J Sci Comput (2009) 40: 4–36
DOI 10.1007/s10915-009-9293-1

Uniformly Convergent Iterative Methods for Discontinuous Galerkin Discretizations

Blanca Ayuso de Dios · Ludmil Zikatanov

Received: 23 June 2008 / Revised: 2 April 2009 / Accepted: 3 April 2009 / Published online: 15 April 2009

Abstract We present iterative and preconditioning techniques for the solution of the linear systems resulting from several discontinuous Galerkin (DG) Interior Penalty (IP) discretizations of elliptic problems. We analyze the convergence properties of these algorithms for both symmetric and non-symmetric IP schemes. The iterative methods are based on a "natural" decomposition of the first order DG finite element space as a direct sum of the Crouzeix-Raviart non-conforming finite element space and a subspace that contains functions discontinuous at interior faces. We also present numerical examples confirming the theoretical results.

Keywords Discontinuous Galerkin finite element methods · Subspace correction methods · Interior Penalty methods · Iterative methods for non-symmetric problems

1 Introduction

The problem for efficient solution of the systems of equations arising from discontinuous (DG) discretizations has drawn a lot of attention, in the last years. Discontinuous Galerkin methods have many advantages over other types of finite element methods (e.g., flexibility in handling non-matching grids and in designing hp-refinement strategies, conservation properties and wide range of applicability). However, the fact that they result in a larger number of degrees of freedom as compared to a conforming method has been regarded as one main obstacle in their efficient implementation. Indeed, the computationally extensive part when using DG discretizations is the solution of the resulting ill-conditioned linear systems, which are of much larger size than the ones obtained via conforming method. As a

B. Ayuso de Dios
Departamento de Matemáticas, Universidad Autónoma de Madrid, Madrid 28049, Spain
e-mail: blanca.ayuso@uam.es

L. Zikatanov (✉)
Department of Mathematics, Penn State University, University Park, PA 16802, USA
e-mail: ludmil@psu.edu

result, efficient implementation of a DG discretization requires the use of advanced iterative methods for the solution of the resulting systems.

In the recent years, the efforts in the development of efficient iterative methods and solvers for the linear systems arising from DG discretizations of elliptic problems have been centered around the design of optimal multilevel solution methods, such as: domain decomposition and Schwarz methods (see [3–5, 27]); two-level and two grid techniques (see [23]); multigrid methods (see [16, 28]). Most of these works provide solution techniques and convergence analysis applicable to symmetric DG methods for second and fourth order elliptic problems. However, the attempts to construct and analyze preconditioners and iterative methods for the corresponding non-symmetric DG schemes encounters several barriers, which seem to be not easy to circumvent. Some of the difficulties originate from the fact that for the non-symmetric DG approximation of elliptic problems, the non-symmetric part of the bilinear form is not a compact, lower order perturbation of the symmetric part. For example, in [3] it was numerically demonstrated that commonly used sufficient conditions for the convergence of the Generalized Minimal Residual (GMRES) method [25] are not satisfied for the preconditioned system (with Schwarz method) arising from a non-symmetric DG discretization of an elliptic problem. Such negative results also show that the analysis of convergence of preconditioned GMRES for non-symmetric DG discretizations could be, to say the least, hard to do and non-standard to a great extend.

In this paper we propose and analyze uniformly convergent iterative methods and preconditioners for both symmetric and non-symmetric DG schemes. The construction of the preconditioners and the iterative algorithms is done using the framework of space decomposition and subspace correction methods (see [39, 40]). As underlying class of DG discretizations we use symmetric or non-symmetric Interior Penalty (IP) DG finite element methods (see [6, 7, 22, 32, 33, 37, 38], and [24]), in which the inter-element continuity is weakly enforced by penalizing discontinuities of the function or of its derivatives across inter-element boundaries. We restrict our considerations here to discretizations with piece-wise linear elements. In addition to the classical family of IP methods, which we call of Type-1: symmetric (SIPG), non-symmetric (NIPG) and incomplete (IIPG); we also consider the family of IP methods that result by evaluating the integrals in the bilinear forms numerically. These methods, which we call of Type-0, are similar to those introduced in [13, 14].

A key ingredient in our construction of uniform preconditioners is a decomposition of the underlying DG linear finite element space into two non-conforming subspaces. One of them is the well-known Crouzeix-Raviart finite element space, originally introduced by Strang in [35, 36, p. 178] and later successfully used for the Stokes problem [21]. Its complementary space contains functions that have non-zero jump at each internal face. In a different context, such decomposition has been proposed and used recently in [19] for obtaining a priori error bounds. We prove that for the Type-0 IP methods, this decomposition is orthogonal in an appropriate energy inner-product, thus leading to much simpler linear systems with block lower triangular structure. We further explore this orthogonal decomposition to design uniform preconditioners and uniformly convergent iterative methods for the classical Type-1 IP methods, as well as for the over-penalized IP methods recently introduced in [13] and [14]. In fact, we show that for symmetric interior penalty methods (SIPG, [6]), the Type-0 methods provide uniform preconditioners for the classical SIPG (Type-1) method. For non-symmetric NIPG and IIPG methods we propose an iterative method which uses as iterator the symmetric part of the underlying stiffness matrix. We prove that under a mild restriction on the penalty parameter such method is uniformly convergent in energy norm. To our knowledge, results on uniformly convergent iterative methods for the non-symmetric IIPG and NIPG were not available in the literature, and the convergence analysis presented here provides the first such results.

The rest of the paper is organized as follows. In Sect. 2 we introduce notation and present the IP methods of Type-0 and Type-1 that we consider. In Sect. 3 we introduce a decomposition of the DG finite element space and discuss its properties. A solver for the IP methods of Type-0 is introduced and analyzed in Sect. 4. In Sect. 5 we propose iterative methods and preconditioners for the IP methods of Type-1, and we present their analysis. Finally, in Sect. 6 we present some numerical examples that validate and confirm the presented theory. The paper is closed with an Appendix containing the proof of two auxiliary results.

2 Interior Penalty Discontinuous Galerkin Methods

In this section, we introduce the model problem and the basic notation. We also present the two families of Interior Penalty (IP) discontinuous Galerkin methods considered in this work, Type-1 and Type-0. The former type is the classical IP family [6, 22, 32]. The latter type (Type-0) is what results by computing numerically the integrals that appear in the classic IP family. In the last part of the section, we show that these two families of methods are spectrally equivalent.

2.1 Preliminaries

Let Ω be a bounded connected domain in $\mathbb{R}^d$ with Lipschitz boundary $\partial\Omega$ and let $f \in L^2(\Omega)$. For the sake of simplicity and easy presentation of the main ideas, we restrict ourselves to the model problem (see Remark 2.1):

$$-\Delta u = f \quad \text{in } \Omega, \qquad u = 0 \quad \text{on } \partial\Omega. \tag{2.1}$$

Remark 2.1 All the results presented here remain also valid for more general second order elliptic problems—in divergence form and with piecewise constant coefficient.

Throughout this paper, we use the standard notation for Sobolev spaces (see [1]). For a bounded domain $D \subset \mathbb{R}^d$, $d = 1, 2, 3$, we denote by $H^m(D)$ the standard Sobolev space of order $m \geq 0$, and by $\|\cdot\|_{m,D}$ and $|\cdot|_{m,D}$ the usual Sobolev norm and semi norm, respectively. We also use the notation $x_1 \lesssim y_1$ ($x_2 \gtrsim y_2$) whenever there exist constants C_1 and C_2 independent of the mesh size, but possibly dependent on the mesh regularity and such that $x_1 \leq C_1 y_1$ ($x_2 \geq C_2 y_2$). If two sided inequalities hold, then we write $x_1 \approx x_2$. In most places, when we use such notation, we will indicate the explicit dependence on the penalty parameter.

Domain Partitioning Let $\mathcal{T}_h$ be a shape-regular family of partitions of Ω into d-dimensional simplices T (triangles if $d = 2$ and tetrahedrons if $d = 3$). We denote by h_T the diameter of T and we set $h = \max_{T\in\mathcal{T}_h} h_T$. We also assume that $\mathcal{T}_h$ is conforming in the sense it does not contain hanging nodes. A face (shared by two neighboring elements or being part of the boundary) is denoted in general by E. Clearly, such face is a $(d-1)$ dimensional simplex, that is, an edge in two dimensions and a triangular face in three dimensions. We denote by $\mathcal{E}_h^o$ and $\mathcal{E}_h^\partial$ the collection of all interior faces and boundary faces, respectively. The set of all faces (the skeleton of the triangulation) is denoted by $\mathcal{E}_h$.

Trace Operators Following [7], we recall the definition of the *average* and *jump* trace operators for scalar and for vector-valued functions. Let T^+ and T^- be two neighboring elements, and $\mathbf{n}^+$, $\mathbf{n}^-$ be their outward normal unit vectors, respectively ($\mathbf{n}^\pm = \mathbf{n}_{T^\pm}$). Let $\zeta^\pm$ and $\boldsymbol{\tau}^\pm$ be the restriction of ζ and $\boldsymbol{\tau}$ to $T^\pm$. We set:

$$\begin{aligned} \{\zeta\} &= \frac{1}{2}(\zeta^+ + \zeta^-), \quad [\![\zeta]\!] = \zeta^+\mathbf{n}^+ + \zeta^-\mathbf{n}^- \quad \text{on } E \in \mathcal{E}_h^o, \\ \{\boldsymbol{\tau}\} &= \frac{1}{2}(\boldsymbol{\tau}^+ + \boldsymbol{\tau}^-), \quad [\![\boldsymbol{\tau}]\!] = \boldsymbol{\tau}^+\cdot\mathbf{n}^+ + \boldsymbol{\tau}^-\cdot\mathbf{n}^- \quad \text{on } E \in \mathcal{E}_h^o. \end{aligned} \tag{2.2}$$

For $E \in \mathcal{E}_h^\partial$, we set

$$[\![\zeta]\!] = \zeta\mathbf{n}, \quad \{\boldsymbol{\tau}\} = \boldsymbol{\tau} \quad \text{on } E \in \mathcal{E}_h^\partial. \tag{2.3}$$

We will also use the notation

$$(u,w)_{\mathcal{T}_h} = \sum_{T\in\mathcal{T}_h}\int_T uw dx \quad \forall u,w \in L^2(\Omega), \qquad \langle u,w\rangle_{\mathcal{E}_h} = \sum_{E\in\mathcal{E}_h}\int_E uw \quad \forall u,w, \in L^2(\mathcal{E}_h).$$

Finite Element Spaces We restrict our attention to piecewise linear approximation. Let V^{DG} denote the discontinuous finite element space defined by:

$$V^{DG} = \left\{u \in L^2(\Omega) : u_{|_T} \in \mathbb{P}^1(T)\ \forall T \in \mathcal{T}_h\right\}, \tag{2.4}$$

where $\mathbb{P}^1(T)$ denotes the space of linear polynomials on T.

2.2 Interior Penalty Methods

In what follows (unless it is explicitly specified), since we will only be concerned with discrete functions we denote also by u the DG finite element approximation to (2.1). All the methods we consider, can be recast in the following form: Find $u \in V^{DG}$ such that

$$\mathcal{A}^{DG}(u,w) = (f,w)_{\mathcal{T}_h}, \quad \forall w \in V^{DG}. \tag{2.5}$$

Here $\mathcal{A}^{DG}$ looks different for different IP methods. We focus on two types of $\mathcal{A}^{DG}(\cdot,\cdot)$– Type-0 IP methods and Type-1 IP methods. The Type-1 methods are the classical Interior Penalty methods. To each Type-1 method there corresponds a Type-0 method obtained by evaluating the penalty term in the bilinear form approximately.

1. For the IP methods of Type-1, we set $\mathcal{A}^{DG}(\cdot,\cdot) = \mathcal{A}(\cdot,\cdot)$ and define

$$\begin{aligned} \mathcal{A}(u,w) &= (\nabla u, \nabla w)_{\mathcal{T}_h} - \langle\{\nabla u\}, [\![w]\!]\rangle_{\mathcal{E}_h} + \theta\langle[\![u]\!], \{\nabla w\}\rangle_{\mathcal{E}_h} \\ &\quad + \langle S_E[\![u]\!], [\![w]\!]\rangle_{\mathcal{E}_h} \quad \forall u,w \in V^{DG}. \end{aligned} \tag{2.6}$$

2. For the IP methods of Type-0, we set $\mathcal{A}^{DG}(\cdot,\cdot) = \mathcal{A}_0(\cdot,\cdot)$ and define

$$\begin{aligned} \mathcal{A}_0(u,w) &= (\nabla u, \nabla w)_{\mathcal{T}_h} - \langle\{\nabla u\}, [\![w]\!]\rangle_{\mathcal{E}_h} + \theta\langle[\![u]\!], \{\nabla w\}\rangle_{\mathcal{E}_h} \\ &\quad + \langle S_E\mathcal{P}_E^0([\![u]\!]), \mathcal{P}_E^0([\![w]\!])\rangle_{\mathcal{E}_h} \quad \forall u,w \in V^{DG}, \end{aligned} \tag{2.7}$$

where $\mathcal{P}_E^0 : L^2(E) \to \mathbb{P}^0(E)$ is for each $E \in \mathcal{E}_h$, the L^2-orthogonal projection onto the constants, defined by:

$$\mathcal{P}_E^0(u) := \frac{1}{|E|} \int_E u \quad \forall u \in L^2(E). \tag{2.8}$$

Clearly, $\mathcal{P}_E^0$ satisfies

$$\|\mathcal{P}_E^0(u)\|_{0,E} \le \|u\|_{0,E} \quad \forall u \in L^2(E),\ \forall E \in \mathcal{E}_h. \tag{2.9}$$

Furthermore, notice that, denoting by m_E the mass center of E, numerical integration gives

$$\mathcal{P}_E^0(u) = u(m_E) \qquad \text{whenever } u \text{ is a linear polynomial on } E. \tag{2.10}$$

Therefore, in view of (2.10), the approximation of Type-0 defined in (2.7) can also be regarded as using the lowest order Gaussian quadrature rule (or the midpoint integration rule) for the evaluation of the integrals over $\mathcal{E}_h$ in the classical IP family methods. Since $u, w \in V^{DG}$ are piecewise linear polynomials, $\mathcal{A}$ and $\mathcal{A}_0$ differ only in the computation of the $\langle S_E[\![u]\!], [\![w]\!]\rangle_{\mathcal{E}_h}$ term (the other integrals over $\mathcal{E}_h$ are reproduced exactly).

For both Type-0 and Type-1 methods, $\theta = -1$ gives the symmetric (SIPG) (see [6]); $\theta = 1$ and $\theta = 0$ lead to the non-symmetric NIPG (see [32]) and IIPG (see [37]) schemes, respectively. The symmetric and skew-symmetric parts of $\mathcal{A}(\cdot,\cdot)$ are given by

$$\begin{aligned}
\widetilde{\mathcal{A}}(u, w) &= (\nabla u, \nabla w)_{\mathcal{T}_h} + \frac{\theta - 1}{2}\left[\langle\{\nabla u\}, [\![w]\!]\rangle_{\mathcal{E}_h} + \langle[\![u]\!], \{\nabla w\}\rangle_{\mathcal{E}_h}\right] \\
&\quad + \langle S_E[\![u]\!], [\![w]\!]\rangle_{\mathcal{E}_h} \quad \forall u, w \in V^{DG}, \\
\mathcal{S}(u, w) &= \frac{\theta + 1}{2}\left[\langle[\![u]\!], \{\nabla w\}\rangle_{\mathcal{E}_h} - \langle\{\nabla u\}, [\![w]\!]\rangle_{\mathcal{E}_h}\right] \quad \forall u, w \in V^{DG},
\end{aligned} \tag{2.11}$$

so that $\mathcal{A}(u, w) = \widetilde{\mathcal{A}}(u, w) + \mathcal{S}(u, w)$ and obviously for SIPG $\mathcal{S}(u, w) = 0\ \forall u, w \in V^{DG}$.

The symmetric and skew-symmetric parts of $\mathcal{A}_0(\cdot,\cdot)$ are defined in an analogous way. Notice that, while $\widetilde{\mathcal{A}}_0(\cdot,\cdot)$ differs from $\widetilde{\mathcal{A}}(\cdot,\cdot)$ in the last term, we always have that $\mathcal{S}_0(u, w) = \mathcal{S}(u, w)$ for all $u, w \in V^{DG}$ (the quadrature is exact for the integrals involved in the definition of the skew-symmetric part $\mathcal{S}(\cdot,\cdot)$).

In the definitions (2.6) and (2.7) we take

$$S_E = \alpha |E|^{\frac{-1}{d-1}} \approx \alpha h_E^{-1} \quad \forall E \in \mathcal{E}_h, \quad \alpha > 0, \tag{2.12}$$

where h_E denotes the length of the edge E in $d = 2$ and the diameter of the face in $d = 3$ and α is the *penalty parameter* on each face $E \in \mathcal{E}_h$.

For all methods of Type-1 (2.6), and for sufficiently large value of the penalty parameter α, *continuity* and *coercivity* can be shown in the following energy norm (see [7] for details),

$$|\!|\!| u |\!|\!|^2 := \sum_{T \in \mathcal{T}_h} \|\nabla u\|_{0,T}^2 + \sum_{E \in \mathcal{E}_h} h_E^{-1} \|[\![u]\!]\|_{0,E}^2 \quad \text{where } \forall u \in V^{DG}. \tag{2.13}$$

For Type-0 methods (2.7), similar properties can also be easily shown to hold, following standard arguments (see [7]), albeit in a different energy norm, defined as:

$$|\!|\!| u |\!|\!|_{\mathcal{N}}^2 := \sum_{T \in \mathcal{T}_h} \|\nabla u\|_{0,T}^2 + \sum_{E \in \mathcal{E}_h} h_E^{-1} \|\mathcal{P}_E^0([\![u]\!])\|_{0,T}^2. \tag{2.14}$$

In summary, the following estimates hold, with constants depending on the shape regularity of the mesh, the domain Ω and α:

$$\mathcal{A}(u,w) \lesssim |||u|||\,|||w|||, \quad \mathcal{A}_0(u,w) \lesssim |||u|||_{\mathcal{N}}\,|||w|||_{\mathcal{N}}, \quad \forall u, w \in V^{DG}, \tag{2.15}$$

$$\mathcal{A}(u,u) \gtrsim |||u|||^2, \quad \mathcal{A}_0(u,u) \gtrsim |||u|||^2_{\mathcal{N}}, \quad \forall u \in V^{DG}. \tag{2.16}$$

On the Choice of Penalty Parameter Since the penalty parameter α enters many of the estimates that we prove or use in our proofs, we would like to comment some more on the choice of its value. First, to simplify the notation, we take the same value of the penalty parameter for all edges/faces $E \in \mathcal{E}_h$ in most of the considerations that follow. Thus, the value of the penalty parameter α will be independent of the mesh size. Only two exceptions to this convention are found in Remarks 2.2 and 4.5, when we comment on the application of our results to the weakly over-penalized IP methods introduced in [13] and [14]. The penalty parameter in these methods depends on the mesh size and we specify this dependence and outline the changes then.

We always set $\alpha \geq \alpha^*$ where α^* is a *fixed* value of the penalty parameter that gives positive definite symmetric part $\widetilde{\mathcal{A}}^{DG}(\cdot,\cdot)$ of the bilinear form $\mathcal{A}^{DG}(\cdot,\cdot)$; i.e., ensures coercivity of $\mathcal{A}^{DG}(\cdot,\cdot)$. The results that we prove below hold for *any* value of α^* such that $\widetilde{\mathcal{A}}^{DG}(\cdot,\cdot)$ is positive definite. To avoid confusion, we point out here, that we have in mind the non-trivial cases when α^* is "small". In another word, α^* is close to (but strictly larger than) the value of the penalty parameter for which the symmetric part of the method in hand $\widetilde{\mathcal{A}}^{DG}(\cdot,\cdot)$ is *only* positive *semi*-definite. One may certainly take different values of α^* for Type-0 methods and Type-1 methods. However, such distinction is not essential at all (as Lemma 2.3 shows), and we will refer to one and the same minimal value α^* for both Type-0 and Type-1 methods.

In accordance with the discussion above and for easier reference later on we shall call α^*, *the minimal value of the penalty parameter for which* $\mathcal{A}^{DG}(\cdot,\cdot)$ *is coercive*.

It follows then that, for $\alpha \geq \alpha^*$ the bilinear forms corresponding to the symmetric parts of $\mathcal{A}(\cdot,\cdot)$ and $\mathcal{A}_0(\cdot,\cdot)$ induce norms equivalent to $|||\cdot|||$ and $|||\cdot|||_{\mathcal{N}}$,

$$\|u\|^2_{\widetilde{\mathcal{A}}} := \widetilde{\mathcal{A}}(u,u) \approx |||u|||^2, \qquad \|u\|^2_{\widetilde{\mathcal{A}}_0} := \widetilde{\mathcal{A}}_0(u,u) \approx |||u|||^2_{\mathcal{N}}.$$

Finally, whenever we will need to distinguish between the different IP methods, we shall use superscripts to denote the bilinear forms corresponding to SIPG, NIPG and IIPG Type-1 methods, i.e., $\mathcal{A}^{s,\alpha}(\cdot,\cdot)$ for SIPG, $\mathcal{A}^{n,\alpha}(\cdot,\cdot)$ for NIPG, and $\mathcal{A}^{I,\alpha}(\cdot,\cdot)$ for IIPG. The superscript α indicates the actual value of the penalty parameter taken in the definition of the method.

Weighted Residual Form In several proofs, we will also use an equivalent form of the DG methods obtained via the *weighted residual* approach introduced in [17] (see also [9]). Using this approach the Type-0 and Type-1 methods are as follows:

$$\mathcal{A}(u,w) = (-\Delta u, w)_{\mathcal{T}_h} + \langle [\![\nabla u]\!], \{w\}\rangle_{\mathcal{E}_h^o} + \langle [\![u]\!], \mathcal{B}_1(w)\rangle_{\mathcal{E}_h} \quad \forall u, w \in V^{DG}, \tag{2.17}$$

$$\mathcal{A}_0(u,w) = (-\Delta u, w)_{\mathcal{T}_h} + \langle [\![\nabla u]\!], \{w\}\rangle_{\mathcal{E}_h^o} + \langle [\![u]\!], \mathcal{P}^0_E(\mathcal{B}_1(w))\rangle_{\mathcal{E}_h} \quad \forall u, w \in V^{DG}. \tag{2.18}$$

Here $\mathcal{B}_1(w)$ is defined as

$$\mathcal{B}_1(w) = \theta\{\nabla w\} + \alpha_E |E|^{\frac{-1}{(d-1)}} [\![w]\!] \quad \forall E \in \mathcal{E}_h. \tag{2.19}$$

Throughout the paper, both the weighted residual and classical forms will be used interchangeably.

Remark 2.2 By allowing the penalty parameter to depend on h_E (and so to vary from one face to another, i.e.; $\alpha \mapsto \alpha_E$) and by setting $\alpha_E \approx h_E^{-2}$, we recover the methods proposed in [13] and [14] as a particular case of the Type-0 methods. Most of the results that we present for the Type-0 methods, with the exception of Lemma 2.3, apply almost verbatim to the weakly over penalized non-symmetric (WOPNIP) and symmetric (WOPSIP) interior penalty methods. One difference, that we need to note here, is that the systems of equations arising from WOPNIP and WOPSIP discretizations have condition number of order $O(h^{-4})$ (in contrast to condition numbers of order h^{-2} for the standard Type-0 IP methods). Nevertheless, the solution techniques that we propose in section 4 can be applied in a straightforward way to solve such systems.

2.3 Equivalence Relations between Type-0 and Type-1 Interior Penalty Methods

Next result shows that the symmetric parts of the bilinear forms associated with Type-0 and Type-1 methods are equivalent.

Lemma 2.3 *Let $\mathcal{A}(\cdot,\cdot)$ be a bilinear form corresponding to a Type-1 method and let $\mathcal{A}_0$ be the corresponding Type-0 bilinear form. Then, there exist a positive constant $c_0 = c_0(\alpha)$, depending only on the shape regularity of the mesh and the penalty parameter α, such that,*

$$\mathcal{A}_0(u,u) \le \mathcal{A}(u,u) \le c_0(\alpha)\mathcal{A}_0(u,u) \quad \forall u \in V^{DG}. \tag{2.20}$$

Proof The lower bound follows immediately from the fact that the projection operator $\mathcal{P}_E^0$ is bounded in $L^2(E)$ and has norm 1 (see (2.9)). We only need to show the upper bound. Comparing the terms in the definitions of $\mathcal{A}_0$ and $\mathcal{A}$ it is easy to see that to get the upper bound in (2.20) it will be enough to show that

$$\sum_E S_E \|[\![u]\!]\|_{0,E}^2 \le C\left(\|\nabla v\|_{0,\mathcal{T}_h}^2 + \sum_E S_E \|\mathcal{P}_E^0 [\![u]\!]\|_{0,E}^2 \right).$$

Adding and subtracting $\mathcal{P}_E^0[\![u]\!]$ in the term on the left side above and using that $\mathcal{P}_E^0$ is $L_2(E)$-orthogonal projection we have for each face E,

$$\|[\![u]\!]\|_{0,E}^2 = \|\mathcal{P}_E^0([\![u]\!])\|_{0,E}^2 + \|[\![u]\!] - \mathcal{P}_E^0([\![u]\!])\|_{0,E}^2. \tag{2.21}$$

Hence, we only need to estimate the last term on the right side of the above inequality. Observe, that on each $E \in \mathcal{E}$, $E = \partial T^+ \cap \partial T^-$,

$$\|[\![u]\!] - \mathcal{P}_E^0([\![u]\!])\|_{0,E}^2 \lesssim \|u^+ - \mathcal{P}_E^0(u^+)\|_{0,E}^2 + \|u^- - \mathcal{P}_E^0(u^-)\|_{0,E}^2.$$

We will estimate the first term on the right side. The second one is estimated in an analogous way. By a standard scaling argument, mapping T^+ to the reference simplex $\widehat{T}$, and denoting the function on $\widehat{T}$ corresponding to u^+ with $\hat{u}$, we have

$$h_E^{-1}\|u^+ - \mathcal{P}_E^0(u^+)\|_{0,E}^2 \lesssim h^{d-1} h_E^{-1} \|[\hat{u} - \mathcal{P}_{\widehat{E}}^0(\hat{u})]\|_{0,\widehat{E}}^2.$$

We fix $\hat{u}$ and define a linear functional $F : H^1(\widehat{T}) \mapsto \mathbb{R}$ by

$$F(\hat{v}) := \int_{\widehat{E}} [\hat{u} - \mathcal{P}^0_{\widehat{E}}(\hat{u})][\hat{v} - \mathcal{P}^0_{\widehat{E}}(\hat{v})].$$

By Cauchy-Schwarz inequality and the Trace Theorem on $\widehat{T}$,

$$|F(\hat{v})| \lesssim \|\hat{u} - \mathcal{P}^0_{\widehat{E}}(\hat{u})\|_{0,\widehat{E}} \|\hat{v}\|_{1,\widehat{T}},$$

and hence $F(\hat{v})$ is bounded. Moreover $F(\hat{v}) = 0$, $\forall\ v \in \mathbb{P}^0(\widehat{T})$. By Bramble-Hilbert Lemma it follows that we can replace the full norm $\|v\|_{1,\widehat{T}}$ on the right side with a semi-norm to get that

$$|F(\hat{v})| \lesssim \|\hat{u} - \mathcal{P}^0_{\widehat{E}}(\hat{u})\|_{0,\widehat{E}} |\hat{v}|_{1,\widehat{T}}.$$

Taking $\hat{v} = \hat{u}$ and using the definition of $F(\cdot)$ then leads to

$$|F(\hat{u})| = \|\hat{u} - \mathcal{P}^0_{\widehat{E}}(\hat{u})\|^2_{0,\widehat{E}} \lesssim \|\hat{u} - \mathcal{P}^0_{\widehat{E}}(\hat{u})\|_{0,\widehat{E}} |\hat{u}|_{1,\widehat{T}}.$$

Using this inequality and mapping back to T^+ then gives

$$\begin{aligned} h_E^{-1} \|u^+ - \mathcal{P}^0_E(u^+)\|^2_{0,E} &\lesssim h^{d-1} h_E^{-1} \|\hat{u} - \mathcal{P}^0_{\widehat{E}}(\hat{u})\|^2_{0,\widehat{E}} \\ &\lesssim h^{d-1} h_E^{-1} |\hat{u}|^2_{1,\widehat{T}} \lesssim |u^+|^2_{1,T^+}. \end{aligned}$$

Hence, substituting into (2.21) and summing over all the faces, we finally have

$$\begin{aligned} \sum_E \alpha h_E^{-1} \|[\![u]\!]\|^2_{0,E} &\le \sum_E \alpha h_E^{-1} \|\mathcal{P}^0_E([\![u]\!])\|^2_{0,E} + C(\alpha) \sum_{T \in \mathcal{T}_h} \|\nabla u\|^2_{0,T} \\ &\le c_0(\alpha) \mathcal{A}_0(u,u) \quad \forall u \in V^{DG} \end{aligned}$$

and the proof is complete. □

We now focus on the description of the decomposition of the DG space into subspaces, needed later in the construction of iterative methods.

3 A Subspace Decomposition of V^{DG}

In this section we show a subspace decomposition for the V^{DG} space. In a different context, such decomposition has been also obtained recently in [19]. Although there are slight differences in the allocation of the boundary degrees of freedom, the main idea on how to decompose the DG finite element space is essentially the same. Since it is essential to our subsequent analysis, we have included the derivation of the space decomposition in Proposition 3.1.

We start by reviewing first the definition of the well known Crouzeix-Raviart nonconforming finite element space V^{CR} [21] (see also [15, 18]) and also introduce another subspace of V^{DG}, denoted here by $\mathcal{Z}$. As we shall see later, these form a non-overlapping decomposition of V^{DG}. These spaces are defined via the local (on each face) L^2-orthogonal projections onto the constant function, $\mathcal{P}_E^0$ (see (2.8)). The Crouzeix-Raviart finite element space is

$$V^{CR} = \left\{ v \in L^2(\Omega) : v_{|_T} \in \mathbb{P}^1(T)\ \forall T \in \mathcal{T}_h \text{ and } \mathcal{P}_E^0 [\![v]\!] = 0\ \forall E \in \mathcal{E}_h^o \right\}. \tag{3.1}$$

Note that $v(m_E) = 0$ at all boundary faces $E \in \partial\Omega$. We also define the space

$$\mathcal{Z} = \left\{ z \in L^2(\Omega) : z_{|_T} \in \mathbb{P}^1(T)\ \forall T \in \mathcal{T}_h \text{ and } \mathcal{P}_E^0 \{v\} = 0\ \forall E \in \mathcal{E}_h^o \right\}. \tag{3.2}$$

Observe that any function from the space $\mathcal{Z}$ has the property that if it is nonzero on an internal face, then it also has a nonzero jump on this face. In another word, these functions are highly oscillatory.

We denote by n_T, n_E and n_V the number of elements, faces (edges in $d = 2$) and vertices of the partition and by n_{BE} the number of boundary faces. Since the restriction of a function in V^{DG} on every element $T \in \mathcal{T}_h$ is a linear polynomial, its representation on T depends on the basis chosen for the linear polynomials on a simplex. A canonical choice will be a basis dual to degrees of freedom located at the vertices of the simplex. For our analysis, more convenient choice is a basis dual to degrees of freedom located at the mass centers of the faces of T. Let $\varphi_{E,T}$ denote the canonical Crouzeix-Raviart (CR) basis function on T, dual to the degree of freedom at the mass center of the face E, and extended as zero outside T. Hence, $\varphi_{E,T}$ satisfies

$$\varphi_{E,T} \in \mathbb{P}^1(T) : \varphi_{E,T}(x) = 0 \quad \forall x \notin T, \quad \varphi_{E,T}(m_{E'}) = \begin{cases} 1 & \text{if } E = E', \\ 0 & \text{otherwise.} \end{cases}$$

For all $u \in V^{DG}$ we then have

$$u(x) = \sum_{T \in \mathcal{T}_h} \sum_{E \in \partial T} u_T(m_E) \varphi_{E,T}(x) = \sum_{E \in \mathcal{E}_h} u^+(m_E) \varphi_E^+(x) + \sum_{E \in \mathcal{E}_h^o} u^-(x_E) \varphi_E^-(x), \tag{3.3}$$

where in the last identity we have just changed the order of summation and used the short hand notation $\varphi_E^{\pm}(x) := \varphi_{E,T^{\pm}}(x)$ together with

$$u^{\pm}(m_E) := u_{T^{\pm}}(m_E) = \frac{1}{|E|} \int_E u^{\pm} ds \quad \forall E \in \mathcal{E}_h^o, : E = \partial T^+ \cap \partial T^-,$$

$$u(m_E) := u_T(m_E) = \frac{1}{|E|} \int_E u_T ds \quad \forall E \in \mathcal{E}_h^\partial, : E = \partial T \cap \partial \Omega.$$

Clearly, $(d+1)n_T = 2n_E - n_{BE}$ and hence, a basis for the DG space is given by the set of functions $\{\varphi_E^+, \varphi_E^-\}_{E \in \mathcal{E}_h^o} \cup \{\varphi_E\}_{E \in \mathcal{E}_h^\partial}$. We now show that this considerations lead naturally to a splitting of V^{DG} as a direct sum of the spaces V^{CR} and $\mathcal{Z}$ defined in (3.1) and (3.2), respectively.

Proposition 3.1 *For any $u \in V^{DG}$ there exists a unique $v \in V^{CR}$ and a unique $z \in \mathcal{Z}$ such that $u = v + z$, that is*

$$V^{DG} = V^{CR} \oplus \mathcal{Z}. \tag{3.4}$$

Proof From the definition of the jump and average operators (2.2), it follows that $\forall u, w \in V^{DG}$ and $\forall x, y \in E \in \mathcal{E}_h^o$,

$$
\begin{aligned}
u^+(y)w^+(x) + u^-(y)w^-(x) &= 2\{u(y)\, w(x)\} \\
&= 2\{u\}(y)\{w\}(x) + \frac{1}{2}[\![u]\!](y)[\![w]\!](x) \\
&= 2\frac{(u^+(y)+u^-(y))}{2}\frac{(w^+(x)+w^-(x))}{2} + \frac{(u^+(y)-u^-(y))}{2}(w^+(x)-w^-(x)).
\end{aligned}
$$

Then, using the above expression and taking into account the definition of the jump operator on boundary faces (2.3) we can re-arrange the representation (3.3) of $u \in V^{DG}$. More precisely, we find that any $u \in V^{DG}$ can be expressed as

$$
\begin{aligned}
u(x) &= \sum_{E\in\mathcal{E}_h^o} u^+(m_E)\varphi_E^+(x) + \sum_{E\in\mathcal{E}_h^o} u^-(m_E)\varphi_E^-(x) + \sum_{E\in\mathcal{E}_h^\partial} u(m_E)\varphi_E(x) = \sum_{E\in\mathcal{E}_h^\partial} u(m_E)\varphi_E(x) \\
&\quad + \sum_{E\in\mathcal{E}_h^o} \frac{(u^+(m_E)+u^-(m_E))}{2}(\varphi_E^+(x)+\varphi_E^-(x)) \\
&\quad + \sum_{E\in\mathcal{E}_h^o} \frac{(u^+(m_E)-u^-(m_E))}{2}(\varphi_E^+(x)-\varphi_E^-(x)) \\
&= \sum_{E\in\mathcal{E}_h^o}\left(\frac{1}{|E|}\int_E \{u\}ds\right)(\varphi_E^+(x)+\varphi_E^-(x)) \\
&\quad + \sum_{E\in\mathcal{E}_h^o}\left(\frac{1}{|E|}\int_E [\![u]\!]\frac{\mathbf{n}^+}{2}ds\right)(\varphi_E^+(x)-\varphi_E^-(x)) + \sum_{E\in\mathcal{E}_h^\partial}\left(\frac{1}{|E|}\int_E u ds\right)\mathbf{n}[\![\varphi_E]\!](x) \\
&= v^{CR}(x) + z(x).
\end{aligned}
$$

By defining now

$$
\varphi_E^{CR} = \varphi_{E,T^+} + \varphi_{E,T^-} = 2\{\varphi_{E,T^\pm}\} \quad \forall E \in \mathcal{E}_h^o,\ E = T^+\cap T^-, \tag{3.5}
$$

and

$$
\begin{cases}
\psi^z_{E,T^+} = \varphi_{E,T^+} - \varphi_{E,T^-} = [\![\varphi_{E,T^\pm}]\!]\mathbf{n}^+ & \forall E \in \mathcal{E}_h^o, E = T^+\cap T^-, \\
\psi^z_{E,T^-} = \varphi_{E,T^-} - \varphi_{E,T^+} = [\![\varphi_{E,T^\pm}]\!]\mathbf{n}^- & \forall E \in \mathcal{E}_h^o, E = T^+\cap T^-, \\
\psi^z_{E,T} = \varphi_{E,T} = [\![\varphi_{E,T}]\!]\mathbf{n} \quad \forall E \in \mathcal{E}_h^\partial, E = T\cap \partial\Omega, &
\end{cases} \tag{3.6}
$$

we have two set of basis functions; $\{\varphi_E^{CR}\}_{E\in\mathcal{E}_h^o}$ which is just the set of canonical Crouzeix-Raviart basis functions for non-conforming linear finite elements and so, it is continuous at m_E for all E; and, $\{\psi^z_{E,T}\}_{E\in\mathcal{E}_h^o} \cup \{\psi^z_{E,T}\}_{E\in\mathcal{E}_h^\partial}$ which are discontinuous across each $E \in \mathcal{E}_h^o$. Therefore, for all $u \in V^{DG}$ there exist

$$
\begin{aligned}
v^{CR} &= \sum_{E\in\mathcal{E}_h^o}\left(\frac{1}{|E|}\int_E \{u\}ds\right)\varphi_E^{CR}(x) \in V^{CR}, \\
z &= \sum_{E\in\mathcal{E}_h}\left(\frac{1}{|E|}\int_E [\![u]\!]\frac{\mathbf{n}^+}{2}ds\right)\psi^z_{E,T^+}(x) \in \mathcal{Z},
\end{aligned}
$$

such that $u = v^{CR} + z$ and so (3.4) is shown and the proof is complete. □

Note that from the proof one can easily see that the set

$$\{\psi^{z}_{E,T}\}_{E\in\mathcal{E}_h} \cup \{\varphi^{CR}_{E,T}\}_{E\in\mathcal{E}_h^o}, \tag{3.7}$$

provides a natural basis for the DG finite element space V^{DG}.

Next result is simple but crucial for the subsequent analysis and shows that the splitting of Proposition 3.1 is orthogonal in the energy norm defined by $\mathcal{A}_0(\cdot,\cdot)$.

Lemma 3.2 *Let $u \in V^{DG}$ be such that $u = v + z$ with $v \in V^{CR}$ and $z \in \mathcal{Z}$. Let $\mathcal{A}_0(\cdot,\cdot)$ be the bilinear form defined in (2.7). Then,*

$$\mathcal{A}_0(v,z) = 0 \quad \forall v \in V^{CR},\ \forall z \in \mathcal{Z}. \tag{3.8}$$

Furthermore if $\mathcal{A}_0(\cdot,\cdot)$ is symmetric, then $\mathcal{A}_0(v,z) = \mathcal{A}_0(z,v) = 0$ $\forall v \in V^{CR}$, $\forall z \in \mathcal{Z}$ and the decomposition (3.4) is $\mathcal{A}_0$-orthogonal; $V^{CR} \perp_{\mathcal{A}_0} \mathcal{Z}$.

Proof Here, we use the equivalent weighted residual definition of $\mathcal{A}_0(\cdot,\cdot)$ given in (2.18). We have

$$\begin{aligned}\mathcal{A}_0(v,z) &= (-\Delta v, z)_{\mathcal{T}_h} + \langle [\![\nabla v]\!], \{z\}\rangle_{\mathcal{E}_h^o} + \langle [\![v]\!], \mathcal{P}^0_E(\mathcal{B}_1(z))\rangle_{\mathcal{E}_h} \\ &= 0 \quad \forall v \in V^{CR},\ \forall z \in \mathcal{Z},\end{aligned}$$

where the first term is zero since v is linear in each T; the second vanishes from the definition of the space $\mathcal{Z}$ and last term also vanishes independently of the choice of θ (and hence $\mathcal{B}_1(v)$), from the definition of the CR space. Moreover if $\mathcal{A}_0(\cdot,\cdot)$ is symmetric, then $\mathcal{A}_0(v,z) = \mathcal{A}_0(z,v)$ $\forall v \in V^{CR}$, $\forall z \in \mathcal{Z}$ and so $V^{CR} \perp_{\mathcal{A}_0} \mathcal{Z}$ and the proof is complete. □

Comments on the Boundary Conditions We note here that unlike in [19], the degrees of freedom corresponding to the boundary faces correspond to functions in $\mathcal{Z}$. This choice may now seem somewhat arbitrary, since the basis functions in $\mathcal{Z}$ corresponding to boundary faces could equally be considered as elements of a Crouzeix-Raviart finite element space. However, there are several reasons for this selection. First of all, in the DG solution $u \in V^{DG}$ of (2.5), the boundary conditions are weakly imposed. Therefore it seems natural to seek for a decomposition $u = z + v$ in which this is reflected appropriately. Noting that for the approximation of (2.1) with Crouzeix-Raviart finite elements, the Dirichlet boundary conditions are imposed strongly, within the space, it follows that for $v \in V^{CR}$ there are no degrees of freedom (v is already prescribed there) on the boundary. Thus, the way to weakly impose the boundary conditions is to have the boundary degrees of freedom of u in its component in the $\mathcal{Z}$ space. Yet another reason is that (as we shall see later in Sect. 4), this choice together with Lemma 3.2 leads to space decomposition with nice properties. For example, for Type-0 methods the stiffness matrix is block diagonal (or block lower triangular). In addition, again for Type-0 methods the restriction of the bilinear form $\mathcal{A}_0(\cdot,\cdot)$ to the Crouzeix-Raviart space is exactly the bilinear form corresponding to the Dirichlet problem on V^{CR} (see e.g. [15]).

The case of Neumann conditions on a part or the whole of $\partial\Omega$ is different and requires a more subtle approach and we will not consider it here. We would like to point out, however, that the extension to the case of Neumann boundary conditions and systems of partial differential equations, such as linear elasticity, of the preconditioning techniques and the iterative algorithms that we analyze here, is also possible (see [8] for details).

In the following two sections, we shall show how this decomposition can be used for the design and analysis of uniformly convergent DG solvers.

4 Solvers for the Type-0 Interior Penalty Methods

In this section, we show how the space decomposition (3.4) can be used for the construction of solvers for the solution of (2.5). We consider first the algebraic formulation of the DG methods when represented in the new basis (3.7) and propose an exact solver. Then, we study the main features related with the proposed algorithm; in particular the solution steps in the V^{CR} and $\mathcal{Z}$ spaces.

4.1 Matrix Notation and the Block form of the Stiffness Matrix

We consider the discretized problem (2.5) with $\mathcal{A}^{DG}(\cdot,\cdot) = \mathcal{A}_0(\cdot,\cdot)$. In view of the splitting (3.4), let $\forall u, w \in V^{DG}$ be such that $u = z + v,\ w = \psi + \varphi$ with $z, \psi \in \mathcal{Z},\ v, \varphi \in V^{CR}$. Then,

$$\mathcal{A}_0(u,w) = \mathcal{A}_0((z,v);(\psi,\varphi)) = \mathcal{A}_0(z+v,\psi+\varphi) = \mathcal{A}_0(z,\psi) + \mathcal{A}_0(z,\varphi) + \mathcal{A}_0(v,\varphi), \tag{4.1}$$

where in the last identity we have also used that the term $\mathcal{A}_0(v,\psi)$ vanishes by virtue of Lemma 3.2. Furthermore, for the SIPG method $\mathcal{A}_0(z,\varphi)$ is also zero and therefore, we have as a direct consequence of Proposition 3.1 and Lemma 3.2 that the decomposition (3.4) is stable (with constant 1) in the energy norm defined by the Type-0 SIPG method; i.e.,

$$\mathcal{A}_0^s(u,u) = \mathcal{A}_0^s(z,z) + \mathcal{A}_0^s(v,v) \quad \forall u \in V^{DG},\ u = z + v.$$

Next, we denote by A_0 the discrete operators defined by $\mathcal{A}_0(u,w) = (A_0 u, w)$. Let $\mathbb{A}_0$ be the matrix representation of the operator A_0 in the new basis (3.7) and let $\mathbf{u} = [\mathbf{z}, \mathbf{v}]^T$, $\mathbf{f} = [\mathbf{f}_z, \mathbf{f}_v]^T$ be the vector representation of the unknown function u and of the right hand side f, respectively, in this new basis. From (4.1) it follows immediately that $\mathbb{A}_0$ has the following block structure:

$$\mathbb{A}_0 = \begin{bmatrix} \mathbb{A}_0^{zz} & \mathbf{0} \\ \mathbb{A}_0^{vz} & \mathbb{A}_0^{vv} \end{bmatrix}, \tag{4.2}$$

where we have denoted by $\mathbb{A}_0^{zz}$, $\mathbb{A}_0^{vv}$ the matrix representation of the operator A_0 restricted to the subspaces $\mathcal{Z}$ and V^{CR}, respectively, and by $\mathbb{A}_0^{vz}$ the matrix representation of the term that accounts for the coupling (or non-symmetry) $\mathcal{A}_0(\psi^z, \varphi^{CR})$. Therefore, the Type-0 methods lead in all cases ($\theta = 0, \pm 1$) to operators that admit block lower triangular matrix representation. Furthermore, the stiffness matrix $\mathbb{A}_0$ associated with Type-0 SIPG method turns out to be *block-diagonal* (see Remark 4.1). To confirm these observations, we have represented in Fig. 2, the sparsity patterns (non-zero patterns) of $\mathbb{A}_0$ for SIPG, NIPG and IIPG methods, respectively. Notice that for the IIPG method the $\mathbb{A}_0^{zz}$-block is diagonal, as we shall prove in Lemma 4.7 at the end of this section.

An important feature of this splitting is that matrix representations of the bilinear forms in the basis (3.7) are much sparser. The number of non-zeros in the stiffness matrix in this basis is approximately half the number of non-zeros in the matrix representation in the standard nodal basis (see Figs. 1 and 2).

Remark 4.1 As we have already mentioned, the fact that the decomposition is done in a way that the boundary degrees of freedom of $u \in V^{DG}$ correspond to the functions in the space $\mathcal{Z}$ (see the comments after the proof of Proposition 3.1) made possible this reduction to block lower triangular (or even block diagonal for SIPG) structure of the matrices $\mathbb{A}_0$.

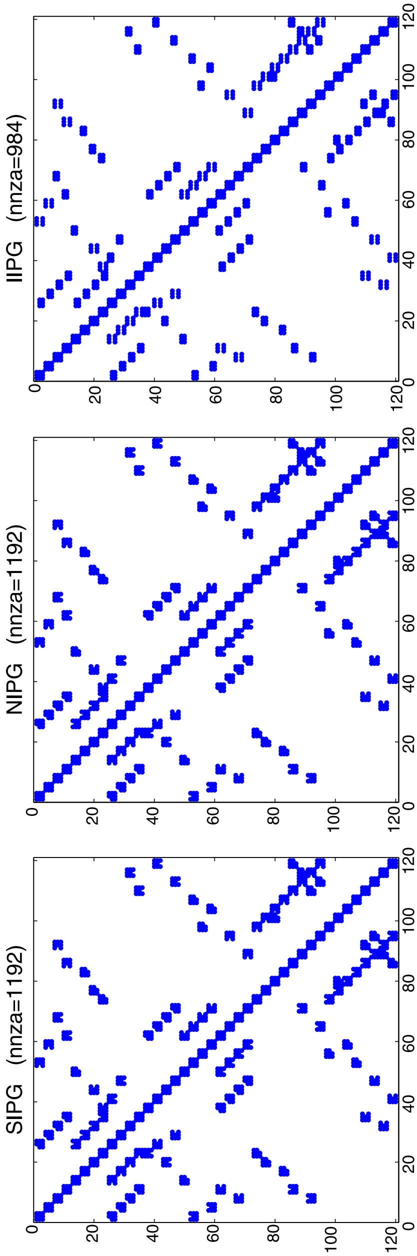

Fig. 1 Non-zero pattern of the matrix representation in the standard nodal basis of the operators associated with Type-0 IP methods. From *left* to *right*: SIPG, NIPG and IIPG methods

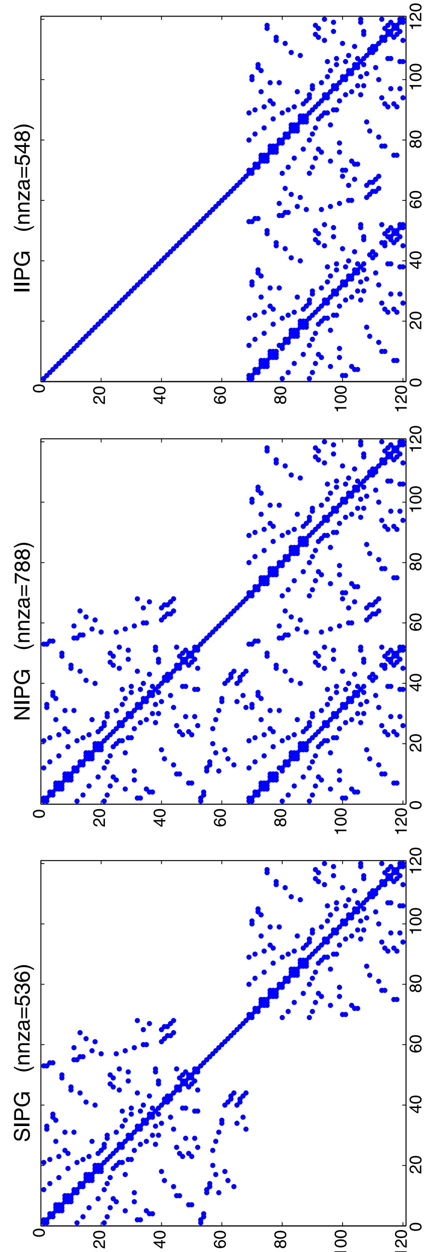

Fig. 2 Non-zero pattern of the matrix representations in the basis (3.7) of the operators associated with Type-0 IP methods; i.e., $\mathbb{A}_0$. From *left* to *right*: SIPG, NIPG and IIPG methods

Exact Solver Since we have 2×2 block lower triangular form of the stiffness matrix, a simple block forward substitution will give the solution of (2.5). Denote $F(w)=(f,w)_{\mathcal{T}_h}$ $\forall w\in V^{DG}$, then the algorithm is as follows:

Algorithm 4.2

1. *Find* $z\in\mathcal{Z}$ *such that* $\mathcal{A}_0(z,\psi)=F(\psi)$ *for all* $\psi\in\mathcal{Z}$.
2. *Find* $v\in V^{CR}$ *such that* $\mathcal{A}_0(v,\varphi)=F(\varphi)-\mathcal{A}_0(z,\varphi)$ *for all* $\varphi\in V^{CR}$.
3. *Set* $u=z+v$.

A simple argument shows that the u defined via the above algorithm solves (2.5) of Type-0 (in both symmetric or non-symmetric case). Indeed, Proposition 3.1 and Lemma 3.2 guarantee that for any $\eta\in V^{\rm DG}$ there exist unique $\varphi\in V^{CR}$ and $\psi\in\mathcal{Z}$ such that $\eta=\varphi+\psi$ and $\mathcal{A}(\varphi,\psi)=0$. Hence:

$$\begin{aligned}
F(\eta)&=F(\varphi)+F(\psi) && \text{[by linearity of } F]\\
&=\mathcal{A}_0(z,\varphi)+\mathcal{A}_0(v,\varphi)+F(\psi) && \text{[from step 2 of algorithm]}\\
&=\mathcal{A}_0(u,\varphi)+\mathcal{A}_0(z,\psi) && \text{[because } u=v+z]\\
&=\mathcal{A}_0(u,\varphi)+\mathcal{A}_0(v+z,\psi)=\mathcal{A}_0(u,\eta). && \text{[from (3.8)]}
\end{aligned}$$

Since $\eta\in V^{\rm DG}$ was arbitrary, this proves that $u=v+z$ is a solution to (2.5).

In matrix notation this algorithm is:

1. Solve for $\mathbf{u_z}$: $\mathbb{A}_0^{zz}\mathbf{u_z}=\mathbf{f_z}$.
2. Solve for $\mathbf{u_v}$: $\mathbb{A}_0^{vv}\mathbf{u_v}=\mathbf{f_v}-\mathbb{A}_0^{vz}\mathbf{u_z}$.
3. Set $\mathbf{u}=[\mathbf{u_z},\mathbf{u_v}]^T$.

What we discuss next are efficient algorithms for the solution of problems with $\mathbb{A}_0^{vv}$ and $\mathbb{A}_0^{zz}$ (step 1 and step 2 of the Algorithm 4.2). This amounts to the solution of a problem in V^{CR} and the solution of a problem in $\mathcal{Z}$ and we consider the relevant techniques next.

4.2 Solution in V^{CR}

As we mentioned already, restricting any of the Type-0 IP methods to the V^{CR} space gives the standard non-conforming CR finite element method for the solution of the model problem (2.1) (see [15]).

$$\mathcal{A}_0(v,\varphi)=(\nabla v,\nabla\varphi)_{\mathcal{T}_h}=\sum_{T\in\mathcal{T}_h}(\nabla v,\nabla\varphi)_T\quad\forall v,\varphi\in V^{CR}. \tag{4.3}$$

There are many works devoted to the optimal solution and preconditioning for the algebraic system resulting from discretizations of second order elliptic problems with Crouzeix-Raviart elements. We will not delve into the details about such algorithms here, since the results related to their convergence are well known. For instance, the Schwarz method proposed and analyzed in [10] or the Multigrid algorithms studied in [11, 12] and [30] are all proved to be uniformly convergent (with respect to the mesh size) for the problems corresponding to the bilinear form given in (4.3). Thus, the solution in the step 2 of the Algorithm 4.2 can be obtained by these techniques in optimal computational time.

4.3 Solution in $\mathcal{Z}$

In this section we study in more detail the restriction of $\mathcal{A}_0(\cdot,\cdot)$ on the space $\mathcal{Z}$. Most of the properties (with the only exception of Lemma 4.7) that we prove in this section, such as the symmetry of $\mathbb{A}_0^{zz}$ and the uniform bounds on the condition number of $\mathbb{A}_0^{zz}$ hold also for the corresponding restriction of Type-1 bilinear form to $\mathcal{Z}$.

We begin by showing that the restrictions of the bilinear forms (of both Type-0 and Type-1) on $\mathcal{Z}$ are symmetric.

Lemma 4.3 *Let $\mathcal{A}(\cdot,\cdot)$ be the bilinear form of a non-symmetric Type-1 IP method as defined in (2.6) ($\theta=0$ or $\theta=1$) and let $\mathcal{A}_0(\cdot,\cdot)$ be the corresponding Type-0 bilinear form. Then both $\mathcal{A}_0(\cdot,\cdot)$ and $\mathcal{A}(\cdot,\cdot)$ restricted to $\mathcal{Z}$ are symmetric*:

$$\mathcal{A}_0(z,\phi)=\mathcal{A}_0(\phi,z)\quad \forall z,\phi\in\mathcal{Z},\ \theta=0,1,$$
$$\mathcal{A}(z,\phi)=\mathcal{A}(\phi,z)\quad \forall z,\ \phi\in\mathcal{Z},\ \theta=0,1.$$

Proof From the definition (3.2) of the $\mathcal{Z}$ space, it follows that

$$(\nabla\psi,\nabla z)_{\mathcal{T}_h}=\langle\{\nabla\psi\},[\![z]\!]\rangle_{\mathcal{E}_h}=\langle\{\nabla z\},[\![\psi]\!]\rangle_{\mathcal{E}_h}\quad \forall z,\psi\in\mathcal{Z}. \tag{4.4}$$

Substituting the above identity in the definition of the methods (2.7) gives

$$\begin{aligned}\mathcal{A}_0(z,\psi)&=\theta\langle[\![z]\!],\{\nabla\psi\}\rangle_{\mathcal{E}_h}+\langle[\![z]\!],\mathcal{P}_E^0([\![\psi]\!])\rangle_{\mathcal{E}_h}\\&=\theta(\nabla\psi,\nabla z)_{\mathcal{T}_h}+\langle[\![\psi]\!],\mathcal{P}_E^0([\![z]\!])\rangle_{\mathcal{E}_h}=\mathcal{A}_0(\psi,z),\end{aligned}$$

which shows the symmetry of $\mathcal{A}_0(\cdot,\cdot)$. Finally, the symmetry for $\mathcal{A}(\cdot,\cdot)$ follows from that of $\mathcal{A}_0(\cdot,\cdot)$, since the difference between these bilinear forms is obviously symmetric. □

We now prove bounds on the eigenvalues of $\mathcal{A}_0(\cdot,\cdot)$ and $\mathcal{A}(\cdot,\cdot)$, when restricted to $\mathcal{Z}$. Since these restrictions are symmetric, the Lemma below gives such bounds.

Lemma 4.4 *Let $\mathcal{Z}$ be the space defined in (3.2). Then for all $z\in\mathcal{Z}$, the following estimates hold*

$$\alpha h^{-2}\|z\|_0^2\lesssim\mathcal{A}_0(z,z)\lesssim\alpha h^{-2}\|z\|_0^2, \tag{4.5}$$

and also,

$$\alpha h^{-2}\|z\|_0^2\lesssim\mathcal{A}(z,z)\lesssim\alpha h^{-2}\|z\|_0^2, \tag{4.6}$$

where α denotes as usual the value of the penalty parameter of the method.

Proof In both cases, the upper bounds follow in a straightforward fashion from a trace inequality [2]

$$h_E^{-1}\|\mathcal{P}_E^0[\![z]\!]\|_{0,E}^2\le h_E^{-1}\|[\![z]\!]\|_{0,E}^2\lesssim h_E^{-1}\|z\|_{0,E}^2\lesssim h_E^{-2}\|z\|_{0,T}^2+\|\nabla z\|_{0,T}^2\quad \forall z\in H^1(T),$$

together with the standard inverse inequality (see, for example, [20, Theorem 17.2, p. 135]):

$$\|\nabla w\|_{0,T}^2\lesssim h_T^{-2}\|w\|_{0,T}^2\quad \forall w\in\mathbb{P}^k(T),k\ge 1\ \forall T\in\mathcal{T}_h.$$

We now prove the lower bounds. From the definition of the basis functions for the space $\mathcal{Z}$ (3.6), we have

$$\forall z \in \mathcal{Z}, \quad z = \sum_{E\in\mathcal{E}_h} z_E = \sum_{E\in\mathcal{E}_h} z_E(m_E)\psi_E.$$

Using this expression we have

$$\|z\|_0^2 = \sum_{T\in\mathcal{T}_h}\int_T |z|^2 dx = \sum_{T\in\mathcal{T}_h}\sum_{E\subset\partial T}\int_T |z_E|^2 dx.$$

We shall estimate first $\|z\|_{0,T}^2$ for a fixed T, say $T = T^+$. Denoting by z_E^+ the trace of z_E on $E = T^+ \cap T^-$, from the interior of T^+ and using standard scaling arguments we find

$$\begin{aligned}\|z\|_{0,T}^2 &= \int_T |z|^2 dx = \sum_{E\subset\partial T}\int_T |z_E|^2 dx \simeq \sum_{\widehat{E}\subset\partial\widehat{T}} h_T^d \int_{\widehat{T}} |\hat{z}_{\widehat{E}}|^2 dx \\ &\approx h_T^d \sum_{\widehat{E}\subset\partial\widehat{T}} [\hat{z}_{\widehat{E}}(m_{\widehat{E}})]^2 \approx h_T^d \sum_{E\subset\partial T} [z_E(m_E)]^2 = h_T^d \sum_{E\subset\partial T} [z_E^+(m_E)]^2,\end{aligned}$$

where we have also used the fact that in finite dimension all norms are equivalent (we have used that on $\widehat{T}$ the mass matrix for the basis $\{\hat{\psi}_{\widehat{E}}\}$ is spectrally equivalent to its diagonal). Then, from the definition of the space $\mathcal{Z}$ (3.2) we have

$$\int_E \{z_E\} = 0 \quad \Longrightarrow \quad z_E^+(m_E) = \frac{1}{|E|}\int_E z_E^+ = -\frac{1}{|E|}\int_E z_E^- = -z_E^-(m_E) \quad \forall E \in \mathcal{E}_h^o,$$

which also implies,

$$[\mathcal{P}_E^0([\![z_E]\!])]^2 = [[\![z_E]\!](m_E)]^2 = \begin{cases} [z_E^+(m_E) - z_E^-(m_E)]^2 = 4[z_E^+(m_E)]^2 & \forall E \in \mathcal{E}_h^o, \\ [z_E^+(m_E)]^2 & \forall E \in \mathcal{E}_h^\partial. \end{cases} \tag{4.7}$$

Thus, defining $\mathfrak{c}_E = 1/4$ for $E \in \mathcal{E}_h^o$ and $\mathfrak{c}_E = 1$ for $E \in \mathcal{E}_h^\partial$, these observations yield to,

$$\begin{aligned}\|z\|_{0,T}^2 &\approx h_T^d \sum_{E\subset\partial T} [z_E^+(m_E)]^2 = h_T^d \sum_{E\subset\partial T} \mathfrak{c}_E [\mathcal{P}_E^0([\![z_E]\!])]^2 \\ &= h_T^d \sum_{E\subset\partial T} \frac{\mathfrak{c}_E}{|E|}\int_E [\mathcal{P}_E^0([\![z_E]\!])]^2 = \sum_{E\subset\partial T} \frac{\mathfrak{c}_E h_T^d h_E}{|E|}\frac{1}{h_E}\int_E [\mathcal{P}_E^0([\![z_E]\!])]^2,\end{aligned}$$

and so, noting that $\frac{\mathfrak{c}_E h_T^d h_E}{|E|} \approx Ch^2$ and summing over all triangles we finally have,

$$\begin{aligned}\|z\|_0^2 &\lesssim h^2 \sum_{T\in\mathcal{T}_h}\sum_{E\subset\partial T}\frac{1}{h_E}\int_E [\mathcal{P}_E^0([\![z_E]\!])]^2 \lesssim h^2 \sum_{E\in\mathcal{E}_h}\frac{\alpha}{\alpha h_E}\int_E [\mathcal{P}_E^0([\![z_E]\!])]^2 \\ &\lesssim \alpha^{-1} h^2 \mathcal{A}_0(z, z),\end{aligned}$$

which proves (4.5).

Following an argument analogous to the proof of Lemma 2.3, the inequality (4.6) follows as well. □

This Lemma guarantees that the $\mathbb{A}_0^{zz}$-block, and correspondingly the $\mathbb{A}^{zz}$-block, are well-conditioned. More precisely, denoting by κ_2 the condition number w.r.t. the ℓ^2-norm for a quasi-uniform mesh we have $\kappa_2(\mathbb{A}_0^{zz}) = O(1)$ and correspondingly $\kappa_2(\mathbb{A}^{zz}) = O(1)$, independently of mesh size. Therefore, Lemma 4.4 guarantees that systems with $\mathbb{A}_0^{zz}$ (resp. $\mathbb{A}^{zz}$) can be solved to any precision by the method of Conjugate Gradients (CG) and the number of CG iterations needed is independent of the size of the problem. This is a simple consequence of the well known estimate on the convergence of CG (see, e.g., [29, 34]).

Remark 4.5 A result similar to the one in Lemma 4.4 can be easily shown to hold for the WOPNIP or WOPSIP methods of [13] and [14]. For these methods,

$$\alpha h^{-4}\|z\|_0^2 \lesssim \mathcal{A}_0^{WO}(z,z) \lesssim \alpha h^{-4}\|z\|_0^2, \quad \forall z \in \mathcal{Z}. \tag{4.8}$$

Hence, for such discretizations the solver proposed and described in Algorithm 4.2 can be applied without any change to the solution of linear systems resulting from WOPNIP and WOPSIP discretizations.

Remark 4.6 Notice that restricted to $\mathcal{Z}$, the semi-norm

$$|||z|||_*^2 = \sum_{E\in\mathcal{E}_h} \|h_E^{-1/2}\mathcal{P}_E^0([\![z]\!])\|_{0,E}^2 \quad \forall z \in \mathcal{Z}, \tag{4.9}$$

is actually a norm. This follows from the fact that if (4.9) is zero for some $z \in \mathcal{Z}$, then each term in the sum has to be zero and so reasoning as in the proof of Lemma 4.4 it is easy to see that such a z must be an element of the space V^{CR}. Since $V^{CR} \cap \mathcal{Z} = \{0\}$, we conclude that $z \equiv 0$ and therefore the only element in $\mathcal{Z}$ for which the norm $|||\cdot|||_*$ vanishes is the 0 function. That is,

$$\begin{aligned}\text{Let } z \in \mathcal{Z} : |||z|||_*^2 = 0 \quad &\Leftrightarrow \quad \mathcal{P}_E^0([\![z]\!]) = 0 \quad \forall E \in \mathcal{E}_h \\ &\Leftrightarrow \quad z_E^+(m_E) = z_E^-(m_E) \quad \forall E \in \mathcal{E}_h \quad \Leftrightarrow \quad z \in V^{CR} \quad \Leftrightarrow \quad z \equiv 0.\end{aligned}$$

We conclude this section by showing that for the IIPG Type-0 method the block $\mathbb{A}_0^{zz}$ turns out to be diagonal.

Lemma 4.7 *Let $\mathcal{A}_0^I(\cdot,\cdot)$ be the bilinear form of the non-symmetric IIPG Type-*0 *method. Let $\{\psi_{E,T}^z\}_{E\in\mathcal{E}_h}$ be the basis for the space $\mathcal{Z}$ as defined in* (3.6). *Let $\mathbb{A}_0^{zz}$ be the matrix representation in this basis of the restriction to the subspace $\mathcal{Z}$ of the operator associated to $\mathcal{A}_0^I(\cdot,\cdot)$. Then, $\mathbb{A}_0^{zz}$ is diagonal.*

Proof Note that from the definition (2.7) of the method ($\theta = 0$) together with (4.4) we have

$$\begin{aligned}\mathcal{A}_0^I(z,\psi) &= (\nabla z, \nabla\psi)_{\mathcal{T}_h} - \langle\{\nabla z\}, [\![\psi]\!]\rangle_{\mathcal{E}_h} + \langle S_E\mathcal{P}_E^0([\![z]\!]), \mathcal{P}_E^0([\![\psi]\!])\rangle_{\mathcal{E}_h} \\ &= \langle S_E\mathcal{P}_E^0([\![z]\!]), \mathcal{P}_E^0([\![\psi]\!])\rangle_{\mathcal{E}_h} \quad \forall z, \psi \in \mathcal{Z}.\end{aligned}$$

Let $\{\psi_{E,T}^z\}$ be the basis functions (3.6). To show that $\mathbb{A}_0^{zz}$ is diagonal it is enough to show that for the basis functions (3.6), the following relation holds

$$\mathcal{A}_0^I(\psi_E^z, \psi_{E'}^z) = c_E\delta_{E,E'}, \quad c_E \neq 0,\ \forall E \in \mathcal{E}_h, \tag{4.10}$$

where $\delta_{E,E'}$ is the delta function associated with the edge/face E. Observe that the above relation readily implies that the off-diagonal terms of $\mathbb{A}_0^{zz}$ are zero.

We next show (4.10). Note that the supports of ψ_E^z and $\psi_{E'}^z$ are disjoint unless $E, E' \subset T$ for some $T \in \mathcal{T}_h$. We first prove the result when $T \cap \partial\Omega = \emptyset$. From (4.7) we readily get

$$\begin{aligned}\mathcal{A}_0^I(\psi_E^z, \psi_{E'}^z) &= S_E \int_E \mathcal{P}_E^0([\![\psi_E^z]\!])\mathcal{P}_E^0([\![\psi_{E'}^z]\!]) = S_E[2\psi_E^z(m_E)][2\psi_{E'}^z(m_E)] \\ &= S_E 2 \cdot 2\delta_{E,E'}, \quad E, E' \subset \partial T,\ E, E' \in \mathcal{E}_h^o,\end{aligned}$$

which gives (4.10) with $c_E = 4S_E$.

For boundary edges/faces (4.10) also follows (with $c_E = S_E$) in a straightforward and similar fashion and the details are omitted here. □

5 Preconditioning and Iterative Methods for the Type-1 Interior Penalty Methods

In this section we focus on the iterative methods for the solution of problem (2.5) with $\mathcal{A}^{DG}(\cdot,\cdot) = \mathcal{A}(\cdot,\cdot)$. We first introduce the needed matrix notation and propose a general iterative algorithm for the non-symmetric methods, commenting briefly on the difference between *preconditioners* for the symmetric (SIPG) methods and convergent iterations for the non-symmetric methods. We then discuss and analyze two preconditioners for the symmetric SIPG method. The section is closed with the analysis for the case of the iterative algorithm for non-symmetric methods.

5.1 Matrix Representations

We rewrite the problem (2.5) in the new basis (3.7). From the splitting (3.4) we have that $\forall u, w \in V^{DG}$, $u = z + v$, $w = \psi + \varphi$ with $z, \psi \in \mathcal{Z}$ and $v, \varphi \in V^{CR}$ so that $\mathcal{A}(u, w) = \mathcal{A}((z, v); (\psi, \varphi))$. We denote by A and $\widetilde{A}$ the discrete operators defined by $\mathcal{A}(u, w) = (Au, w)$ and their symmetric parts $\widetilde{\mathcal{A}}(u, w) = (\widetilde{A}u, w)$. Let $\mathbb{A}$ and $\widetilde{\mathbb{A}}$ be the matrix representations of the operators A and $\widetilde{A}$ respectively in the basis (3.7). Then $\mathbb{A}$ has the following block-structure

$$\mathbb{A} = \begin{bmatrix} \mathbb{A}^{zz} & \mathbb{A}^{zv} \\ \mathbb{A}^{vz} & \mathbb{A}^{vv} \end{bmatrix}, \tag{5.1}$$

where the blocks $\mathbb{A}^{vz}$ and $\mathbb{A}^{zv}$ are the matrix representations of the action of the operator A, associated to the bilinear forms

$$\begin{aligned}&\mathbb{A}^{zv} \to \mathcal{A}(v, \psi) = \langle S_E[\![v]\!], [\![\psi]\!]\rangle_{\mathcal{E}_h^o}, \\ &\mathbb{A}^{vz} \to \mathcal{A}(z, \varphi) = (\theta + 1)(\nabla z, \nabla\varphi)_{\mathcal{T}_h} + \langle S_E[\![z]\!], [\![\varphi]\!]\rangle_{\mathcal{E}_h^o},\end{aligned} \tag{5.2}$$

with θ and S_E defined in the usual way. In view of Lemma 4.3, it follows that the block $\mathbb{A}^{zz}$ is symmetric (and positive definite, provided $\alpha \geq \alpha^*$). Notice also that it is easy to verify that $\mathbb{A}^{vv}$ is also symmetric and positive definite. In fact, by setting $u = v \in V^{CR}$ and $w = \varphi \in V^{CR}$ in (2.6) and using the definition (3.1) of the V^{CR}-space we have

$$\mathcal{A}(v, \varphi) = (\nabla v, \nabla\varphi)_{\mathcal{T}_h} + \langle S_E[\![v]\!], [\![\varphi]\!]\rangle_{\mathcal{E}_h^o} \quad \forall v, \varphi \in V^{CR},$$

which is obviously symmetric. We next deal with the symmetric part $\widetilde{A}(\cdot,\cdot)$. Its matrix form $\widetilde{\mathbb{A}} = (\mathbb{A} + \mathbb{A}^T)/2$ is

$$\widetilde{\mathbb{A}} = \begin{bmatrix} \mathbb{A}^{zz} & \widetilde{\mathbb{A}}^{zv} \\ \widetilde{\mathbb{A}}^{vz} & \mathbb{A}^{vv} \end{bmatrix} = \begin{bmatrix} \mathbb{A}^{zz} & \frac{1}{2}(\mathbb{A}^{zv} + \mathbb{A}^{vz}) \\ \frac{1}{2}(\mathbb{A}^{zv} + \mathbb{A}^{vz}) & \mathbb{A}^{vv} \end{bmatrix}, \tag{5.3}$$

where we have used that $\mathbb{A}^{zz}$ and $\mathbb{A}^{vv}$ are symmetric. Note that the off-diagonal blocks represent

$$\widetilde{\mathbb{A}}^{zv} \to \widetilde{\mathcal{A}}(v,\psi) = \frac{(\theta+1)}{2}(\nabla v, \nabla\psi)_{\mathcal{T}_h} + \langle S_E[\![v]\!], [\![\psi]\!]\rangle_{\mathcal{E}_h^o} \quad \forall v \in V^{CR},\ \psi \in \mathcal{Z},$$

$$\widetilde{\mathbb{A}}^{vz} \to \widetilde{\mathcal{A}}(z,\varphi) = \frac{(\theta+1)}{2}(\nabla z, \nabla\varphi)_{\mathcal{T}_h} + \langle S_E[\![z]\!], [\![\varphi]\!]\rangle_{\mathcal{E}_h^o} \quad \forall z \in \mathcal{Z},\ \varphi \in V^{CR}.$$

As noted earlier (see (2.11))

$$\mathcal{S}(u,w) = \mathcal{S}_0(u,w) \quad \forall u,\ w \in V^{DG}. \tag{5.4}$$

This identity will play a key role in our analysis of non-symmetric methods. Consider the following decomposition of elements of V^{DG} (3.4); $u = z + v,\ w = \psi + \varphi$ with $z, \psi \in \mathcal{Z}$ and $v, \varphi \in V^{CR}$. Integration by parts and the definition of the spaces V^{CR} and $\mathcal{Z}$, then give another useful identity:

$$(\nabla z, \nabla\varphi)_{\mathcal{T}_h} = \langle [\![z]\!], \{\nabla\varphi\}\rangle_{\mathcal{E}_h} \quad \forall z \in \mathcal{Z},\ \varphi \in V^{CR}.$$

Thus, we have that

$$\mathcal{S}(u,w) = \mathcal{S}((z,v);(\varphi,\psi)) = \frac{(\theta+1)}{2}\left[(\nabla z, \nabla\varphi)_{\mathcal{T}_h} - (\nabla v, \nabla\psi)_{\mathcal{T}_h}\right] \quad \forall u,\ w \in V^{DG}.$$

5.2 Preconditioners and Iterators

We now briefly discuss the general settings related to the solution of the systems obtained from symmetric or non-symmetric Type-1 IP DG discretizations. In both cases (symmetric or non-symmetric) we design the solvers using a bilinear form $\mathcal{B}(\cdot,\cdot)$ that is an approximation to $\mathcal{A}(\cdot,\cdot)$.

For the symmetric methods (SIPG Type-1 methods), $\mathcal{B}(\cdot,\cdot)$ is a *preconditioner* and is used in conjunction with Conjugate Gradient Method. The analysis of the convergence of the resulting Preconditioned Conjugate Gradient Method (PCG) boils down to showing that $\mathcal{B}(\cdot,\cdot)$ is spectrally equivalent to $\mathcal{A}(\cdot,\cdot)$ with constants independent of the problem size.

For the non-symmetric Type-1 IP methods $\mathcal{B}(\cdot,\cdot)$ is an *iterator*, and is used in a linear iterative method (see Algorithm 5.3 below). The convergence analysis in this case amounts to proving that, in a suitable norm, the error is reduced uniformly on each iteration.

5.3 Uniform Preconditioning for Type-1 Symmetric IP Methods

We first describe the preconditioners $\mathcal{B}$ for the Type-1 SIPG method. Note that by virtue of Lemma 2.3, the norm defined by $\mathcal{A}^s(\cdot,\cdot)$ is equivalent to the norm defined by the bilinear form $\mathcal{A}_0^s(\cdot,\cdot)$, and therefore

$$\mathcal{B}(\cdot,\cdot) = \mathcal{A}_0^s(\cdot,\cdot) \tag{5.5}$$

is a uniform preconditioner for $\mathcal{A}^s(\cdot,\cdot)$. Moreover, note that the matrix representation of the preconditioner given by $\mathcal{A}_0^s$ is block-diagonal.

Another preconditioner can be obtained, by using the decomposition (3.4) and setting

$$\mathcal{B}(u,w) = \mathcal{A}(z,\psi) + \mathcal{A}(v,\varphi), \tag{5.6}$$

where $u = z + v$ and $w = \psi + \varphi$ with $v, \varphi \in V^{CR}$ and $z, \psi \in \mathcal{Z}$. This is just the additive (parallel) subspace correction method [39] based on the splitting $V^{DG} = \mathcal{Z} + V^{CR}$. In matrix representation, this will be a block-Jacobi preconditioner. To show that such preconditioner is spectrally equivalent to $\mathcal{A}(\cdot,\cdot)$, following [39] we need to show that the decomposition given in (3.4) is stable in the energy norm defined by the Type-1 SIPG bilinear form.

Lemma 5.1 *Let $\mathcal{A}(\cdot,\cdot)$ be the symmetric bilinear form defined for $\theta = -1$. Then, for every $u \in V^{DG}$ such that $u = v + z$ with $v \in V^{CR}$ and $z \in \mathcal{Z}$, the following inequality holds*

$$\mathcal{A}(z,z) + \mathcal{A}(v,v) \le c_0(\alpha)\mathcal{A}(v+z, v+z) \quad \forall z \in \mathcal{Z},\ v \in V^{CR}, \tag{5.7}$$

where $c_0(\alpha)$ is the same constant as in Lemma 2.3.

Proof Set $u = v + z$ and let $\mathcal{A}_0(\cdot,\cdot)$ be the bilinear form corresponding to the Type-0 SIPG method ($\theta = -1$). From Lemmas 2.3 and 3.2 we have

$$\mathcal{A}(u,u) \ge \mathcal{A}_0(u,u) = [\mathcal{A}_0(z,z) + \mathcal{A}_0(v,v)] \ge \frac{1}{c_0}[\mathcal{A}(z,z) + \mathcal{A}(v,v)],$$

which is the inequality that we need and so the proof is complete. □

According to the abstract theory given in [39], to show that the condition number of the preconditioned system is uniformly bounded, we only needed to show the above estimate. We then have the following theorem, whose proof we omit, since it is a direct consequence of Fundamental Theorem I in [39].

Theorem 5.2 *Let the penalty parameter $\alpha \ge \alpha^*$ be large enough so that $\mathcal{A}^{s,\alpha}(\cdot,\cdot)$ and $\mathcal{A}_0^{s,\alpha}(\cdot,\cdot)$ are coercive. Let $\mathcal{B}(\cdot,\cdot)$ be defined as in (5.5) or via (5.6). Then $\mathcal{B}$ gives a uniform preconditioner for $\mathcal{A}^s$, namely, the following estimates hold, with positive constants c_1 and c_2, which depend on the penalty parameter α:*

$$c_1(\alpha)\mathcal{A}^s(u,u) \le \mathcal{B}(u,u) \le c_2(\alpha)\mathcal{A}^s(u,u) \quad \forall u \in V^{DG}.$$

Following the standard estimates for Preconditioned Conjugate Gradient method [29, 31, 34], the spectral equivalence results that we proved in this section show that a Preconditioned Conjugate Gradient (PCG) with either $\mathcal{B}(\cdot,\cdot) = \mathcal{A}_0(\cdot,\cdot)$, or $\mathcal{B}(\cdot,\cdot)$ defined in (5.6) will converge to the solution of (2.5) uniformly in the norm defined by $\mathcal{A}(\cdot,\cdot)$ at a rate bounded by $\frac{\sqrt{c_2/c_1}-1}{\sqrt{c_2/c_1}+1}$. As it is well known, to apply the preconditioners we proposed here, one needs algorithms for efficient solution of variational problems with $\mathcal{B}(\cdot,\cdot)$. By the construction of $\mathcal{B}(\cdot,\cdot)$ (in both cases with block diagonal matrix representation) one may use the methods suggested in Sects. 4.2 and 4.3.

5.4 Iterative Methods for Type-1 Non-symmetric IP Methods

In this section we address the issue of iterative solvers for the classical (Type-1) NIPG and IIPG methods. We use the following basic linear iterative algorithm with an iterator $\mathcal{B}(\cdot,\cdot)$:

Algorithm 5.3 *Given initial guess u_0, for $k=0,1,\dots$ until convergence*:

1. *Solve* $\mathcal{B}(e_k,w)=(f,w)_{\mathcal{T}_h}-\mathcal{A}^{DG}(u_k,w)\quad \forall w\in V^{DG}$,
2. *Update* $u_{k+1}=u_k+e_k$.

For both IIPG and NIPG methods as an iterator $\mathcal{B}(\cdot,\cdot)$ in Algorithm 5.3 we take the corresponding symmetric part, that is

$$\mathcal{B}(\cdot,\cdot)=\widetilde{\mathcal{A}}(\cdot,\cdot).$$

Below we prove uniform convergence of the iterates obtained via Algorithm 5.3 to the solution of (2.5) for both IIPG and NIPG, under a technical restriction on the penalty parameter.

To show uniform convergence of such an iterative method, it is sufficient to show that $\exists\,\Lambda,\, 0\le\Lambda<1$ such that

$$\|u-u_{k+1}\|_{\widetilde{\mathcal{A}}}\le\Lambda\|u-u_k\|_{\widetilde{\mathcal{A}}}.$$

For uniform convergence we also need to show that Λ is uniformly (with respect to the mesh size) bounded away from 1. We would like to point out that in the convergence proof for both Type-1 IIPG and NIPG methods, we use a strengthened CBS inequality in a norm defined by the symmetric part of the Type-0 IIPG. As a consequence, the hypothesis on the penalty parameter required by our results for the Type-1 NIPG method is somewhat more restrictive than for the corresponding IIPG.

We present in detail the analysis for the iterative method for Type-1 IIPG discretizations. Then, we state the convergence result for Type-1 NIPG discretizations, commenting on the main changes in the proof. We begin with stating a general result on a CBS inequality given in Lemma A.1.

Lemma A.1 *Let $\mathcal{A}_0^{I,\alpha^*}(\cdot,\cdot)$ be the Type-0 non-symmetric IIPG method with penalty parameter α^* and let $\widetilde{\mathcal{A}}_0^{I,\alpha^*}(\cdot,\cdot)$ be its symmetric part. Let α^* be the minimal value of the penalty parameter for which $\mathcal{A}_0^{I,\alpha^*}(\cdot,\cdot)$ is coercive. Then, the following CBS inequality holds for all $v\in V^{CR}$ and $z\in\mathcal{Z}$*:

$$\exists\gamma^*<1:\left[\frac{1}{2}(\nabla z,\nabla v)_{\mathcal{T}_h}\right]^2\le[\gamma^*]^2\widetilde{\mathcal{A}}_0^{I,\alpha^*}(z,z)\widetilde{\mathcal{A}}_0^{I,\alpha^*}(v,v). \tag{A.1}$$

The proof of this Lemma is given in the Appendix.

Next result follows from the previous Lemma and the particular structure of $\mathcal{A}_0^I(\cdot,\cdot)$ on the subspace $\mathcal{Z}$. We state it as a Lemma since it will be used frequently in the subsequent analysis.

Lemma 5.4 *Let $K>1$ be a fixed constant and let α^* be the minimal value of the penalty parameter for which $\mathcal{A}_0^{I,\alpha^*}(\cdot,\cdot)$ is coercive. Let $\mathcal{A}_0^{I,K\alpha^*}(\cdot,\cdot)$ be the Type-0 non-symmetric IIPG method with penalty parameter $\alpha=K\alpha^*$, and let $\widetilde{\mathcal{A}}_0^{I,K\alpha^*}(\cdot,\cdot)$ be its symmetric part.*

Then, the following inequality holds for all $v \in V^{CR}$ *and* $z \in \mathcal{Z}$:

$$|(\nabla z, \nabla v)_{\mathcal{T}_h}| \le \frac{2\gamma^*}{\sqrt{K}} \sqrt{\widetilde{\mathcal{A}}_0^{I,K\alpha^*}(z,z)} \sqrt{\widetilde{\mathcal{A}}_0^{I,K\alpha^*}(v,v)},$$

where $\gamma^* < 1$ *is the CBS constant given in* (A.1) *associated with the inner product defined by* $\widetilde{\mathcal{A}}_0^{I,\alpha^*}(\cdot,\cdot)$.

Proof For the Type-0 IIPG method and for any value of the penalty parameter α we have that

$$\begin{cases} \widetilde{\mathcal{A}}_0^{I,\alpha}(z,z) = \sum_{E\in\mathcal{E}_h} \alpha |E|^{-1/(d-1)} \|\mathcal{P}_E^0([\![z]\!])\|_{0,E}^2 \quad \forall z \in \mathcal{Z}, \\ \widetilde{\mathcal{A}}_0^{I,\alpha}(v,v) = \sum_{T\in\mathcal{T}_h} \|\nabla v\|_{0,T}^2 \quad \forall v \in V^{CR}. \end{cases} \tag{5.8}$$

That is, $\|z\|^2_{\widetilde{\mathcal{A}}^{I,\alpha}}$ depends linearly on the penalty parameter while $\|v\|^2_{\widetilde{\mathcal{A}}^{I,\alpha}}$ is independent of it. Therefore, by setting $\alpha = K\alpha^*$, we have for $z \in \mathcal{Z}$ and $v \in V^{CR}$

$$\widetilde{\mathcal{A}}_0^{I,K\alpha^*}(z,z) = K\widetilde{\mathcal{A}}_0^{I,\alpha^*}(z,z) \quad \text{and} \quad \widetilde{\mathcal{A}}_0^{I,K\alpha^*}(v,v) = \widetilde{\mathcal{A}}_0^{I,\alpha^*}(v,v) \quad \forall K \ge 1.$$

Using now estimate (A.1) from Lemma A.1 together with the above identities we arrive at

$$\begin{aligned} |(\nabla z, \nabla v)| &\le 2\gamma^* \sqrt{\widetilde{\mathcal{A}}_0^{I,\alpha^*}(z,z)} \sqrt{\widetilde{\mathcal{A}}_0^{I,\alpha^*}(v,v)} \\ &= 2\frac{\gamma^*}{\sqrt{K}} \sqrt{K\widetilde{\mathcal{A}}_0^{I,\alpha^*}(z,z)} \sqrt{\widetilde{\mathcal{A}}_0^{I,\alpha^*}(v,v)} \\ &= 2\frac{\gamma^*}{\sqrt{K}} \sqrt{\widetilde{\mathcal{A}}_0^{I,K\alpha^*}(z,z)} \sqrt{\widetilde{\mathcal{A}}_0^{I,K\alpha^*}(v,v)}, \end{aligned}$$

which concludes the proof. □

Next Lemma shows a stability estimate.

Lemma 5.5 *Let* $K > 1$ *be a fixed constant and let* α^* *be the minimal value of the penalty parameter for which* $\mathcal{A}_0^{I,\alpha^*}(\cdot,\cdot)$ *is coercive. Let* $\widetilde{\mathcal{A}}_0^{I,K\alpha^*}(\cdot,\cdot)$ *be the symmetric part of the bilinear form corresponding to Type-0 IIPG method with penalty parameter* $\alpha = K\alpha^*$. *Then, for all* $u \in V^{DG}$ *such that* $u = z + v$ *with* $v \in V^{CR}$ *and* $z \in \mathcal{Z}$, *we have*

$$\left(\widetilde{\mathcal{A}}_0^{I,K\alpha^*}(z,z) + \widetilde{\mathcal{A}}_0^{I,K\alpha^*}(v,v)\right) \le \left[\frac{\sqrt{K}}{\sqrt{K} - \gamma^*}\right] \widetilde{\mathcal{A}}_0^{I,K\alpha^*}(z+v, z+v). \tag{5.9}$$

Proof Let $u \in V^{DG}$ such that $u = z + v$ with $z \in \mathcal{Z}$ and $v \in V^{CR}$ (this can always be done in view of Proposition 3.1). From the definition of $\widetilde{\mathcal{A}}_0^{I,\alpha}(\cdot,\cdot)$ we have

$$\begin{aligned} \widetilde{\mathcal{A}}_0^{I,\alpha}(v+z, v+z) &= \widetilde{\mathcal{A}}_0^{I,\alpha}(z,z) + \widetilde{\mathcal{A}}_0^{I,\alpha}(v,v) + \widetilde{\mathcal{A}}_0^{I,\alpha}(z,v) + \widetilde{\mathcal{A}}_0^{I,\alpha}(v,z) \\ &\ge \widetilde{\mathcal{A}}_0^{I,\alpha}(z,z) + \widetilde{\mathcal{A}}_0^{I,\alpha}(v,v) - |\widetilde{\mathcal{A}}_0^{I,\alpha}(z,v)| - |\widetilde{\mathcal{A}}_0^{I,\alpha}(v,z)| \\ &\ge \widetilde{\mathcal{A}}_0^{I,\alpha}(z,z) + \widetilde{\mathcal{A}}_0^{I,\alpha}(v,v) - |(\nabla z, \nabla v)|\,. \end{aligned}$$

We now set $\alpha = K\alpha^*$ and apply Lemma 5.4 followed by an obvious identity to obtain that

$$\widetilde{\mathcal{A}}_0^{I,K\alpha^*}(v+z, v+z) \geq \widetilde{\mathcal{A}}_0^{I,K\alpha^*}(z,z) + \widetilde{\mathcal{A}}_0^{I,K\alpha^*}(v,v) - 2\frac{\gamma^*}{\sqrt{K}}\sqrt{\widetilde{\mathcal{A}}_0^{I,K\alpha^*}(z,z)}\sqrt{\widetilde{\mathcal{A}}_0^{I,K\alpha^*}(v,v)}$$
$$= \left(1 - \frac{\gamma^*}{\sqrt{K}}\right)(\widetilde{\mathcal{A}}_0^{I,K\alpha^*}(z,z) + \widetilde{\mathcal{A}}_0^{I,K\alpha^*}(v,v))$$
$$+ \frac{\gamma^*}{\sqrt{K}}\left(\sqrt{\widetilde{\mathcal{A}}_0^{I,K\alpha^*}(z,z)} - \sqrt{\widetilde{\mathcal{A}}_0^{I,K\alpha^*}(v,v)}\right)^2,$$

and so, noting that the second term on the right side can be dropped since it is non-negative, we reach (5.9) and the proof is complete. □

Next result provides a bound on the skew-symmetric part of the bilinear form corresponding to Type-1 IIPG method.

Lemma 5.6 *Let $K > 1$ be a fixed constant and let $\mathcal{A}^{I,\alpha}(\cdot,\cdot)$ be the bilinear form of the IIPG method with penalty parameter $\alpha = K\alpha^*$, where α^* is the minimal value of the penalty parameter for which $\mathcal{A}_0^{I,\alpha^*}(\cdot,\cdot)$ is coercive. Let $\widetilde{\mathcal{A}}^{I,\alpha}(\cdot,\cdot)$ and $\mathcal{S}(\cdot,\cdot)$ denote its symmetric and skew symmetric parts. Then, there exists $\Lambda_0 > 0$ such that*

$$|\mathcal{S}(u,w)| \leq \Lambda_0[\widetilde{\mathcal{A}}^{I,\alpha}(u,u)]^{1/2}[\widetilde{\mathcal{A}}^{I,\alpha}(w,w)]^{1/2} \quad \forall u, w \in V^{DG}. \tag{5.10}$$

Furthermore, for $K \geq 4$ it follows that $\Lambda_0 < 1$.

Proof To estimate $|\mathcal{S}(u,w)|$ we first decompose u and w by means of the splitting (3.4) in the form:

$$u = z + v, \quad w = \psi + \varphi, \quad z, \psi \in \mathcal{Z}, \quad v, \varphi \in V^{CR}. \tag{5.11}$$

Then, by applying Lemma 5.4 together with a Cauchy-Schwarz inequality[1] and Lemma 5.5 we obtain

$$|(Su,w)| = \frac{1}{2}|(\nabla z, \nabla\varphi) - (\nabla\psi, \nabla v)|$$
$$\leq \frac{1}{2}(|(\nabla z, \nabla\varphi)| + |(\nabla\psi, \nabla v)|)$$
$$\leq \frac{\gamma^*}{\sqrt{K}}\left(\sqrt{\widetilde{\mathcal{A}}_0^{I,K\alpha^*}(z,z)}\sqrt{\widetilde{\mathcal{A}}_0^{I,K\alpha^*}(\varphi,\varphi)} + \sqrt{\widetilde{\mathcal{A}}_0^{I,K\alpha^*}(v,v)}\sqrt{\widetilde{\mathcal{A}}_0^{I,K\alpha^*}(\psi,\psi)}\right)$$
$$\leq \frac{\gamma^*}{\sqrt{K}}\left(\widetilde{\mathcal{A}}_0^{I,K\alpha^*}(z,z) + \widetilde{\mathcal{A}}_0^{I,K\alpha^*}(v,v)\right)^{1/2}\left(\widetilde{\mathcal{A}}_0^{I,K\alpha^*}(\varphi,\varphi) + \widetilde{\mathcal{A}}_0^{I,K\alpha^*}(\psi,\psi)\right)^{1/2}$$
$$\leq \frac{\gamma^*}{\sqrt{K} - \gamma^*}\sqrt{\widetilde{\mathcal{A}}_0^{I,K\alpha^*}(u,u)}\sqrt{\widetilde{\mathcal{A}}_0^{I,K\alpha^*}(w,w)},$$

[1] For $x = [x_1, x_2]^T$, $y = [y_1, y_2]^T$ we use $x_1y_1 + x_2y_2 \leq \sqrt{x_1^2 + x_2^2}\sqrt{y_1^2 + y_2^2}$, with $x_1 = \sqrt{\widetilde{\mathcal{A}}_0^{I,K\alpha^*}(z,z)}$, $x_2 = \sqrt{\widetilde{\mathcal{A}}_0^{I,K\alpha^*}(v,v)}$, $y_1 = \sqrt{\widetilde{\mathcal{A}}_0^{I,K\alpha^*}(\varphi,\varphi)}$ and $y_2 = \sqrt{\widetilde{\mathcal{A}}_0^{I,K\alpha^*}(\psi,\psi)}$.

and so by setting $\Lambda_0 = \frac{\gamma^*}{\sqrt{K}-\gamma^*}$ we reach (5.10). Taking $K \geq 4$ we obtain $\Lambda_0 < 1$ since $\gamma^* < 1$. □

We now have all the ingredients needed to show the main result of this section. In the statement of the theorems for iterative methods for IIPG and NIPG discretizations, we take u to be the exact solution to the DG problem (solution of (2.5)).

Theorem 5.1 *Let α^* be the minimal value of the penalty parameter for which $\widetilde{\mathcal{A}}_0^{I,\alpha^*}(\cdot,\cdot)$ is coercive. Let $\mathcal{A}^{I,\alpha}(\cdot,\cdot)$ be the bilinear form of a Type-1 IIPG method (with $\theta = 0$), and penalty parameter satisfying $\alpha \geq 4\alpha^*$. Let $\mathcal{B}(\cdot,\cdot) = \widetilde{\mathcal{A}}^{I,\alpha}(\cdot,\cdot)$ in Algorithm 5.3 and let u_k and u_{k+1} be two consecutive iterates obtained via this algorithm. Then, there exists a positive constant $\Lambda < 1$ such that*

$$\|u - u_{k+1}\|_{\widetilde{\mathcal{A}}} \leq \Lambda \|u - u_k\|_{\widetilde{\mathcal{A}}}. \tag{5.12}$$

Proof From step 2 in Algorithm 5.3 we have that $e_k = u_{k+1} - u_k$. From step 1 in Algorithm 5.3, with $w = u - u_{k+1}$ we have that

$$\mathcal{B}(e_k, u - u_{k+1}) - (f, u - u_{k+1})_{\mathcal{T}_h} + \mathcal{A}(u_k, u - u_{k+1}) = 0.$$

Hence,

$$\begin{aligned}
\mathcal{B}(u - u_{k+1}, u - u_{k+1}) &= \mathcal{B}(u - u_{k+1}, u - u_{k+1}) + \mathcal{B}(e_k, u - u_{k+1}) \\
&\quad - (f, u - u_{k+1})_{\mathcal{T}_h} + \mathcal{A}(u_k, u - u_{k+1}) \\
&= \mathcal{B}(u, u - u_{k+1}) - \mathcal{B}(u_k, u - u_{k+1}) \\
&\quad - \mathcal{A}(u, u - u_{k+1}) + \mathcal{A}(u_k, u - u_{k+1}) \\
&= \mathcal{B}(u - u_k, u - u_{k+1}) - \mathcal{A}(u - u_k, u - u_{k+1}) \\
&= [\mathcal{B} - \mathcal{A}](u - u_k, u - u_{k+1}).
\end{aligned}$$

Taking $\mathcal{B}(\cdot,\cdot) = \widetilde{\mathcal{A}}(\cdot,\cdot)$ and recalling the definition of the skew-symmetric part $\mathcal{S}(\cdot,\cdot)$ of $\mathcal{A}(\cdot,\cdot)$, we obtain

$$\|u - u_{k+1}\|_{\widetilde{\mathcal{A}}}^2 = -\mathcal{S}(u - u_k, u - u_{k+1}) = \mathcal{S}(u - u_{k+1}, u - u_k). \tag{5.13}$$

Then, by setting $u = u - u_k$ and $w = u - u_{k+1}$ in (5.10) of Proposition 5.6 and substituting in (5.13) we find that

$$\|u - u_{k+1}\|_{\widetilde{\mathcal{A}}}^2 \leq \Lambda_0 \|u - u_{k+1}\|_{\widetilde{\mathcal{A}}} \|u - u_k\|_{\widetilde{\mathcal{A}}},$$

and therefore (5.12) holds with $\Lambda = \Lambda_0$, which is obviously less than 1, since $\alpha \geq 4\alpha^*$. □

In a similar way the following convergence result holds for the iterative method for the solution of the linear system resulting from NIPG discretization.

Theorem 5.2 *Let α^* be the minimal value of the penalty parameter for which $\widetilde{\mathcal{A}}_0^{I,\alpha^*}(\cdot,\cdot)$ is coercive. Let $\mathcal{A}^{n,\alpha}(\cdot,\cdot)$ be the bilinear form of the NIPG method with $\alpha \geq 16\alpha^*$. Let $\mathcal{B}(\cdot,\cdot) = \widetilde{\mathcal{A}}^{n,\alpha}(\cdot,\cdot)$ in Algorithm 5.3 and let u_k and u_{k+1} be two consecutive iterates obtained via this algorithm. Then, there exists a positive constant $\Lambda_1 < 1$ such that*

$$\|u - u_{k+1}\|_{\widetilde{\mathcal{A}}} \leq \Lambda_1 \|u - u_k\|_{\widetilde{\mathcal{A}}}.$$

We omit the proof, since it is essentially the same as that of Theorem 5.1 provided the relation

$$\widetilde{\mathcal{A}}_0^{I,\alpha}(z,z) \leq \widetilde{\mathcal{A}}_0^{n,\alpha}(z,z) \quad \forall z \in \mathcal{Z}, \qquad \widetilde{\mathcal{A}}_0^{I,\alpha}(v,v) = \widetilde{\mathcal{A}}_0^{n,\alpha}(v,v) \quad \forall v \in V^{CR}, \tag{5.14}$$

is taken into account and Proposition 5.6 and Lemma 5.5 are replaced respectively, by Proposition 5.7 and Lemma 5.8 given below.

Proposition 5.7 *Let $K > 1$ be a fixed constant and let α^* be the minimal value of the penalty parameter for which $\mathcal{A}_0^{I,\alpha^*}(\cdot,\cdot)$ is coercive. Let $\mathcal{A}^{n,K\alpha^*}(\cdot,\cdot)$ be the bilinear form of the Type-1 NIPG method with penalty parameter $\alpha = K\alpha^*$ and let $\widetilde{\mathcal{A}}^{n,\alpha}(\cdot,\cdot)$ and $\mathcal{S}(\cdot,\cdot)$ denote its symmetric and skew symmetric parts. Then, there exists $\Lambda_1 > 0$ such that*

$$|\mathcal{S}(u,w)| \leq \Lambda_1 [\widetilde{\mathcal{A}}^{n,\alpha}(u,u)]^{1/2} [\widetilde{\mathcal{A}}^{n,\alpha}(w,w)]^{1/2} \quad \forall v, w \in V^{DG}, \ \Lambda_1 = \frac{2\gamma^*}{\sqrt{K} - 2\gamma^*}, \tag{5.15}$$

where $\gamma^* < 1$ is the CBS constant given in (A.1) associated with the inner product defined by $\widetilde{\mathcal{A}}_0^{I,\alpha^*}(\cdot,\cdot)$.

Furthermore, for $K \geq 16$ it follows that $\Lambda_1 < 1$.

The proof of this Proposition follows essentially the same lines as for Proposition 5.6 and therefore it is omitted. For its proof instead of Lemma 5.5 one has to use:

Lemma 5.8 *Let $K > 1$ be a fixed constant and let α^* be the minimal value of the penalty parameter for which $\mathcal{A}_0^{I,\alpha^*}(\cdot,\cdot)$ is coercive. Let $\mathcal{A}^{n,K\alpha^*}(\cdot,\cdot)$ be the bilinear form of the Type-1 NIPG method with $\alpha = K\alpha^*$. Then, the decomposition $V^{DG} = \mathcal{Z} \oplus V^{CR}$ is stable in the energy norm defined by $\widetilde{\mathcal{A}}_0^{n,\alpha}(\cdot,\cdot)$. More precisely, for all $u \in V^{DG}$ such that $u = z + v$ with $z \in \mathcal{Z}$ $v \in V^{CR}$,*

$$\left(\widetilde{\mathcal{A}}_0^{n,K\alpha^*}(z,z) + \widetilde{\mathcal{A}}_0^{n,K\alpha^*}(v,v)\right) \leq \left[\frac{\sqrt{K}}{\sqrt{K} - 2\gamma^*}\right] \widetilde{\mathcal{A}}_0^{n,\alpha}(u,u) \quad \forall u \in V^{DG}, \tag{5.16}$$

where $\gamma^* < 1$ is the CBS constant given in (A.1) associated with the inner product defined by $\widetilde{\mathcal{A}}_0^{I,\alpha^*}(\cdot,\cdot)$.

Proof The proof is similar to that of Lemma 5.5. Let $z \in \mathcal{Z}$ and $v \in V^{CR}$, then

$$\widetilde{\mathcal{A}}_0^{n,\alpha}(v+z, v+z) \geq \widetilde{\mathcal{A}}_0^{n,\alpha}(z,z) + \widetilde{\mathcal{A}}_0^{n,\alpha}(v,v) - 2\,|(\nabla z, \nabla v)|\,.$$

By setting $\alpha = K\alpha^*$, applying Lemma 5.4 and using relation (5.14) together with the arithmetic-geometric inequality we get

$$\begin{aligned}
\widetilde{\mathcal{A}}_0^{n,K\alpha^*}(v+z, v+z) &\geq \widetilde{\mathcal{A}}_0^{n,K\alpha^*}(z,z) + \widetilde{\mathcal{A}}_0^{n,K\alpha^*}(v,v) - 4\frac{\gamma^*}{\sqrt{K}}\sqrt{\widetilde{\mathcal{A}}_0^{I,K\alpha^*}(z,z)}\sqrt{\widetilde{\mathcal{A}}_0^{I,K\alpha^*}(v,v)} \\
&\geq \widetilde{\mathcal{A}}_0^{n,K\alpha^*}(z,z) + \widetilde{\mathcal{A}}_0^{n,K\alpha^*}(v,v) - 4\frac{\gamma^*}{\sqrt{K}}\sqrt{\widetilde{\mathcal{A}}_0^{n,K\alpha^*}(z,z)}\sqrt{\widetilde{\mathcal{A}}_0^{n,K\alpha^*}(v,v)} \\
&\geq \left(1 - \frac{2\gamma^*}{\sqrt{K}}\right)(\widetilde{\mathcal{A}}_0^{n,K\alpha^*}(z,z) + \widetilde{\mathcal{A}}_0^{n,K\alpha^*}(v,v)),
\end{aligned}$$

which concludes the proof. □

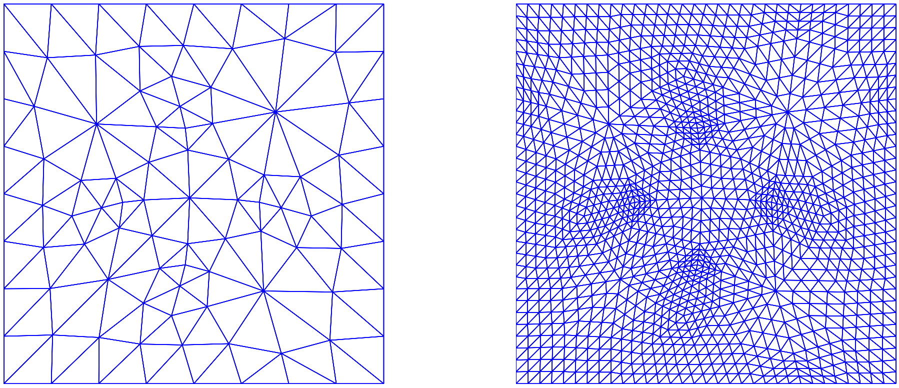

Fig. 3 (**a**) Plot of coarsest triangulation. (**b**) Computational grid (triangulation) obtained after two refinements

6 Numerical Examples

In this section we present a set of numerical experiments to assess and validate the theory developed for the proposed iterative methods and preconditioning techniques. We consider the model problem (2.1) and its discretization with the Type-1 Interior Penalty DG methods on the unit square in $\mathbb{R}^2$. In our experiments, we use 6 computational unstructured triangulations. Similar results, although not reported here, were obtained for structured partitions. In the tables and figures below, the coarsest grid corresponds to level $J = 0$. Each refined triangulation on level J, $J = 1, \ldots, 5$ is obtained by subdividing each of the triangles forming the grid on level $(J - 1)$ into four congruent triangles (see Fig. 3). The number of degrees of freedom on each level is denoted by n_J and satisfies that $n_J = 4^J n_0$, $J = 0, \ldots, 5$. The coarsest grid has 160 triangular elements and $n_0 = 480$ degrees of freedom. The calculation of the spectral equivalence constants, strengthened Cauchy-Schwarz constants and convergence rates which we report below was done using the ARPACK sparse eigensolver routines in MATLAB.

6.1 Preconditioning for SIPG

This first example is devoted to illustrate the conclusions of Theorem 5.2 on preconditioning Type-1 SIPG, namely that the Type-0 SIPG method is a uniform preconditioner for the Type-1 SIPG method. In both methods we have taken for the penalty parameter $\alpha = 10$. For each level of refinement, we have computed numerically the (ℓ^2) condition number of the stiffness matrix, $\kappa_2(\mathbb{A})$, and that of the preconditioned stiffness matrix, $\kappa_2(\mathbb{A}_0^{-1}\mathbb{A})$. The numerical values are reported in Table 1 together with the number of degrees of freedom on each level. From these results it can be easily seen that $\kappa_2(\mathbb{A})$ grows quadratically with n_J, as expected from a well known estimate $\kappa_2(\mathbb{A}) \approx O(h^{-2})$. Observe that for all the levels of refinement the matrix $(\mathbb{A}_0^{-1}\mathbb{A})$ which defines the preconditioned system has a significantly smaller condition number and $\kappa_2(\mathbb{A}_0^{-1}\mathbb{A}) = O(1)$.

Although there is a slight growth in $\kappa_2(\mathbb{A}_0^{-1}\mathbb{A})$ for the first two refinements, from the third refinement onward the condition number of the preconditioned system seems to stabilize to a constant and the growth is negligible, confirming the uniform bounds given in Theorem 5.2.

Table 1 Tabulated values of $\kappa_2(\mathbb{A})$ and $\kappa_2(\mathbb{A}_0^{-1}\mathbb{A})$ for Type-1 SIPG and preconditioned (with Type-0 SIPG preconditioner) Type-1 SIPG methods, respectively

J	n_J	$\kappa_2(\mathbb{A})$	$\kappa_2(\mathbb{A}_0^{-1}\mathbb{A})$
0	480	6.18×10^2	10.02
1	1920	2.46×10^3	11.83
2	7680	9.86×10^3	13.70
3	30720	3.94×10^4	14.70
4	122880	1.57×10^5	15.08
5	491520	6.31×10^5	15.21

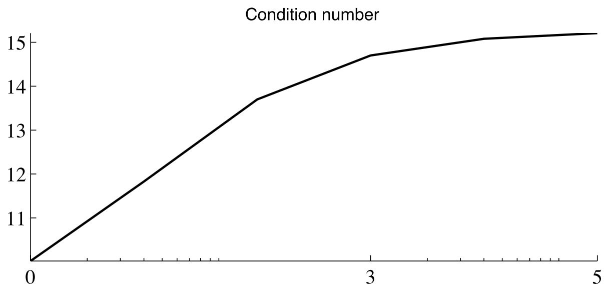

Fig. 4 Plot of $\kappa(A_0^{-1}A)$ versus the level number

To illustrate better the asymptotic behavior of the condition number $\kappa_2(\mathbb{A}_0^{-1}\mathbb{A})$ the graph in Fig. 4 shows the variation of the $\kappa_2(\mathbb{A}_0^{-1}\mathbb{A})$ with the level of refinement. The slope of the curve plotted in Fig. 4 clearly approaches zero, indicating that that the condition number of $\mathbb{A}_0^{-1}\mathbb{A}$ remains uniformly bounded with respect to the degrees of freedom (equivalently the mesh size) for a fixed value of the penalty parameter α.

6.2 Uniformly Convergent Iterative Methods for IIPG and NIPG

The next set of examples is aimed at validating the theory developed for the linear iterative method given in Algorithm 5.3 for the non-symmetric Type-1 schemes IIPG and NIPG. In order to verify Theorems 5.1 and 5.2 we have computed numerically the convergence rates for both iterative methods, for different levels of refinement $J = 0, \ldots, 5$, and for increasing values of the penalty parameter of the form $\alpha = K\alpha^*$ with $K = 1, 2, 4, 8, 16$ and 32. For each level of refinement $J = 0, \ldots, 5$ we have set $\alpha^* = 0.9, 1.3, 1.3, 1.3, 1.6$ and 1.6, respectively. These values of α^* are selected so that for a given method we have: (a) the corresponding bilinear forms are coercive; (b) they are close to the critical value of the penalty parameter, for which the method of Type-0 becomes unstable (semi-definite). Type-1 method is stable whenever the corresponding Type-0 method is stable (by Lemma 2.3). In Table 2 are given the computed rates of convergence for the iteration for the IIPG. Observe that for $K = 1$ the iteration may diverge, as also the theory predicts. For larger values of K it can be seen that the iteration is convergent and the convergence rate improves as K is increased. Note also that for each fixed value of K, the rate of convergence of the proposed iterative method is uniformly bounded with respect to the number of degrees of freedom n_J, as predicted by Theorem 5.1. From these results we can see that the technical hypothesis $K \geq 4$ in Theorem 5.1 is a mild restriction. The corresponding convergence rates for the iterative method for the NIPG are given in Table 3. Notice that for $K = 1, 2$ the iteration diverge, in agreement with Theorem 5.2. As for the IIPG, the rate of convergence improves as K is increased. From the results in this table we also observe that for each fixed K the

Table 2 Convergence rate of Algorithm 5.3 for the Incomplete Interior Penalty method (IIPG) for different levels and different values of the Penalty parameter $\alpha = K\alpha^*$

K	J					
	0	1	2	3	4	5
1	2.42	1.37	1.56	1.79	1.34	1.38
2	0.84	0.68	0.73	0.78	0.68	0.69
4	0.50	0.43	0.45	0.47	0.43	0.44
8	0.33	0.29	0.30	0.32	0.29	0.30
16	0.22	0.20	0.21	0.21	0.20	0.21
32	0.15	0.14	0.14	0.15	0.14	0.14

Table 3 Convergence rate of Algorithm 5.3 for the Non-symmetric Interior Penalty method (NIPG) for different levels and different values of the Penalty parameter $\alpha = K\alpha^*$

K	J					
	0	1	2	3	4	5
1	1.75	1.52	1.60	1.68	1.57	1.60
2	1.23	1.08	1.13	1.19	1.11	1.13
4	0.87	0.76	0.80	0.85	0.79	0.80
8	0.61	0.54	0.57	0.60	0.56	0.57
16	0.43	0.38	0.40	0.42	0.39	0.40
32	0.30	0.27	0.28	0.30	0.28	0.28

Table 4 Values of γ^* for Type-0 IIPG scheme

J	0	1	2	3	4	5
α^*	0.90	1.30	1.30	1.30	1.60	1.60
γ^*	0.95	0.73	0.79	0.84	0.70	0.71
$\frac{\gamma^*}{\sqrt{4-\gamma^*}}$	0.91	0.58	0.65	0.73	0.54	0.55

rate of convergence is uniform with respect the level of refinement. Furthermore, notice that as predicted by our theory, the iteration for NIPG requires for convergence, a higher value of the penalty parameter than that required by the iterative method for the IIPG. We also note that the estimate in Theorem 5.2 provides an upper bound on the convergence rate, although $K \geq 16$, which is needed by our theory, may be somewhat restrictive for the numerical examples that we consider here.

Finally, we have computed numerically the constant γ^* in the strengthened CBS inequality (A.1) from Lemma A.1, together with our estimates of the convergence rates for the iteration of the IIPG and NIPG methods. The results are displayed in Table 4. In the second row we present for each level of refinement the values of α^* taken in our computations. For each level we have used these values of the penalty parameter to compute numerically the constant γ^* from Lemma A.1. In the third row we list the values of γ^*. In the fourth row of the table, the numerical values of the estimate of the convergence rate for the iterative method for the IIPG (see Theorem 5.1), $\Lambda_0 \leq \frac{\gamma^*}{\sqrt{K-\gamma^*}}$ for $K = 4$. Comparing these values with the computed convergence rates for IIPG from Table 2, for $K = 4$, verifies that the estimate given in Theorem 5.1 is an appropriate upper bound. Notice that for the iterative method for NIPG with $K = 16$, the same upper bound applies (since $\frac{2\gamma^*}{\sqrt{16-2\gamma^*}} = \frac{\gamma^*}{\sqrt{4-\gamma^*}}$).

Acknowledgements The authors are thankful to the referees for the useful comments and suggestions which helped to improve the presentation, and also for pointing to us the related article [19]. Part of this work was completed while the authors were visiting the Istituto di Matematica Applicata e Tecnologie Informatiche (IMATI) of the CNR in Pavia, Italy in 2007. Thanks go to the IMATI for the kind hospitality and support. B. Ayuso also thanks the Department of Mathematics at Penn-State University for the hospitality and support. The first author has also been partially supported by MEC under project MTM2008-03541 and by CAM under project S0505/ESP-0158. The work of the second author is supported in part by the National Science Foundation, DMS-058110, DMS-0810982, and by the Lawrence Livermore National Laboratory through subcontract #B574151.

Appendix: Proof of Lemma A.1

We prove here the strengthened Cauchy-Bunyakowskii-Schwarz (CBS) inequalities for the skew-symmetric part of the non-symmetric Type-0 IIPG method, used in the convergence analysis of the iterative methods for Type-1 non-symmetric IP methods. For the sake of completeness we state also in Lemma A.2 the analogous result for NIPG, although it has not been used in our analysis

Lemma A.1 *Let $\mathcal{A}_0^{I,\alpha^*}(\cdot,\cdot)$ be the Type-0 non-symmetric IIPG method with penalty parameter α^* and let $\widetilde{\mathcal{A}}_0^{I,\alpha^*}(\cdot,\cdot)$ be its symmetric part. Let α^* be the minimal value of the penalty parameter for which $\mathcal{A}_0^{I,\alpha^*}(\cdot,\cdot)$ is coercive. Then, the following CBS inequality holds for all $v \in V^{CR}$ and $z \in \mathcal{Z}$:*

$$\exists \gamma^* < 1 : \left[\frac{1}{2}(\nabla z, \nabla v)\right]^2 \leq [\gamma^*]^2 \widetilde{\mathcal{A}}_0^{I,\alpha^*}(z,z)\widetilde{\mathcal{A}}_0^{I,\alpha^*}(v,v). \tag{A.1}$$

Proof Let $\mathcal{A}_0^{s,\beta}(\cdot,\cdot) = \widetilde{\mathcal{A}}_0^{s,\beta}(\cdot,\cdot)$ be the bilinear form of the symmetric Type-0 SIPG method, with penalty parameter β chosen so that $\mathcal{A}_0^{s,\beta}(\cdot,\cdot)$ is coercive. Recall that the continuity and coercivity properties of the Type-0 SIPG and IIPG methods imply that there exists two positive constants c_1, c_2 possibly depending on the penalty parameters α^* and β such that

$$c_1(\alpha^*,\beta)\mathcal{A}_0^{s,\beta}(u,u) \leq \mathcal{A}_0^{I,\alpha^*}(u,u) \leq c_2(\alpha^*,\beta)\mathcal{A}_0^{s,\beta}(u,u), \quad \forall u \in V^{DG}. \tag{A.2}$$

It is easy to observe that we can always choose β so that the quotient $\frac{c_2(\alpha^*,\beta)}{c_1(\alpha^*,\beta)} > 1$.

Let now $P_z : V^{DG} \longrightarrow \mathcal{Z}$ be the orthogonal projection into $\mathcal{Z}$ be defined by:

$$P_z : V^{DG} \longrightarrow \mathcal{Z}, \quad \mathcal{A}_0^{s,\beta}(P_z(u),\psi) = \mathcal{A}_0^{s,\beta}(u,\psi) \quad \forall \psi \in \mathcal{Z}. \tag{A.3}$$

Note that by definition and the Cauchy-Schwarz inequality it follows that

$$\begin{aligned}\mathcal{A}_0^{s,\beta}(P_z(u),P_z(u)) &= \mathcal{A}_0^{s,\beta}(u,P_z(u)) \\ &\leq [\mathcal{A}_0^{s,\beta}(u,u)]^{1/2}[\mathcal{A}_0^{s,\beta}(P_z(u),P_z(u))]^{1/2}, \quad \forall u \in V^{DG},\end{aligned}$$

and therefore,

$$\mathcal{A}_0^{s,\beta}(P_z(u),P_z(u)) \leq \mathcal{A}_0^{s,\beta}(u,u) \quad \forall u \in V^{DG}. \tag{A.4}$$

Then, by taking into account the equivalence in (A.2), we define

$$\mathcal{K}_0 := \sup_{u \in V^{DG}} \frac{\widetilde{\mathcal{A}}_0^{I,\alpha^*}(P_z(u),P_z(u))}{\widetilde{\mathcal{A}}_0^{I,\alpha^*}(u,u)} \leq \frac{c_2}{c_1}\frac{\mathcal{A}_0^{s,\beta}(P_z(u),P_z(u))}{\mathcal{A}_0^{s,\beta}(u,u)} \leq \frac{c_2}{c_1}.$$

Next, let $z \in \mathcal{Z}$ and $v \in V^{CR}$ be fixed and observe that by definition $P_z(v) = 0\ \forall\, v \in V^{CR}$. Following [26] we define the quadratic polynomial

$$q_0(t) := \widetilde{\mathcal{A}}_0^{I,\alpha^*}(z + tv, z + tv) - \frac{1}{\mathcal{K}_0}\widetilde{\mathcal{A}}_0^{I,\alpha^*}(P_z(z + tv), P_z(z + tv)) \geq 0 \quad \forall t \geq 0,$$

where the non-negativeness of $q_0(t)$ follows from the definition of $\mathcal{K}_0$. Note that

$$\begin{aligned} q_0(t) &= \widetilde{\mathcal{A}}_0^{I,\alpha^*}(z + tv, z + tv) - \frac{1}{\mathcal{K}_0}\widetilde{\mathcal{A}}_0^{I,\alpha^*}(z, z) \\ &= \left(1 - \frac{1}{\mathcal{K}_0}\right)\widetilde{\mathcal{A}}_0^{I,\alpha^*}(z, z) + t^2\widetilde{\mathcal{A}}_0^{I,\alpha^*}(v, v) + t[\widetilde{\mathcal{A}}_0^{I,\alpha^*}(z, v) + \widetilde{\mathcal{A}}_0^{I,\alpha^*}(v, z)] \geq 0 \quad \forall t \geq 0, \end{aligned}$$

implies that the discriminant of this quadratic polynomial is non positive

$$\left[\frac{\widetilde{\mathcal{A}}_0^{I,\alpha^*}(z, v) + \widetilde{\mathcal{A}}_0^{I,\alpha^*}(v, z)}{2}\right]^2 - \left[1 - \frac{1}{\mathcal{K}_0}\right]\widetilde{\mathcal{A}}_0^{I,\alpha^*}(z, z) \cdot \widetilde{\mathcal{A}}_0^{I,\alpha^*}(v, v) \leq 0,$$

and therefore

$$\left[\frac{\widetilde{\mathcal{A}}_0^{I,\alpha^*}(z, v) + \widetilde{\mathcal{A}}_0^{I,\alpha^*}(v, z)}{2}\right]^2 \leq [\gamma^*]^2\widetilde{\mathcal{A}}_0^{I,\alpha^*}(z, z)\widetilde{\mathcal{A}}_0^{I,\alpha^*}(v, v),$$

where $[\gamma^*]^2 = [1 - \frac{1}{\mathcal{K}_0}] < 1$. Substituting the actual definition of the off-diagonal terms $\widetilde{\mathcal{A}}_0^{I,\alpha^*}(z, v)$ and $\widetilde{\mathcal{A}}_0^{I,\alpha^*}(z, v)$ we reach (A.1). □

Lemma A.2 *Let $\mathcal{A}_0^{n,\alpha^*}(\cdot, \cdot)$ be the Type-0 non-symmetric NIPG method with penalty parameter α^* and let $\widetilde{\mathcal{A}}_0^{n,\alpha^*}(\cdot, \cdot)$ be its symmetric part. Let α^* be any value of the penalty parameter for which $\mathcal{A}_0^{n,\alpha^*}(\cdot, \cdot)$ is coercive. Then,*

$$\exists \gamma_1^* < 1 : [(\nabla z, \nabla v)]^2 \leq [\gamma_1^*]^2\widetilde{\mathcal{A}}_0^{n,\alpha^*}(z, z)\widetilde{\mathcal{A}}_0^{n,\alpha^*}(v, v). \tag{A.5}$$

Proof The proof of this lemma follows the same lines as the previous proof just by replacing $\widetilde{\mathcal{A}}_0^{I,\alpha^*}(\cdot, \cdot)$ with $\widetilde{\mathcal{A}}_0^{n,\alpha^*}(\cdot, \cdot)$. We omit the details. □

References

1. Adams, R.A.: Sobolev spaces. Pure and Applied Mathematics, vol. 65. Academic Press [Harcourt Brace Jovanovich], San Diego (1975)
2. Agmon, S.: Lectures on elliptic boundary value problems. Van Nostrand Mathematical Studies, vol. 2. Van Nostrand, Princeton (1965). Prepared for publication by B. Frank Jones, Jr. with the assistance of George W. Batten, Jr.
3. Antonietti, P.F., Ayuso, B.: Schwarz domain decomposition preconditioners for discontinuous Galerkin approximations of elliptic problems: non-overlapping case. Math. Model. Numer. Anal. **41**(1), 21–54 (2007)
4. Antonietti, P.F., Ayuso, B.: Multiplicative Schwarz methods for discontinuous Galerkin approximations of elliptic problems. Math. Model. Numer. Anal. **42**(3), 443–469 (2008)
5. Antonietti, P.F., Ayuso, B.: Two-level Schwarz preconditioners for super penalty discontinuous Galerkin methods. Commun. Comput. Phys. **5**, 398–412 (2009)

6. Arnold, D.N.: An interior penalty finite element method with discontinuous elements. SIAM J. Numer. Anal. **19**(4), 742–760 (1982)
7. Arnold, D.N., Brezzi, F., Cockburn, B., Marini, L.D.: Unified analysis of discontinuous Galerkin methods for elliptic problems. SIAM J. Numer. Anal. **39**(5), 1749–1779 (2001/02) (electronic)
8. Ayuso, B., Georgiev, I., Kraus, J., Zikatanov, L.: Preconditioning techniques for discontinuous Galerkin methods discretizations for linear elasticity equations (2009, in preparation)
9. Ayuso, B., Marini, L.D.: Discontinuous Galerkin methods for advection-diffussion-reaction problems. SIAM J. Numer. Anal. **47**(2), 1391–1420 (2009)
10. Brenner, S.C.: Two-level additive Schwarz preconditioners for nonconforming finite element methods. Math. Comput. **65**(215), 897–921 (1996)
11. Brenner, S.C.: Convergence of nonconforming multigrid methods without full elliptic regularity. Math. Comput. **68**(225), 25–53 (1999)
12. Brenner, S.C.: Convergence of nonconforming V-cycle and F-cycle multigrid algorithms for second order elliptic boundary value problems. Math. Comput. **73**(247), 1041–1066 (2004) (electronic)
13. Brenner, S.C., Owens, L.: A weakly over-penalized non-symmetric interior penalty method. JNAIAM J. Numer. Anal. Ind. Appl. Math. **2**(1–2), 35–48 (2007)
14. Brenner, S.C., Owens, L., Sung, L.-Y.: A weakly over-penalized symmetric interior penalty method. Electron. Trans. Numer. Anal. **30**, 107–127 (2008)
15. Brenner, S.C., Ridgway Scott, L.: The Mathematical Theory of Finite Element Methods. Texts in Applied Mathematics, vol. 15. Springer, New York (1994)
16. Brenner, S.C., Zhao, J.: Convergence of multigrid algorithms for interior penalty methods. Appl. Numer. Anal. Comput. Math. **2**(1), 3–18 (2005)
17. Brezzi, F., Cockburn, B., Marini, L.D., Süli, E.: Stabilization mechanisms in discontinuous Galerkin finite element methods. Comput. Methods Appl. Mech. Eng. **195**(25–28), 3293–3310 (2006)
18. Brezzi, F., Fortin, M.: Mixed and Hybrid Finite Element Methods. Springer Series in Computational Mathematics, vol. 15. Springer, New York (1991)
19. Burman, E., Stamm, B.: Low order discontinuous Galerkin methods for second order elliptic problems. SIAM J. Numer. Anal. **47**(1), 508–533 (2008)
20. Ciarlet, P.G.: Basic error estimates for elliptic problems. In: Handbook of Numerical Analysis, vol. II, pp. 17–351. North-Holland, Amsterdam (1991)
21. Crouzeix, M., Raviart, P.-A.: Conforming and nonconforming finite element methods for solving the stationary Stokes equations. I. Rev. Fr. Autom. Inform. Rech. Opér. Sér. Rouge **7**(R-3), 33–75 (1973)
22. Dawson, C., Sun, S., Wheeler, M.F.: Compatible algorithms for coupled flow and transport. Comput. Methods Appl. Mech. Eng. **193**(23–26), 2565–2580 (2004)
23. Dobrev, V.A., Lazarov, R.D., Vassilevski, P.S., Zikatanov, L.T.: Two-level preconditioning of discontinuous Galerkin approximations of second-order elliptic equations. Numer. Linear Algebra Appl. **13**(9), 753–770 (2006)
24. Douglas Jr., J., Dupont, T.: Interior penalty procedures for elliptic and parabolic Galerkin methods. In: Computing Methods in Applied Sciences, Second Internat. Sympos., Versailles, 1975. Lecture Notes in Phys., vol. 58, pp. 207–216. Springer, Berlin (1976)
25. Eisenstat, S.C., Elman, H.C., Schultz, M.H.: Variational iterative methods for nonsymmetric systems of linear equations. SIAM J. Numer. Anal. **20**(2), 345–357 (1983)
26. Falgout, R.D., Vassilevski, P.S., Zikatanov, L.T.: On two-grid convergence estimates. Numer. Linear Algebra Appl. **12**(5–6), 471–494 (2005)
27. Feng, X., Karakashian, O.A.: Two-level additive Schwarz methods for a discontinuous Galerkin approximation of second order elliptic problems. SIAM J. Numer. Anal. **39**(4), 1343–1365 (2001) (electronic)
28. Gopalakrishnan, J., Kanschat, G.: A multilevel discontinuous Galerkin method. Numer. Math. **95**(3), 527–550 (2003)
29. Hestenes, M.R., Stiefel, E.: Methods of conjugate gradients for solving linear systems. J. Res. Natl. Bur. Stand. **49**, 409–436 (1953)
30. Kraus, J., Margenov, S., Synka, J.: On the multilevel preconditioning of Crouzeix-Raviart elliptic problems. Numer. Linear Algebra Appl. **15**(5), 395–416 (2008)
31. Luenberger, D.G.: Linear and Nonlinear Programming, 2nd edn. Kluwer Academic, Boston (2003)
32. Rivière, B., Wheeler, M.F., Girault, V.: Improved energy estimates for interior penalty, constrained and discontinuous Galerkin methods for elliptic problems. I. Comput. Geosci. **3**(3–4), 337–360 (2000) 1999
33. Rivière, B., Wheeler, M.F., Girault, V.: A priori error estimates for finite element methods based on discontinuous approximation spaces for elliptic problems. SIAM J. Numer. Anal. **39**(3), 902–931 (2001) (electronic)
34. Saad, Y.: Iterative Methods for Sparse Linear Systems, 2nd edn. Society for Industrial and Applied Mathematics, Philadelphia (2003)

35. Strang, G.: Variational crimes in the finite element method. In: The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations, Proc. Sympos. Univ. Maryland, Baltimore, Md., 1972, pp. 689–710. Academic Press, San Diego (1972)
36. Strang, G., Fix, G.J.: An Analysis of the Finite Element Method. Prentice-Hall Series in Automatic Computation. Prentice-Hall, Englewood Cliffs (1973)
37. Sun, S., Wheeler, M.F.: Symmetric and nonsymmetric discontinuous Galerkin methods for reactive transport in porous media. SIAM J. Numer. Anal. **43**(1), 195–219 (2005) (electronic)
38. Wheeler, M.F.: An elliptic collocation-finite element method with interior penalties. SIAM J. Numer. Anal. **15**(1), 152–161 (1978)
39. Xu, J.: Iterative methods by space decomposition and subspace correction. SIAM Rev. **34**(4), 581–613 (1992)
40. Xu, J., Zikatanov, L.: The method of alternating projections and the method of subspace corrections in Hilbert space. J. Am. Math. Soc. **15**(3), 573–597 (2002) (electronic)