{qu}} h_j \leq h_K \leq h_j. \quad (6.3)$$



\mathcal{T}_j obtained by a bisection of \mathcal{T}_{j-1}
 new vertex after refinement \mathbf{b}_{j2}
 neighboring vertices on the
 refinement edge $\mathbf{b}_{j1}, \mathbf{b}_{j3}$
 $\mathcal{B}_j = \{\mathbf{b}_{j1}, \mathbf{b}_{j2}, \mathbf{b}_{j3}\} \subset \mathcal{V}_j$

Figure 2: Illustration of the set \mathcal{B}_j ; the refinement \mathcal{T}_j (dotted lines) of mesh \mathcal{T}_{j-1} (full lines).

In the second setting, we work with a hierarchy generated from a quasi-uniform coarse grid by a series of bisections, e.g. newest vertex bisection cf. Sewell [27]. In this case, one refinement edge of \mathcal{T}_{j-1} , for $j \in \{1, \dots, J\}$, gives us a new finer mesh \mathcal{T}_j . We denote by $\mathcal{B}_j \subset \mathcal{V}_j$ the set consisting of the new vertex obtained after the bisection together with its two neighbors on the refinement edge, illustrated in Figure 2, for $d = 2$. We also denote by $h_{\mathcal{B}_j}$ the maximal diameter of elements having a vertex in the set \mathcal{B}_j , for $j \in \{1, \dots, J\}$. This setting is described by the following assumption:

Assumption 6.3 (Local quasi-uniformity of bisection-generated meshes). *\mathcal{T}_0 is a conforming quasi-uniform mesh with parameter C_{qu}^0 . The graded conforming mesh \mathcal{T}_J is generated from \mathcal{T}_0 by a series of bisections. There exists a fixed positive real number $C_{\text{loc,qu}}$ such that for any $j \in \{1, \dots, J\}$, there holds*

$$C_{\text{loc,qu}} h_{\mathcal{B}_j} \leq h_K \leq h_{\mathcal{B}_j} \quad \forall K \in \mathcal{T}_j \text{ such that a vertex of } K \text{ belongs to } \mathcal{B}_j. \quad (6.4)$$

In the third setting, we assume we have a higher regularity of the solution by:

Assumption 6.4 (H^2 -regularity). *In addition to Assumption 6.2, we moreover assume that the domain Ω is convex and that $\mathcal{K} = \text{Id}$, so that the solution of (2.1) has full-elliptic regularity $u \in H^2(\Omega)$.*

6.2 Main theorems

For the solver it holds:

Theorem 6.5 (*p*-robust error contraction of the multilevel solver). *Let $u_J \in V_J^p$ be the (unknown) solution of (2.3) and let $u_J^i \in V_J^p$ be arbitrary, $i \geq 0$. Take u_J^{i+1} to be constructed from u_J^i using one step of the multilevel solver of Definition 4.1. Under Assumption 6.1 and either of Assumptions 6.2, 6.3, or 6.4 there holds*

$$\|\mathcal{K}^{\frac{1}{2}} \nabla(u_J - u_J^{i+1})\| \leq \alpha \|\mathcal{K}^{\frac{1}{2}} \nabla(u_J - u_J^i)\|, \quad (6.5)$$

where $0 < \alpha < 1$ only depends on the space dimension d , the mesh shape regularity parameter $\kappa_{\mathcal{T}}$, the diffusion coefficient \mathcal{K} , and the parameters associated to the employed Assumption 6.2, 6.3, or 6.4. If either of Assumptions 6.2, or 6.3 holds, then α additionally depends at most linearly on the number of mesh levels J .

In the above theorem, α represents an estimation of the algebraic error contraction factor at each step i . In particular, this means that the solver of Definition 4.1 contracts the algebraic error at each iteration step in a robust way both with respect to the number of mesh elements in \mathcal{T}_J (to the mesh size h) and with respect to the polynomial degree p .

For the estimator, in turn, we have:

Theorem 6.6 (*p*-robust reliable and efficient bound on the algebraic error). *Let $u_J \in V_J^p$ be the (unknown) solution of (2.3) and let $u_J^i \in V_J^p$ be arbitrary, $i \geq 0$. Let η_{alg}^i be given by Definition 5.1. Let Assumption 6.1 and either of the Assumptions 6.2, 6.3, or 6.4 hold. Then, in addition to $\|\nabla(u_J - u_J^i)\| \geq \eta_{\text{alg}}^i$ of (5.2), there holds*

$$\eta_{\text{alg}}^i \geq \beta \|\mathcal{K}^{\frac{1}{2}} \nabla(u_J - u_J^i)\|, \quad (6.6)$$

where $0 < \beta < 1$ has the same dependencies as α in Theorem 6.5.

The theorem allows to write η_{alg}^i as a two-sided bound of the algebraic error (up to the generic constant β for the upper bound), meaning that the estimator is robustly efficient with respect to the polynomial degree p .

6.3 Additional results

Theorems 6.6 and 6.5 are actually equivalent, similarly as in [19].

Corollary 6.7 (Equivalence of the *p*-robust estimator efficiency and *p*-robust solver contraction). *Let the assumptions of Theorems 6.6 and 6.5 be satisfied. Then (6.6) holds if and only if (6.5) holds, and $\beta = \sqrt{1 - \alpha^2}$.*

Proof. We give the proof for completeness Starting from (6.5) with $0 < \alpha < 1$,

$$\begin{aligned} \|\mathcal{K}^{\frac{1}{2}} \nabla(u_J - u_J^{i+1})\|^2 &\leq \alpha^2 \|\mathcal{K}^{\frac{1}{2}} \nabla(u_J - u_J^i)\|^2 \\ \stackrel{(4.6)}{\Leftrightarrow} \|\mathcal{K}^{\frac{1}{2}} \nabla(u_J - u_J^i)\|^2 - \sum_{j=0}^J (\lambda_j^i \|\mathcal{K}^{\frac{1}{2}} \nabla \rho_j^i\|)^2 &\leq \alpha^2 \|\mathcal{K}^{\frac{1}{2}} \nabla(u_J - u_J^i)\|^2 \\ \stackrel{(5.1)}{\Leftrightarrow} \|\mathcal{K}^{\frac{1}{2}} \nabla(u_J - u_J^i)\|^2 - (\eta_{\text{alg}}^i)^2 &\leq \alpha^2 \|\mathcal{K}^{\frac{1}{2}} \nabla(u_J - u_J^i)\|^2 \\ \Leftrightarrow \|\mathcal{K}^{\frac{1}{2}} \nabla(u_J - u_J^i)\|^2 (1 - \alpha^2) &\leq (\eta_{\text{alg}}^i)^2. \end{aligned}$$

□

Finally, the following corollary formulates a three-part equivalence.

Corollary 6.8 (Equivalence error–estimator–localized contributions). *If Assumption 6.4 holds, or if either of Assumptions 6.2, or 6.3 holds together with $\lambda_j^i \leq R$, $1 \leq j \leq J$, where $R \geq 1$ is a real parameter, then*

$$\|\mathcal{K}^{\frac{1}{2}} \nabla(u_J - u_J^i)\|^2 \approx \|\mathcal{K}^{\frac{1}{2}} \nabla \rho_0^i\|^2 + \sum_{j=1}^J \sum_{\mathbf{a} \in \mathcal{V}_j} \|\mathcal{K}^{\frac{1}{2}} \nabla \rho_{j,\mathbf{a}}^i\|^2 \approx (\eta_{\text{alg}}^i)^2. \quad (6.7)$$

The constants involved in the equivalence have the same dependency as α in (6.5), and additionally depend on R for Assumptions 6.2, or 6.3.

7 Adaptive number of smoothing steps

We consider here a simple and practical way to make the solver described in Definition 4.1 choose autonomously and adaptively the number of smoothing steps on each mesh level. We remind that the non-adaptive version can be seen as a V-cycle multigrid with no pre- and one post-smoothing step only. The idea of the adaptive version is to make more post-smoothing steps *if needed*. While with the non-adaptive version we constructed level-wise contributions that defined the a posteriori error estimator (proved above to be equivalent to the error), with the adaptive version we aim to add more contributions from levels that contribute more to the algebraic error. Thus, we rely on the computable level-wise contributions we have already at our disposal and whenever it is a dominant factor (via a Dörfler [10] type condition) in the a posteriori estimator, another smoothing step is employed.

Definition 7.1 (Adaptive multilevel solver). *Let ν_{\max} be the maximum number of smoothing steps, and $0 < \theta < 1$.*

1. Initialize $u_J^0 \in V_0$ as the zero function.
2. Let $i \geq 0$ be the iteration counter, and $u_J^i \in V_J^p$ the current approximation.
 - (a) Let ρ_0^i be defined by (3.2), let $\lambda_0^i := 1$ and $u_{J,0}^i := u_J^i + \lambda_0^i \rho_0^i$.
 - (b) For $j \in \{1, \dots, J\}$:
 - i. Construct from $u_{J,j-1}^i: \rho_j^i, \lambda_j^i$, and $u_{J,j}^i$ by (4.2), (4.3), and (4.4), respectively.
Set $\rho_{j,1}^i := \rho_j^i, \lambda_{j,1}^i := \lambda_j^i, u_{J,j,1}^i := u_{J,j}^i$.
 - ii. $\nu = 2$;
while $\left[\nu \leq \nu_{\max} \&\& (\lambda_{j,\nu-1}^i \|\mathcal{K}^{\frac{1}{2}} \nabla \rho_{j,\nu-1}^i\|)^2 \geq \theta^2 \left(\sum_{k=0}^{j-1} (\lambda_k^i \|\mathcal{K}^{\frac{1}{2}} \nabla \rho_k^i\|)^2 + \sum_{\ell=1}^{\nu-2} (\lambda_{j,\ell}^i \|\mathcal{K}^{\frac{1}{2}} \nabla \rho_{j,\ell}^i\|)^2 \right) \right]$
do
Construct from $u_{J,j,\nu-1}^i: \rho_j^i, \lambda_j^i$, and $u_{J,j}^i$ by (4.2), (4.3), and (4.4), respectively.
Set $\rho_{j,\nu}^i := \rho_j^i, \lambda_{j,\nu}^i := \lambda_j^i, u_{J,j,\nu}^i := u_{J,j}^i$.
 $\nu := \nu + 1$;
endwhile
 - iii. Let $\rho_j^i = \rho_{j,\nu}^i$; $\lambda_j^i = \lambda_{j,\nu}^i$; $u_{J,j}^i = u_{J,j,\nu}^i$ when $j < J$, $u_J^{i+1} = u_{J,J,\nu}^i$ otherwise.
 - iv. If $u_J^{i+1} = u_J^i$, then stop the solver.

Remark 7.2 (Adaptive substep). *Note that if we skip the adaptive substep 2(b)ii in Definition 7.1, we obtain the non-adaptive version of the solver described in Definition 4.1.*

8 Numerical experiments

We shall consider in this section various problems: a smooth solution ‘‘Sine’’, a smooth solution with a ‘‘Peak’’ behavior; a singular solution in ‘‘L-shape’’ domain; a ‘‘Checkerboard’’ solution whose diffusion coefficient jump is of order 10^2 ; a ‘‘Skyscraper’’ solution whose diffusion coefficient jump is of order 10^5 . In this section, we use the following stopping criterion: the residual drops below 10^{-5} times the initial residual. This criterion makes the robustness study easier, as we expect the number of iterations i_{stop} needed to reach it, to be similar for different polynomial degrees (indicating p -robustness), different number of mesh levels (indicating J -robustness), and different orders of the jump in the diffusion coefficient (indicating \mathcal{K} -robustness). The behavior of the multilevel solver given in Definition 4.1 is presented in Table 1.

J	p	DoF	Sine				Peak				L-shape				Checkerboard				Skyscraper			
			111p		1ppp		111p		1ppp		111p		1ppp		111p		1ppp		111p		1ppp	
			i_{end}	$\bar{\alpha}$																		
3	1	$2e^4$	19	0.53	19	0.53	19	0.51	19	0.51	21	0.57	21	0.57	18	0.51	18	0.51	19	0.53	19	0.53
	3	$1e^5$	29	0.64	13	0.34	28	0.62	14	0.35	29	0.68	11	0.35	28	0.67	11	0.36	31	0.69	13	0.43
	6	$6e^5$	30	0.64	13	0.34	30	0.63	14	0.34	26	0.67	9	0.33	25	0.65	10	0.34	28	0.68	11	0.41
	9	$1e^6$	31	0.65	14	0.35	30	0.63	14	0.34	23	0.65	9	0.33	23	0.64	9	0.33	26	0.67	10	0.40
4	1	$6e^4$	21	0.54	21	0.54	20	0.52	20	0.52	21	0.57	21	0.57	19	0.53	19	0.53	19	0.53	19	0.53
	3	$6e^5$	29	0.64	13	0.32	29	0.63	14	0.32	28	0.68	11	0.35	27	0.66	11	0.36	30	0.69	11	0.41
	6	$2e^6$	31	0.65	13	0.33	30	0.63	14	0.33	25	0.66	9	0.32	24	0.64	9	0.33	27	0.67	10	0.40
	9	$5e^6$	32	0.65	14	0.33	31	0.63	15	0.33	23	0.64	9	0.32	23	0.63	9	0.33	25	0.66	9	0.38

Table 1: Illustration of the convergence of the solver.

9 Proof of Theorem 6.6

Our approach to proving Theorem 6.6 consists in studying level-wise the uncomputable exact residual lifting $\tilde{\rho}_{J,\text{alg}}^i$ given by (3.1). We do so by relying on the polynomial degree robust stable decomposition result of Schöberl et al. [26]. This will allow us to exploit the similarities of the local contributions (4.1) (used to build our a posteriori estimator η_{alg}^i of Definition 5.1) and the global ones (3.3) that constitute $\tilde{\rho}_{J,\text{alg}}^i$.

We will first present the proof of p -robust efficiency of the estimator stated in Theorem 6.6 in the general setting. Then we give the proof of p -robust and J -robust efficiency of the estimator under the additional assumptions stated in the second part of Theorem 6.6.

Hereafter, we will use the notation $x_1 \lesssim x_2$ when there exists c , a positive real constant only depending on the mesh shape regularity parameter $\kappa_{\mathcal{T}}$ and the space dimension d such that $x_1 \leq cx_2$. Similarly, $x_1 \gtrsim x_2$ means $x_2 \lesssim x_1$ and $x_1 \approx x_2$ means that $x_1 \lesssim x_2$ and $x_2 \lesssim x_1$ simultaneously. If these constants additionally depend on the number of mesh levels J , we use the notations \lesssim_J , \gtrsim_J , and \approx_J , respectively.

9.1 General properties of the estimator η_{alg}^i

We first present some general properties of the estimator η_{alg}^i needed for the proof.

Lemma 9.1 (Estimation of $\|\mathcal{K}^{\frac{1}{2}} \nabla \rho_j^i\|$ by local contributions). *Let ρ_j^i and $\rho_{j,\mathbf{a}}^i$ for $j \in \{1, \dots, J\}$, $\mathbf{a} \in \mathcal{V}_j$, be constructed as in Definition 4.1. There holds*

$$\|\mathcal{K}^{\frac{1}{2}} \nabla \rho_j^i\|^2 \leq (d+1) \sum_{\mathbf{a} \in \mathcal{V}_j} \|\mathcal{K}^{\frac{1}{2}} \nabla \rho_{j,\mathbf{a}}^i\|_{\omega_j^{\mathbf{a}}}^2. \quad (9.1)$$

Proof. Since $\rho_j^i = \sum_{\mathbf{a} \in \mathcal{V}_j} \rho_{j,\mathbf{a}}^i$ by construction (4.2), and the inequality $|\sum_{k=1}^{d+1} a_k|^2 \leq (d+1) \sum_{k=1}^{d+1} |a_k|^2$ lead to

$$\begin{aligned} \|\mathcal{K}^{\frac{1}{2}} \nabla \rho_j^i\|^2 &= \sum_{K \in \mathcal{T}_j} \|\mathcal{K}^{\frac{1}{2}} \nabla \rho_j^i\|_K^2 = \sum_{K \in \mathcal{T}_j} \left\| \sum_{\mathbf{a} \in \mathcal{V}_K} \mathcal{K}^{\frac{1}{2}} \nabla \rho_{j,\mathbf{a}}^i \right\|_K^2 \\ &\leq (d+1) \sum_{K \in \mathcal{T}_j} \sum_{\mathbf{a} \in \mathcal{V}_K} \|\mathcal{K}^{\frac{1}{2}} \nabla \rho_{j,\mathbf{a}}^i\|_K^2 = (d+1) \sum_{\mathbf{a} \in \mathcal{V}_j} \|\mathcal{K}^{\frac{1}{2}} \nabla \rho_{j,\mathbf{a}}^i\|_{\omega_j^{\mathbf{a}}}^2. \end{aligned}$$

□

Lemma 9.2 (Estimation of local contributions by $\|\mathcal{K}^{\frac{1}{2}} \nabla \rho_j^i\|$). *Let ρ_j^i , $\rho_{j,\mathbf{a}}^i$, λ_j^i for $j \in \{1, \dots, J\}$, be constructed as in Definition 4.1. There holds*

$$\sum_{\mathbf{a} \in \mathcal{V}_j} \|\mathcal{K}^{\frac{1}{2}} \nabla \rho_{j,\mathbf{a}}^i\|_{\omega_j^{\mathbf{a}}}^2 \leq (d+1) (\lambda_j^i \|\mathcal{K}^{\frac{1}{2}} \nabla \rho_j^i\|)^2. \quad (9.2)$$

Proof. We start by using the expression of local problems (4.1) with $\rho_{j,\mathbf{a}}^i \in V_{j,\mathbf{a}}^i$, $j \in \{1, \dots, J\}$, used as test function

$$\sum_{\mathbf{a} \in \mathcal{V}_j} \|\mathcal{K}^{\frac{1}{2}} \nabla \rho_{j,\mathbf{a}}^i\|_{\omega_j^{\mathbf{a}}}^2 = \sum_{\mathbf{a} \in \mathcal{V}_j} \left((f, \rho_{j,\mathbf{a}}^i)_{\omega_j^{\mathbf{a}}} - (\mathcal{K} \nabla u_{J,j-1}^i, \nabla \rho_{j,\mathbf{a}}^i)_{\omega_j^{\mathbf{a}}} \right) \stackrel{(4.2)}{=} (f, \rho_j^i) - (\mathcal{K} \nabla u_{J,j-1}^i, \nabla \rho_j^i).$$

If $\rho_j^i = 0$, this means $\sum_{\mathbf{a} \in \mathcal{V}_j} \|\mathcal{K}^{\frac{1}{2}} \nabla \rho_{j,\mathbf{a}}^i\|_{\omega_j^{\mathbf{a}}}^2 = 0$ and thus (9.2) holds trivially. Otherwise, we have

$$\begin{aligned} \sum_{\mathbf{a} \in \mathcal{V}_j} \|\mathcal{K}^{\frac{1}{2}} \nabla \rho_{j,\mathbf{a}}^i\|_{\omega_j^{\mathbf{a}}}^2 &= \left((f, \rho_j^i) - (\mathcal{K} \nabla u_{J,j-1}^i, \nabla \rho_j^i) \right) \stackrel{(4.3)}{=} \lambda_j^i \|\mathcal{K}^{\frac{1}{2}} \nabla \rho_j^i\|^2 \\ &\stackrel{(9.1)}{\leq} \lambda_j^i \|\mathcal{K}^{\frac{1}{2}} \nabla \rho_j^i\| \left((d+1) \sum_{\mathbf{a} \in \mathcal{V}_j} \|\mathcal{K}^{\frac{1}{2}} \nabla \rho_{j,\mathbf{a}}^i\|_{\omega_j^{\mathbf{a}}}^2 \right)^{\frac{1}{2}}. \end{aligned}$$

After simplifying the expression and squaring, we obtain the desired result (9.2). \square

9.2 General properties of the exact residual lifting $\tilde{\rho}_{J,\text{alg}}^i$

Two other important properties of $\tilde{\rho}_j^i$ will be useful for the proof. First, the relations of orthogonality of a given mesh error contribution $\tilde{\rho}_j^i$, $j \in \{1, \dots, J\}$, with respect to previous mesh level functions. And secondly, the local properties of $\tilde{\rho}_j^i$ with respect to local functions of the same mesh. In particular, the local properties will allow the transition from the uncomputable $\tilde{\rho}_j^i$ to the available local contributions of ρ_j^i .

Lemma 9.3 (Inter-level properties of $\tilde{\rho}_j^i$). *Consider the hierarchical construction of the error $\tilde{\rho}_{J,\text{alg}}^i$ given in (3.1). For $j \in \{1, \dots, J\}$ and $k \in \{0, \dots, j-1\}$, there holds*

$$(\mathcal{K} \nabla \tilde{\rho}_j^i, \nabla v_k) = 0 \quad \forall v_k \in V_k. \quad (9.3)$$

Proof. Take $v_k \in V_k$. Note that since $k \leq j-1$, we have $v_k \in V_{j-1} \subset V_j$. The definition given in (3.3) applied to $\tilde{\rho}_j^i$ and $\tilde{\rho}_{j-1}^i$, allows us to write

$$\begin{aligned} (\mathcal{K} \nabla \tilde{\rho}_j^i, \nabla v_k) &= (f, v_k) - (\mathcal{K} \nabla u_J^i, \nabla v_k) - \sum_{l=0}^{j-2} (\mathcal{K} \nabla \tilde{\rho}_l^i, \nabla v_k) - (\mathcal{K} \nabla \tilde{\rho}_{j-1}^i, \nabla v_k) \\ &= (\mathcal{K} \nabla \tilde{\rho}_{j-1}^i, \nabla v_k) - (\mathcal{K} \nabla \tilde{\rho}_{j-1}^i, \nabla v_k) = 0. \end{aligned}$$

\square

Below, we present the relation between $\tilde{\rho}_j^i$ and ρ_j^i locally on patches, more precisely when tested against functions of the local spaces $V_{j,\mathbf{a}}^i$ given by (2.7).

Lemma 9.4 (Local relation between $\tilde{\rho}_j^i$ and $\rho_{j,\mathbf{a}}^i$). *Let $j \in \{1, \dots, J\}$. Let $\tilde{\rho}_j^i$, ρ_j^i be given by (3.3), (4.2), respectively. For all $\mathbf{a} \in \mathcal{V}_{j-s}$ and all $v_{j,\mathbf{a}} \in V_{j,\mathbf{a}}^i$, we have*

$$(\mathcal{K} \nabla \tilde{\rho}_j^i, \nabla v_{j,\mathbf{a}})_{\omega_j^{\mathbf{a}}} = (\mathcal{K} \nabla \rho_{j,\mathbf{a}}^i, \nabla v_{j,\mathbf{a}})_{\omega_j^{\mathbf{a}}} - \sum_{k=1}^{j-1} (\mathcal{K} \nabla (\tilde{\rho}_k^i - \lambda_k^i \rho_k^i), \nabla v_{j,\mathbf{a}})_{\omega_j^{\mathbf{a}}}, \quad (9.4)$$

where $\rho_{j,\mathbf{a}}^i \in V_{j,\mathbf{a}}^i$ is defined as the solution of a local problem by (4.1). We use the convention that the sum in the relation above is zero when $j=1$.

Proof. We take $v_{j,\mathbf{a}} \in V_{j,\mathbf{a}}^i$. This implies that $v_{j,\mathbf{a}}$ is zero on the boundary of the patch domain $\omega_j^{\mathbf{a}}$. Since $v_{j,\mathbf{a}} \in V_J^p$ when $j \in \{1, \dots, J-1\}$ and $v_{j,\mathbf{a}} \in V_J^p$ otherwise, we can use it as a test function in the definition of $\tilde{\rho}_j^i$ in (3.3) as well as in the definition of $\rho_{j,\mathbf{a}}^i$ in (4.1). We conclude by subtracting the two following identities

$$\begin{aligned} (\mathcal{K} \nabla \tilde{\rho}_j^i, \nabla v_{j,\mathbf{a}})_{\omega_j^{\mathbf{a}}} &= (f, v_{j,\mathbf{a}})_{\omega_j^{\mathbf{a}}} - (\mathcal{K} \nabla u_J^i, \nabla v_{j,\mathbf{a}})_{\omega_j^{\mathbf{a}}} - \sum_{k=0}^{j-1} (\mathcal{K} \nabla \tilde{\rho}_k^i, \nabla v_{j,\mathbf{a}})_{\omega_j^{\mathbf{a}}}, \\ (\mathcal{K} \nabla \rho_{j,\mathbf{a}}^i, \nabla v_{j,\mathbf{a}})_{\omega_j^{\mathbf{a}}} &= (f, v_{j,\mathbf{a}})_{\omega_j^{\mathbf{a}}} - (\mathcal{K} \nabla u_J^i, \nabla v_{j,\mathbf{a}})_{\omega_j^{\mathbf{a}}} - \sum_{k=0}^{j-1} \lambda_k^i (\mathcal{K} \nabla \rho_k^i, \nabla v_{j,\mathbf{a}})_{\omega_j^{\mathbf{a}}}. \end{aligned}$$

□

9.3 Proof of Theorem 6.6 (p -robust estimator efficiency in low-regularity setting)

We begin by presenting here a result given in [19], obtained by a combination of a one-level p -robust stable decomposition proven in Schöberl *et al.* [26] and a multilevel stable decomposition for piecewise linear functions given in Xu *et al.* [34].

Lemma 9.5 (p -robust multilevel stable decomposition). *Let $v_J \in V_J^p$. Under Assumption 6.1, and either of the Assumptions 6.2 or Assumption 6.3, we have*

$$v_J = \sum_{j=0}^J \sum_{\mathbf{a} \in \mathcal{V}_j} v_{j,\mathbf{a}}, \quad v_{j,\mathbf{a}} \in V_j^{\mathbf{a}} \quad (9.5)$$

$$\sum_{j=0}^J \sum_{\mathbf{a} \in \mathcal{V}_j} \|\mathcal{K}^{\frac{1}{2}} \nabla v_{j,\mathbf{a}}\|_{\omega_j^{\mathbf{a}}}^2 \leq C_S \|\mathcal{K}^{\frac{1}{2}} \nabla v_J\|^2, \quad (9.6)$$

where $C_S \geq 1$ only depends on the space dimension d , and the mesh shape regularity parameter $\kappa_{\mathcal{T}}$. If Assumption 6.2 is satisfied, C_S also depends on the maximum strength of refinement parameter C_{ref} , and C_{qu} . If Assumption 6.3 is satisfied, C_{ref} also depends on the local quasi-uniformity parameter $C_{\text{loc,qu}}$ and quasi-uniformity parameter of the coarse mesh C_{qu}^0 .

The previous results and properties allow us to give a concise proof in the following.

Proof of Theorem 6.6 (p -robust estimator efficiency in the general setting). First, note that by (3.5), we have $\|\mathcal{K}^{\frac{1}{2}} \nabla (u_J - u_J^i)\| = \|\mathcal{K}^{\frac{1}{2}} \nabla \tilde{\rho}_{J,\text{alg}}^i\|$. Thus, we work with the exact algebraic residual lifting $\tilde{\rho}_{J,\text{alg}}^i$. We begin by first applying Lemma 9.5

$$\tilde{\rho}_{J,\text{alg}}^i = \sum_{j=0}^J \sum_{\mathbf{a} \in \mathcal{V}_j} e_{j,\mathbf{a}}, \quad e_{j,\mathbf{a}} \in V_j^{\mathbf{a}} \quad (9.7)$$

$$\sum_{j=0}^J \sum_{\mathbf{a} \in \mathcal{V}_j} \|\mathcal{K}^{\frac{1}{2}} \nabla e_{j,\mathbf{a}}\|_{\omega_j^{\mathbf{a}}}^2 \leq C_S \|\mathcal{K}^{\frac{1}{2}} \nabla \tilde{\rho}_{J,\text{alg}}^i\|^2, \quad (9.8)$$

Taking into account the variations of the diffusion coefficient \mathcal{K} , we have

$$\sum_{j=0}^J \sum_{\mathbf{a} \in \mathcal{V}_j} \|\mathcal{K}^{\frac{1}{2}} \nabla e_{j,\mathbf{a}}\|_{\omega_j^{\mathbf{a}}}^2 \leq C_S \kappa \|\mathcal{K}^{\frac{1}{2}} \nabla \tilde{\rho}_{J,\text{alg}}^i\|^2. \quad (9.9)$$

We use this decomposition to estimate

$$\begin{aligned} & \|\mathcal{K}^{\frac{1}{2}} \nabla \tilde{\rho}_{J,\text{alg}}^i\|^2 \stackrel{(9.7)}{=} \left(\mathcal{K} \nabla \tilde{\rho}_{J,\text{alg}}^i, \sum_{j=0}^J \sum_{\mathbf{a} \in \mathcal{V}_j} \nabla e_{j,\mathbf{a}} \right) \stackrel{(3.2)}{=} \left(\mathcal{K} \nabla \rho_0^i, \sum_{\mathbf{a} \in \mathcal{V}_0} \nabla e_{0,\mathbf{a}} \right) + \sum_{j=1}^J \sum_{\mathbf{a} \in \mathcal{V}_j} \left(\mathcal{K} \nabla \tilde{\rho}_{J,\text{alg}}^i, \nabla e_{j,\mathbf{a}} \right)_{\omega_j^{\mathbf{a}}} \\ & \stackrel{(3.4)}{=} \left(\mathcal{K} \nabla \rho_0^i, \sum_{\mathbf{a} \in \mathcal{V}_0} \nabla e_{0,\mathbf{a}} \right) + \sum_{j=1}^J \sum_{\mathbf{a} \in \mathcal{V}_j} \left((f, e_{j,\mathbf{a}})_{\omega_j^{\mathbf{a}}} - \left(\mathcal{K} \nabla u_J, \nabla e_{j,\mathbf{a}} \right)_{\omega_j^{\mathbf{a}}} \right) \\ & \stackrel{(4.1)}{\stackrel{(4.4)}{=}} \left(\mathcal{K} \nabla \rho_0^i, \sum_{\mathbf{a} \in \mathcal{V}_0} \nabla e_{0,\mathbf{a}} \right) + \sum_{j=1}^J \sum_{\mathbf{a} \in \mathcal{V}_j} \left(\left(\mathcal{K} \nabla \rho_{j,\mathbf{a}}^i, \nabla e_{j,\mathbf{a}} \right)_{\omega_j^{\mathbf{a}}} + \sum_{k=0}^{j-1} \left(\lambda_k^i \mathcal{K} \nabla \rho_k^i, \nabla e_{j,\mathbf{a}} \right)_{\omega_j^{\mathbf{a}}} \right) \\ & = \left(\mathcal{K} \nabla \rho_0^i, \sum_{\mathbf{a} \in \mathcal{V}_0} \nabla e_{0,\mathbf{a}} \right) + \sum_{j=1}^J \sum_{\mathbf{a} \in \mathcal{V}_j} \left(\mathcal{K} \nabla \rho_{j,\mathbf{a}}^i, \nabla e_{j,\mathbf{a}} \right)_{\omega_j^{\mathbf{a}}} + \sum_{j=1}^J \sum_{k=0}^{j-1} \left(\lambda_k^i \mathcal{K} \nabla \rho_k^i, \sum_{\mathbf{a} \in \mathcal{V}_j} \nabla e_{j,\mathbf{a}} \right). \end{aligned}$$

We will now estimate each of the above three terms using Young's inequality and patch overlap arguments as done in the proof of Lemma 9.1. First, we have, using the fact that $\lambda_0^i = 1$

$$\left(\mathcal{K}\nabla\rho_0^i, \sum_{\mathbf{a}\in\mathcal{V}_0} \nabla e_{0,\mathbf{a}}\right) \leq \frac{(d+1)C_{S,\mathcal{K}}}{2} \left(\lambda_0^i \|\mathcal{K}^{\frac{1}{2}}\nabla\rho_0^i\|\right)^2 + \frac{1}{2(d+1)C_{S,\mathcal{K}}} (d+1) \sum_{\mathbf{a}\in\mathcal{V}_0} \|\mathcal{K}^{\frac{1}{2}}\nabla e_{0,\mathbf{a}}\|_{\omega_0^{\mathbf{a}}}^2.$$

For the second term, we similarly obtain

$$\begin{aligned} \sum_{j=1}^J \sum_{\mathbf{a}\in\mathcal{V}_j} \left(\mathcal{K}\nabla\rho_{j,\mathbf{a}}^i, \nabla e_{j,\mathbf{a}}\right)_{\omega_j^{\mathbf{a}}} &\leq C_{S,\mathcal{K}} \sum_{j=1}^J \sum_{\mathbf{a}\in\mathcal{V}_j} \|\mathcal{K}^{\frac{1}{2}}\nabla\rho_{j,\mathbf{a}}^i\|_{\omega_j^{\mathbf{a}}}^2 + \frac{1}{4C_{S,\mathcal{K}}} \sum_{j=1}^J \sum_{\mathbf{a}\in\mathcal{V}_j} \|\mathcal{K}^{\frac{1}{2}}\nabla e_{j,\mathbf{a}}\|_{\omega_j^{\mathbf{a}}}^2 \\ &\stackrel{(9.1)}{\leq} C_{S,\mathcal{K}}(d+1) \sum_{j=1}^J \left(\lambda_j^i \|\mathcal{K}^{\frac{1}{2}}\nabla\rho_j^i\|\right)^2 + \frac{1}{4C_{S,\mathcal{K}}} \sum_{j=1}^J \sum_{\mathbf{a}\in\mathcal{V}_j} \|\mathcal{K}^{\frac{1}{2}}\nabla e_{j,\mathbf{a}}\|_{\omega_j^{\mathbf{a}}}^2 \end{aligned}$$

Finally, for the third term we have

$$\begin{aligned} \sum_{j=1}^J \sum_{k=0}^{j-1} \left(\lambda_k^i \mathcal{K}\nabla\rho_k^i, \sum_{\mathbf{a}\in\mathcal{V}_j} \nabla e_{j,\mathbf{a}}\right) &\leq \frac{2(d+1)C_{S,\mathcal{K}}J}{2} \sum_{j=1}^J \sum_{k=0}^{j-1} \left(\lambda_k^i \|\mathcal{K}^{\frac{1}{2}}\nabla\rho_k^i\|\right)^2 + \frac{\sum_{j=1}^J \sum_{k=0}^{j-1} \left\|\mathcal{K}^{\frac{1}{2}} \sum_{\mathbf{a}\in\mathcal{V}_j} \nabla e_{j,\mathbf{a}}\right\|^2}{2(2(d+1)C_{S,\mathcal{K}}J)} \\ &\leq (d+1)C_{S,\mathcal{K}}J^2 \sum_{k=0}^J \left(\lambda_k^i \|\mathcal{K}^{\frac{1}{2}}\nabla\rho_k^i\|\right)^2 + \frac{1}{4C_{S,\mathcal{K}}} \sum_{j=1}^J \sum_{\mathbf{a}\in\mathcal{V}_j} \|\mathcal{K}^{\frac{1}{2}}\nabla e_{j,\mathbf{a}}\|_{\omega_{j,0}^{\mathbf{b}}}^2 \end{aligned}$$

Summing these components together, we can now pursue our main estimate

$$\begin{aligned} \|\mathcal{K}^{\frac{1}{2}}\nabla\tilde{\rho}_{J,\text{alg}}^i\|^2 &\leq 2(d+1)C_{S,\mathcal{K}}J^2 \sum_{j=0}^J \left(\lambda_j^i \|\mathcal{K}^{\frac{1}{2}}\nabla\rho_j^i\|\right)^2 + \frac{1}{2C_{S,\mathcal{K}}} \sum_{j=0}^J \sum_{\mathbf{a}\in\mathcal{V}_j} \|\mathcal{K}^{\frac{1}{2}}\nabla e_{j,\mathbf{a}}\|_{\omega_{j,0}^{\mathbf{b}}}^2 \\ &\stackrel{(5.1)}{\leq} 2(d+1)C_{S,\mathcal{K}}J^2 \left(\eta_{\text{alg}}^i\right)^2 + \frac{1}{2} \|\mathcal{K}^{\frac{1}{2}}\nabla\tilde{\rho}_{J,\text{alg}}^i\|^2 \end{aligned}$$

After subtracting on both sides $\frac{1}{2}\|\mathcal{K}^{\frac{1}{2}}\nabla\tilde{\rho}_{J,\text{alg}}^i\|^2$, we finally obtain the desired result.

$$\|\mathcal{K}^{\frac{1}{2}}\nabla\tilde{\rho}_{J,\text{alg}}^i\|^2 \leq 4(d+1)C_{S,\mathcal{K}}J^2 \left(\eta_{\text{alg}}^i\right)^2 \quad (9.10)$$

□

Proof of Corollary 6.8. First note that by $\lambda_j^i \leq R$, $1 \leq j \leq J$, and an overlapping argument we have

$$\left(\lambda_j^i \|\mathcal{K}^{\frac{1}{2}}\nabla\rho_j^i\|\right)^2 \leq R^2 \|\mathcal{K}^{\frac{1}{2}}\nabla\rho_j^i\|^2 \leq R^2(d+1) \sum_{\mathbf{a}\in\mathcal{V}_j} \|\mathcal{K}^{\frac{1}{2}}\nabla\rho_{j,\mathbf{a}}^i\|_{\omega_j^{\mathbf{a}}}^2. \quad (9.11)$$

Now, from the main proof we can conclude

$$\begin{aligned} \left(\eta_{\text{alg}}^i\right)^2 &\stackrel{(5.2)}{\leq} \left\|\mathcal{K}^{\frac{1}{2}}\nabla(u_J - u_J^i)\right\|^2 \stackrel{(9.10)}{\leq} 4(d+1)C_{S,\mathcal{K}}J^2 \left(\eta_{\text{alg}}^i\right)^2 \stackrel{(5.1)}{=} 4(d+1)C_{S,\mathcal{K}}J^2 \sum_{j=0}^J \left(\lambda_j^i \|\mathcal{K}^{\frac{1}{2}}\nabla\rho_j^i\|\right)^2 \\ &\stackrel{(9.11)}{\leq} 4(d+1)^2 C_{S,\mathcal{K}}J^2 R^2 \sum_{j=0}^J \sum_{\mathbf{a}\in\mathcal{V}_j} \|\mathcal{K}^{\frac{1}{2}}\nabla\rho_{j,\mathbf{a}}^i\|_{\omega_j^{\mathbf{a}}}^2 \stackrel{(9.2)}{\leq} 4(d+1)^3 C_{S,\mathcal{K}}J^2 R^2 \left(\eta_{\text{alg}}^i\right)^2 \end{aligned}$$

□

9.4 Proof of Theorem 6.6 (p - and J -robust estimator efficiency in H^2 -regularity setting)

We now work in the setting of Assumption 6.4, we can prove that the result of Theorem 6.6 holds not only p -robustly but also J -robustly.

Remark 9.6 (Relation of consecutive mesh sizes). *Putting together Assumptions 6.4, we have*

$$C_{\text{qu}}C_{\text{ref}}h_{j-1} \leq h_j. \quad (9.12)$$

In order to obtain J -robustness in addition to p -robustness of our main result, we proceed by exhibiting another level-wise stable decomposition of the algebraic error components. Firstly, we define the piecewise linear component of the stable decomposition via a H^1 -orthogonal projection. Then we use a duality type argument to obtain the properties of the decomposition that will allow to estimate the algebraic error p - and J -robustly by our a posteriori error estimator.

Definition 9.7 (H^1 -orthogonal lowest-order projection of error components). *For any $j \in \{1, \dots, J\}$, $\tilde{\rho}_j^i$ given by (3.3), let $c_j^i \in V_j^1$ be the solution of*

$$(\nabla c_j^i, \nabla v_j) = (\nabla \tilde{\rho}_j^i, \nabla v_j), \quad \forall v_j \in V_j^1. \quad (9.13)$$

Remark 9.8 (Orthogonality properties of c_j^i). *Note that due to Definition 9.7, for any $j \in \{1, \dots, J\}$, c_j^i satisfies the following orthogonality with piecewise linear functions of previous levels*

$$(\nabla c_j^i, \nabla v_k) \stackrel{(9.13)}{=} (\nabla \tilde{\rho}_j^i, \nabla v_k) \stackrel{(9.3)}{=} 0, \quad \forall v_k \in V_k^1, \quad \forall k < j. \quad (9.14)$$

Lemma 9.9 (H^2 -regularity result). *Under Assumptions 6.4, 6.3, for any $j \in \{1, \dots, J\}$, $\tilde{\rho}_j^i$ given by (3.3), and c_j^i given by Definition 9.7, there holds*

$$\|\tilde{\rho}_j^i\| \lesssim h_j \|\nabla \tilde{\rho}_j^i\|, \quad (9.15)$$

$$\|c_j^i\| \lesssim h_j \|\nabla c_j^i\|. \quad (9.16)$$

Proof. To prove the first result, we proceed by a standard duality argument.

We consider the following problem: find $\xi_j \in H_0^1(\Omega)$ such that

$$(\nabla \xi_j, \nabla v) = (\tilde{\rho}_j^i, v) \quad \forall v \in H_0^1(\Omega). \quad (9.17)$$

Following Grisvard [14, Theorem 4.3.1.4], under Assumption 6.4 $\xi_j \in H^2(\Omega)$ and

$$|\xi_j|_{H^2(\Omega)} = \|\Delta \xi_j\| = \|\tilde{\rho}_j^i\|. \quad (9.18)$$

Consider $\mathbb{I}_{j-1}^1(\xi_j)$ the \mathbb{P}^1 -Lagrange interpolation of ξ_j on mesh level $j-1$. Since $\xi_j \in H^2(\Omega)$, following, e.g., Ern and Guermond [12, Corollary 1.110], we obtain

$$\|\nabla(\xi_j - \mathbb{I}_{j-1}^1(\xi_j))\| \leq C_{\text{app}}h_{j-1}|\xi_j|_{H^2(\Omega)}. \quad (9.19)$$

In particular: $\mathbb{I}_{j-1}^1(\xi_j) \in V_{j-1}$, so by the orthogonality relation (9.3)

$$(\nabla \mathbb{I}_{j-1}^1(\xi_j), \nabla \tilde{\rho}_j^i) = 0. \quad (9.20)$$

We have now all the elements to conclude

$$\begin{aligned} \|\tilde{\rho}_j^i\|^2 &\stackrel{(9.17)}{=} (\nabla \xi_j, \nabla \tilde{\rho}_j^i) \stackrel{(9.20)}{=} (\nabla(\xi_j - \mathbb{I}_{j-1}^1(\xi_j)), \nabla \tilde{\rho}_j^i) \leq \|\nabla(\xi_j - \mathbb{I}_{j-1}^1(\xi_j))\| \|\nabla \tilde{\rho}_j^i\| \\ &\stackrel{(9.19)}{\leq} C_{\text{app}}h_{j-1}|\xi_j|_{H^2(\Omega)} \|\nabla \tilde{\rho}_j^i\| \stackrel{(9.18)}{=} C_{\text{app}}h_{j-1}\|\tilde{\rho}_j^i\| \|\nabla \tilde{\rho}_j^i\| \stackrel{(9.12)}{\leq} \frac{C_{\text{app}}}{C_{\text{qu}}C_{\text{ref}}}h_j\|\tilde{\rho}_j^i\| \|\nabla \tilde{\rho}_j^i\| \end{aligned}$$

which gives us $\|\tilde{\rho}_j^i\| \lesssim h_j \|\nabla \tilde{\rho}_j^i\|$.

To obtain (9.16), we use the same argument. It is enough to appropriately modify the right hand side of the dual problem (9.17), and replace the orthogonality relation (9.20) by (9.14). \square

We can now present the stable decomposition that will help us prove the efficiency of the a posteriori error estimator.

Lemma 9.10 (Stable decomposition of the error level-wise components). *For all $j \in \{1, \dots, J\}$, $\tilde{\rho}_j^i$ given by (3.3) and c_j^i given by Definition 9.7, there exist $\tilde{\rho}_{j,\mathbf{a}}^i \in V_j^{\mathbf{a}}$, such that*

$$\tilde{\rho}_j^i = c_j^i + \sum_{\mathbf{a} \in \mathcal{V}_j} \tilde{\rho}_{j,\mathbf{a}}^i, \quad (9.21)$$

$$\|\nabla c_j^i\|^2 + \sum_{\mathbf{a} \in \mathcal{V}_j} \|\nabla \tilde{\rho}_{j,\mathbf{a}}^i\|_{\omega_j^{\mathbf{a}}}^2 \leq C_{\text{SD}} \|\nabla \tilde{\rho}_j^i\|^2. \quad (9.22)$$

Proof. We now rely on the stable decomposition result of Schöberl *et al.* [26] for one-level setting. We will first show that c_j^i satisfies

$$\|\nabla c_j^i\|^2 + \|\nabla(\tilde{\rho}_j^i - c_j^i)\|^2 + \sum_{K \in \mathcal{T}_j} h_K^{-2} \|(\tilde{\rho}_j^i - c_j^i)\|_K^2 \lesssim \|\nabla \tilde{\rho}_j^i\|^2. \quad (9.23)$$

Then as following [26] this will give us

$$\|\nabla c_j^i\|^2 + \sum_{\mathbf{a} \in \mathcal{V}_j} \|\nabla \tilde{\rho}_{j,\mathbf{a}}^i\|_{\omega_j^{\mathbf{a}}}^2 \leq C_{\text{SD}} \|\nabla \tilde{\rho}_j^i\|^2.$$

To show (9.23), we first use the Definition 9.7 of c_j^i

$$\|\nabla c_j^i\|^2 = (\nabla c_j^i, \nabla c_j^i) \stackrel{(9.13)}{=} (\nabla c_j^i, \nabla \tilde{\rho}_j^i) \leq \|\nabla c_j^i\| \|\nabla \tilde{\rho}_j^i\|.$$

This allows to estimate the first and second term (after using the triangle inequality) of (9.23). The third term is then estimated by

$$\begin{aligned} \sum_{K \in \mathcal{T}_j} h_K^{-2} \|(\tilde{\rho}_j^i - c_j^i)\|_K^2 &\stackrel{\text{Assumption 6.3}}{\leq} C_{\text{qu}}^{-2} h_j^{-2} \sum_{K \in \mathcal{T}_j} \|(\tilde{\rho}_j^i - c_j^i)\|_K^2 \leq 2C_{\text{qu}}^{-2} h_j^{-2} (\|\tilde{\rho}_j^i\|^2 + \|c_j^i\|^2) \\ &\stackrel{(9.15)}{\leq} 2 \left(\frac{C_{\text{app}}}{C_{\text{ref}} C_{\text{qu}}^2} \right)^2 (\|\nabla \tilde{\rho}_j^i\|^2 + \|\nabla c_j^i\|^2) \leq 4 \left(\frac{C_{\text{app}}}{C_{\text{ref}} C_{\text{qu}}^2} \right)^2 \|\nabla \tilde{\rho}_j^i\|^2. \end{aligned}$$

□

Remark 9.11 (Localized writing of level-wise components). *Note that since $\tilde{\rho}_j^i = c_j^i + \sum_{\mathbf{a} \in \mathcal{V}_j} \tilde{\rho}_{j,\mathbf{a}}^i$, we can decompose the piecewise linear $c_j^i \in V_j^1$ using nodal basis functions. We can then write*

$$\tilde{\rho}_j^i = c_j^i + \sum_{\mathbf{a} \in \mathcal{V}_j} \tilde{\rho}_{j,\mathbf{a}}^i = \sum_{\mathbf{a} \in \mathcal{V}_j} (c_{j,\mathbf{a}}^i \psi_{j,\mathbf{a}} + \tilde{\rho}_{j,\mathbf{a}}^i), \quad (9.24)$$

where $c_{j,\mathbf{a}}^i$ is the nodal value on vertex $\mathbf{a} \in \mathcal{V}_j$, and $c_{j,\mathbf{a}}^i \psi_{j,\mathbf{a}} + \tilde{\rho}_{j,\mathbf{a}}^i \in V_j^{\mathbf{a}}$.

Lemma 9.12 (L^2 -stability of nodal decomposition). *For all $j \in \{1, \dots, J\}$, and any $v_j \in V_j^1$ decomposed into basis hat functions $v_j = \sum_{\mathbf{a} \in \mathcal{V}_j} v_{j,\mathbf{a}} \psi_{j,\mathbf{a}}$, we have*

$$\|v_j\|^2 \leq (d+1) \sum_{\mathbf{a} \in \mathcal{V}_j} \|v_{j,\mathbf{a}} \psi_{j,\mathbf{a}}\|_{\omega_j^{\mathbf{a}}}^2, \quad \text{and} \quad \sum_{\mathbf{a} \in \mathcal{V}_j} \|v_{j,\mathbf{a}} \psi_{j,\mathbf{a}}\|_{\omega_j^{\mathbf{a}}}^2 \leq C_{\text{nd}} \|v_j\|^2, \quad (9.25)$$

where C_{nd} only depends on space dimension d and the mesh shape regularity parameter $\kappa_{\mathcal{T}}$.

Proof. For the first estimate, we apply the usual overlapping argument as done for (9.1). As for the second estimate, consider a patch $\omega_j^{\mathbf{a}}$ and K element contained in the patch. Since $v_j \in V_j^1$ and by equivalence of norms in finite dimension, we have

$$\|v_{j,\mathbf{a}} \psi_{j,\mathbf{a}}\|_{\omega_j^{\mathbf{a}}} \approx \|v_{j,\mathbf{a}} \psi_{j,\mathbf{a}}\|_K \leq \|v_{j,\mathbf{a}} \psi_{j,\mathbf{a}}\|_{\infty} |K|^{\frac{1}{2}} \leq \left\| \sum_{\mathbf{a} \in \mathcal{V}_K} v_{j,\mathbf{a}} \psi_{j,\mathbf{a}} \right\|_{\infty} |K|^{\frac{1}{2}} = \|v_j\|_{\infty, K} |K|^{\frac{1}{2}} \lesssim \|v_j\|_K.$$

The result is obtained by summing both sides over all vertices. □

Lemma 9.13 (Level-wise estimation of c_j^i). *Let $j \in \{1, \dots, J\}$, let $c_j^i = \sum_{\mathbf{a} \in \mathcal{V}_j} c_{j,\mathbf{a}}^i \psi_{j,\mathbf{a}}$ be given by Definition 9.7. There holds*

$$\sum_{\mathbf{a} \in \mathcal{V}_j} \|c_{j,\mathbf{a}}^i \nabla \psi_{j,\mathbf{a}}\|_{\omega_j^{\mathbf{a}}}^2 \leq C_{\text{stab}} \|\nabla c_j^i\|^2, \quad (9.26)$$

where $C_{\text{stab}, \kappa}$ only depends on the space dimension d , the mesh shape regularity parameter $\kappa_{\mathcal{T}}$, quasi-uniformity parameter C_{qu} , and the strength refinement parameter C_{ref} .

Proof. We start by using an inverse inequality, we denote by $h_{\omega_j^{\mathbf{a}}}$ the diameter of patch $\omega_j^{\mathbf{a}}$, then use the quasi-uniformity Assumption 6.3

$$\begin{aligned} \sum_{\mathbf{a} \in \mathcal{V}_j} \|c_{j,\mathbf{a}}^i \nabla \psi_{j,\mathbf{a}}\|_{\omega_j^{\mathbf{a}}}^2 &\leq C_{\text{inv}} \sum_{\mathbf{a} \in \mathcal{V}_j} h_{\omega_j^{\mathbf{a}}}^{-2} \|c_{j,\mathbf{a}}^i \psi_{j,\mathbf{a}}\|_{\omega_j^{\mathbf{a}}}^2 \leq C_{\text{qu}}^{-2} C_{\text{inv}} h_j^{-2} \sum_{\mathbf{a} \in \mathcal{V}_j} \|c_{j,\mathbf{a}}^i \psi_{j,\mathbf{a}}\|_{\omega_j^{\mathbf{a}}}^2 \\ &\stackrel{(9.25)}{\leq} C_{\text{qu}}^{-2} C_{\text{inv}} C_{\text{nd}} h_j^{-2} \|c_j^i\|^2 \stackrel{(9.16)}{\leq} C_{\text{stab}} \|\nabla c_j^i\|^2, \quad C_{\text{stab}} := \frac{C_{\text{inv}} C_{\text{nd}} C_{\text{app}}^2}{C_{\text{qu}}^4 C_{\text{ref}}^2} \end{aligned}$$

□

Lemma 9.14 (p -robust level-wise error estimation). *Let $j \in \{1, \dots, J\}$, let $\tilde{\rho}_j^i$ and ρ_j^i be defined by (3.3) and (4.2), respectively. There holds*

$$\|\nabla \tilde{\rho}_j^i\|^2 \lesssim \sum_{\mathbf{a} \in \mathcal{V}_j} \|\nabla \rho_{j,\mathbf{a}}^i\|_{\omega_j^{\mathbf{a}}}^2. \quad (9.27)$$

Proof. Let $j \in \{1, \dots, J\}$, we begin by using the splitting (9.21)

$$\begin{aligned} \|\nabla \tilde{\rho}_j^i\|^2 &\stackrel{(9.24)}{=} \sum_{\mathbf{a} \in \mathcal{V}_j} (\nabla \tilde{\rho}_j^i, \nabla (c_{j,\mathbf{a}}^i \psi_{j,\mathbf{a}} + \tilde{\rho}_{j,\mathbf{a}}^i))_{\omega_j^{\mathbf{a}}} \\ &\stackrel{(9.4)}{=} \sum_{\mathbf{a} \in \mathcal{V}_j} \left((\nabla \rho_{j,\mathbf{a}}^i, \nabla (c_{j,\mathbf{a}}^i \psi_{j,\mathbf{a}} + \tilde{\rho}_{j,\mathbf{a}}^i))_{\omega_j^{\mathbf{a}}} - \sum_{k=1}^{j-1} (\nabla (\tilde{\rho}_k^i - \lambda_k^i \rho_k^i), \nabla (c_{j,\mathbf{a}}^i \psi_{j,\mathbf{a}} + \tilde{\rho}_{j,\mathbf{a}}^i))_{\omega_j^{\mathbf{a}}} \right) \\ &\stackrel{(9.24)}{=} \sum_{\mathbf{a} \in \mathcal{V}_j} (\nabla \rho_{j,\mathbf{a}}^i, \nabla (c_{j,\mathbf{a}}^i \psi_{j,\mathbf{a}} + \tilde{\rho}_{j,\mathbf{a}}^i))_{\omega_j^{\mathbf{a}}} - \sum_{k=1}^{j-1} (\nabla (\tilde{\rho}_k^i - \lambda_k^i \rho_k^i), \nabla \tilde{\rho}_j^i) \\ &\stackrel{(9.3)}{=} \sum_{\mathbf{a} \in \mathcal{V}_j} (\sqrt{2C_{\text{SD}} C_{\text{stab}} (d+1)} \nabla \rho_{j,\mathbf{a}}^i, \frac{1}{\sqrt{2C_{\text{SD}} C_{\text{stab}} (d+1)}} \nabla (c_{j,\mathbf{a}}^i \psi_{j,\mathbf{a}} + \tilde{\rho}_{j,\mathbf{a}}^i))_{\omega_j^{\mathbf{a}}} - 0 \\ &\leq C_{\text{SD}} C_{\text{stab}} (d+1) \sum_{\mathbf{a} \in \mathcal{V}_j} \|\nabla \rho_{j,\mathbf{a}}^i\|_{\omega_j^{\mathbf{a}}}^2 + \frac{1}{4C_{\text{SD}} C_{\text{stab}} (d+1)} \sum_{\mathbf{a} \in \mathcal{V}_j} \left\| \nabla (c_{j,\mathbf{a}}^i \psi_{j,\mathbf{a}} + \tilde{\rho}_{j,\mathbf{a}}^i) \right\|_{\omega_j^{\mathbf{a}}}^2 \\ &\leq C_{\text{SD}} C_{\text{stab}} (d+1) \sum_{\mathbf{a} \in \mathcal{V}_j} \|\nabla \rho_{j,\mathbf{a}}^i\|_{\omega_j^{\mathbf{a}}}^2 + \frac{1}{2C_{\text{SD}} C_{\text{stab}}} \sum_{\mathbf{a} \in \mathcal{V}_j} (\|c_{j,\mathbf{a}}^i \nabla \psi_{j,\mathbf{a}}\|_{\omega_j^{\mathbf{a}}}^2 + \|\nabla \tilde{\rho}_{j,\mathbf{a}}^i\|_{\omega_j^{\mathbf{a}}}^2) \\ &\stackrel{(9.26)}{\leq} C_{\text{SD}} C_{\text{stab}} (d+1) \sum_{\mathbf{a} \in \mathcal{V}_j} \|\nabla \rho_{j,\mathbf{a}}^i\|_{\omega_j^{\mathbf{a}}}^2 + \frac{1}{2C_{\text{SD}} C_{\text{stab}}} \left(C_{\text{stab}} \|\nabla c_j^i\|^2 + \sum_{\mathbf{a} \in \mathcal{V}_j} \|\nabla \tilde{\rho}_{j,\mathbf{a}}^i\|_{\omega_j^{\mathbf{a}}}^2 \right) \\ &\stackrel{(9.22)}{\leq} C_{\text{SD}} C_{\text{stab}} (d+1) \sum_{\mathbf{a} \in \mathcal{V}_j} \|\nabla \rho_{j,\mathbf{a}}^i\|_{\omega_j^{\mathbf{a}}}^2 + \frac{1}{2} \|\nabla \tilde{\rho}_j^i\|^2, \end{aligned}$$

which finally leads to the result

$$\|\nabla \tilde{\rho}_j^i\|^2 \leq 2C_{\text{SD}} C_{\text{stab}} (d+1) \sum_{\mathbf{a} \in \mathcal{V}_j} \|\nabla \rho_{j,\mathbf{a}}^i\|_{\omega_j^{\mathbf{a}}}^2 \quad (9.28)$$

□

We can now give a concise proof of Theorem 6.6.

Proof of Theorem 6.6 (p-robust estimator efficiency under additional assumptions). To estimate the algebraic error we use the level-wise decomposition (3.5). Each level's contribution was estimated in Lemma 9.14. We achieve the result by summing these estimates on different levels.

$$\begin{aligned} \|\nabla(u_J - u_J^i)\|^2 &\stackrel{(3.5)}{=} \|\nabla\rho_0^i\|^2 + \sum_{j=1}^J \|\nabla\tilde{\rho}_j^i\|^2 \stackrel{(9.28)}{\leq} \|\nabla\rho_0^i\|^2 + 2C_{\text{stab}}(d+1) \sum_{j=1}^J \sum_{\mathbf{a} \in \mathcal{V}_j} \|\nabla\rho_{j,\mathbf{a}}^i\|^2 \\ &\stackrel{(9.2)}{\leq} 2C_{\text{stab}}(d+1)^2 \left((\lambda_0^i \|\nabla\rho_0^i\|)^2 + \sum_{j=1}^J (\lambda_j^i \|\nabla\rho_j^i\|)^2 \right) = 2C_{\text{stab}}(d+1)^2 (\eta_{\text{alg}}^i)^2 \end{aligned} \quad (9.29)$$

Thus we showed $\eta_{\text{alg}}^i \geq \beta \|\nabla(u_J - u_J^i)\|$, for $\beta := \frac{1}{(d+1)\sqrt{2C_{\text{stab}}}} > 0$, depending only on the space dimension d , the mesh shape regularity parameter $\kappa_{\mathcal{T}}$, quasi-uniformity parameter C_{qu} , and the strength refinement parameter C_{ref} . \square

Proof of Corollary 6.8. As a consequence of the Proof of Theorem 6.6, we have

$$\begin{aligned} (\eta_{\text{alg}}^i)^2 &\stackrel{(5.2)}{\leq} \|\nabla(u_J - u_J^i)\|^2 \stackrel{(9.29)}{\leq} 2C_{\text{SD}}C_{\text{stab}}(d+1) \left(\|\nabla\rho_0^i\|^2 + \sum_{j=1}^J \sum_{\mathbf{a} \in \mathcal{V}_j} \|\nabla\rho_{j,\mathbf{a}}^i\|^2 \right) \\ &\stackrel{(9.2)}{\leq} 2C_{\text{SD}}C_{\text{stab}}(d+1)^2 (\eta_{\text{alg}}^i)^2 \end{aligned}$$

\square

10 Conclusions and future work

In this work we presented a multilevel algebraic solver, whose procedure naturally constructs an a posteriori estimator of the algebraic error on each iteration. The solver can be seen as a geometric multigrid solver relying on V-cycles with zero pre- and one post-smoothing, where the smoother is additive Schwarz associated to patches of elements (block Jacobi). A crucial difference compared to classic multigrid solvers is the use of an optimal step-size in the error correction stage that takes on each level of the mesh hierarchy. This significantly improves the behavior of the solver and conveniently enough, makes the analysis easier. We also present a simple and efficient way for the solver to pick adaptively the number of post-smoothing steps on each level. We show that the non-adaptive version of the solver (with only one post-smoothing step) contracts the error in each iteration robustly with respect to the polynomial degree p of the underlying finite element discretization. This result is moreover equivalent to showing the p -robust efficiency of the a posteriori error estimate. If we additionally assume H^2 -regularity of the solution, we can show that these results are also robust with respect to the number of mesh levels J . In this case, the error estimator is equivalent to a sum of highly-localized computable contributions. Future work will explore how to incorporate this information in the solver in order to make it adaptive by only tackling problematic regions which contribute most to the algebraic error. Numerical results indicate that even for singular test cases, the solver behaves robustly with respect to the polynomial degree p , number of levels J and also the diffusion coefficient \mathcal{K} .

References

- [1] P. F. ANTONIETTI AND G. PENNESI, *V-cycle multigrid algorithms for discontinuous Galerkin methods on non-nested polytopic meshes*, J. Sci. Comput., 78 (2019), pp. 625–652.
- [2] R. E. BANK, T. F. DUPONT, AND H. YSERENTANT, *The hierarchical basis multigrid method*, Numer. Math., 52 (1988), pp. 427–458.
- [3] J. H. BRAMBLE AND J. E. PASCIAK, *New convergence estimates for multigrid algorithms*, Math. Comp., 49 (1987), pp. 311–329.

- [4] J. H. BRAMBLE, J. E. PASCIAK, AND A. H. SCHATZ, *The construction of preconditioners for elliptic problems by substructuring. I*, Math. Comp., 47 (1986), pp. 103–134.
- [5] A. BRANDT, *Multi-level adaptive solutions to boundary-value problems*, Math. Comp., 31 (1977), pp. 333–390.
- [6] A. BRANDT, S. MCCORMICK, AND J. RUGE, *Algebraic multigrid (AMG) for sparse matrix equations*, in Sparsity and its applications (Loughborough, 1983), Cambridge Univ. Press, Cambridge, 1985, pp. 257–284.
- [7] C. CANUTO AND A. QUARTERONI, *Preconditioned minimal residual methods for Chebyshev spectral calculations*, J. Comput. Phys., 60 (1985), pp. 315–337.
- [8] T. F. CHAN AND W. L. WAN, *Robust multigrid methods for nonsmooth coefficient elliptic linear systems*, vol. 123, 2000, pp. 323–352. Numerical analysis 2000, Vol. III. Linear algebra.
- [9] L. CHEN, M. HOLST, J. XU, AND Y. ZHU, *Local multilevel preconditioners for elliptic equations with jump coefficients on bisection grids*, Comput. Vis. Sci., 15 (2012), pp. 271–289.
- [10] W. DÖRFLER, *A convergent adaptive algorithm for Poisson’s equation*, SIAM J. Numer. Anal., 33 (1996), pp. 1106–1124.
- [11] M. DRYJA, M. V. SARKIS, AND O. B. WIDLUND, *Multilevel Schwarz methods for elliptic problems with discontinuous coefficients in three dimensions*, Numer. Math., 72 (1996), pp. 313–348.
- [12] A. ERN AND J.-L. GUERMOND, *Theory and practice of finite elements*, vol. 159 of Applied Mathematical Sciences, Springer-Verlag, New York, 2004.
- [13] I. G. GRAHAM AND M. J. HAGGER, *Unstructured additive Schwarz-conjugate gradient method for elliptic problems with highly discontinuous coefficients*, SIAM J. Sci. Comput., 20 (1999), pp. 2041–2066.
- [14] P. GRISVARD, *Elliptic problems in nonsmooth domains*, vol. 24 of Monographs and Studies in Mathematics, Pitman (Advanced Publishing Program), Boston, MA, 1985.
- [15] W. HACKBUSCH, *Multi-grid methods and applications*, vol. 4 of Springer Series in Computational Mathematics, Springer, Berlin, 2003.
- [16] W. HEINRICHS, *Line relaxation for spectral multigrid methods*, J. Comput. Phys., 77 (1988), pp. 166–182.
- [17] B. JANSSEN AND G. KANSCHAT, *Adaptive multilevel methods with local smoothing for H^1 - and H^{curl} -conforming high order finite element methods*, SIAM J. Sci. Comput., 33 (2011), pp. 2095–2114.
- [18] J. P. LUCERO LORCA AND G. KANSCHAT, *Multilevel Schwarz preconditioners for singularly perturbed symmetric reaction-diffusion systems*. arXiv:1811.03839 preprint, 2018. URL <https://arxiv.org/abs/1811.03839v1>.
- [19] A. MIRAÇI, J. PAPEŽ, AND M. VOHRALÍK, *A multilevel algebraic error estimator and the corresponding iterative solver with p -robust behavior*. HAL preprint 02070981, 2019. URL <https://hal.inria.fr/hal-02070981>.
- [20] P. OSWALD, *Multilevel finite element approximation*, Teubner Skripten zur Numerik. [Teubner Scripts on Numerical Mathematics], B. G. Teubner, Stuttgart, 1994. Theory and applications.
- [21] J. PAPEŽ, U. RÜDE, M. VOHRALÍK, AND B. WOHLMUTH, *Sharp algebraic and total a posteriori error bounds for h and p finite elements via a multilevel approach. Recovering mass balance in any situation*. HAL preprint 01662944, 2019. URL <https://hal.inria.fr/hal-01662944>.
- [22] U. RÜDE, *Mathematical and computational techniques for multilevel adaptive methods*, vol. 13 of Frontiers in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1993.

- [23] U. RÜDE, *Error estimates based on stable splittings*, in Domain decomposition methods in scientific and engineering computing (University Park, PA, 1993), vol. 180 of Contemp. Math., Amer. Math. Soc., Providence, RI, 1994, pp. 111–118.
- [24] J. W. RUGE AND K. STÜBEN, *Algebraic multigrid*, in Multigrid methods, vol. 3 of Frontiers Appl. Math., SIAM, Philadelphia, PA, 1987, pp. 73–130.
- [25] R. SCHEICHL, P. S. VASSILEVSKI, AND L. T. ZIKATANOV, *Multilevel methods for elliptic problems with highly varying coefficients on nonaligned coarse grids*, SIAM J. Numer. Anal., 50 (2012), pp. 1675–1694.
- [26] J. SCHÖBERL, J. M. MELENK, C. PECHSTEIN, AND S. ZAGLMAYR, *Additive Schwarz preconditioning for p -version triangular and tetrahedral finite elements*, IMA J. Numer. Anal., 28 (2008), pp. 1–24.
- [27] E. G. SEWELL, *Automatic generation of triangulations for piecewise polynomial approximation*, ProQuest LLC, Ann Arbor, MI, 1972. Thesis (Ph.D.)—Purdue University.
- [28] N. SPILLANE, V. DOLEAN, P. HAURET, F. NATAF, C. PECHSTEIN, AND R. SCHEICHL, *A robust two-level domain decomposition preconditioner for systems of PDEs*, C. R. Math. Acad. Sci. Paris, 349 (2011), pp. 1255–1259.
- [29] N. SPILLANE, V. DOLEAN, P. HAURET, F. NATAF, C. PECHSTEIN, AND R. SCHEICHL, *Abstract robust coarse spaces for systems of PDEs via generalized eigenproblems in the overlaps*, Numer. Math., 126 (2014), pp. 741–770.
- [30] H. SUNDAR, G. STADLER, AND G. BIROS, *Comparison of multigrid algorithms for high-order continuous finite element discretizations*, Numer. Linear Algebra Appl., 22 (2015), pp. 664–680.
- [31] A. THEKALE, T. GRADL, K. KLAMROTH, AND U. RÜDE, *Optimizing the number of multigrid cycles in the full multigrid algorithm*, Numer. Linear Algebra Appl., 17 (2010), pp. 199–210.
- [32] P. S. VASSILEVSKI, *Hybrid V -cycle algebraic multilevel preconditioners*, Math. Comp., 58 (1992), pp. 489–512.
- [33] J. XU, *Iterative methods by space decomposition and subspace correction*, SIAM Rev., 34 (1992), pp. 581–613.
- [34] J. XU, L. CHEN, AND R. H. NOCHETTO, *Optimal multilevel methods for $H(\text{grad})$, $H(\text{curl})$, and $H(\text{div})$ systems on graded and unstructured grids*, in Multiscale, nonlinear and adaptive approximation, Springer, Berlin, 2009, pp. 599–659.
- [35] J. XU AND Y. ZHU, *Uniform convergent multigrid methods for elliptic problems with strongly discontinuous coefficients*, Math. Models Methods Appl. Sci., 18 (2008), pp. 77–105.