Multilevel Preconditioned Augmented Lagrangian Techniques for 2nd Order Mixed Problems

R. Hiptmair B. Wohlmuth T. Schiekofer *

Summary. We are concerned with the efficient solution of saddle point problems arising from the mixed discretization of 2nd order elliptic problems in two dimensions. We consider the mixed discretization of the boundary value problem by means of lowest order Raviart–Thomas elements. This leads to a saddle point problem, which can be tackled by Uzawa–like iterative solvers. We suggest a prior modification of the saddle point problem according to the augmented Lagrangian approach (cf. [16]) in order to make it more amenable to the iterative procedure. In order to boost the speed of iterative methods, we additionally employ a multilevel preconditioner first presented by Vassilevski and Wang in [27]. It is based on a special splitting of the space of vector valued fluxes, which exploits the close relationship between peicewise linear continuous finite element functions and divergence free fluxes. We prove that this splitting gives rise to an optimal preconditioner: it achieves condition numbers bounded independently on the depth of refinement. The proof is set in the framework of Schwarz methods (cf. [29, 30]). It relies on established results about standard multilevel methods as well as a strengthened Cauchy–Schwarz inequality for RT_0–spaces.

Key words. augmented Lagrangian algorithm, saddle point problems, mixed finite elements, multilevel preconditioners.

AMS(MOS) subject classifications. Primary 65F10; Secondary 65N20, 65N30.

Zusammenfassung. Gegenstand der Arbeit ist die effiziente Lösung von Sattelpunktproblemen, wie sie bei der gemischten Diskretisierung elliptischer Randwertprobleme zweiter Ordnung in 2D entstehen. Wir betrachten die gemischte Diskretisierung der Randwertprobleme mit Hilfe von Raviart–Thomas Elementen niedrigster Ordnung. Man erhält ein Sattelpunktproblem das sich durch vom Uzawa–Algorithmus abgeleitete iterative Verfahren lösen läßt. Wir schlagen vor, das Sattelpunktproblem zunächst mit Hilfe der Technik der Erweiterten Lagrangeschen Multiplikatoren zu modifizieren, um die iterative Lösung zu erleichtern. Um die Konvergenz der iterativen Verfahren zu beschleunigen, setzen wir einen Multilevel Vorkonditionierer ein, der zuerst von Vassilevski und Wang in [27] vorgeschlagen wurde. Er stützt sich auf eine spezielle Zerlegung des Raums der vektorwertigen Finite Elemente Ansatzfunktionen, die die enge Beziehung zwischen stückweise linearen, stetigen Finite Elemente Funktionen und divergenzfreien Flüssen ausnützt. Wir zeigen, daß diese Zerlegung auf einen optimalen Vorkonditionierer führt; er erzielt Konditionszahlen die nicht von der Verfeinerungstiefe abhängen. Der Beweis folgt dem üblichen Vorgehen im Fall additiver Schwarz Methoden (vgl. [29, 30]). Er stützt sich auf wohlbekannte Tatsachen über Standard Multilevelverfahren sowie auf eine verschärfte Cauchy–Schwarzsche Ungleichung.

*Universität Augsburg, Institut für Mathematik, Lehrstuhl für Angewandte Mathematik I, Universitätsstraße 14, e–mail: hiptmair@math.uni-augsburg.de. The first and the second author were supported by FORTWIHR, Bavarian Consortium for High Performance Scientific Computing

1 Introduction. For many application the core numerical task boils down to solving a 2nd order elliptic boundary value problem of the form

$$\begin{array}{rcll} -\mathrm{div}\,(a\nabla u) & = & f & \text{in } \Omega \subset \mathbb{R}^2 \\ u & = & g & \text{on } \Gamma_D \subset \partial\Omega \\ \langle a\nabla u, \boldsymbol{\nu}\rangle & = & 0 & \text{on } \Gamma_N \subset \partial\Omega \end{array} \tag{1}$$

where $f \in L^2(\Omega)$ and $a \in L^\infty(\Omega)$ with $\overline{\alpha} \geq a(\mathbf{x}) \geq \underline{\alpha} > 0$ a.e. in Ω. Especially, for irregularly shaped domains Ω or if adaptive techniques are contemplated, finite element methods (cf. [7]) offer a versatile and powerful tool for discretizing (1). The conventional approach starts with the primal variational formulation of (1). It leads to a positive definite linear system, for which highly efficient iterative solution methods have been developed over the last decades. Yet, in some cases *mixed finite element* schemes (see [8]) might be preferred, either for the sake of superior accuracy or compliance with basic physical laws. They are based on a minimization problem with a linear constraint:

$$J(\mathbf{j}) = \inf_{\mathrm{div}\,\mathbf{v}=-f} J(\mathbf{v})\;; \quad J(\mathbf{v}) := \frac{1}{2}\int_\Omega a^{-1}\,\langle \mathbf{v},\mathbf{v}\rangle\,d\mathbf{x} \tag{2}$$

Enforcing the constraint by means of a Lagrangian multiplier we arrive at a saddle point problem. The *dual variational formulation* of (1) characterizes its stationary point.

In a more lucid way, the dual formulation can also be obtained by first introducing the *flux* $\mathbf{j} = a(\mathbf{x})\nabla u$ and thus converting (1) into a first order system

$$\begin{array}{rcll} a(\mathbf{x})^{-1}\mathbf{j} - \nabla u & = & 0 & \text{in } \Omega \\ \mathrm{div}\,\mathbf{j} & = & -f & \text{in } \Omega \\ u & = & g & \text{on } \Gamma_D \\ \langle \mathbf{j}, \boldsymbol{\nu}\rangle & = & 0 & \text{on } \Gamma_N \end{array} \quad . \tag{3}$$

Then, using Green's formula, the individual equations are cast in weak form. Thus we end up with a variational problem posed over the Hilbert spaces $\boldsymbol{\mathcal{V}} := \boldsymbol{H}(\mathrm{div};\Omega,\Gamma_N)$ and $\mathcal{W} := L^2(\Omega)$: Seek $(\mathbf{j},u) \in \boldsymbol{H}(\mathrm{div};\Omega,\Gamma_N) \times L^2(\Omega)$ such that

$$\begin{array}{llcll} a(\mathbf{j},\mathbf{v}) & +\quad b(\mathbf{v},u) & = & g(\mathbf{v}) & \forall \mathbf{v} \in \boldsymbol{H}(\mathrm{div};\Omega,\Gamma_N) \\ b(\mathbf{j},w) & & = & f(w) & \forall w \in L^2(\Omega) \end{array} \tag{4}$$

where $a(\mathbf{j},\mathbf{v}) := \int_\Omega a^{-1}\,\langle \mathbf{j},\mathbf{v}\rangle\,d\mathbf{x}$ and $b(\mathbf{v},u) := \int_\Omega u\,\mathrm{div}\,\mathbf{v}\,d\mathbf{x}$. The linear forms f, g, given by $f(w) := -\int_\Omega f w\,d\mathbf{x}$ and $g(\mathbf{v}) := \int_{\Gamma_D} g\,\langle \mathbf{v},\boldsymbol{\nu}\rangle\,d\Gamma$, provide the source term and Dirichlet boundary data, respectively. Moreover we stick to the conventional notation

$$\boldsymbol{H}(\mathrm{div};\Omega,\Gamma_N) := \{\mathbf{v} \in \boldsymbol{L}^2(\Omega); \mathrm{div}\,\mathbf{v} \in L^2(\Omega), \langle \mathbf{v},\boldsymbol{\nu}\rangle_{|\Gamma_N} = 0\}$$

with the inner product

$$(\mathbf{j}, \mathbf{v})_{\boldsymbol{H}(\mathrm{div};\Omega)} := (\mathbf{j}, \mathbf{v})_{\boldsymbol{L}^2(\Omega)} + (\operatorname{div} \mathbf{v}, \operatorname{div} \mathbf{j})_{L^2(\Omega)} \quad .$$

For the sake of lucidity, we introduce an operator notation of (4):

$$\begin{array}{rcrcl} A\,\mathbf{j} & + & B^* u & = & g \\ B\,w & & & = & f \end{array} \quad . \tag{5}$$

Here $A : \boldsymbol{\mathcal{V}} \to \boldsymbol{\mathcal{V}}'$ and $B : \boldsymbol{\mathcal{V}} \to \mathcal{W}'$ are the operators, naturally associated with the bilinear forms $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$, respectively. The condition on $a(\mathbf{x})$ makes A a selfadjoint, positive definite operator. B is easily verified to be surjective, hence rendering B^* injective.

A great deal of research has been devoted to the efficient iterative solution of (5). This can be achieved by exploiting the close relationship between mixed schemes and standard finite element discretizations (see e.g. [6]). Conversely, Rusten and Winther in [26] tackled the system (5) as a whole and developed algebraic preconditioning techniques for the minimal residual method. The algebraic perspective also prevailed in the work of Ewing et al. ([13]). A different approach was pursued by Ewing and Wang ([14, 15]) and Vassilevski and Wang (cf. [27]). Their idea basically amounts to reducing (5) to a problem exclusively set in flux space $\boldsymbol{\mathcal{V}}$. This is achieved in [14] by a sort of preprocessing which converts (5) into a problem posed over the space of divergence free fluxes. For this an optimal multilevel precoditioner is proposed based upon schemes for standard finite element spaces. Conversely, in [27] a penalty approach is used. It makes possible the elimination of the scalar unknown u. A multilevel splitting of the flux space is then investigated and it is shown that it gives rise to an additive Schwarz method of nearly optimal performance: For successive uniform refinement of the mesh a $O(\log h^{-1})$ bound for the condition number of the preconditioned system is established. The penalty method was also the starting point for the development of a hierarchical basis preconditioner in [9]. Hoppe and Wohlmuth later showed in [21] that it yields condition numbers behaving like $O((\log h^{-1})^2)$.

This paper aims to extend the method of Vassilevski and Wang to an adaptive setting where local refinement is admitted. An improved condition number estimate is proved. In addition we point out that the preconditioner from [27] is well suited to saddle point problems arising from augmented Lagrangian methods.

The paper is organized in the following way: The next section gives a brief survey of the construction and basic properties of lowest order Raviart–Thomas elements (cf. [8, 25]). Then we derive the augmented Lagrangian system and discuss some of its properties. We point out why preconditioning is indispensable and which variational problem has to be targeted. In section 4 we present the multilevel splitting of Vassilevski and Wang and explore it in the case of local refinement. Section 5 contains the proof that the smallest eigenvalue of the preconditioned problem is actually bounded from below independently on the number of refinement levels and the size of the augmented Lagrangian parameter. We do so by confirming the assumptions of Lion's lemma (cf. [3], Lemma 1). The following section is devoted to establishing a meshsize independent bound for the largest eigenvalue. The proof makes use of a strengthened Cauchy–Schwarz inequality for the splitting under consideration.

The details of the additive preconditioner are the subject of the following section. There also an improvement of the generic algorithm is suggested. The final section illustrates the behavior of the preconditioner when applied to several model problems.

2 Finite element spaces. We chose finite element spaces that are based upon a simplicial triangulation $\mathcal{T}_h$ of a polygonally bounded , simply connected Ω. Moreover we assume the elements to be shape regular, that is, for all triangles the smallest angle is uniformly bounded from below. In particular our discretization of (4) relies on lowest order Raviart–Thomas spaces, first presented in [25]: For each element $T \in \mathcal{T}_h$ the approximation of vector fields in $\boldsymbol{H}(\text{div}; \Omega, \Gamma_N)$ is given by

$$\boldsymbol{\mathcal{V}}_h(T) := \{\mathbf{x} \rightarrow \mathbf{a} + \beta\mathbf{x}; \mathbf{a} \in \mathbb{R}^2, \beta \in \mathbb{R}\}$$

and $L^2(\Omega)$ is replaced by the space $\mathcal{W}_h$ of locally constant functions. To ensure conformity in $\boldsymbol{H}(\text{div}; \Omega)$ the continuity of normal components of discrete vector fields across interelement boundaries has to be enforced. This is accomplished by choosing the following degrees of freedom:

$$\boldsymbol{\mathcal{V}}_h(T) \rightarrow \mathbb{R} \; ; \; \mathbf{v}_h \rightarrow \int_e \langle \mathbf{v}_h, \boldsymbol{\nu} \rangle \, d\Gamma \; ; \; e \in \partial T \text{ is edge of } T \quad . \tag{6}$$

Observing that $\langle \mathbf{v}_h, \boldsymbol{\nu} \rangle_{|e} = \text{const.}$ and $\dim \boldsymbol{\mathcal{V}}_h(T) = 3$, we see that these degrees of freedom uniquely determine a global finite element function and guarantee conformity.

Since $\operatorname{div} \boldsymbol{\mathcal{V}}_h = \mathcal{W}_h$ the essential Babŭska–Brezzi stability condition (see [5])

$$\sup_{\mathbf{v}_h \in \boldsymbol{\mathcal{V}}_h} \frac{b(\mathbf{v}_h, w_h)}{\|\mathbf{v}_h\|_{\boldsymbol{H}(\text{div};\Omega)}} \geq \beta \, \|w_h\|_{L^2(\Omega)} \quad \forall w_h \in \mathcal{W}_h \tag{7}$$

holds. Next we note that $\boldsymbol{\mathcal{V}}_h^0 := \{\mathbf{v}_h \in \boldsymbol{\mathcal{V}}_h; \operatorname{div} \mathbf{v}_h = 0\}$ coincides with the space of locally constant vector fields. Owing to theorem 3.1 from [17] applied to an element T, for any $\mathbf{v}_h^0 \in \boldsymbol{\mathcal{V}}_h^0(T)$ there exists a $\varphi_h^T \in H(\mathbf{curl}; T)$, such that $\mathbf{v}_h^0 = \mathbf{curl}\,\varphi$ in T. As $\mathbf{v}_{h|T}^0$ is constant, φ_h^T must be linear. A closer scrutiny reveals, that those linear functions φ_h^T can be patched together to yield a continuous function. In sum, any vector field in $\boldsymbol{\mathcal{V}}_h^0$ possesses a representation as $\mathbf{curl}\,\varphi_h$ with φ_h from the space $\mathcal{S}_h$ of piecewise linear, continuous finite element functions over $\mathcal{T}_h$. Besides, zero outflow conditions on Γ_N are matched by vanishing traces of functions in $\mathcal{S}_h$ on Γ_N.

Plugging the spaces $\boldsymbol{\mathcal{V}}_h, \mathcal{W}_h$ into the variational problem (4), we obtain an indefinite linear system with a block structure exactly like (5). Unless ambiguous, we will keep the notation (5) for the discrete problem as well.

In later sections we are going to make extensive use of the following lemma, stating the stability of the canonical basis of $\boldsymbol{\mathcal{V}}_h$.

LEMMA 2.1. *For an arbitrary element T let $\{\boldsymbol{\iota}^1, \boldsymbol{\iota}^2, \boldsymbol{\iota}^3\}$ be the set of canonical basis functions of $\boldsymbol{\mathcal{V}}_h(T)$ belonging to the degrees of freedom (6). Then for all $\alpha_i \in \mathbb{R}$, $1 \leq i \leq 3$*

$$\frac{1}{3C_1} \sum_{i=1}^{3} \alpha_i^2 \leq \quad \left\| \sum_{i=1}^{3} \alpha_i \boldsymbol{\iota}^i \right\|_{\boldsymbol{L}^2(T)}^2 \leq C_1 \sum_{i=1}^{3} \alpha_i^2 \tag{8}$$

where

$$C_1 := \frac{\sin\gamma_{\max} + \cos\gamma_{\min}\sin\gamma_{\min}}{\sin^2\gamma_{\min}}$$

with $\gamma_{\min}, \gamma_{\max}$ *the smallest and largest angle of* T*, respectively.*

Proof. The proof largely relies on the Piola transform (see [8]) mapping vector fields over the reference triangle $\hat{T} := \text{convex}\left\{\binom{0}{0}, \binom{1}{0}, \binom{0}{1}\right\}$ to those in $\boldsymbol{\mathcal{V}}_h(T)$. Let

$$\Phi : \hat{T} \mapsto T\ ,\ \Phi(\hat{\mathbf{x}}) := \mathrm{Q}\hat{\mathbf{x}} + \mathbf{t}$$

be the affine mapping linking $\hat{T}$ and T. Since the Piola transform preserves the values of degrees of freedom of RT_0 we have

$$\int_T \left|\sum_{i=1}^{3} \alpha_i \boldsymbol{\iota}^i\right|^2 d\mathbf{x} = \int_{\hat{T}} \left\langle \frac{\mathrm{Q}^T\mathrm{Q}}{|\det \mathrm{Q}|} \sum_{i=1}^{3} \alpha_i \hat{\boldsymbol{\iota}}^i, \sum_{i=1}^{3} \alpha_i \hat{\boldsymbol{\iota}}^i \right\rangle d\hat{\mathbf{x}} \quad .$$

where $\{\hat{\boldsymbol{\iota}}^1, \hat{\boldsymbol{\iota}}^2, \hat{\boldsymbol{\iota}}^3\}$ is the set of canonical basis functions on the reference element. Q may be chosen as $\mathrm{Q} = \left[\vec{\mathbf{d}}_2, \vec{\mathbf{d}}_0\right]$, where $\vec{\mathbf{d}}_0$ is the vector of the shortest edge of T, $\vec{\mathbf{d}}_2$ that of the longest. Without loss of generality we may assume $|\vec{\mathbf{d}}_2| = 1$.

$$\mathrm{Q}^T\mathrm{Q} = \begin{pmatrix} 1 & |\vec{\mathbf{d}}_0|\cos\gamma_1 \\ |\vec{\mathbf{d}}_0|\cos\gamma_1 & |\vec{\mathbf{d}}_0|^2 \end{pmatrix} \quad ; \quad \left(\mathrm{Q}^T\mathrm{Q}\right)^{-1} = \frac{1}{|\vec{\mathbf{d}}_0|^2\sin^2\gamma_1} \begin{pmatrix} |\vec{\mathbf{d}}_0|^2 & -|\vec{\mathbf{d}}_0|\cos\gamma_1 \\ -|\vec{\mathbf{d}}_0|\cos\gamma_1 & 1 \end{pmatrix}$$

From Gershgorin's theorem applied to both matrices we have

$$\begin{aligned} \lambda_{\max}(\mathrm{Q}^T\mathrm{Q}) &< 1 + |\vec{\mathbf{d}}_0|\cos\gamma_1 \\ \lambda_{\min}(\mathrm{Q}^T\mathrm{Q}) &> \frac{|\vec{\mathbf{d}}_0|^2\sin^2\gamma_1}{1 + |\vec{\mathbf{d}}_0|\cos\gamma_1} \end{aligned}$$

A simple computation shows $|\det \mathrm{Q}| = |\vec{\mathbf{d}}_0|\sin\gamma_1$. This immediately yields

$$\begin{aligned} \frac{\lambda_{\max}}{|\det \mathrm{Q}|} &\le \frac{1 + |\vec{\mathbf{d}}_0|\cos\gamma_1}{|\vec{\mathbf{d}}_0|\sin\gamma_1} \le \frac{1 + |\vec{\mathbf{d}}_0|\cos\gamma_{\min}}{|\vec{\mathbf{d}}_0|\sin\gamma_{\min}} \\ \frac{\lambda_{\min}}{|\det \mathrm{Q}|} &\ge \frac{|\vec{\mathbf{d}}_0|\sin\gamma_1}{1 + |\vec{\mathbf{d}}_0|\cos\gamma_1} \ge \frac{|\vec{\mathbf{d}}_0|\sin\gamma_{\min}}{1 + |\vec{\mathbf{d}}_0|\cos\gamma_{\min}} \end{aligned}$$

Elementary trigonometry tells $|\vec{\mathbf{d}}_0| = \sin\gamma_{min}/\sin\gamma_{\max}$ and after a few computations we arrive at

$$\frac{1}{C_1}\lambda_{\max}(\mathrm{Q}^T\mathrm{Q}) \le |\det \mathrm{Q}| \le C_1\lambda_{\min}(\mathrm{Q}^T\mathrm{Q})\ .$$

To finish the proof we rely on the formulae

$$\hat{\boldsymbol{\iota}}^1 = \binom{\hat{x}_1}{\hat{x}_2 - 1} \quad , \quad \hat{\boldsymbol{\iota}}^2 = \binom{\hat{x}_1}{\hat{x}_2} \quad , \quad \hat{\boldsymbol{\iota}}^3 = \binom{\hat{x}_1 - 1}{\hat{x}_2}\ .$$

Using them we get for $\mathbf{j} = \alpha_1 \hat{\boldsymbol{\iota}}^1 + \alpha_2 \hat{\boldsymbol{\iota}}^2 + \alpha_2 \hat{\boldsymbol{\iota}}^2$

$$\|\mathbf{j}\|_{\boldsymbol{L}^2(\hat{T})} = \frac{1}{3}\left(\left|\mathbf{j}\begin{pmatrix}\frac{1}{2}\\0\end{pmatrix}\right|^2 + \left|\mathbf{j}\begin{pmatrix}\frac{1}{2}\\\frac{1}{2}\end{pmatrix}\right|^2 + \left|\mathbf{j}\begin{pmatrix}0\\\frac{1}{2}\end{pmatrix}\right|^2\right) = \frac{1}{3}\left(\alpha_1^2+\alpha_2^2+\alpha_3^2+(\alpha_1-\alpha_3)^2\right)$$

From this we can conclude

$$\frac{1}{3}\left(\alpha_1^2+\alpha_2^2+\alpha_3^2\right) \leq \|\mathbf{j}\|_{\boldsymbol{L}^2(\hat{T})} \leq \alpha_1^2+\alpha_2^2+\alpha_3^2; .$$

Combining the estimates completes the proof. □

Simply summing over all elements gives a global version of the previous lemma:

$$\frac{1}{3C_1}\sum_e \alpha_e^2 \leq \quad \left\|\sum_{i=1}^3 \alpha_i \boldsymbol{\iota}^i\right\|^2_{\boldsymbol{L}^2(\Omega)} \leq 2C_1 \sum_e \alpha_e^2 \,, \tag{9}$$

where the sums cover all edges of the mesh.

The following lemma informs about the norm of the discrete divergence operator, underscoring the necessity of preconditioning.

LEMMA 2.2. *Let $B : \boldsymbol{\mathcal{V}}_h \mapsto \mathcal{W}_h'$ be the operator related to the bilinear from $b(\cdot,\cdot)$ from (4) restricted to the lowest order Raviart–Thomas spaces $\boldsymbol{\mathcal{V}}_h, \mathcal{W}_h$ defined over the triangulation $\mathcal{T}_h$. Then we have*

$$\|B\|_{L^2(\Omega)\mapsto L^2(\Omega)} > C\,h^{-1}\,, \;\; \textit{where } h := \min_{T\in\mathcal{T}_h} \operatorname{diam} T \quad .$$

with $C > 0$ not depending on the meshwidth.

Proof. Pick an element $\tilde{T}$ of $\mathcal{T}_h$ with minimal diameter $\tilde{h}$. Let $\tilde{w}_h$ be the canonical basis function of $\mathcal{W}_h$ with support in $\tilde{T}$ and $\tilde{\mathbf{v}}_h$ be a canonical basis function of $\boldsymbol{\mathcal{V}}_h$ attached to an edge of $\tilde{T}$. The previous lemma gives

$$\|\tilde{\mathbf{v}}_h\|_{L^2(\Omega)} = O(1)$$

and from definition of $\tilde{w}_h$ we immediately deduce

$$\|\tilde{w}_h\|_{L^2(\Omega)} = \sqrt{|\tilde{T}|} = O(\tilde{h})$$

since $\mathcal{T}$ is assumed to be shape regular. Using the definition of the bilinear form $b(\cdot,\cdot)$ a straightforward computation finishes the proof:

$$\begin{aligned}\|B\|_{L^2(\Omega)\mapsto L^2(\Omega)} &\geq \frac{b(\tilde{\mathbf{v}}_h,\tilde{w}_h)}{\|\tilde{\mathbf{v}}_h\|_{L^2(\Omega)}\,\|\tilde{w}_h\|_{L^2(\Omega)}}\\ &\geq Ch^{-1}\cdot 1\end{aligned}$$

where $C > 0$ does not depend on the size of the elements, but only on the smallest angle. □

3 Augmented Lagrangian algorithm. The augmented Lagrangian approach seeks to exploit the advantages of the penalty method and to steer clear of its drawbacks at the same time. To this end a two track strategy is pursued in the treatment of the linear constraint in (2): Both a Lagrangian multiplier and an additional penalty term are introduced. The real parameter $r > 0$ governing the strength of penalization is called the *augmented Lagrangian parameter.*

On the algebraic level the equations of the augmented Lagrangian approach are easily derived: It boils down to simply multiplying the second equation of (5) by $rB^*\iota^{-1}$ and adding the result to the first. There ι^{-1} denotes the inverse of the Riesz isomorphism $\iota : \boldsymbol{\mathcal{V}} \mapsto \boldsymbol{\mathcal{V}}'$ in the Hilbert space $\boldsymbol{\mathcal{V}}$. This yields the equivalent system

$$\begin{array}{rcccl} (A + rB^*\iota^{-1}B)\,\mathbf{j} & + & B^*u & = & g + rB^*\iota^{-1}f \\ B\,\mathbf{j} & & & = & f \end{array} \tag{10}$$

Unlike the classical penalty method this scheme does not affect the solution; the exact solution of (5) can be recovered from (10), no matter what the value of r might be. Furthermore for moderate values of r the condition of the overall linear operator of (10) does not deteriorate severely. Thus the only consideration guiding the choice of r is the objective to make (10) as amenable as possible to an iterative solver.

The following two theorems help assess the impact of a particular choice of r on important properties of (10):

THEOREM 3.1. *Let $A_r := A + rB^*\iota^{-1}B$. Then A_r is invertible $\forall r \geq 0$ and its spectral condition number $\kappa(A_r)$ satisfies*

$$\kappa(A_r) \simeq r\frac{1}{\sigma}\|B\|_{L^2(\Omega)\mapsto L^2(\Omega)}$$

for large $r > 0$, where σ denotes the smallest eigenvalue of A.

Proof. For the proof we refer to [16]. □

In light of lemma 2.2 the above theorem means that for $r \to \infty$

$$\kappa(A_r) \sim r \cdot h^{-1}$$

because the operator A is a kind of scaled identity mapping and hence its spectrum is neatly bounded, no matter what the meshwidth may be.

THEOREM 3.2. *Let $\lambda_{\max}$ and $\lambda_{\min}$ denote the largest and smallest eigenvalue of the symmetric positive definite operator $\iota^{-1}BA^{-1}B^*$. Then the spectral condition number $\kappa(\iota^{-1}BA_r^{-1}B^*)$ of the operator $\iota^{-1}BA_r^{-1}B^* : \mathcal{W} \mapsto \mathcal{W}'$ is given by*

$$\kappa(\iota^{-1}BA_r^{-1}B^*) = \frac{\lambda_{\max}}{\lambda_{\min}} \cdot \frac{1 + r\lambda_{\min}}{1 + r\lambda_{\max}}$$

Proof. For arbitrary operators $T_1 : \mathcal{X} \mapsto \mathcal{Y}$, $T_2 : \mathcal{Y} \mapsto \mathcal{X}$ on Banach spaces $\mathcal{X}, \mathcal{Y}$ holds

$$\sigma\left(T_1T_2\right)/\{0\} = \sigma\left(T_2T_1\right)/\{0\}\ . \tag{11}$$

Consequently

$$
\begin{aligned}
\sigma\left(\iota^{-1} B A_r^{-1} B^*\right) &= \sigma\left(\underbrace{\iota^{-1} B A^{-\frac{1}{2}}\left(Id + rA^{-\frac{1}{2}} B^* \iota^{-1} B A^{-\frac{1}{2}}\right)^{-1}}_{=:T_1} \underbrace{A^{-\frac{1}{2}} B^*}_{=:T_2}\right) \\
&= \sigma\left(A^{-\frac{1}{2}} B^* \iota^{-1} B A^{-\frac{1}{2}}\left(Id + rA^{-\frac{1}{2}} B^* \iota^{-1} B A^{-\frac{1}{2}}\right)^{-1}\right) / \{0\} \\
&= \left\{\frac{\lambda}{1 + r\cdot\lambda}\,; \lambda \in \sigma\left(A^{-\frac{1}{2}} B^* \iota^{-1} B A^{-\frac{1}{2}}\right)\right\} / \{0\} \qquad .
\end{aligned}
$$

Again we can refer to (11) for the next step of the proof:

$$
\left(A^{-\frac{1}{2}} B^* \iota^{-1} B A^{-\frac{1}{2}}\right) / \{0\} = \sigma\left(\iota^{-1} B A^{-1} B^*\right) / \{0\}
$$

We point out that the operator $\iota^{-1} B A_r^{-1} B^*$ is positive definite. Since $\lambda \mapsto \frac{\lambda}{1+r\cdot\lambda}$ is nondecreasing for $\lambda \geq 0$, we are finished. □

The principal message of theorem (3.2) is that $\kappa(\iota^{-1} B A_r^{-1} B^*) \to 1$ as $r \to \infty$. This accounts for the appeal of the augmented Lagrangian approach. For iterative, Uzawa like, solvers for saddle point problems (see e.g. [2]) $\kappa(B A_r^{-1} B^*)$ determines the speed of the outer iteration. It will converge significantly faster for large values of r. The overall performance does not necessarily improve; as explained above, the condition of A_r becomes worse for larger r and finer meshes, crippling the inner iteration. In other words, the inner iteration significantly slows down when meshes are refined or r is increased.

For these reasons, the augmented Lagrangian scheme does not pay off, unless an effective preconditioner of A_r is available. Among others, it must meet two requirements

1. The condition number of CA_r must not increase as $r \to \infty$.
2. The condition number of CA_r must remain bounded independently of the number of refinement levels involved.

4 Multilevel splitting. The bilinear form $a_r(\cdot,\cdot) : \boldsymbol{\mathcal{V}}_h \times \boldsymbol{\mathcal{V}}_h \mapsto \mathbb{R}$ related to the operator A_r of the augmented Lagrangian approach has the form

$$
a_r(\mathbf{j}_h, \mathbf{v}_h) := \int_\Omega \langle \mathbf{j}_h, \mathbf{v}_h \rangle \, d\mathbf{x} + r \int_\Omega \operatorname{div} \mathbf{j}_h \cdot \operatorname{div} \mathbf{v}_h \, d\mathbf{x} \tag{12}
$$

Obviously, for $r > 0$, $a_r(\cdot,\cdot)$ is symmetric, continuous and $\boldsymbol{H}(\operatorname{div};\Omega)$–coercive. In this sense, it bears strong resemblance to bilinear forms occurring in the standard variational formulation of elliptic problems. This observation hints that it also shares some of their behavior, especially the decoupling of frequencies and the pronounced amplification of all high frequencies. As this property is the ultimate foundation of multilevel preconditioning, this approach might seem promising in this case as well.

Multilevel schemes owe their name to the ladders of approximation spaces of ever increasing resolution employed. As it is customarily done, we construct them by means of a sequence of nested triangulations $\{\mathcal{T}_l\}_l$, obtained by successive refinement of a coarse initial triangulation $\mathcal{T}_0$. To generate that stack of meshes the popular red–green refinement scheme is used (cf. [1, 3, 12]). It creates the next finer mesh by dividing some elements of the current triangulation. The algorithm maintains compatibility of the meshes and preserves shape regularity.

Since we permit refinement confined to only a part of the domain, the construction of the finite element spaces associated with each level is somewhat involved. Let us fix some terminology first: Each element T is assigned a unique level $l(T)$. It is zero for all elements of the initial triangulation. If an element T emerged from the division of another, its father T', we set $l(T) := l(T') + 1$. The meshwidth h_l on level l is defined as the longest diameter of all elements T with $l(T) = l$. This make sense, since shape regularity guarantees that all those elements are of almost the same size, provided this held true for the initial triangulation. Moreover the rules for refinement force the meshwidth to decrease almost geometrically; we have $h_l \approx \frac{1}{2} h_{l-1}\,, l \geq 1$ (The $\approx$ is to blame on green refinements). By normalization we can set $h_0 := 1$.

Further denote by $\mathsf{N}_{[l]}$ the set of all vertices that belong to some triangle T with $l(T) = l$. We define the triangulation $\mathcal{T}_l$ to consist of all those elements whose children belong to level $l+1$. The related Raviart–Thomas finite element spaces are denoted by $\boldsymbol{\mathcal{V}}_l$ and $\mathcal{W}_l$, respectively.

Finally, meticulous examination reveals that the refinement scheme also makes sure a "gradual" transition at the edge of strongly refined zones. In particular for any vertex $\mathbf{p}$ we have

$$\delta_{\mathrm{ref}} := \max\{l(T) - l(T'); T, T' \in \mathcal{T}_l\,,\, \mathbf{p} \in T \wedge \mathbf{p} \in T'\} \leq \left[\left[\frac{2\pi}{\gamma_0} - 2\right]/2\right] \tag{13}$$

with γ_0 the smallest angle of elements of the initial triangulation. For uniform refinement obviously $\delta_{\mathrm{ref}} = 0$.

Let $\{\mathcal{T}_l\}_{l=0}^L$ be a stack of nested triangulations, created as explained above. The main idea behind the multilevel splitting of $\boldsymbol{\mathcal{V}}_h := \boldsymbol{\mathcal{V}}_L$ is to treat the divergence free part of vectorfields and the remainder separately. Basically the decomposition looks like

$$\boldsymbol{\mathcal{V}}_h := \boldsymbol{\mathcal{V}}_0 + \sum_{l=1}^{L} \boldsymbol{\mathcal{H}}_l + \boldsymbol{\mathcal{D}}_h \tag{14}$$

with $\operatorname{div} \boldsymbol{\mathcal{D}}_h = \{0\}$ and

$$\boldsymbol{\mathcal{H}}_l := \{\mathbf{v}_l \in \boldsymbol{\mathcal{V}}_l; \langle \mathbf{v}_l, \boldsymbol{\nu}\rangle_{|\partial T} = 0, \forall T \in \mathcal{T}_{l-1}\} \quad (l \geq 1) \tag{15}$$

In a second step a certain decomposition of $\boldsymbol{\mathcal{D}}_h$ has to be fixed. At this stage the fact that $\boldsymbol{\mathcal{D}}_h = \mathbf{curl}\, \mathcal{S}_h$ can be exploited: Decompositions of $\mathcal{S}_h$ immediately give rise to such of $\boldsymbol{\mathcal{D}}_h$.

Yet, we have to investigate the feasibility of the splitting (14) first. Following [27], let us start with an arbitrary $\mathbf{v}_h \in \boldsymbol{\mathcal{V}}_h$. Denoting by Q_l the $L^2(\Omega)$–orthogonal projection onto the space $\mathcal{W}_l$ and setting $Q_{-1} = 0$ we can represent $\operatorname{div} \mathbf{v}_h$ in the following way:

$$\operatorname{div} \mathbf{v}_h := \sum_{l=0}^{L} (Q_l - Q_{l-1}) \operatorname{div} \mathbf{v}_h \tag{16}$$

For simplicity we write $z_l := (Q_l - Q_{l-1})\,\mathrm{div}\,\mathbf{v}_h$. Recall that, by definition, functions from $\mathcal{W}_{l-1}$ are constant on each element of $\mathcal{T}_{l-1}$. This fact, along with the properties of the L^2–projection, ensures that

$$\int_T z_l \, d\mathbf{x} = 0 \; ; \; \forall T \in \mathcal{T}_{l-1} \; , \; l \geq 1 \; . \tag{17}$$

(17) means a consistency condition, which allows to determine a $\mathbf{h}_l \in \boldsymbol{\mathcal{H}}_l$ such that $\mathrm{div}\,\mathbf{h}_l = z_l$. We give an explicit construction: Pick an arbitrary element T of $\mathcal{T}_{l-1}$ that has been subdivided. On this occasion new edges are created in the interior of T, whose associated nodal flux basis functions are contained in $\boldsymbol{\mathcal{H}}_l$. Depending on the kind of subdivision the construction of $\mathbf{q}_{l|T}$ is carried out as follows:

1. In the case of *red refinement* (see figure 1 for notations) choose

$$\mathbf{q}_l(\mathbf{x}) = \sum_{i=1}^{3} \int_{T_i} z_l \, d\mathbf{x} \cdot \boldsymbol{\iota}^i(\mathbf{x}) \; , \; \mathbf{x} \in T \; . \tag{18}$$

   As the mean value of the divergence vanishes for every local basis function not located on the boundary, (17) bears out the validity of the construction. Hence (18) gives the desired vector field.

2. If a triangle has suffered *green refinement* (as sketched in figure 2) we just set

$$\mathbf{q}_l(\mathbf{x}) = \int_{T_1} z_l \, d\mathbf{x} \cdot \boldsymbol{\iota}^*(\mathbf{x}) \; , \; \mathbf{x} \in T \; . \tag{19}$$

   As above, we see that this is a proper choice.

A $\mathbf{q}_0 \in \boldsymbol{\mathcal{V}}_0$ with $\mathrm{div}\,\mathbf{q}_0 = z_0$ cannot be determined locally; on the coarsest grid the following global variabtional problem has to be solved to get $\mathbf{q}_0$: Seek $(\mathbf{q}_0, p) \in \boldsymbol{\mathcal{V}}_0 \times \mathcal{W}_0$ such that

$$\begin{array}{rcll} (\mathbf{q}_0, \mathbf{v}_h)_{L^2(\Omega)} + b(\mathbf{v}_h, p) & = & g(\mathbf{v}) & , \forall \mathbf{v}_h \in \boldsymbol{\mathcal{V}}_0 \\ b(\mathbf{q}_0, w_h) & = & \int_\Omega z_0 w_h \, d\mathbf{x} & , \forall w_h \in \mathcal{W}_0 \end{array} \tag{20}$$

One may imagine vectorfields of $\boldsymbol{\mathcal{H}}_l$ $(l \geq 1)$ to be composed of several isolated "bumps". Clearly they are rapidly oscillating. Loosely speaking, this makes it possible to bound their $L^2(\Omega)$–norm by means of their divergence. This statement is made more precise in the following lemma:

LEMMA 4.1. *For* $\mathbf{q}_l \in \boldsymbol{\mathcal{H}}_l$, $(l \geq 1)$ *we have*

$$\|\mathbf{q}_l\|_{\boldsymbol{L}^2(\Omega)} \leq C_2 h_l \, \|\mathrm{div}\,\mathbf{q}_l\|_{L^2(\Omega)} \tag{21}$$

with a constant $C > 0$ *depending only on the shape regularity of* $\mathcal{T}_0$.

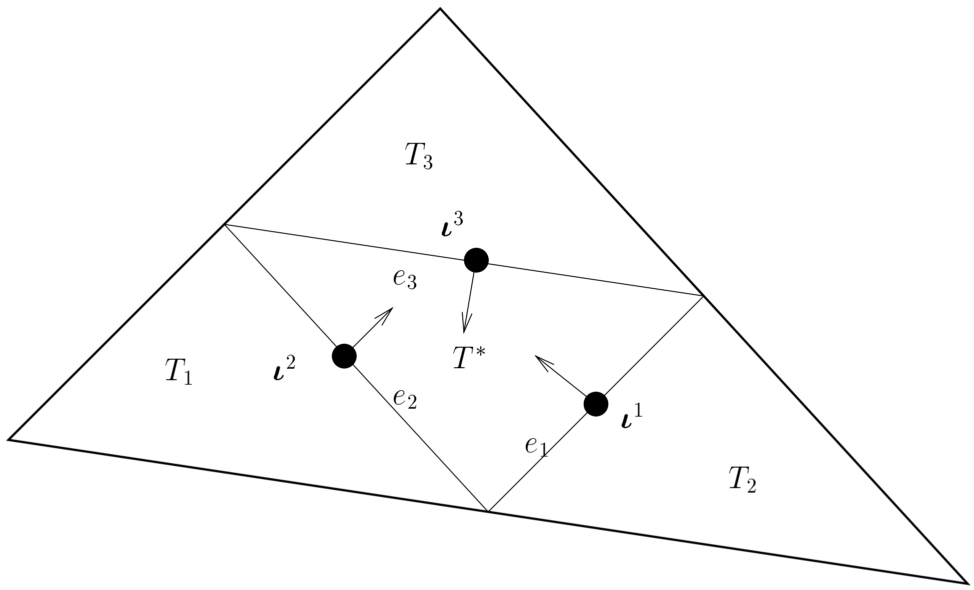

FIGURE 1: *Local calculation of* $\mathbf{q}_l$ *in the case of red refinement. The filled disks show the nodes for* RT_0 *fluxes, the attached arrows the orientation of basis fields.*

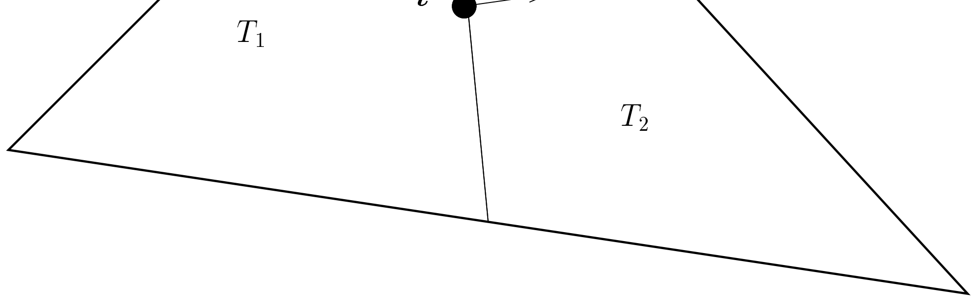

FIGURE 2: *Local calculation of* $\mathbf{q}_l$ *in the case of green refinement*

Proof. To begin with, note that $\mathbf{q}_{l|T} = 0$ in all T shared by $\mathcal{T}_l$ and $\mathcal{T}_{l-1}$. So we need only consider elements of $\mathcal{T}_{l-1}$ that have been subdivided. Let $\bar{T} \in \mathcal{T}_{l-1}$ be such a *macroelement* and assume that it has undergone red refinement as in figure 1. $\{e_1, e_2, e_3\}$ are the freshly created edges on level l in the interior of $\bar{T}$. Denote by $\{\boldsymbol{\iota}^1, \boldsymbol{\iota}^2, \boldsymbol{\iota}^3\}$ the canonical basis functions of $\boldsymbol{\mathcal{H}}_l$ attached to these edges. The divergence theorem readily yields

$$\operatorname{div} \mathbf{q}_{l|T_i} = \frac{1}{|T_i|} \int\limits_{e_i} \langle \mathbf{q}_l, \boldsymbol{\nu} \rangle \; d\Gamma \quad . \tag{22}$$

Thanks to lemma 2.1 we can set up a series of estimates:

$$\begin{aligned} \|\mathbf{q}_l\|^2_{\boldsymbol{L}^2(\bar{T})} &\leq 2C_1 \sum_{i=1}^{3} \left| \int\limits_{e_i} \langle \mathbf{q}_l, \boldsymbol{\nu} \rangle \; d\Gamma \right|^2 = 2C_1 \sum_{i=1}^{3} |T_i|^2 \left| \operatorname{div} \mathbf{q}_{l|T_i} \right|^2 = 2C_1 \sum_{i=1}^{3} |T_i| \, \|\operatorname{div} \mathbf{q}_l\|^2_{L^2(T_i)} \\ &\leq \frac{6}{4} C_1 |\bar{T}| \|\operatorname{div} \mathbf{q}_l\|^2_{L^2(\bar{T})} \leq 3C_1 h_l^2 \|\operatorname{div} \mathbf{q}_l\|^2_{L^2(\bar{T})} \end{aligned}$$

Analogously the case of green refinement can be treated (see figure 2):

$$\begin{aligned} \|\mathbf{q}_l\|^2_{\boldsymbol{L}^2(\bar{T})} &\leq 2C_1 \left| \int\limits_{e^*} \langle \mathbf{q}_l, \boldsymbol{\nu} \rangle \; d\Gamma \right|^2 = 2C_1 |T_1|^2 \left| \operatorname{div} \mathbf{q}_{l|T_1} \right|^2 = 2C_1 \, |T_1| \, \|\operatorname{div} \mathbf{q}_l\|^2_{L^2(T_1)} \\ &\leq \frac{2}{4} C_1 |\bar{T}| \|\operatorname{div} \mathbf{q}_l\|^2_{L^2(\bar{T})} \leq C_1 h_l^2 \|\operatorname{div} \mathbf{q}_l\|^2_{L^2(\bar{T})} \end{aligned}$$

Summing over all relevant elements gives the statement of the lemma with $C_2 = \sqrt{3C_1}$. □

A similar estimate is required for the component $\mathbf{q}_0$. Owing to the global nature of its definition (20) the global stability constant from (7) enters the bound:

LEMMA 4.2. *For $\mathbf{q}_0$ from (20) we can estimate*

$$\|\mathbf{q}_0\|_{L^2(\Omega)} \leq \max \left\{ \left(\frac{2}{\beta} - 1 \right), 1 \right\} \|\operatorname{div} \mathbf{q}_0\|_{L^2(\Omega)}$$

with β defined in (7).

Proof. $\mathbf{q}_0$ is defined via the saddle point problem (20). Standard stability estimates (cf. [8], Prop. 1.3) immediately yield the bound above. □

Next we have to take care of the divergence free part $\boldsymbol{\mathcal{D}}_h$. Taking into account the trivial identity

$$a_r(\mathbf{curl}\, \varphi, \mathbf{curl}\, \psi) = \int\limits_{\Omega} a^{-1} \langle \nabla \varphi, \nabla \psi \rangle \, d\mathbf{x} \qquad \forall \varphi, \psi \in \mathcal{S}_h \tag{23}$$

we arrive at a simple strategy: Assume we have a viable multilevel splitting of the space $\mathcal{S}_h$ of continuous, piecewise linear finite element functions with respect to the bilinear form on the right hand side of (23):

$$\mathcal{S}_h := \sum_k \mathcal{X}_k \tag{24}$$

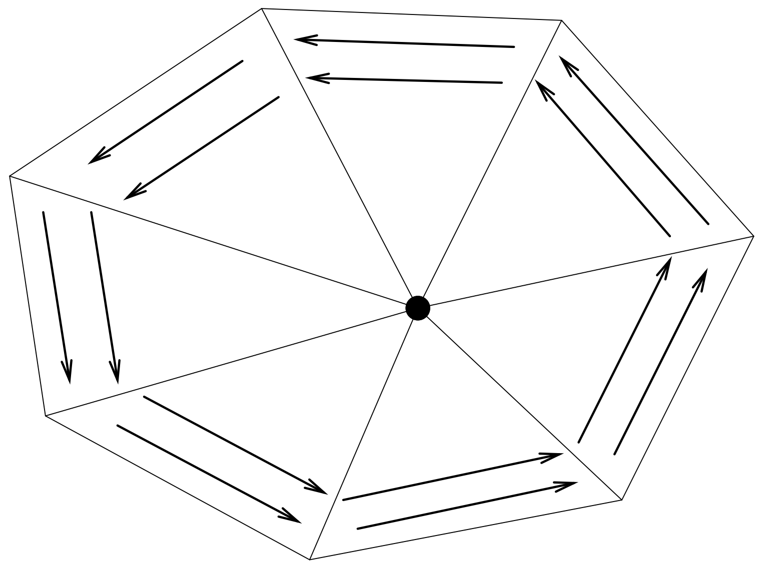

FIGURE 3: *Divergence free vector field obtained as the* **curl** *of a linear "hat function" located at the central vertex*

Then a splitting of $\boldsymbol{\mathcal{D}}_h$ of equal quality is given by

$$\boldsymbol{\mathcal{D}}_h = \sum_k \mathbf{curl}\, \mathcal{X}_k \ . \tag{25}$$

An excellent splitting of $\mathcal{S}_h$ has been discovered by Bramble et al. (cf. [4]). They chose the $\mathcal{X}_k$ as one–dimensional subspaces spanned by the nodal basis functions on all levels. Writing $\psi^l_{\mathbf{p}}$ for the linear "hat function" perched upon the vertex $\mathbf{p}$ on level l the decomposition reads

$$\mathcal{S}_h = \sum_{l=1}^{L} \sum_{\mathbf{p}\in \mathsf{N}_{[l]}} \mathrm{Span}\{\psi^l_{\mathbf{p}}\} + \mathcal{S}_0 \ , \tag{26}$$

where $\mathcal{S}_0$ is the space of C^0 linear finite element functions on the coarsest mesh. Following the rule (25) we readily have a suitable splitting of $\boldsymbol{\mathcal{D}}_h$ into one-dimensional subspaces. They are naturally associated with the vertices on each level. Figure 3 shows a basis vector field for one such subspace.

REMARK I. An alternative choice for the splitting of $\mathcal{S}_h$ is the hierarchical basis decomposition. Though inferior to (26) in terms of asymptotic behavior, it is still widely used in practice and achieves a satisfactory reduction of condition numbers for reasonable meshes. In analogy to (26) we can describe it by

$$\mathcal{S}_h = \sum_{l=1}^{L} \sum_{\mathbf{p}\in \mathsf{N}_{[l]}/\mathsf{N}_{[l-1]}} \mathrm{Span}\{\psi^l_{\mathbf{p}}\} + \mathcal{S}_0 \ . \tag{27}$$

In [21] the hierarchical basis preconditioner of Cai et al. from [9] was shown to be based on the very splitting (27). This finding offers an elegant way to establish the condition number estimates for the hierarchical basis method: given slight adjustments, the proofs in the following section cover them.

5 Stability. We intend to establish the stability of the splitting (14), where $\boldsymbol{\mathcal{V}}_h^0$ is further decomposed according to (25) and (26). The proof parallels the construction of the decomposition; we separately take care of the divergence free part of the flux space and the components in the spaces $\boldsymbol{\mathcal{H}}_l$ (see (14)). Fortunately we do not need to worry about the stability of (26), since it has already been dealt with in [11]. Using (23) the fundamental result can be put as

THEOREM 5.1. *(cf. [11]) With the notations and the terminology of section 4 let*

$$\boldsymbol{\mathcal{D}}_{\mathbf{p}}^l := \mathrm{Span}\left\{\mathbf{curl}\,\psi_{\mathbf{p}}^l\right\} \quad (l \geq 1, \mathbf{p} \in \mathsf{N}_{[l]}) \quad , \quad \boldsymbol{\mathcal{D}}_0 := \boldsymbol{\mathcal{V}}_0$$

Then we have

$$\boldsymbol{\mathcal{D}}_h := \sum_{l=1}^{L} \sum_{\mathbf{p}\in\mathsf{N}_{[l]}} \boldsymbol{\mathcal{D}}_{\mathbf{p}}^l + \boldsymbol{\mathcal{D}}_0 \tag{28}$$

and the splitting is stable *in the sense that there is a positive constant C_{BPX}, only depending on the shape regularity of the initial triangulation $\mathcal{T}_0$, such that*

$$\inf_{\substack{\mathbf{v}^0+\sum_{\mathbf{p},l}\mathbf{v}_{\mathbf{p}}^l=\mathbf{v}_0\\ \mathbf{v}_{\mathbf{p}}^l\in\boldsymbol{\mathcal{D}}_{\mathbf{p}}^l;\mathbf{v}^0\in\boldsymbol{\mathcal{D}}_0}} \sum_{l=0}^{L} \sum_{\mathbf{p}\in\mathsf{N}_{[l]}} \left\|\mathbf{v}_{\mathbf{p}}^l\right\|_{A_r}^2 \leq C_{\mathrm{BPX}} \left\|\mathbf{v}_0\right\|_{A_r}^2$$

for all $\mathbf{v}_0 \in \boldsymbol{\mathcal{V}}_h^0$ with $\|\cdot\|_{A_r}$ being the norm induced by the symmetric positive definite bilinear form $a_r(\cdot,\cdot)$ from (12).

Proof. A detailed proof for $\mathcal{S}_h$, making use of the theory of Besov spaces, can be found in [11]. The result for uniform refinement is due to Oswald [24]. Zhang in [31] deduced the same result by more elementary means. □

More effort has to be spent on harnessing the remaining components of the decomposition. It turns out that the $L^2(\Omega)$–orthogonality of the divergence of the individual components $\mathbf{h}_l$ from (18), (19) plays an important role. It paves the way for bounding the energy norm of the non–divergence free part of a vector field in $\boldsymbol{\mathcal{V}}_h$. This is explicitly done in the following lemma, restating a result from [27].

LEMMA 5.2. *([27], Lemma 3.3) There exists a constant $C_4 > 0$, not depending on $\mathbf{v}_h \in \boldsymbol{\mathcal{V}}_h$ or the number of refinement levels, such that*

$$\left\|\mathbf{q}_h\right\|_{A_r}^2 \leq \left(\frac{C_4}{\underline{\alpha}} + r\right) \left\|\mathrm{div}\,\mathbf{v}\right\|_{L^2(\Omega)}^2$$

for any $\mathbf{v} \in \boldsymbol{\mathcal{V}}_h$. There we set

$$\mathbf{q}_h := \sum_{l=0}^{L} \mathbf{q}_l$$

with the vector fields $\mathbf{q}_l \in \boldsymbol{\mathcal{H}}_l$ *being defined according to (18), (19).*

Proof. Our task essentially boils down to bounding the $L^2(\Omega)$–norm of $\mathbf{q}$ by means of its divergence. To begin with, we employ the triangle inequality and lemmas 4.1, 4.2:

$$\|\mathbf{q}_h\|_{\boldsymbol{L}^2(\Omega)} \leq \sum_{l=0}^{L} \|\mathbf{q}_l\|_{\boldsymbol{L}^2(\Omega)} \leq \underbrace{\max\left\{C_2, \frac{2}{\beta} - 1, 1\right\}}_{=:C_3} \sum_{l=0}^{L} h_l \, \|\operatorname{div} \mathbf{q}_l\|_{L^2(\Omega)}$$

and continue by means of the Cauchy–Schwarz inequality.

$$\|\mathbf{q}_h\|_{\boldsymbol{L}^2(\Omega)} \leq C_3 \left(\sum_{l=0}^{L} h_l^2\right)^{\frac{1}{2}} \left(\sum_{l=0}^{L} \|\operatorname{div} \mathbf{q}_l\|^2_{L^2(\Omega)}\right)^{\frac{1}{2}}$$

We exploit the geometric decrease of mesh width

$$\|\mathbf{q}_h\|_{\boldsymbol{L}^2(\Omega)} \leq C_3 \frac{2}{\sqrt{3}} h_0 \left(\sum_{l=0}^{L} \|\operatorname{div} \mathbf{q}_l\|^2_{L^2(\Omega)}\right)^{\frac{1}{2}}$$

and the $L^2(\Omega)$ orthogonality of the divergences.

$$\|\mathbf{q}_h\|_{\boldsymbol{L}^2(\Omega)} \leq \underbrace{C_3 \frac{2}{\sqrt{3}} h_0}_{=:\sqrt{C_4}} \|\operatorname{div} \mathbf{q}_h\|_{L^2(\Omega)}$$

Recalling the definition of the energy norm we can carry on

$$\|\mathbf{q}_h\|^2_{A_r} = \left(a^{-1}\mathbf{q}_h, \mathbf{q}_h\right)_{\boldsymbol{L}^2(\Omega)} + r \, \|\operatorname{div} \mathbf{q}_h\|^2_{L^2(\Omega)} \leq \frac{1}{\underline{\alpha}} \|\mathbf{q}_h\|^2_{\boldsymbol{L}^2(\Omega)} + r \, \|\operatorname{div} \mathbf{q}_h\|^2_{L^2(\Omega)}$$

and using the above estimate the proof is complete. □

The essential statement is made in the following theorem

THEOREM 5.3. *For the decomposition*

$$\boldsymbol{\mathcal{V}}_h = \boldsymbol{\mathcal{V}}_0 + \sum_{l=1}^{L} \boldsymbol{\mathcal{H}}_l + \sum_{l=1}^{L} \sum_{\mathbf{p} \in \mathsf{N}_{[l]}} \operatorname{Span}\{\mathbf{curl}\, \psi^l_{\mathbf{p}}\} := \sum_k \boldsymbol{\mathcal{Z}}_k \tag{29}$$

of the RT_0 *flux space we have the stability estimate*

$$\inf_{\substack{\sum_k \mathbf{v}_k = \mathbf{v} \\ \mathbf{v}_k \in \boldsymbol{\mathcal{Z}}_k}} \sum_k \|\mathbf{v}_k\|^2_{A_r} \leq C_{\text{stab}}(r) \|\mathbf{v}\|^2_{A_r} \tag{30}$$

for all $\mathbf{v} \in \boldsymbol{\mathcal{V}}_h$. *The constant* $C_{\text{stab}}(r) > 0$ *does not deteriorate on finer meshes and remains bounded as* $r \to \infty$

Proof. We start with an arbitrary $\mathbf{v}_h \in \boldsymbol{\mathcal{V}}_h$ and follow the prescription from section 4 to obtain the decomposition

$$\mathbf{v}_h = \mathbf{q} + \mathbf{v}_h^0 \text{ with } \operatorname{div} \mathbf{v}_h^0 = 0 \text{ and } \mathbf{q} \text{ from lemma 5.2.} \tag{31}$$

According to theorem 5.1 we can find a further splitting

$$\mathbf{v}_h^0 = \sum_{l=0}^{L} \sum_{\mathbf{p} \in \mathsf{N}_{[l]}} \mathbf{v}_{\mathbf{p}}^l + \mathbf{v}_0^0 \text{ with } \mathbf{v}_{\mathbf{p}}^l \in \boldsymbol{\mathcal{D}}_{\mathbf{p}}^l, \ \mathbf{v}_0^0 \in \boldsymbol{\mathcal{D}}_{\mathbf{p}}^l \tag{32}$$

which satisfies

$$\sum_{l=1}^{L} \sum_{\mathbf{p} \in \mathsf{N}_{[l]}} \|\mathbf{v}_{\mathbf{p}}^l\|_{A_r}^2 + \|\mathbf{v}_0^0\|_{A_r}^2 \leq C_{\mathrm{BPX}} \|\mathbf{v}_h^0\|_{A_r}^2 \ . \tag{33}$$

Still we have to cope with the "wrong" bound in (33); to confirm the desired stability for the divergence free part, we have to bound $\|\mathbf{v}_h^0\|_{A_r}$ by $\|\mathbf{v}_h\|_{A_r}$. Thanks to lemma 5.2 this is quickly done, since $\mathbf{v}_h^0 = \mathbf{v}_h - \mathbf{q}_h$:

$$\begin{aligned} a_r(\mathbf{v}_h^0, \mathbf{v}_h^0) &\leq 2\left(a_r(\mathbf{v}_h, \mathbf{v}_h) + a_r(\mathbf{q}_h, \mathbf{q}_h)\right) \\ &\leq 2\left(\|\mathbf{v}_h\|_{A_r}^2 + \left(\frac{C_4}{\underline{\alpha}} + r\right)\|\operatorname{div} \mathbf{v}_h\|_{L^2(\Omega)}^2\right) \\ &\leq 2\left(2 + \frac{C_4}{\underline{\alpha} r}\right)\|\mathbf{v}_h\|_{A_r}^2 \end{aligned}$$

The next step focuses on the component in the spaces $\boldsymbol{\mathcal{H}}_l\,(0 \leq l \leq L)$. First we apply the inequalities of lemmata 4.1, 4.2:

$$\sum_{l=0}^{L} a_r(\mathbf{q}_l, \mathbf{q}_l) \leq \sum_{l=0}^{L}\left(\frac{1}{\underline{\alpha}}\|\mathbf{q}_l\|_{\mathbf{L}^2(\Omega)}^2 + r\,\|\operatorname{div} \mathbf{q}_l\|_{L^2(\Omega)}^2\right) \leq \sum_{l=0}^{L}\left(\frac{1}{\underline{\alpha}} C_3^2 h_l^2\,\|\operatorname{div} \mathbf{q}_l\|_{L^2(\Omega)}^2 + r\,\|\operatorname{div} \mathbf{q}_l\|_{L^2(\Omega)}^2\right)$$

and then make use of the $L^2(\Omega)$–orthogonality:

$$\sum_{l=0}^{L} a_r(\mathbf{q}_l, \mathbf{q}_l) \leq \left(\frac{1}{\underline{\alpha}} C_3^2 + r\right) \sum_{l=0}^{L} \|\operatorname{div} \mathbf{q}_l\|_{L^2(\Omega)}^2 = \left(\frac{C_3^2}{\underline{\alpha}} + r\right)\|\operatorname{div} \mathbf{v}_h\|_{L^2(\Omega)}^2 \leq \left(1 + \frac{C_3^2}{\underline{\alpha} r}\right)\|\mathbf{v}_h\|_{A_r}^2$$

Merging both estimates yields the proposition of the theorem with

$$C_{\mathrm{stab}}(r) := (4C_{\mathrm{BPX}} + 1) + \frac{1}{\underline{\alpha} r}\left(2C_4 C_{\mathrm{BPX}} + C_3^2\right) \ .$$

In the limit we have

$$\lim_{r \to \infty} C_{\mathrm{stab}}(r) = 4C_{\mathrm{BPX}} + 1 \ .$$

□

REMARK II. For the hierarchical basis decomposition (27) all the proofs of the current section carry over unchanged. Only the analogue of theorem 5.1 (see [30]) provides a poorer estimate with $O(L^2)$ dependency of the constant. Ultimately this involves a $O(L^2)$ growth of the stability constant of the hierarchical splitting.

6 Quasi–orthogonality. Beside issues of stability we need to explore the *approximate orthogonality* of the multilevel splitting (14),(25),(26). This is conveniently gauged by means of a strengthened Cauchy–Schwarz inequality (see [29]). Guided by the general pattern of the splitting (14) we tackle its basic components separately. First we establish a local precursor of the strengthened Cauchy–Schwarz inequality for components with nonzero divergence:

LEMMA 6.1. *(cf. lemma 4.4 from [28]) Let $\tilde{T}$ be an arbitrary element of $\mathcal{T}_m$ ($m \in \{0,\dots,L-1\}$) and $L \geq k > m$. Then we have for any $\mathbf{z}^m \in \boldsymbol{\mathcal{V}}_m$ and $\mathbf{q}^k \in \boldsymbol{\mathcal{H}}_k$*

$$a_{r|\tilde{T}}(\mathbf{q}^k, \mathbf{z}^m) \leq \frac{2C_2}{\underline{\alpha}} \left(\frac{1}{2}\right)^{k-m} \left|\tilde{T}\right|^{\frac{1}{2}} \left\|\operatorname{div} \mathbf{q}^k\right\|_{L^2(\tilde{T})} \left\|\mathbf{z}^m\right\|_{\boldsymbol{L}^2(\tilde{T})}$$

with C_2 from lemma 4.1

Proof. Initially we point out that $\operatorname{div} \mathbf{z}^m$ is constant on $\tilde{T}$. For $\mathbf{q}^k \in \boldsymbol{\mathcal{H}}_k$ we have

$$\int_{\tilde{T}} \operatorname{div} \mathbf{q}^k \cdot \operatorname{div} \mathbf{z}^m \, d\mathbf{x} = (\operatorname{div} \mathbf{z}^m)_{|\tilde{T}} \int_{\partial\tilde{T}} \left\langle \mathbf{q}^k, \boldsymbol{\nu} \right\rangle \, d\Gamma = 0$$

since the normal components of vector fields from $\boldsymbol{\mathcal{H}}_k$ vanish on the boundaries of elements on a coarser level. This implies

$$a_{r|\tilde{T}}(\mathbf{q}^k, \mathbf{z}^m) = \left(a^{-1}\mathbf{q}^k, \mathbf{z}^m\right)_{\boldsymbol{L}^2(\tilde{T})} \leq \frac{1}{\underline{\alpha}} \left\|\mathbf{q}^k\right\|_{\boldsymbol{L}^2(\tilde{T})} \cdot \left\|\mathbf{z}^m\right\|_{L^2(\tilde{T})} \ . \tag{34}$$

The mesh $\mathcal{T}_k$, restricted to $\tilde{T}$, forms a valid triangulation of $\tilde{T}$. Therefore we can apply lemma 4.1 to $\tilde{T}$. Observe that in this case the coarse grid meshwidth is given by h_m. Then lemma 4.1 furnishes the estimate

$$\left\|\mathbf{q}^k\right\|_{\boldsymbol{L}^2(\tilde{T})} \leq C_2 \left(\frac{h_k}{h_m}\right) \left|\tilde{T}\right|^{\frac{1}{2}} \left\|\operatorname{div} \mathbf{q}^k\right\|_{L^2(\tilde{T})} \ . \tag{35}$$

We know $h_k/h_m \leq 2\left(\frac{1}{2}\right)^{k-m}$ and so the proposition emerges by merely joining the estimates (34) and (35). □

Our treatment of the divergence free parts of (14) takes the cue from the approach in [3]. We divide $\mathsf{N}_{[l]}$ into four subsets $\mathsf{N}^1_{[l]}, \mathsf{N}^2_{[l]}, \mathsf{N}^3_{[l]}, \mathsf{N}^4_{[l]}$, such that vertices with a joint edge belong to different subsets. This can be achieved by a four coloring of the vertex graph of the mesh. Based on this arrangement we define (see also (28))

$$\boldsymbol{\mathcal{D}}^l_i := \sum_{\mathbf{p} \in \mathsf{N}^i_{[l]}} \boldsymbol{\mathcal{D}}^l_{\mathbf{p}} \ , \ i \in \{1,2,3,4\} \tag{36}$$

The supports of “hat” functions attached to distinct vertices in a set $\mathsf{N}^i_{[l]}$ are disjoint, rendering their **curl**s orthogonal with respect to the $a_r(\cdot,\cdot)$ inner product. Consequently the decomposition

$$\boldsymbol{\mathcal{D}}_h := \sum_{l=0}^{L} \sum_{i=1}^{4} \boldsymbol{\mathcal{D}}^l_i \tag{37}$$

fully preserves the orthogonality properties of (28). In addition, (37) is more easily investigated due to the smaller number of subspaces involved, enabling us to obtain a local strengthened Cauchy–Schwarz inequality.

LEMMA 6.2. *Pick an arbitrary triangle $\tilde{T} \in \mathcal{T}_m$ $(m \in \{0, \dots, L-1\})$ and a $L \geq k > m$. For any $\mathbf{d}^k \in \boldsymbol{\mathcal{D}}_i^k$, $(1 \leq i \leq 4)$ and $\mathbf{z}^m \in \boldsymbol{\mathcal{V}}_m$ holds*

$$a_{r|\tilde{T}}(\mathbf{d}^k, \mathbf{z}^m) \leq \frac{1}{\underline{\alpha}} \left\|\mathbf{d}^k\right\|_{\boldsymbol{L}^2(\tilde{T})} \left\|\mathbf{z}^m\right\|_{\boldsymbol{L}^2(\tilde{T})} \cdot \begin{cases} \sqrt{20 \cdot 2^{\delta_{\mathrm{ref}}} \left(\frac{1}{2}\right)^{k-m}} & \text{if } k - m - \delta_{\mathbf{ref}} > 2 \\ 1 & \text{if } k - m - \delta_{\mathbf{ref}} \leq 2 \end{cases}$$

with the "quasiuniformity measure" δ_{ref} from (13).

Proof. (2) As $\operatorname{div} \mathbf{d}^k = 0$ there is a $\phi_i^k \in \mathcal{S}(\tilde{T})$ with $\mathbf{curl}\, \phi_i^k = \mathbf{d}^k$. By $\mathcal{S}(\tilde{T})$ we denote the space of piecewise linear continuous finite element functions over $\mathcal{T}_k$ restricted to $\tilde{T}$. Recalling the definition (36) we can provide the basis representation

$$\phi_i^k = \sum_{\mathbf{p} \in \overline{\tilde{T}}\,;\, \mathbf{p} \in \mathsf{N}_{[l]}^i} \phi_i^k(\mathbf{p}) \psi_{\mathbf{p}}^k \,.$$

The classical idea from [3] is to isolate an internal part of ϕ_i^k and a boundary part:

$$\phi_i^k = \underbrace{\sum_{\mathbf{p} \in \mathsf{N}_{[k]}^i\,;\, \mathbf{p} \in \partial\tilde{T}} \phi_i^k(\mathbf{p}) \psi_{\mathbf{p}}^k}_{=:\phi_{i,bd}^k} + \underbrace{\sum_{\mathbf{p} \in \mathsf{N}_{[k]}^i\,;\, \mathbf{p} \in \tilde{T}} \phi_i^k(\mathbf{p}) \psi_{\mathbf{p}}^k}_{=:\phi_{i,int}^k} \tag{38}$$

This approach is motivated by an interesting feature of the RT_0 flux ansatz; denoting by curl^* the adjoint of the $\mathbf{curl}$ operator, that is $\operatorname{curl}^* \mathbf{u} = \frac{\partial u_2}{\partial x} - \frac{\partial u_1}{\partial y}$, we have $\operatorname{curl}^*(\mathbf{a} + \beta \mathbf{x}) = 0$. This permits to drop the internal part of ϕ_i^k, because of

$$\int_{\tilde{T}} \left\langle \mathbf{curl}\, \phi_{i,int}^k, \mathbf{z}^m \right\rangle d\mathbf{x} = \underbrace{\int_{\tilde{T}} \phi_{i,int}^k \cdot \operatorname{curl}^* \mathbf{z}\, d\mathbf{x}}_{=0, \text{ as } \operatorname{curl}^* \mathbf{z} = 0} - \underbrace{\int_{\partial\tilde{T}} \phi_{i,int}^k \left\langle \mathbf{z}, \boldsymbol{\tau} \right\rangle d\mathbf{x}}_{=0, \text{ as } \phi_{i,int}^k = 0 \text{ on } \partial\tilde{T}} = 0 \,.$$

($\boldsymbol{\tau}$ is a unit tangential vector.) How we benefit from this is illustrated by

$$\int_{\tilde{T}} \left\langle \mathbf{curl}\, \phi_i^k, \mathbf{z}^m \right\rangle d\mathbf{x} = \int_{\tilde{T}} \left\langle \mathbf{curl}\, \phi_{i,bd}^k, \mathbf{z}^m \right\rangle d\mathbf{x} = \int_{\Gamma} \left\langle \mathbf{curl}\, \phi_{i,bd}^k, \mathbf{z}^m \right\rangle d\mathbf{x}$$

where

$$\Gamma := \bigcup \{\operatorname{supp} \psi_{\mathbf{p}}^k ; \mathbf{p} \in \mathsf{N}_{[k]}^i, \mathbf{p} \in \partial\tilde{T}\}$$

is a sort of narrow fringe along the boundary of $\tilde{T}$. In sum, integration can be confined to a small part of $\tilde{T}$. As a preliminary result we have by the Cauchy–Schwarz inequality

$$\int_{\tilde{T}} \left\langle \mathbf{curl}\, \phi_{i,bd}^k, \mathbf{z}^m \right\rangle d\mathbf{x} \leq \left\|\mathbf{curl}\, \phi_{i,bd}^k\right\|_{\boldsymbol{L}^2(\Gamma)} \left\|\mathbf{z}^m\right\|_{L^2(\Gamma)} \tag{39}$$

The first factor can be easily bounded, thanks to the orthogonality of $\mathbf{curl}\,\phi^k_{i,int}$ and $\mathbf{curl}\,\phi^k_{i,bd}$ (Remember that the supports of these functions are disjoint by construction):

$$\left\|\mathbf{curl}\,\phi^k_{i,bd}\right\|_{\boldsymbol{L}^2(\Gamma)} \leq \left\|\mathbf{curl}\,\phi^k_i\right\|_{\boldsymbol{L}^2(\tilde{T})}$$

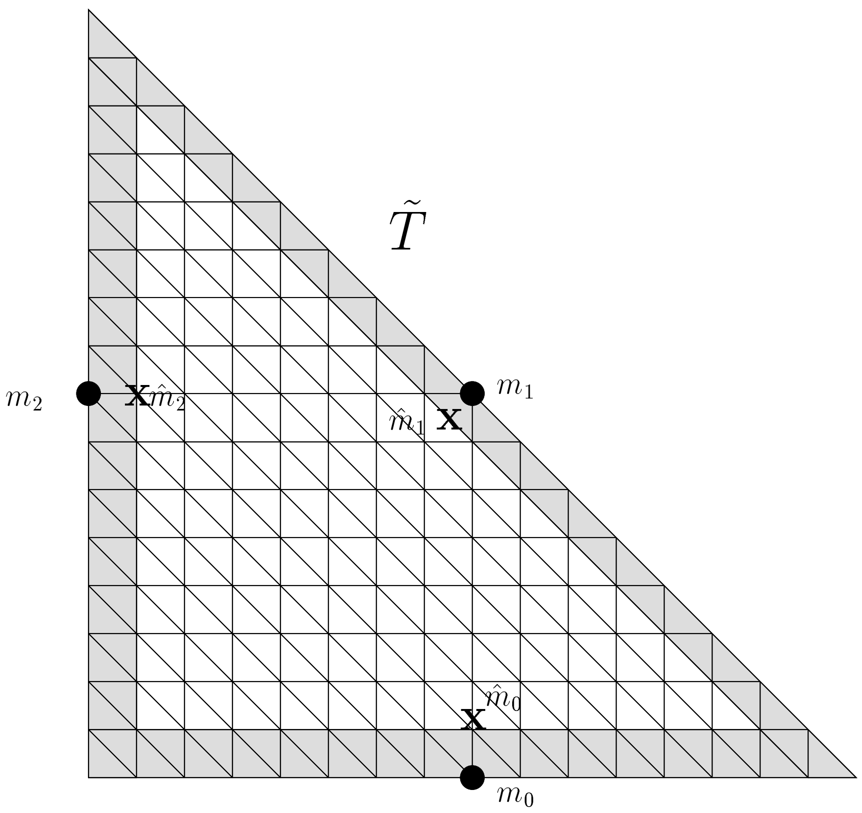

FIGURE 4: *$\tilde{T}$ with virtual mesh $\tilde{\mathcal{T}}_{\mathrm{virt}}$. Shaded boundary strip Γ' contains Γ.*

A bound for the second factor in (39) is determined making use of the fact that the area of Γ is only a fraction of $|\tilde{T}|$ if k and m differ widely. According to the quasiuniformity property (13) all elements in Γ are at least on level $\max\{m, k-\delta_{\mathrm{ref}}\}$. If $k-\delta_{\mathrm{ref}} \leq m-2$ we resort to the trivial estimate $\|\mathbf{q}\|_{L^2(\Gamma)} \leq \|\mathbf{q}\|_{L^2(\tilde{T})}$. Otherwise we carry out $s := k-m-\delta_{\mathrm{ref}} > 2$ steps of virtual *uniform* refinement of $\tilde{T}$, resulting in a triangulation $\tilde{\mathcal{T}}_{\mathrm{virt}}$ of $\tilde{T}$. The previous considerations demonstrated that

$$\Gamma \subset \Gamma' := \bigcup \left\{ t \in \tilde{\mathcal{T}}_{\mathrm{virt}}; \bar{t} \cap \partial\tilde{T} \neq \emptyset \right\} \ .$$

The situation is sketched in figure 4. Consider a linear function $f \in \mathcal{P}_1(\tilde{T})$. Standard Gaussian quadrature gives

$$\int_{\Gamma'} |f|^2 \, d\mathbf{x} = \int_{\tilde{T}} |f|^2 \, d\mathbf{x} - \int_{\tilde{T}/\Gamma'} |f|^2 \, d\mathbf{x} = \frac{1}{3}|\tilde{T}| \left(\sum_{j=1}^{3} f(\mathbf{m}_j)^2 \right) - \frac{1}{3}|\tilde{T}/\Gamma'| \left(\sum_{j=1}^{3} f(\hat{\mathbf{m}}_j)^2 \right)$$

Merely counting the small triangles in figure 4, we see that

$$\gamma := \frac{|\Gamma'|}{|\tilde{T}|} = 12\left(\left(\frac{1}{2}\right)^{s} - \left(\frac{1}{2}\right)^{2s}\right) \leq 12\rho \text{ with } \rho := \left(\frac{1}{2}\right)^{s} \tag{40}$$

(40) permits to estimate

$$\begin{aligned}\int\limits_{\Gamma'} |f|^2\,d\mathbf{x} &= \frac{1}{3}|\tilde{T}|\left(\sum_{j=1}^{3} f(\mathbf{m}_j)^2 - (1-\gamma)\sum_{j=1}^{3} f(\hat{\mathbf{m}}_j)^2\right)\\ &\leq |\tilde{T}|\sum_{j=1}^{3} f(\mathbf{m}_j)^2 \cdot (1-(1-\gamma)(1-2\rho)(1-6\rho))\end{aligned}$$

$$\int\limits_{\Gamma'} |f|^2\,d\mathbf{x} \leq 20\rho \int\limits_{\tilde{T}} |f|^2\,d\mathbf{x}\ . \tag{41}$$

To obtain the last estimate, we express f by means of its values at the points $\mathbf{m}_i$, $i = 1, 2, 3$ and barycentric coordinates. Then tedious computations lead to

$$\sum_{i=1}^{3} f(\hat{\mathbf{m}}_i)^2 = \left(2\rho^2 + (1-2\rho)^2\right)\sum_{i=1}^{3} f(\mathbf{m}_i)^2 + \left(4\rho(1-2\rho) + 2\rho^2\right)\sum_{k\neq l} f(\mathbf{m}_k)f(\mathbf{m}_l)\ ,$$

which can be bounded by

$$\sum_{i=1}^{3} f(\hat{\mathbf{m}}_i)^2 \geq (1-2\rho)(1-6\rho)\sum_{i=1}^{3} f(\mathbf{m}_i)^2\ .$$

(41) means for $\mathbf{z}^m$, since its components are linear over $\tilde{T}$:

$$\|\mathbf{z}^m\|_{\boldsymbol{L}^2(\Gamma')} \leq \sqrt{20\left(\frac{1}{2}\right)^{s}}\,\|\mathbf{z}^m\|_{\boldsymbol{L}^2(\tilde{T})}$$

Now the proof is complete. □

For the sake of an unified treatment we change notations

$$\boldsymbol{\mathcal{Y}}_0^l := \boldsymbol{\mathcal{H}}_l \quad , \quad \boldsymbol{\mathcal{Y}}_i^l := \boldsymbol{\mathcal{D}}_i^l \tag{42}$$

and we are now concerned with the splitting

$$\boldsymbol{\mathcal{V}}_h = \sum_{l=0}^{L}\sum_{i=0}^{4} \boldsymbol{\mathcal{Y}}_i^l \tag{43}$$

THEOREM 6.3 (STRENGTHENED CAUCHY–SCHWARZ INEQUALITY). *For the splitting (43) of $\boldsymbol{\mathcal{V}}_h$ holds a strengthened Chauchy–Schwarz inequality of the form*

$$a_r(\mathbf{y}_i^k, \mathbf{y}_l^m) \leq \underbrace{\frac{1}{\underline{\alpha}}\max\left\{\frac{2C_2\sqrt{\overline{\alpha}}}{\sqrt{r}}, \sqrt{20\cdot 2^{\delta_{\mathrm{ref}}}}\right\}\left(\frac{1}{\sqrt{2}}\right)^{|k-m|}}_{e_{i,l,k,m}} \left\|\mathbf{y}_i^k\right\|_{A_r}\ \left\|\mathbf{y}_l^m\right\|_{A_r}$$

for all $\mathbf{y}_i^k \in \boldsymbol{\mathcal{Y}}_i^k, \mathbf{y}_l^m \in \boldsymbol{\mathcal{Y}}_l^m$, $l, i \in \{0, \dots, 4\}$. *The constant* $C > 0$ *does neither depend on* $\mathbf{y}^k, \mathbf{y}^m$ *nor on the number* L *of refinement levels.*

Proof. Without loss of generality we may assume $k > m$. First consider a $\mathbf{y}_0^k \in \boldsymbol{\mathcal{Y}}_0^k$ and apply lemma 6.1 to each element of $\mathcal{T}_m$:

$$\begin{aligned} a_r(\mathbf{y}_0^k, \mathbf{z}^m) &\le \sum_{T \in \mathcal{T}_m} a_{r|T}(\mathbf{y}_0^k, \mathbf{z}^m) \\ &\le \frac{2C_2}{\underline{\alpha}} \left(\frac{1}{2}\right)^{k-m} \sum_{T \in \mathcal{T}_m} |T|^{\frac{1}{2}} \left\|\operatorname{div} \mathbf{y}_0^k\right\|_{L^2(T)} \left\|\mathbf{z}^m\right\|_{\boldsymbol{L}^2(T)} \\ &\le \frac{2C_2}{\underline{\alpha}} \left(\frac{1}{2}\right)^{k-m} \left(\sum_{T \in \mathcal{T}_m} \left\|\operatorname{div} \mathbf{y}_0^k\right\|_{L^2(T)}^2\right)^{\frac{1}{2}} \left(\sum_{T \in \mathcal{T}_m} \left\|\mathbf{z}^m\right\|_{\boldsymbol{L}^2(T)}^2\right)^{\frac{1}{2}} \\ &= \frac{2C_2}{\underline{\alpha}} \left(\frac{1}{2}\right)^{k-m} \left\|\operatorname{div} \mathbf{y}_0^k\right\|_{L^2(\Omega)} \left\|\mathbf{z}^m\right\|_{\boldsymbol{L}^2(\Omega)} \quad \forall \mathbf{z}^m \in \boldsymbol{\mathcal{V}}_m \end{aligned}$$

From the definition of the bilinear form a_r in (12) we conclude

$$\begin{aligned} \frac{1}{r} \left\|\mathbf{y}_0^k\right\|_{A_r}^2 &\ge \left\|\operatorname{div} \mathbf{y}_0^k\right\|_{L^2(\Omega)}^2 \\ \overline{\alpha} \left\|\mathbf{z}^m\right\|_{A_r}^2 &\ge \left\|\mathbf{z}^m\right\|_{\boldsymbol{L}^2(\Omega)}^2 \ . \end{aligned} \tag{44}$$

So we end up with

$$a_r(\mathbf{y}_0^k, \mathbf{z}^m) \le \frac{2C_2}{\underline{\alpha}} \frac{\sqrt{\overline{\alpha}}}{\sqrt{r\,\underline{\alpha}}} \left(\frac{1}{2}\right)^{k-m} \left\|\mathbf{y}_0^k\right\|_{A_r} \left\|\mathbf{z}^m\right\|_{A_r} \quad \forall \mathbf{z}^m \in \boldsymbol{\mathcal{V}}_m \tag{45}$$

In a similar fashion lemma 6.2 and (44) yield for a divergence free vector field $\mathbf{y}_i^k \in \boldsymbol{\mathcal{Y}}_i^k$, $i \in \{1, 2, 3, 4\}$:

$$a_r(\mathbf{y}_i^k, \mathbf{z}^m) \le \frac{\overline{\alpha}}{\underline{\alpha}} \sqrt{20 \cdot 2^{\delta_{\text{ref}}} \left(\frac{1}{2}\right)^{k-m}} \left\|\mathbf{y}_i^k\right\|_{A_r} \left\|\mathbf{z}^m\right\|_{A_r} \quad \forall \mathbf{z}^m \in \boldsymbol{\mathcal{V}}_m \tag{46}$$

where C_6 is the constant from lemma 6.2. Combining (45) and (46) we are finished. □

From theorem 6.3 we gain the bound

$$C_{\text{orth}} := \max_{i,m} \sum_{l,k} e_{i,l,m,k} \le C \frac{\overline{\alpha}}{\underline{\alpha} r} \ .$$

7 Additive preconditioner The results of the previous sections laid the theoretical foundation for multilevel preconditioning of the operator A_r ($r > 0$). Here we consider an additive subspace correction scheme based upon the decomposition (43). Then the linear operator C_r, supposed to approximate $A_r{}^{-1}$, is given by

$$C_r := \sum_{l=0}^{L} \sum_{i=0}^{4} \left(A_{r,i}^l\right)^{-1} Q_i^l \tag{47}$$

where $A_{r,i}^l$ is the restriction of A_r to $\boldsymbol{\mathcal{Y}}_i^l$ and $Q_i^l : \boldsymbol{\mathcal{V}}_h \mapsto \boldsymbol{\mathcal{Y}}_i^l$ stands for the $L^2(\Omega)$–orthogonal projection. If $P_i^l : \boldsymbol{\mathcal{V}}_h \mapsto \boldsymbol{\mathcal{Y}}_i^l$ denotes the a_r–orthogonal projection, we may alternatively write

$$C_r := \sum_{l=0}^{L} \sum_{i=0}^{4} P_i^l A_r^{-1} \ . \tag{48}$$

From (48),(6.3),(5.3) and crucial results of multilevel theory (cf. [29], theorem 4.1 and lemma 4.6) we learn

$$\frac{1}{C_{\mathrm{stab}}} (\mathbf{v}_h, \mathbf{v}_h)_{\boldsymbol{L}^2(\Omega)} \leq (C_r A_r \mathbf{v}_h, \mathbf{v}_h)_{\boldsymbol{L}^2(\Omega)} \leq C_{\mathrm{orth}} (\mathbf{v}_h, \mathbf{v}_h)_{\boldsymbol{L}^2(\Omega)} \tag{49}$$

or, equivalently, stated as an estimate for the spectral condition number

$$\kappa(C_r A_r) \leq C_{\mathrm{stab}} \cdot C_{\mathrm{orth}} \ . \tag{50}$$

Keep in mind that both constant can be chosen independent of L. Furthermore, with respect to the augmented Lagrangian parameter $r > 0$, $\kappa(C_r A_r)$ behaves like $O(\frac{1}{r^{3/2}})$ for $r \ll 1$ and decreases to a floor value for $r \to \infty$. The bottom line is that the preconditioner rather benefits from large values of r, in sharp contrast to the deteriorating condition number of A_r (see theorem 3.1).

All the spaces $\boldsymbol{\mathcal{Y}}_i^l$ consist of several parts whose vector fields have disjoint supports. For $\boldsymbol{\mathcal{Y}}_i^l$, $i \in \{1,2,3,4\}$ those are related to individual vertices of $\mathsf{N}_{[l]}$, for $\boldsymbol{\mathcal{Y}}_0^l$ they belong to refined elements of $\mathcal{T}_{l-1}$. They can be dealt with totally separately. In detail, let $\{\boldsymbol{\xi}_{1,i}^l, \ldots, \boldsymbol{\xi}_{D_i^l,i}\} \subset \boldsymbol{\mathcal{Y}}_i^l$ be a basis of such an isolated component with dimensions D_i^l. Then for $\mathbf{v}_h \in \boldsymbol{\mathcal{V}}_h$ the coefficients α_k of the basis representation

$$P_i^l A_r^{-1} \mathbf{v}_h =: \sum_{k=1}^{D_i^l} \alpha_k \boldsymbol{\xi}_{k,i}^l$$

can be computed as the solution of the $D_i^l \times D_i^l$ linear system

$$\left(a_r(\boldsymbol{\xi}_{k,i}^l, \boldsymbol{\xi}_{j,i}^l)\right)_{1 \leq k,j \leq D_i^l} (\alpha_k)_{1 \leq k \leq D_i^l} = \left(\left(\mathbf{v}_h, \boldsymbol{\xi}_{k,i}^l\right)_{\boldsymbol{L}^2(\Omega)}\right)_{1 \leq k \leq D_i^l} \ . \tag{51}$$

Since the decomposition deliberately relied on low–dimensional subspaces, the problems (51) are cheaply solved. Precisely speaking, we have $D_i^l = 1$ for $i \in \{1,2,3,4\}$ and $D_0^l \in \{1,3\}$.

How corrections can efficiently be determined in divergence free subspaces deserves closer attention. In this case (51) reduces to the scalar equation

$$\left(a^{-1} \,\mathbf{curl}\, \psi_{\mathbf{p}}^l, \mathbf{curl}\, \psi_{\mathbf{p}}^l\right)_{\boldsymbol{L}^2(\Omega)} \alpha_0 = (\mathbf{v}_h, \mathbf{curl}\, \psi_{\mathbf{p}})_{\boldsymbol{L}^2(\Omega)} \tag{52}$$

with $\mathbf{p} \in \mathsf{N}_{[l]}$. The representation

$$\mathbf{curl}\, \psi_{\mathbf{p}}^l = \sum_{\text{edges } e, \mathbf{p} \in e} \pm 1 \ \boldsymbol{\iota}_e^l \ , \tag{53}$$

where $\boldsymbol{\iota}_e^l$ is the canonical RT_0 basis function associated with the edge e of $\mathcal{T}_l$, hints at how an explicit transformation to a basis of $\mathcal{S}_l$ can be avoided: We stick to the canonical basis of $\boldsymbol{\mathcal{V}}_l$ and use (53) to collect the values at vertices of $\mathsf{N}_{[l]}$. After a simple scaling according to (52), (53) again offers a recipe how the correction should be distributed back to the flux nodes. Figure 5 illustrates the process.

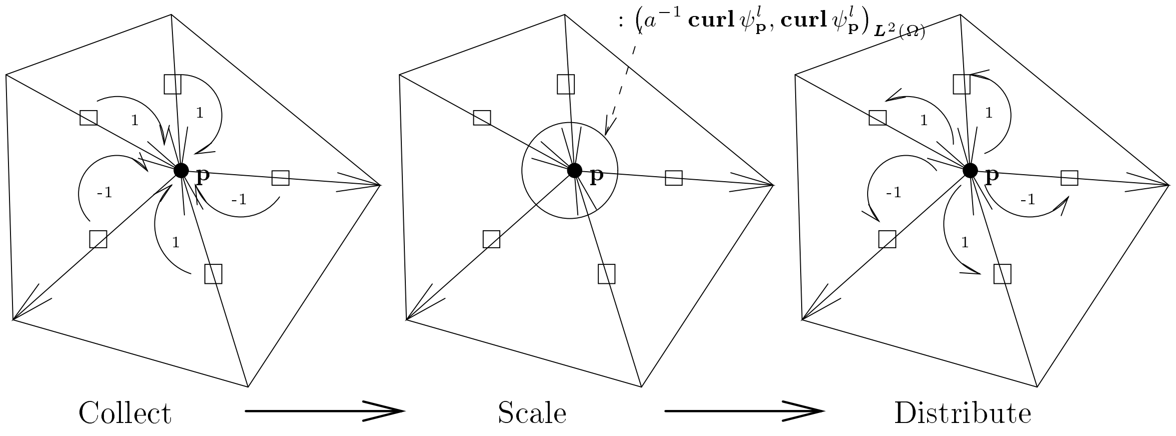

FIGURE 5: *Calculating corrections in divergence free subspaces. The boxes mark the "sites" of nodes of* $\boldsymbol{\mathcal{V}}_h$.

The entire evaluation of the preconditioner for a $\mathbf{r}_h \in \boldsymbol{\mathcal{V}}_h'$ is done in a multigrid like fashion. The transfers of flux functions between different levels are canonical, since we have the embeddings $\boldsymbol{\mathcal{V}}_{l-1} \subset \boldsymbol{\mathcal{V}}_l$ and $\boldsymbol{\mathcal{V}}_l' \subset \boldsymbol{\mathcal{V}}_{l-1}'$. An outline of the algorithm is given in figure 6.

The solution of the local 3×3 problems turns out to be the most expensive part of the algorithm. Thus it becomes desirable to seek a less cumbersome substitute, which, however, must not grossly affect the condition number. The simplest choice already meets the requirements:

LEMMA 7.1. *Let $\bar{T} \in \mathcal{T}_{l-1}$ be an element on level $l \in \{0,\dots,L-1\}$ that has been red refined as in figure 1. Define canonical projections by*

$$I_k \;:\; \begin{cases} \boldsymbol{\mathcal{H}}_l(\bar{T}) & \mapsto \quad \mathrm{Span}\,\{\boldsymbol{\iota}^k\} \\ \mathbf{q}_l := \sum\limits_{m=1}^{3} \alpha_m \boldsymbol{\iota}^m & \mapsto \quad \alpha_k \boldsymbol{\iota}^k \,. \end{cases}$$

Then we have

$$\min\left\{\frac{1}{4}, \frac{1}{6C_1^2}\right\} \|\mathbf{q}\|_{A_r}^2 \le \sum_{k=1}^{3} \|I_k \mathbf{q}_l\|_{A_r}^2 \le \max\left\{2, 6C_1^2\right\} \|\mathbf{q}\|_{A_r}^2 \,.$$

Proof. The proof separately deals with both parts of the A_r–norm:

(1) Applying lemma 2.1 we immediately obtain:

$$\begin{aligned} \sum_{k=1}^{3} \|I_k \mathbf{q}_l\|_{\boldsymbol{L}^2(\bar{T})} &\le 2C_1 \sum_{k=1}^{3} \alpha_k^2 \le 6C_1^2 \|\mathbf{q}_l\|_{\boldsymbol{L}^2(\bar{T})} \\ \|\mathbf{q}_l\|_{\boldsymbol{L}^2(\bar{T})} &\le 2C_1 \sum_{k=1}^{3} \alpha_k^2 \le 6C_1^2 \|I_k \mathbf{q}_l\|_{\boldsymbol{L}^2(\bar{T})} \end{aligned}$$

for $l := L$ **to** 1 **do**

Restrict $\mathbf{r}_l$ to $\boldsymbol{\mathcal{V}}_{l-1}$. This yields $\mathbf{r}_{l-1}$.

Correction in divergence free subspaces:

- Collect values of $\mathbf{r}_l$ from flux nodes in vertices of $\mathsf{N}_{[l]}$.
- Scale values at the vertices.
- Distribute results to flux nodes and add them to $\mathbf{c}_l$.

Solve local problems over refined elements of $\mathcal{T}_l$: Take restrictions of $\mathbf{r}_l$ as right hand sides and add results to $\mathbf{c}_l$

Solve coarse grid problem $A_{r,0}\mathbf{c}_0 = \mathbf{r}_0$

for $l := L-1$ **to** 0 **do**

Add $\mathbf{c}_l \in \boldsymbol{\mathcal{V}}_l$ prolongated to $\boldsymbol{\mathcal{V}}_{l+1}$ to $\mathbf{c}_{l+1}$.

FIGURE 6: *Evaluation of additive preconditioner applied to* $\mathbf{r}_h \in \boldsymbol{\mathcal{V}}_h' =: \mathbf{r}_l$.

(2) Elementary properties of the RT_0 flux ansatz yield

$$\begin{aligned}\sum_{k=1}^{3} \|\operatorname{div} I_k \mathbf{q}_l\|^2_{L^2(\bar{T})} &= \sum_{k=1}^{3} \left(|T_k| |\operatorname{div} I_k \mathbf{q}_{l|T_k}|^2 + |T^*| |\operatorname{div} I_k \mathbf{q}_{l|T^*}|^2\right) \\ &= \frac{1}{2} |\bar{T}| \sum_{k=1}^{3} |\operatorname{div} I_k \mathbf{q}_{l|T_k}|^2 \\ &= 2 \sum_{k=1}^{3} \left\|\operatorname{div} \mathbf{q}_{l|T_k}\right\|^2_{L^2(T_k)} \leq 2 \,\|\operatorname{div} \mathbf{q}_l\|^2_{L^2(\bar{T})}\end{aligned}$$

as well as

$$\begin{aligned}
\|\operatorname{div}\mathbf{q}_l\|^2_{L^2(\bar T)} &= \frac{1}{4}|\bar T|\left(\sum_{k=1}^{3}\left|\operatorname{div}\mathbf{q}_{l|T_k}\right|^2 + \left|\sum_{k=1}^{3}\operatorname{div}\mathbf{q}_l\right|^2\right)\\
&\le |\bar T|\sum_{k=1}^{3}\left|\operatorname{div}\mathbf{q}_{l|T_k}\right|^2\\
&\le 4\sum_{k=1}^{3}\left\|\operatorname{div} I_k\mathbf{q}_{l|T_k}\right\|^2_{L^2(T_k)} \le 4\sum_{k=1}^{3}\|\operatorname{div} I_k\mathbf{q}_l\|^2_{L^2(\bar T)}
\end{aligned}$$

□

Thus we can solve scalar problems instead of linear systems and still achieve asymptotic optimality. Yet the constant might be considerably worse.

8 Numerical experiments. In the preceding sections we established the asymptotic optimality of the additive preconditioner. Nevertheless the bounds were riddled with ominous constant whose values could not be accurately estimated theoretically. As they crucially affect the performance of the overall scheme, numerical experiments are required to gain additional insights.

The calculations were done on a workstation based on the adaptive finite element code GOOFE (see [19]). We resorted to the simplest model problem

$$\begin{aligned} -\Delta u &= f \qquad \text{in } \Omega :=]0;1[^2 \\ u &= 0 \qquad \text{on } \partial\Omega \quad . \end{aligned} \tag{54}$$

We need not specify $f \in L^2(\Omega)$, since the right hand side is irrelevant for the numerical experiments covered below.

The first experiment relied on uniform refinement of an initial triangulation consisting of two triangles. On each of the resulting levels the smallest and largest eigenvalue of C_rA_r for $r = 1$ were determined by means of power iterations within a margin of error of at worst five percent. They are plotted in figure 7. Figure 8 displays the resulting condition numbers. Unfortunately the refinement could not be carried out as far as to reach the asymptotic region. At the levels covered by the experiment, logarithmic growth of $\kappa(C_rA_r)$ still prevails. This is a customary observation with BPX–type preconditioners if not many levels are involved.

How the use of the additive preconditioner accelerates the convergence of a preconditioned CG method for $A_rx = b$ is illustrated by the convergence rates from figure 9. The CG method ran with a random initial guess until the residual dropped below a threshold of 10^{-5}. An estimate for the convergence rate was obtained by

$$\rho \approx \left(\frac{\|r_N\|}{\|r_0\|}\right)^{\frac{1}{N}} \tag{55}$$

where N is the number of CG steps required, r_i the residual after step i and $\|\cdot\|$ denotes the Euclidean norm. The figures plotted represent averages from several runs to offset the

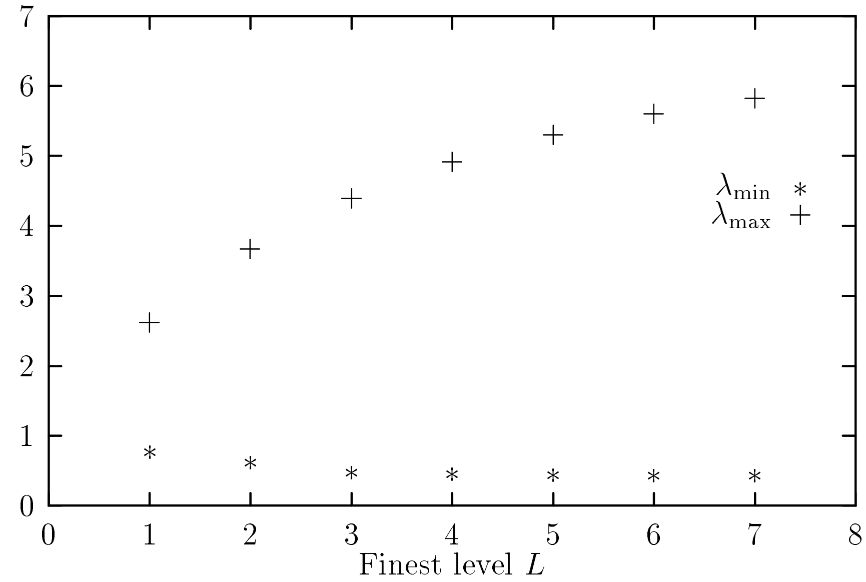

FIGURE 7: $\lambda_{\min}(C_r A_r)$, $\lambda_{\max}(C_r A_r)$ *for* $r = 1$ *in the case of uniform refinement.*

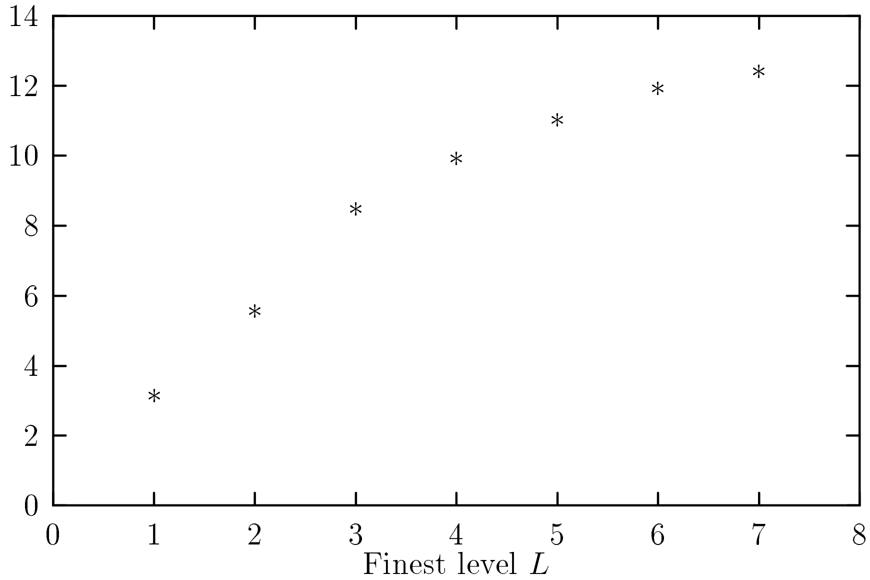

FIGURE 8: $\kappa(C_r A_r) := \frac{\lambda_{\max}}{\lambda_{\min}}$ *for* $r = 1$ *in the case of uniform refinement.*

impact of random x_0. Also displayed are corresponding rates for the CG method without preconditioner. The superior speed of the preconditioned iteration is evident.

We also investigated the dependency of $\kappa(C_r A_r)$ on the value of the augmented Lagrangian parameter r. $\kappa(C_r A_r)$ was obtained as above. The results are shown in figure 10. In agreement with the theoretical bounds the condition of the preconditioned system soars according to a $r^{-\beta}$ $(\beta > 0)$ law as $r \to 0$. Moreover, a comparison of figures 11,12 reveals that a significant drop in the smallest eigenvalue of $C_r A_r$ for small r is the main culprit; the impact on $\lambda_{\max}$ is much weaker.

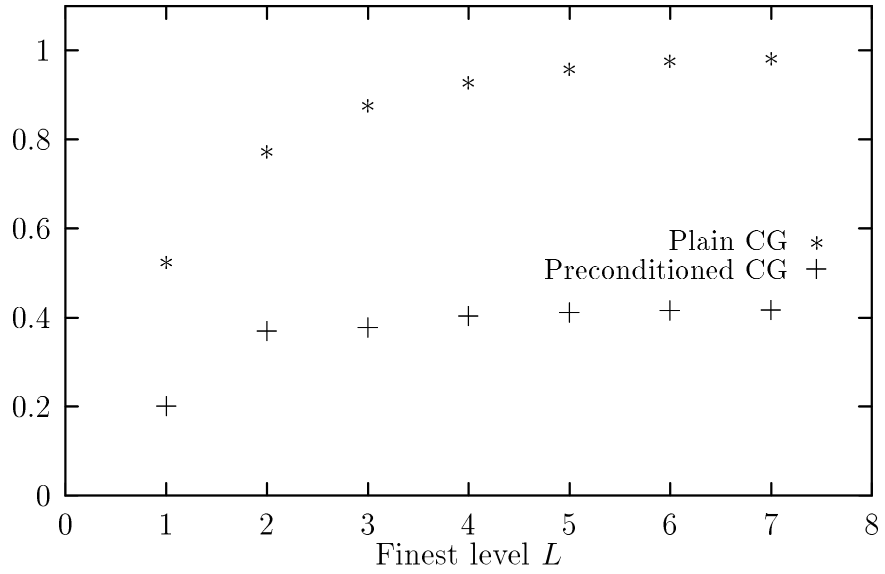

FIGURE 9: *Convergence rates, as defined in for preconditioned CG method and plain CG method.*

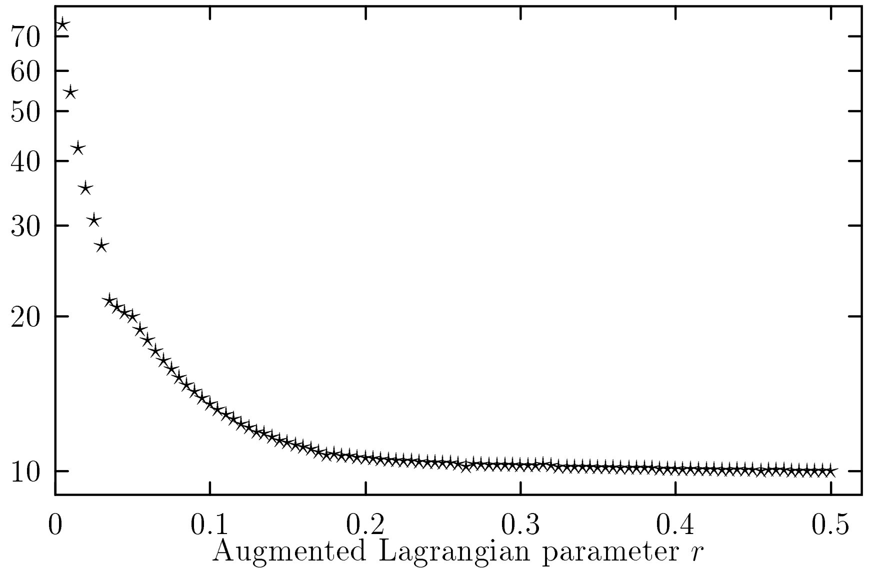

FIGURE 10: *Condition numbers $\kappa(C_r A_r)$ as a function of $r > 0$ (Logarithmic scale).*

We also tested the effect of multilevel preconditioning in the case of local refinement. In figure 13 several of those meshes used for the computations are drawn. They typically arise, for instance, if one tries to zero in on a local singularity guided by a posteriori error estimation as in [20].

In figure 14 the values of $\kappa(C_r A_r)$ for $r = 1$ are plotted: they behave as neatly as for inform refinement.

The final experiment probes the impact of the transition from full local 3×3 problems to

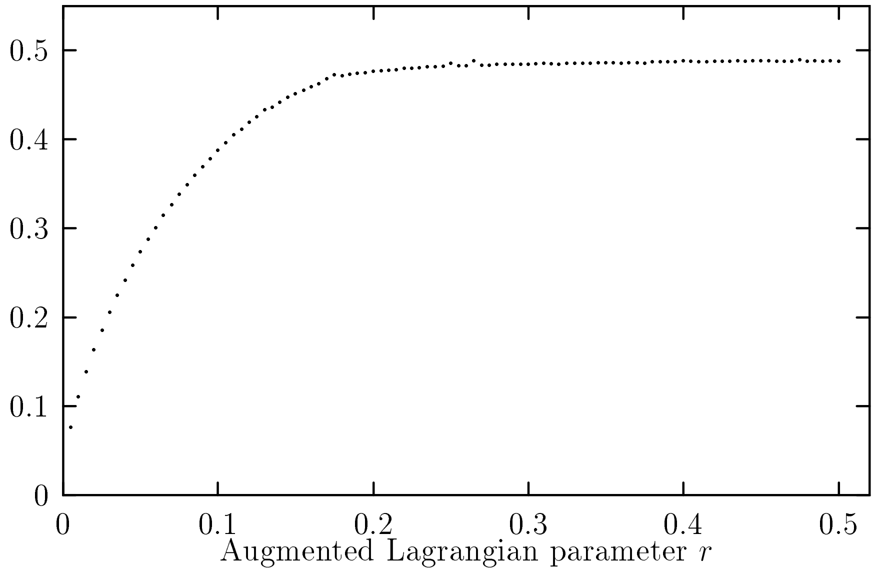

FIGURE 11: $\lambda_{\min}(C_r A_r)$ *depending on the augmented Lagrangian parameter* r.

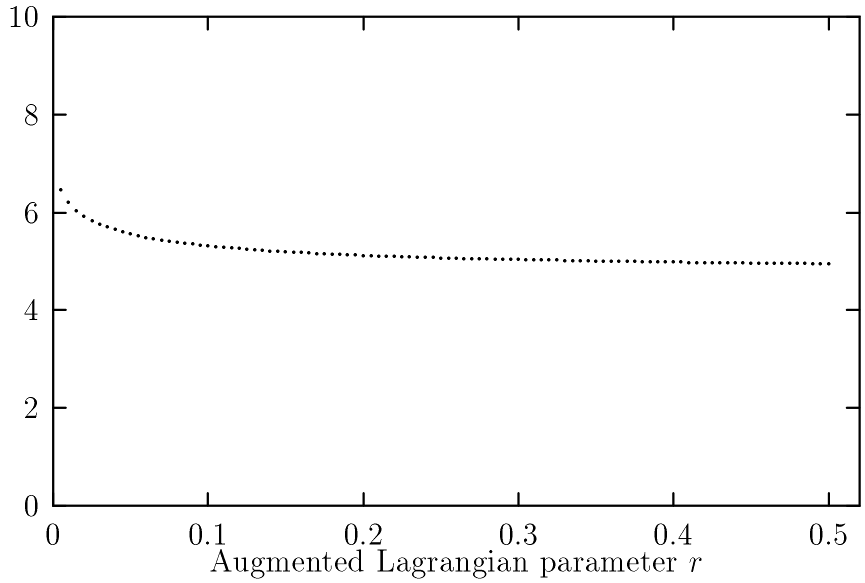

FIGURE 12: $\lambda_{\max}(C_r A_r)$ *depending on the augmented Lagrangian parameter* r.

simpler 1×1 problems, as derived in section 7. The setting was the same as for the first experiment: $r = 1$ and uniform refinement of an initial mesh consisting of two triangles. Figure 15 shows the condition numbers.

Surprisingly enough, ignoring the coupling between the basis functions in the local problem does not seem to wreak much of an havoc to the preconditioner. The figures indicate that the loss in quality is negligible.

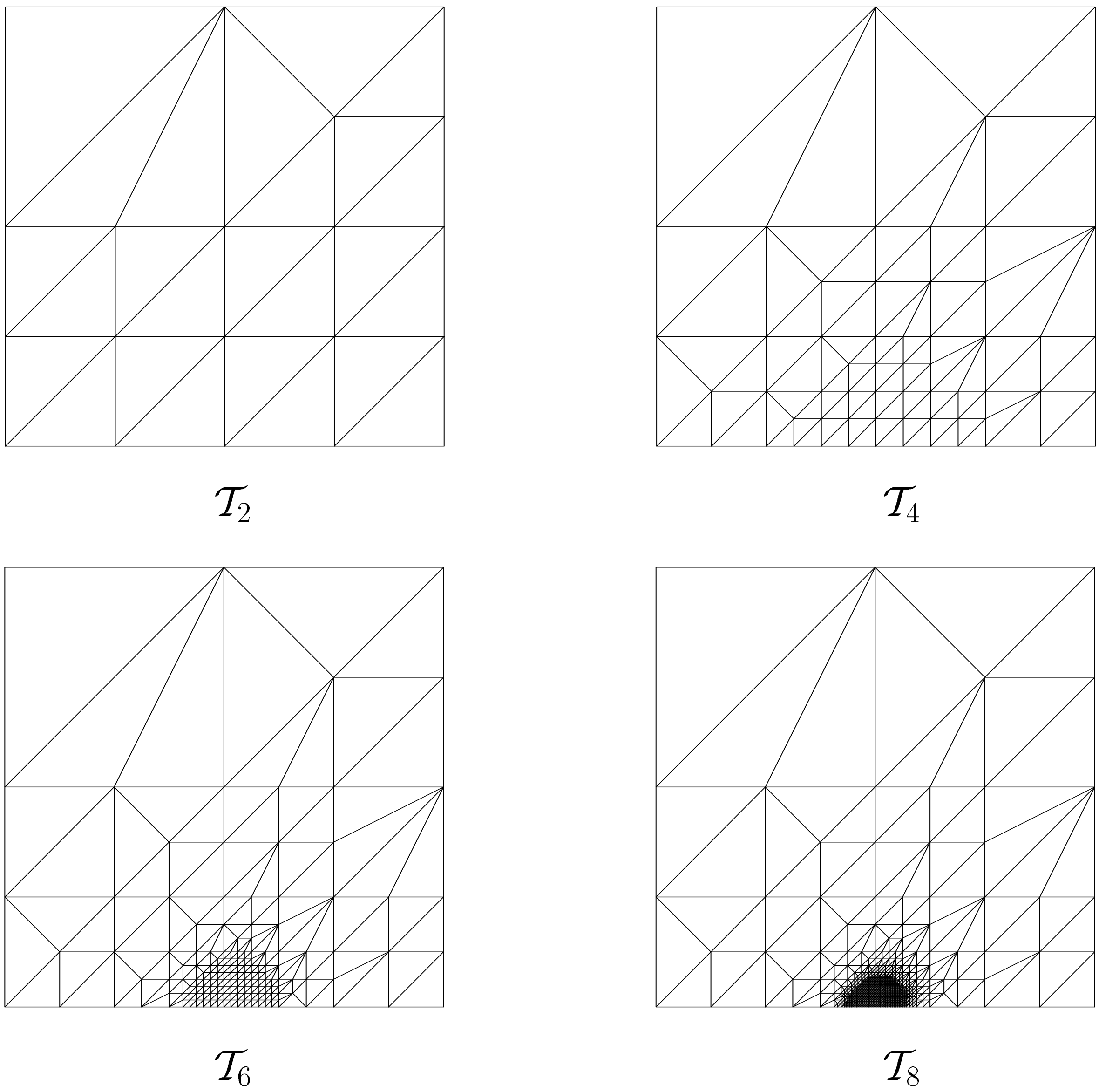

FIGURE 13: *Some locally refined triangulations.*

9 Conclusion The mixed finite element discretization of 2nd order elliptic problems leads to saddle point problems. The augmented Lagrangian approach offers a way to convert them into problems with more favorable characteristics. A positive parameter permits delicate control of the condition numbers of parts of the problems. Mostly those are tackled by means of iterative schemes, which employ inner and outer iterations. To achieve satisfactory speed of the inner iteration, an elliptic operator, set in flux space, requires preconditioning. The paper proved that an additive Schwarz preconditioner based on a multilevel splitting by Vassilevski and Wang yields an optimal method. A bound for the spectral condition of the preconditioned system, independent on the value of the augmented Lagrangian parameter and the depth of refinement, could be established. Numerical results strikingly highlight the efficacy of the scheme.

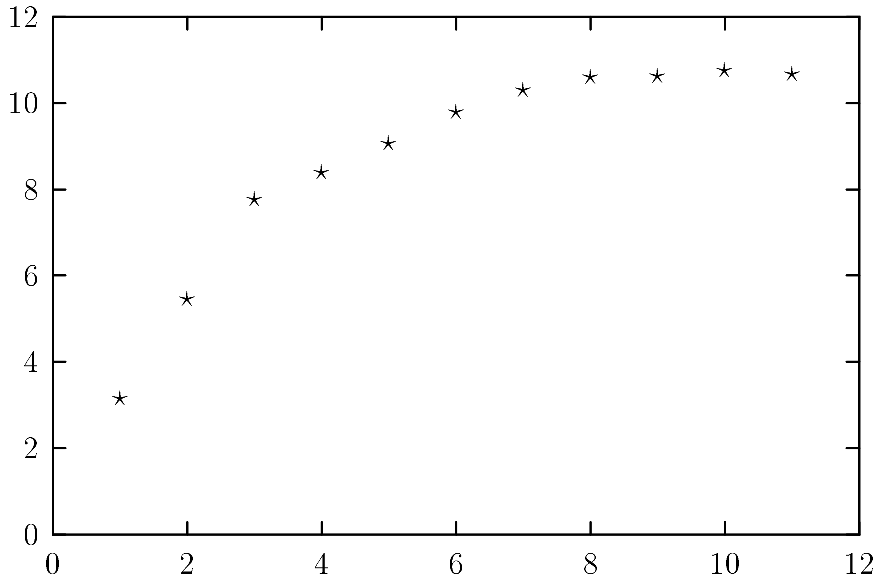

FIGURE 14: *Condition number $\kappa(C_1 A_1)$ as function of the depth of local refinement*

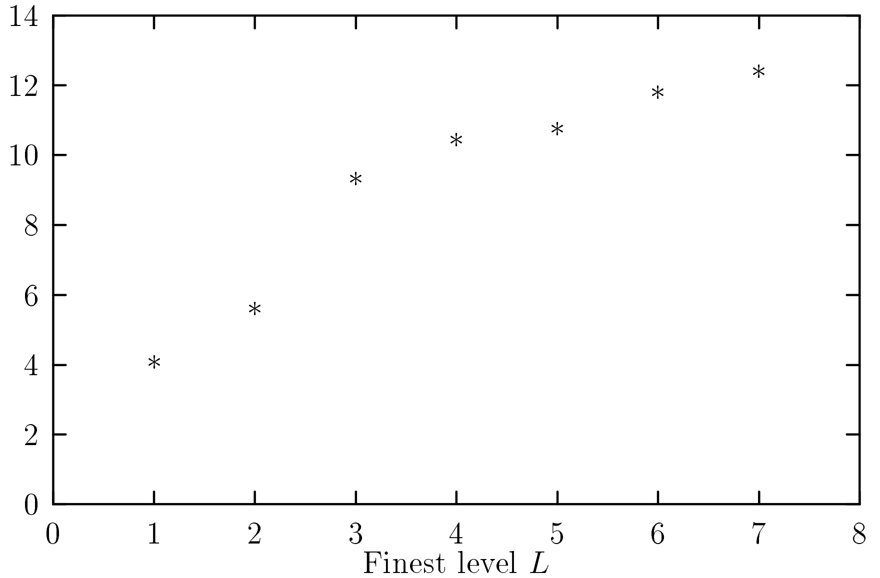

FIGURE 15: *Condition numbers $\kappa(C_1 A_1)$ for the additive preconditioner entirely relying on onedimensional subspaces.*

References

[1] R. Bank, *PLTMG: A Software Package for Solving Elliptic Partial Differential Equations, User's Guide 6.0*, SIAM, Philadelphia, 1990

[2] R. Bank, B. Welfert, H. Yseretant, *A class of iterative methods for solving saddle point problems*, Numer. Math, **56**, pp. 645-666 (1990)

[3] F. Bornemann, *A sharpened condition number estimate for the BPX–precondtioner of elliptic finite element problems on highly nonuniform triangulations*, Preprint SC 91–9 ZIB Berlin (1991)

[4] J.H. Bramble, J.E. Pasciak, J. Xu, *Parallel multilevel preconditioners*, Math. Comp., **55**, pp. 1–22 (1990)

[5] F. Brezzi, *On the existence, uniqueness, and approximation of saddle point problems arising from Lagrangian multipliers*, Math. Comp., **57**, pp. 129–151 (1974)

[6] S. Brenner, *A multigrid algorithm for the lowest order Raviart–Thomas mixed triangular finite element method*, SIAM J. Numer. Anal. 29, 647-678 (1992)

[7] D. Braess, *Finite Elemente*, Springer Verlag Berlin Heidelberg 1992

[8] F. Brezzi, M. Fortin, *Mixed and hybrid finute element methods*, Springer Verlag, New York 1991

[9] Z. Cai, C.I. Goldstein und J. Pasciak, *Multilevel iteration for mixed finite element systems with penalty*, SIAM J. Sci. Comput., **14**, pp. 1072–1088 (1993)

[10] P. Ciarlet, *The finite element method for elliptic problems*, North-Holland, Amsterdam, 1987

[11] W. Dahmen, A. Kunoth, *Multilevel preconditioning*, Numer. Math, **63**, pp. 315–344 (1992)

[12] P. Deuflhard, P. Leinen und H. Yserentant, *Concepts of an adaptive hierarchical finite element code*, IMPACT Comput. Sci. Engrg, **1**, pp. 3–35 (1989)

[13] R. Ewing, D. Lazarov, P. Lu, P. Vassilevski, *Preconditioning indefinte systems arising from mixed finite element discretization of second order elliptic problems*, Lecture Notes in Mathematics, Springer Verlag (1990)

[14] J.R. Ewing, J. Wang, *Analysis of the Schwarz algorithm for mixed finite element methods*, M^2AN, **26**, pp. 739–756 (1992)

[15] ——, *Analysis of multilevel decomposition iterative methods for mixed finite element methods*, M^2AN, **28**, pp. 377–398 (1994)

[16] M. Fortin, R. Glowinski, *Augmented Lagrangian methods: Applications to the numerical solution of boundary value problems*, North Holland, Studies in Mathematics and its Applications, Vol. 15, Amsterdam 1983

[17] V. Girault, P.A. Raviart, *Finite Element Methods for Navier–Stokes Equations: Theory and Algorithms*, Springer Verlag, New York, 1986

[18] R. Glowinski, P. LeTallec, *Augmented Lagrangian and operator–splitting methods in nonlinear mechanics*, SIAM Studies in Applied Mathematics, Philadelphia 1989

[19] R. Hiptmair, *Object Oriented Concepts for an Adaptive Finite Element Code* Report Mathematisches Institut TU München, to appear

[20] R.H.W. Hoppe und B. Wohlmuth, *Adaptive multilevel techniques for mixed finite element discretizations of elliptic boundary value problems*, Report M9411, Mathematisches Institut, TU München 1994 (Submitted to SIAM J. Numer. Anal.)

[21] R.H.W. Hoppe, B. Wohlmuth, *Efficient Numerical Solution of Mixed Finite Element Discretizations by Adaptive Multilevel Methods*, Report M9410, Mathematisches Institut, TU München 1994 (submitted toAppl. Math.)

[22] T.P. Mathew, *Schwarz alternating and iterative refinement methods for mixed formulations of elliptic problems, Part I: Algorithms and numerical results*, Numer. Math. , **65**, pp. 445–468 (1993)

[23] ——, *Schwarz alternating and iterative refinement methods for mixed formulations of elliptic problems. Part II: Convergence theory*, Numer. Math. ,**65**, pp. 469-492 (1993)

[24] P. Oswald, *Two remarks on multilevel preconditioners*, Forsch.-Erg. FSU Jena, Math. /91/1, 1991

[25] P. A. Raviart, J. M. Thomas, *A Mixed Finite Element Method for Second Order Elliptic Problems*, Lecture Notes in Mathematics, **606**, Springer–Verlag 1977

[26] T. Rusten, R. Winther, *A preconditioned iterative method for saddlepoint problems*, SIAM J. Matrix Anal. Appl., **13**, pp. 887–904 (1992)

[27] P.S. Vassilevski, J. Wang, *Multilevel iterative methods for mixed finite element discretizations of elliptic problems*, Numer. Math, **63**, pp. 503–520 (1992)

[28] B. Wohlmuth, *Adaptive Multilevel–Finite–Elemente Methoden zur Lösung elliptischer Randwertprobleme*, Ph.D. Thesis, Mathematisches Institut, TU München 1995

[29] J. Xu, *Iterative methods by space decomposition and subspace correction*, SIAM Review, **34**, pp. 581–613 (1992)

[30] H. Yserentant, *Old and new convergence proofs for multigrid methods*, Acta Nuemerica, pp. 1–100 (1993)

[31] X. Zhang, *Multilevel Schwarz methods*, Numer. Math., **63**, pp. 521–539 (1992)